UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

Merck Sharp & Dohme LLC,
Petitioner,

v.

Halozyme Inc.,
Patent Owner.

Case No. PGR2025-00030

U.S. Patent No. 12,054,758

Declaration of James Naismith, Ph.D.

Table of Contents

I. Introductory Matters....1

A. Background and Qualifications....1

B. Compensation....4

C. Person of Ordinary Skill in the Art....4

D. Terminology Used in this Declaration: "Common Disclosure"....6

E. Materials Considered....9

F. Legal Principles....12

II. Response to Drs. Petsko's and Simpson's Opinions that the Claims Define a Narrow and Structurally Homogeneous Set of Modified PH20 Polypeptides....16

A. The Sequence Identity Claim Parameters Define an Immense and Heterogeneous Set of Amino Acid Sequences....17

1. The Sequence Identity Claim Parameters Define Variants of the Human PH20 Amino Acid Sequence....20
2. The Sequence Identity Claim Parameters Do Not Restrict the Location or Nature of Changes....21
3. The Sequence Identity Claim Parameters Define an Immense Sequence Space....22
4. Drs. Petsko and Simpson Have Differing Opinions About What Set of Polypeptides is "Structurally Homogeneous"....26

B. The Amino Acid Sequences in the Claim Parameter Sequence Space Are Structurally Heterogeneous, Not Homogeneous....31

1. Amino Acid Sequences in the Claim Parameter Sequence Space Will Adopt Many Different Three-Dimensional Structures....32
2. Mutations Increase Obstacles to an Amino Acid Sequence Successfully Folding to Adopt the Native State of a Protein....43
3. Illustration of the Structural Diversity of the Claim Parameter Sequence Space....46

C. Sequence Identity Percentages Are Not a Structural Feature of a Modified PH20 Polypeptide....50

D. The "PH20" Requirement in the Claims Requires the Claimed Modified PH20 Polypeptides to be Folded Proteins 54

1. Modified PH20 Polypeptides are Proteins.....54
2. Defined Terms Used in the Claims ("PH20" and "Modified PH20 Polypeptide") 56
3. POSAs Would Understand the Term "PH20" Requires Modified PH20 Polypeptides to Be Folded Proteins.....59
4. Dr. Simpson's Interpretation of "Modified PH20 Polypeptide" Effectively Removes the Term "PH20" from the Claims 67

E. The Claims as a Whole Define Enzymatically Active Modified PH20 Polypeptides.....69

III. Response to Dr. Petsko's and Dr. Simpson's Opinions on the Common Disclosure's Description of the Claimed Modified PH20 Polypeptides .76

A. It Was Impossible for POSAs to Determine from Information in the Common Disclosure Which Modified PH20 Polypeptides in the Claim Parameter Sequence Space Will Fold and/or Be Enzymatically Active, or Soluble or Able to Increase Delivery of a Therapeutic 77

1. Dr. Petsko's 7 Examples Are Not Representative of the Diversity of the Claimed Modified PH20 Polypeptides.....78
2. Dr. Petsko Confirms the Common Disclosure Does Not Identify the Enzymatically Active Multiply-Modified PH20 Polypeptides Being Claimed.....86

B. "PH20 Alpha-Beta Barrel" Does Not Describe Which Amino Acid Sequences in the Claim Parameter Sequence Space Will Produce It 99

1. Alpha-Beta Barrel Motifs Are Found in Many Enzymes.....101
2. Publications Prior to 2011 Identified the Alpha-Beta Barrel Motif in Hyaluronidases as Having an Unusual Structure 103
3. Dr. Petsko's Description of the PH20 Alpha-Beta Barrel Fold Varies and Is Not an Objective or Precise Description of a Common Structure of the Claimed PH20 Polypeptides 114

4. The Phrase "PH20 Alpha-Beta Barrel Fold" Does Not Identify the Structural Features Shared by Enzymatically Active Forms of PH20 119

5. The Common Disclosure Does Not Portray the Alpha-Beta Barrel Fold as a Common Structural Feature of Modified PH20 Polypeptides and Does Not Describe Representative Examples of Multiply-Modified PH20s 122

C. Multiple Mutations to PH20 Are No More or Less Predictable than in Other Proteins 127

D. Hyaluronidases Share a "PH20 Alpha-Beta Barrel" Motif But Exhibit Distinct Functional Properties.....134

E. The Hyal-EGF Domain Was Known to Be a Key Part of the PH20 Structure But Is Not Discussed in the Common Disclosure.....136

F. Drs. Petsko and Simpson Rely on Information Not Referenced or Contained in the Common Disclosure to Support their Opinions on What the Common Disclosure Would Have Conveyed to a POSA 143

G. Dr. Petsko's and Dr. Simpson's Opinions Contradict the Guidance in the Common Disclosure 145

1. Dr. Petsko's Opinion that POSAs Would Not Use Directed Evolution Techniques to Produce Multiply Modified PH20 Polypeptides Contradicts the Common Disclosure 146

2. The Common Disclosure Does Not Describe Using Tolerance Ratings to Design Modified PH20 Polypeptides 149

3. Dr. Simpson Contradicts the Common Disclosure's Instructions to Not Modify Certain Non-Tolerated Positions in Enzymatically Active Mutants.....150

4. Dr. Petsko's Rescue-Mutation Theory is Not Supported by the Common Disclosure or the Literature 152

5. Dr. Simpson Contradicts the Data and Conclusions in the '457 Publication's Data and the Common Disclosure Regarding PH20 Truncation Mutants.....153

6. Dr. Petsko's Theories About Substrate Binding Sites Is Contrary to the Guidance in the Common Disclosure.....163

IV. Response to Opinions of Drs. Petsko and Simpson Regarding the Ability of a POSA to Make and Use the Claimed Modified PH20 Polypeptides 169

A. POSAs Could Not Identify or Predict Which Multiply Mutated PH20 Polypeptides in the Claimed Sequence Space Can Be Produced or Will Stably Fold.....172

B. Double-Counting and Other Errors in Dr. Petsko's Positional Tolerance Analysis 175

C. Dr. Petsko's Positional Tolerance Analysis Rests on Unreliable Assumptions and Data 179

D. POSAs Could Not Have Predicted Which Combinations of Substitutions Tolerated Individually Would Yield, Enzymatically Active, Soluble Modified PH20 Polypeptides 186

1. Epistasis Associated with Multiple Changes in a PH20 Polypeptide Is Unpredictable.....186
2. The Data in the Common Disclosure and a PH20 Homology Model Cannot Be Used to Predict Which Sequences in the Claim Parameter Sequence Space Will Fold or Exhibit Hyaluronidase Activity, or Solubility or Increased Delivery of a Therapeutic Agent.....194
3. The Active Mutant Experimental Evidence In the Common Disclosure Suggests Epistasis is Likely to Occur in Modified PH20 Polypeptides.....196
4. Many Effects that Influence Folding, Activity and Stability of a Protein are Imperceptible in a Homology Model 201
5. Dr. Petsko's Positional Tolerance Table Does Not Enable Prediction of Multiply Modified PH20 Polypeptides.....205

E. Homology Models Have Limits in Predicting Effects of Mutations 207

F. Dr. Petsko's Positional Tolerance Table Is Unreliable and Cannot Be Used to Predict the Effects of Multiple Changes on a PH20 Polypeptide's Ability to Fold and/or Exhibit Activity 214

1. Drs. Petsko and Simpson Relied on the Common Disclosure's Active/Inactive Mutants Classifications to Form Their Opinions on Positional Tolerance in PH20 216

2. Only Mutants with Activity Above 40% Are Reliable and Provide Insights into Folded PH20 Proteins.....218

G. Dr. Petsko's Positional Tolerance Analysis Overstates What the Data Support.....224

1. Positions Identified in the Prior Art as Essential to Folding, Secretion or Activity.....226

2. Positions that Retain Activity Despite Meaningful Interrogation 230

H. The Tolerance Table Does Not Enable Prediction of Multiply Modified PH20 Variants.....240

I. Dr. Petsko Does Not Account for Protein Aggregation in Multiply Modified PH20 Polypeptides.....247

V. Dr. Park Performed a Rigorous, Unbiased and Reliable Assessment of Single-Amino Acid Substitutions in PH20 Polypeptides as a POSA Would Have in 2011 252

Appendices

A. Materials Considered

B. Tables

B-1: Counts and Classifications of $PH20_{1-447}$ Mutants

B-2: $PH20_{1-447}$ Physiochemical and Activity Data

B-3: Count Comparison $PH20_{1-447}$ Residues

B-4: Solvent Accessible Surface Area Data of PH20

B-5: Histogram of $PH20_{1-447}$ Activity Data

B-6: Activity Data Plot of C-terminal Truncation PH20 Mutants from US-457 and the '429 Patent

B-7: Comparison Table of Dr. Simpson's Activity Classification of C-terminal Truncation PH20 Mutants vs. US-457

B-8: Frequency Chart of Alternative Residues Identified in MSA and Number of MSA Alternatives Tested

B-9: PH20 Positional Alignment: Tested MSA Alternatives

B-10: PH20 Positional Alignment: Domains, Motifs, and Prior Art

B-11: PH20 Residue Characterization

B-12: Compilation of citations on Conventional Tools and Techniques from the Experts

C. C.V. of Dr. James H. Naismith, Ph.D.

I. Introductory Matters

A. Background and Qualifications

1. My educational background, career history, and other relevant qualifications are summarized below. I attach to this Declaration my curriculum vitae (Appendix C), which provides a full and accurate description of my educational background, professional experience, appointments, honors, and publications.

2. I received my Bachelor of Science degree with First Class Honours in Chemistry from the University of Edinburgh in 1989. I received my Ph.D. in Chemistry in 1992 from the University of Manchester for research involving the structural study of concanavalin A and a zinc-dependent aldolase, using X-ray crystallography and enzymological analysis, for which I was awarded the Hibbert Prize. I completed two years of postdoctoral training as a research fellow at the Howard Hughes Medical Institute University of Texas Southwestern Medical Center Dallas, where I continued work in structural biology.

3. I started my own laboratory in structural biology and enzymology at the University of St Andrews in 1995. I earned a Doctor of Science (D.Sc.) degree from the University of St. Andrews in 2016 for sustained contributions to chemical and structural studies of proteins.

4. I currently serve as Professor of Structural Biology in the Nuffield Department of Medicine at the University of Oxford and as Head of the Division of Mathematical, Physical and Life Sciences at Oxford. I am also a member of the University Council of the University of Oxford. From 2017 to 2023, I led the establishment of the Rosalind Franklin Institute as its founding Director and Chief Executive Officer, where I was responsible for scientific vision, faculty recruitment, infrastructure development, and delivery of the Institute's research program. I currently serve as the non-clinical Vice-President of the Academy of Medical Sciences.

5. Prior to my appointment at Oxford, I served as a faculty member at the University of St. Andrews from 1995 to 2017, progressing from Lecturer to Professor and later Bishop Wardlaw Professor of Chemical Biology. During that period, I led the Biomedical Sciences Research Complex, including its design, funding, construction, and scientific direction, and I held senior leadership roles within the University.

6. Throughout my career, I have taught and supervised courses in enzyme kinetics, structural biology, chemistry in biology, and drug discovery, and I have taught undergraduate and graduate-level biochemistry and chemistry for more than two decades. I have supervised over thirty Ph.D. students who have gone on to graduate, and over twenty post-doctoral associates.

7. I have over 30 years of research experience in protein structure and enzymology. My laboratory has investigated the structural and chemical biology of enzymes and molecular recognition complexes, including detailed studies of enzyme mechanisms, substrate specificity, protein–ligand interactions, and the structural basis of catalysis. My work has addressed diverse biological systems, including natural product biosynthesis, carbohydrate-processing enzymes, and mechanisms relevant to antimicrobial resistance.

8. My research contributions have been recognized by election to several scientific academies, including Fellow of the Royal Society of Edinburgh (2005), Fellow of the Academy of Medical Sciences (2012), Fellow of the Royal Society (2014), Fellow of the American Association for the Advancement of Science (2016), Membership of European Molecular Biology Organization (2009), and Membership of Academia Europa (2018). I am a Fellow of the Royal Society of Chemistry, Royal Society of Biology, Institute of Physics, and a Member of the American Chemical Society.

9. I have received numerous awards for research excellence, including the Royal Society Wolfson Research Merit Award (2014), the Tilden Prize (2018), and the Horizon Prize (2022) from the Royal Society of Chemistry. I am the only individual to have received both the Corday-Morgan Medal (2004) for

contributions to experimental chemistry and the Colworth Medal (2004) for research in biochemistry, reflecting the interdisciplinary nature of my work.

10. I have co-authored just under 300 peer-reviewed scientific publications in the areas of protein structure, enzymology, chemical biology, and membrane proteins. A complete list of my publications is included in Appendix C. Google Scholar shows the work has over 28,000 citations and an h-factor of 90.

B. Compensation

11. I am being compensated for my time at the rate of $850 per hour for my work in connection with this matter. I am being reimbursed for reasonable and customary expenses associated with my work in this investigation. This compensation is not dependent in any way on the contents of this Declaration, the substance of any further opinions or testimony that I may provide, or the ultimate outcome of this matter.

C. Person of Ordinary Skill in the Art

12. I understand that my analysis and opinions are to be provided using the perspective of a person of ordinary skill in the art (a "POSA"). I also understand that the timeframe that is relevant to the POSA's perspectives is the period from just prior to December 29, 2011 until December 28, 2012. I will refer to this as the "2011-2012 timeframe" in this declaration. If any of my opinions would have varied within this period, I will note that variance.

13. I agree with Dr. Hecht's description of the scientific field of the patent (*i.e.*, that it concerns proteins, and more particularly, protein structures and modification of them). *See* EX1003 (Hecht Dec.), ¶¶ 33–35. Like Dr. Hecht, I am very familiar with this field, and the individuals who work within it, and was so in the 2011-2012 timeframe.

14. I have been informed by counsel that a POSA is a hypothetical person who is presumed to have the typical skills and knowledge of someone working in the field of the invention. I agree with the description of the POSA provided by Dr. Hecht in his declaration. *See* EX1003, ¶¶ 11–14. As he explained:

> Based on my review of the patent and my experience, I believe a person of ordinary skill in the art (who I may refer to as "a skilled artisan") would have had an undergraduate degree, a Ph.D., and post-doctoral experience in scientific fields relevant to study of protein structure and function (*e.g.*, chemistry, biochemistry, biology, biophysics). From training and experience, the person would have been familiar with factors influencing protein structure, folding and activity, production of modified proteins using recombinant DNA techniques, and use of biological assays to characterize protein function, as well with techniques and tools used to analyze protein structure (*i.e.*, sequence searching and alignments, protein modeling software, etc.).

15. In the 2011-2012 timeframe, I had at least the qualifications of a POSA. The opinions I provide in this declaration are provided from the perspective of a POSA in the 2011-2012 timeframe as I have described it above.

D. Terminology Used in this Declaration: "Common Disclosure"

16. I will use the following abbreviations and terminology in this declaration.[1]

(a) "PH20" refers to the human PH20 protein. In its full-length form (including its 35 amino acid signal sequence) it has 509 amino acids. Its amino acid sequence was first published in 1993.[2] Its sequence also is reported as:

- UniprotKB Accession ID P38567;
- SEQ ID NO:1 in U.S. Patent No. 7,767,429; and
- SEQ ID NO:6 in U.S. Patent No. 12,054,758.

(b) The "mature" protein sequence of PH20 refers to the form of the protein without the signal sequence, while the "precursor" protein sequence includes the signal sequence. For example, the mature form of human PH20 consists of the amino acids from positions 36 to 509 of the precursor form of human PH20 (Uniprot ID: P38567), and

---

[1] Dr. Hecht also used many of these abbreviations. *See* EX1003 (Hecht Dec.), ¶ 15.

[2] EX1029 (Gmachl 1993), 546, Fig. 1.

would have numbering that starts at position 1 and ends at position 474.

(c) "$PH20_{1\text{-}n}$" refers to a human wild-type PH20 polypeptide sequence that starts at position 1 and terminates at position "n" of the PH20 sequence. Unless otherwise indicated, the sequence that is being referred to is the mature form of the protein without the signal sequence. For example, $PH20_{1\text{-}447}$ means the polypeptide starting at position 1 and ending at position 447 of the mature human wild-type PH20 sequence.

(d) "AxxxB" refers to an amino acid substitution at position xxx of the amino acid sequence, where the wild-type residue is A and the residue after the substitution is B.

17. Dr. Petsko uses the term "properly folded" to mean that a protein has adopted the tertiary structure present in the wild type PH20 protein. EX2070 (Petsko Dec.), ¶ 51 ("The POSA would have understood that properly folded PH20s adopt the same tertiary structure-an alpha-beta barrel fold."). The scientific literature more frequently uses the term "native state" or "native fold" to refer to the tertiary structure that exists in the folded form of a naturally occurring protein. I may use either terminology in my declaration. Also, I will use "α/β barrel" when I refer to the "alpha-beta barrel" motif.

18. I was asked to assess U.S. Patent 12,054,758, which I refer to as the '758 Patent (EX1001). I also reviewed a U.S. patent application having the number

13/694,731 filed on December 28, 2012, which I will refer to as the '731 Application (EX1026), and two provisional applications referenced by it.[3] I understand the '731 Application incorporates the contents of the two provisional applications, and adds information (*e.g.*, new exemplary modifications, testing for modified polypeptides with enzymatic activity or stability, examples 11-16).[4]

19. I understand that the disclosure of a patent consists of a narrative section called the specification, and may contain drawings, and sequence listings, which are nucleotide or amino acid sequences. I understand each sequence in the patent is assigned a unique number (*e.g.*, SEQ ID NO: 3). I understand that patents end with claims that define the invention.

20. I reviewed the '758 Patent and the '731 Application. I also reviewed a redline of the two (EX1045). The two disclosures appear to be substantively the same.[5] I understand Dr. Hecht and Dr. Simpson reached the same conclusion.[6] I will refer to that as the "common disclosure." I also used citations to the

---

[3] EX1051 (US 61/631,313); EX1052 (US 61/796,208).

[4] EX1026 ('731 App.), 19:25-26, 28, 140:15-150:26, 311-321; EX1051 (US 61/631,313); EX1052 (US 61/796,208).

[5] *See* EX1003 (Hecht Dec.), ¶ 18.

[6] EX1003 (Hecht Dec.), ¶ 18; EX2068 (Simpson Dec.), ¶ 6.

specification of the ’731 Application. For convenience, a table correlating each citation in the ’731 Application to its corresponding citation in the ’758 patent is provided in EX1154.

E. Materials Considered

21. My opinions are based on my years of education, research, and experience, as well as my investigation and study of relevant materials. I reviewed a number of publications in the course of my assessment, including those listed in the Exhibit List. I also relied on my extensive familiarity with the scientific literature in this field.

22. I reviewed a declaration by Dr. Michael Hecht (EX1003). Dr. Hecht’s declaration provides an accurate description of knowledge and perspectives that would have been held by a POSA in the 2011-2012 timeframe. For example, Dr. Hecht provides an accurate summary of relevant scientific principles, including protein structure principles, protein engineering techniques, factors that influence protein stability, how stability is evaluated experimentally, and information about hyaluronidases. Unless otherwise noted below, I agree with Dr. Hecht’s opinions as they are described in EX1003.

23. I reviewed a declaration by Dr. Sheldon Park (EX1004). Based on my review, Dr. Park’s declaration provides an accurate description of how a POSA in the 2011-2012 timeframe would have approached the task of identifying single

amino acid substitutions in non-essential regions of PH20 that would have been expected to be tolerated by the protein. His analysis is the type of analysis that would have been produced by POSAs when attempting to predict the effects of single mutations in a protein. Dr. Park's declaration also provides an accurate description of the techniques and tools used in the 2011-2012 timeframe to find structurally similar proteins, analyze them using sequence alignment tools, identify conserved and evolutionarily varying positions in the related proteins, produce protein structure models and use these models to assess the tolerance of the protein to individual amino acid substitutions.

24. I reviewed Dr. Park's discussion of the tools he used in his analysis, such as BLAST, Clustal Omega, SWISS-MODEL and PyMol.[7] I am familiar with these tools, and his description of these tools and their capabilities in the 2011-2012 timeframe is accurate. I agree with Dr. Park that while the tools in their modern implementations have different forms and additional capabilities relative to the versions of the tools that existed in the 2011-2012 timeframe, the analyses Dr. Park performed using them relied on features that would have been present in the versions that existed in the 2011-2012 timeframe. For example, sequence alignments performed by the Clustal Omega tool rely on algorithms that were

[7] *See* EX1004 (Park Dec.), ¶¶ 20–24, 139–159.

largely developed in the 1990s and produce outputs that are typically evaluated manually by the user to confirm the accuracy of the alignments.

25. Like Dr. Park, I would have expected that homology models produced by the SWISS-MODEL tool today for PH20 would be very similar to the structural models for PH20 that would have been produced by that tool in the 2011-2012 timeframe.[8] As Dr. Park points out, the template that SWISS-MODEL used to produce the PH20 models was the Hyal-1 structure published by Chao et al. in 2007.[9] (structure ID: 2PE4). I also understand that Dr. Petsko produced and used a PH20 homology model produced in the same manner as Dr. Park's PH20 homology model.[10]

26. I note that many of Dr. Petsko's opinions are consistent with the scientific principles explained by Drs. Hecht and Park in their declarations. I also am familiar with Dr. Petsko's textbook (EX2164), which covers many of these scientific principles in a manner consistent with the descriptions provided by Drs. Hecht and Park.

---

[8] EX1004 (Park Dec.), ¶¶ 139; 151–159.

[9] EX1004 (Park Dec.), ¶ 141; EX1006 (Chao).

[10] *See* EX2070 (Petsko Dec.), ¶ 121.

F. Legal Principles

27. I have been provided a general explanation of some of the legal requirements for obtaining a patent. In addition to these explanations, I have been given guidance about how one evaluates what a patent claim defines and covers, which I address in later sections of this declaration.

28. I have been informed that a patent's disclosure consists of the information in the specification, along with any drawings and sequence listings that accompanied the patent application. When I use the term "patent disclosure," I am using it with this understanding.

29. I have been informed that a patent claiming a set (or "genus") of chemical compounds (*e.g.*, polypeptides) must be supported by the patent disclosure, and that there are two distinct disclosure requirements, as summarized in the following paragraphs.

30. I have been informed that one of the disclosure requirements is that the patent disclosure provide a sufficient written description of the claimed set of polypeptides ("written description" requirement). I have been informed that this can be achieved by a patent disclosure that describes either (i) "a representative number of species falling within the scope of the genus" or (ii) "structural features common to the members of the genus," with either being such that "one of skill in the art can 'visualize or recognize' the members of the genus." I have been

informed that an adequate written description of a genus of polypeptides "requires a precise definition, such as by structure, formula, chemical name, physical properties, or other properties, of species falling within the genus sufficient to distinguish the genus from other materials." I also have been informed that a description that is "merely drawing a fence around the outer limits of a purported genus is not an adequate substitute for describing a variety of materials constituting the genus and showing that one has invented a genus and not just a species."

31. I have been informed a second disclosure requirement is that the patent disclosure provide a description that enables a POSA to practice the claimed invention without engaging in undue experimentation ("enablement requirement").

32. I have been informed that if a patent claims an entire class of compounds (*e.g.*, polypeptides), the patent's specification must enable a person skilled in the art to make and use the entire class, which means the specification must enable the full scope of the invention as defined by its claims. I have also been informed that a research plan that requires a POSA to perform trial and error procedures in order to discover which members of a large class of polypeptides have a desired functional property is not sufficient to satisfy the enablement requirement.

33. I have also been informed that factors such as the breadth of the claims, unpredictability in the field, the degree of guidance in the prior art and in

the patent disclosure, and the level of skill of the POSA are factors, among others, that are considered in assessing the question of enablement.

34. I have been informed that a separate patentability requirement is that an invention must not have been obvious to a POSA in view of what was known in the prior art before the filling date of the patent. I also have been informed that if a patent claim encompasses a compound that would have been obvious in light of the prior art, that claim is unpatentable.

35. I have been informed that for a claimed compound to be found obvious, a POSA must have found a reason in the prior art to make that compound, and must have had a reasonable expectation of success in achieving the claimed invention. I have been informed this does not require the POSA to have absolute certainty about achieving a desired result, and that an invention can be found obvious if a result is expected but still requires some experimentation to confirm.

36. I have been informed that if a particular compound exhibits unexpected properties, that evidence can support a finding that the compound is not obvious. For a claim defining a large class of compounds, the members of the class must share the unexpected property to support a finding that the class of compounds is not obvious. A claim defining a large class of compounds cannot benefit from evidence showing only one compound exhibits a particular

unexpected property. Instead, the evidence must demonstrate that the unexpected property is shared by the entire class of claimed compounds.

37. I have been informed that terms used in the claims should be given their ordinary meaning, unless the terms have been expressly defined to have a particular meaning by the patent disclosure.

38. I have been informed that a patent claim defining an invention is to be interpreted in a manner that is consistent with the description of the claimed invention in the patent disclosure (*i.e.*, the specification, drawings and sequence listings).

39. I have been informed that what a claim defines must be determined by evaluating the claim as a whole, including all of the language used in the claim. I understand that a claim may use combinations of terms or use terms in conjunction with additional language to yield an overall meaning for a claim that is narrower than what a single defined term might alone suggest.

II. Response to Drs. Petsko's and Simpson's Opinions that the Claims Define a Narrow and Structurally Homogeneous Set of Modified PH20 Polypeptides

40. Drs. Petsko and Simpson each portray the claims of the '758 Patent as covering a "narrow" and "structurally homogeneous" set of proteins.[11] But they also each portray the claim language as defining a set of polypeptides that includes any polypeptide having an amino acid sequence that (i) has an overall sequence identity of at least 91% to any one of a list of wild-type PH20 amino acid sequences (*i.e.*, SEQ ID NOS: 3, 7 or between 32 to 66), and (ii) contains a specific amino acid substitution at position 317 (*i.e.*, L to either A, I, K, M, Q or R).[12] These two perspectives cannot be reconciled. If the set of polypeptides being claimed includes any polypeptide with an amino acid sequence that meets only the above two claim requirements, it is both enormous and highly heterogeneous. Even if the set is restricted to enzymatically active modified PH20 polypeptides, a POSA would expect there to be variations in the structural features of that set of proteins, albeit in ways that permit substrate recognition and catalysis. Dr. Hecht illustrated

---

[11] EX2068 (Simpson Dec.), ¶¶ 33–34; EX2070 (Petsko Dec.), ¶ 58 ("[T]he disclosed examples are representative of this homogeneous set, which share common structure-function relationships.").

[12] EX2068 (Simpson Dec.), ¶ 27; EX2070 (Petsko Dec.), ¶ 49.

these points in his declaration.[13] But there is no guidance in the common disclosure or known in the 2011-2012 timeframe that a POSA could use to identify the sequences within this restricted set. The common disclosure also does not identify how many such sequences exist within this restricted set relative to the claimed sequence space.

A. The Sequence Identity Claim Parameters Define an Immense and Heterogeneous Set of Amino Acid Sequences

41. I reviewed the claims of the ’758 Patent. Claim 1 is the independent claim and is reproduced below. Claim 1 concerns “a modified PH20 polypeptide.” Other claims depend from claim 1 and restrict which reference sequence is to be used, the overall sequence identity percentage required, or which substitutions are permitted at position 317.[14]

---

[13] EX1003 (Hecht Dec.), ¶¶ 44-46, 55-61, 156-159.

[14] EX1001 (’758 Patent), 313:18-43, 313:44–316:44. Dr. Hecht addressed the claim parameters in more detail in his declaration. *See e.g.*, EX1003 (Hecht Dec.), ¶¶ 117, 136.

> **1**. A modified PH20 polypeptide, comprising one or more amino acid modifications in an unmodified PH20 polypeptide, wherein:
>
> the unmodified PH20 polypeptide consists of the amino acid sequence selected from the group consisting of SEQ ID NOs: 3, 7 and 32-66;
>
> amino acid modifications are selected from the group consisting of amino acid replacements(s), deletion(s), and/or insertion(s);
>
> the modified PH20 polypeptide comprises an amino acid replacement at a position corresponding to residue 317, with reference to amino acid positions set forth in SEQ ID NO:3;
>
> the replacement at the position corresponding to residue 317 is selected from the group consisting of A, I, K, M, Q, and R;
>
> corresponding amino acid positions are identified by alignment of the PH20 polypeptide with the polypeptide having the amino acid sequence of SEQ ID NO:3; and
>
> the modified PH20 polypeptide has at least 91% sequence identity to a polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NOs: 3, 7 and 32-66.

42. Claim 1 defines a sequence identity parameter (*i.e.*, a minimum percentage of sequence identity a sequence must have to one of a list of reference PH20 amino acid sequences, including one modification at position 317). Claim 1 also requires the modification at position 317 to be a substitution from the wild-type (WT) residue leucine (L) to one of alanine (A), isoleucine (I), lysine (K), methionine (M), glutamine (Q) or arginine (R). I have highlighted the portions of the claim that are associated with these two sets of parameters that govern the amino acid sequence in the polypeptide.

1. A modified PH20 polypeptide, comprising one or more amino acid modifications in an unmodified PH20 polypeptide, wherein:

the unmodified PH20 polypeptide consists of the amino acid sequence selected from the group consisting of SEQ ID NOs: 3, 7 and 32-66;

amino acid modifications are selected from the group consisting of amino acid replacements(s), deletion(s), and/or insertion(s);

the modified PH20 polypeptide comprises an amino acid replacement at a position corresponding to residue 317, with reference to amino acid positions set forth in SEQ ID NO:3;

the replacement at the position corresponding to residue 317 is selected from the group consisting of A, I, K, M, Q, and R;

corresponding amino acid positions are identified by alignment of the PH20 polypeptide with the polypeptide having the amino acid sequence of SEQ ID NO:3; and

the modified PH20 polypeptide has at least 91% sequence identity to a polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NOs: 3, 7 and 32-66.

Position 317 substitution claim parameter

Sequence identity claim parameter (includes one change at position 317)

43. As I explain in §§ II.D and II.E below, and based on what has been explained to me are the legal standards for interpreting patent claims, I do not believe the claims in view of the common disclosure are properly read as capturing every polypeptide with an amino acid sequence that satisfies the sequence identity and position 317 claim parameters. Instead, I read the claims ***as a whole*** as being directed to modified PH20 polypeptides that (i) contain an amino acid sequence that satisfies the sequence identity claim parameters, (ii) meet the position 317

substitution requirement, (iii) have folded to adopt the same tertiary fold found in the native state of the wild-type PH20 protein and (iv) are enzymatically active.[15]

44. While my opinion differs from those of Drs. Petsko and Simpson, I believe they both agree that the claims include within their scope a set of modified PH20 polypeptides that are enzymatically active.[16] Where we appear to disagree is whether the claims also include modified PH20 polypeptides which cannot show detectable activity due to sequences that cannot be produced, cannot fold to the native state of PH20, or aggregate or fold to the native state with such reduced kinetic or thermodynamic stability relative to the native fold that they unfold rapidly.

1. The Sequence Identity Claim Parameters Define Variants of the Human PH20 Amino Acid Sequence

45. All the claims require modified human PH20 polypeptides to have 91% or higher overall amino acid sequence identity to one of a list of "reference"

---

[15] Dr. Petsko uses the phrase "properly folded" with this meaning. *See* ¶ 134; EX2070 (Petsko Dec.), ¶ 51.

[16] EX2070 (Petsko Dec.), ¶ 53 ("Enzymatically active modified PH20s within the scope of the claims…"), ¶ 128 (claims cover "enzymatically active 91% identical PH20 variants…"), ¶ 424 ("The claims encompass enzymatically active PH20s…"); EX2068 (Simpson Dec.), ¶¶ 355-358 ("[T]he enzymatically active claimed variants…").

amino acid sequences of the human PH20 protein (*i.e.*, SEQ ID NOS:3, 7 and 32-66).[17] A POSA would therefore understand the claims to be referring to modified forms of the human PH20 amino acid sequence.

2. The Sequence Identity Claim Parameters Do Not Restrict the Location or Nature of Changes

46. I agree with Dr. Hecht and the other experts that the sequence identity claim parameters do not restrict the location(s) or type(s) of amino acid changes permitted in each modified PH20 polypeptide, other than the substitution required at position 317.[18] Provided that the entire amino acid sequence of a polypeptide satisfies the numerical threshold of the sequence identity claim parameters and has one of the position 317 substitutions, it can contain an amino acid sequence with any (or all) of the following types of changes:

(i) changes at any position in the amino acid sequence (*e.g.*, at highly conserved positions or positions known to be necessary for activity);

(ii) a change that is a substitution can replace any wild-type amino acid with one of the 19 alternative amino acids; and

---

[17] EX1026 (’731 App.), 85:33–86:2; *see also* 51:32–36, 47:20-21, 47:30–34.

[18] EX1003 (Hecht Dec.), ¶ 119, 122; EX1026 (’731 App.), 68:26-34, 86:23-31; EX1130 (Simpson Dep.), 54:2–11, 90:25–91:12; EX1131 (Petsko Dep.), 55:22–56:23.

(iii) changes can delete or add one or several amino acids, and those deletions or additions can occur anywhere within the amino acid sequence.

47. Dr. Simpson confirmed these types of changes are permitted by the claim language.[19]

> Q. Sure. The changes that constitute the difference that translates into the 5% difference from the amino acid sequences of the wild-type PH20, there could be substitutions, deletions, or additions of amino acids. Right?
> A. Yes. Those are possibilities.
> Q. So in the set of those modified PH20 polypeptides, you may have polypeptides that have a deletion of three or five amino acids from the wild-type residue. That's within the set of PH20 polypeptides being claimed. Right?
> A. That is a potential scenario for one of the claimed polypeptides.

3. The Sequence Identity Claim Parameters Define an Immense Sequence Space

48. Dr. Hecht accurately described the size of the sequence space defined by the sequence identity claim parameters (*i.e.*, at least 91% sequence identity to one of SEQ ID NOS: 3, 7 and 32-66, and one position 317 change to A, I, K, M, Q

[19] EX1130 (Simpson Dep.), 90:25-91:12.

or R).[20] Dr. Park calculated the number of distinct amino acid sequences that meet those sequence identity claim parameters.[21] Dr. Park's table is reproduced below. The sequence identity claim parameters define an immense "sequence space" (*e.g.*, between 10^{59} and 10^{112} different amino acid sequences). I may refer to this set of sequences as the "claim parameter sequence space" or "sequence space" in this declaration.

Claim(s)	*PH20 Length*	*# of Changes*	*Pos. 317 Choices*	*Add'l Changes*	*# of Distinct Polypeptides*
1, 3-5, 17-21, 23-24, 27-40	474	42	6	41	3.79×10^{112}
2	474	23	6	22	3.11×10^{66}
6,7	474	42	1	41	6.32×10^{111}
8, 22	465	41	6	40	8.47×10^{109}
9, 10	465	41	1	40	1.41×10^{109}
11	433	38	6	37	6.01×10^{101}
12	430	38	6	37	4.59×10^{101}
13	433	38	1	37	1.00×10^{101}

[20] EX1003 (Hecht Dec.), ¶¶ 115-125.

[21] EX1004 (Park Dec.), ¶¶ 165-169, Appendix F.

14	430	38	1	37	7.66×10^{100}
15	447	40	6	39	8.37×10^{106}
16	447	40	1	39	1.40×10^{106}
25	430	21	1	20	4.40×10^{59}
26	433	21	1	20	5.08×10^{59}

49. Drs. Petsko and Simpson did not identify any errors in Dr. Park's calculations of the number of different amino acid sequences captured by the sequence identity claim parameters.[22] Dr. Simpson also agreed on the scale of numbers of distinct amino acid sequences captured by the sequence identity claim parameters.[23]

50. The vast majority of the sequences within the claim parameter sequence space have multiple changes relative to the wild-type PH20 amino acid sequence. To illustrate, there are 8,480 possible amino acid sequences that have one substitution in the 447 amino acid form of PH20 to 19 alternatives with

---

[22] EX1131 (Petsko Dep.), 58:18-22 ("I wouldn't disagree"), 60:24–61:8 ("I don't disagree"); EX1130 (Simpson Dep.), 90:1–16, 91:18-92:1.

[23] EX1130 (Simpson Dep.), 90:1-16, 91:18–92:1.

position 317 limited to six alternatives.[24] Allowing 2 total substitutions (with one limited to six alternatives) yields approximately 50,850 different sequences. Each additional permitted substitution increases the size and structural diversity of the sequence space exponentially. The potential number of sequences with a total of 15 substitutions is ~10^{44} different sequences, derived using the same approach used by Dr. Park.[25] And when the other changes envisioned by Dr. Simpson are accounted for, the size of the sequence space gets even larger.

51. A requirement of 91% sequence identity to a reference sequence does not mean that all sequences within the claimed set differ from one another by only 9%. For example, if two sequences (A and B) are each 91% identical to the same reference sequence (*e.g.*, SEQ ID NO:32), but share no common substitutions, the A and B sequences may differ from one another at up to 18% of their positions. Thus, compliance with a 91% sequence identity requirement to a reference sequence does not mean there is only a 9% variance in the set of proteins.

---

[24] EX1131 (Petsko Dep.), 112:20-113:11 ("Q. So if you just look at the universe of modified PH20 polypeptides that have one amino acid substitution in the 447 residue sequence, I think there are probably around 9,000 species in that genus, right? A. I don't know. Thousands sounds right….").

[25] EX1004 (Park Dec.), ¶¶ 168-169, Appendix F–1.

4. Drs. Petsko and Simpson Have Differing Opinions About What Set of Polypeptides is "Structurally Homogeneous"

52. Drs. Petsko and Simpson both say that the set of modified PH20 polypeptides defined by the claims is "structurally homogeneous."[26] But what each believes that set of "structurally homogeneous" modified PH20 polypeptides is differs.

53. Dr. Simpson's opinion appears to be that the claims include any polypeptide with an amino acid sequence that meets the "sequence identity" and position 317 requirements.[27]

> Q. Okay. So you believe that any polypeptide that has an amino acid sequence with 95% sequence identity to SEQ ID 3, for example, plus a replacement of aspartic acid at 320 to any of histidine, lysine, arginine, or serine would meet all the requirements of claim 1. Is that right?
> A. Any polypeptide with 95% sequence identity to the reference sequence, SEQ ID 3, for example, with position 320 modified 95% identity, those are the polypeptides covered in claim 1.

---

[26] EX2068 (Simpson Dec.), ¶ 33; *see also* EX2070 (Petsko Dec.), ¶ 58 ("[T]he disclosed examples are representative of this homogeneous set, which share common structure-function relationships.").

[27] EX1130 (Simpson Dep.), 54:2-11. Dr. Simpson's answers during her deposition were directed towards the related U.S. Patent 11,952,600, which has a sequence identity parameter of 95% identity. I refer to the smaller 91% sequence identity parameter of the '758 Patent.

54. She recognizes that set of modified PH20 polypeptides corresponds to the entire claim parameter sequence space of 10^{59} to 10^{112} different sequences.[28] But she still portrays this set of claimed PH20 polypeptides as being "structurally homogeneous."[29] That is not scientifically correct.

> A. I maintain that -- or have expressed that POSAs would have understood the claimed genus to be a narrow set of PH20 variants and subvariants, structurally homogeneous.

> 33. POSAs would have understood that the claimed genus is a narrow set of PH20 variants and sub-variants that are structurally homogeneous. Because the sub-variants all share at least 91% sequence identity with the L317 PH20 variants, there is almost no structural diversity within the claimed genus. POSAs would expect the active sub-variants to retain, for instance, the same alpha-beta barrel fold.

> 317. As I explained in Section X.A, a POSA would have understood that this level of similarity reflects a homogeneous and narrowly defined set of PH20 proteins, and the six disclosed L317 active variants are representative of their 91% identical sub-variants. EX2070, ¶¶58, 389-431.

55. Dr. Simpson recognized that enzymatically active modified PH20 polypeptides are a structurally distinct subset of the larger sequence space because only enzymatically active modified PH20 polypeptides possess an additional

---

[28] EX1130 (Simpson Dep.), 89:3–25.

[29] EX1130 (Simpson Dep.), 89:3–25; EX2068 (Simpson Dec.), ¶¶ 33, 317.

structural attribute she calls the "alpha beta barrel fold."[30] This α/β barrel motif is found only in proteins that have adopted the native fold of the PH20 protein.

56. Dr. Petsko identifies only "folded" proteins as being "structurally homogeneous" (*i.e.*, only a subset of the modified PH20 polypeptides that Dr. Simpson said were "structurally homogeneous").[31] Dr. Petsko also said the "folded" subset contain structural features not found in the larger set Dr. Simpson referenced. For example, at his deposition, Dr. Petsko first identified a set of modified PH20 polypeptides containing amino acid sequences that shared the 91% sequence identity and the position 317 substitution claim requirements.[32] Then, he

---

[30] EX2068 (Simpson Dec.), ¶ 33 ("POSAs would expect the ***active*** sub-variants to retain, for instance, the same alpha beta barrel fold.") (emphasis added). Dr. Simpson's terminology here is incorrect. The α/β barrel is a motif in the tertiary structure of the human PH20 protein; it is not the entire tertiary structure.

[31] EX2070 (Petsko Dec.), ¶ 391 ("POSAs think of the class of modified PH20 proteins sharing 91% of their sequence … as an extremely limited set of proteins. ***Those that fold are a limited set of structurally homogeneous proteins***.") (emphasis added).

[32] EX1131 (Petsko Dep.), 66:13–17 ("Q. Are there other common structural features shared by this set? A. Those are the two that are most common to – that are common to all members of the set."). Dr. Petsko's answers during his

identified a subset that had an additional structural feature that the larger set did not possess, which he referred to as the α/β barrel fold in the PH20 wild type protein.[33]

Q. Once you've looked at that, is another common structural feature of the claimed modified polypeptides that they all have the same alpha-beta barrel fold?
A. No. You don't know that for certain. But you do know that a vast number of them will have the alpha -- that alpha-beta barrel fold because of the requirement of 95 percent sequence identity with proteins that do have the -- that alpha-beta barrel fold.
Q. Are you speaking of the alpha-beta fold that is the one observed in the human PH20 protein?
A. I'm speaking of that specific alpha-beta barrel fold. It's a subset of a common alpha-beta barrel super family.
Q. All right. And just so I'm -- I make sure I am clear about this in my own mind.
Is the alpha-beta barrel fold of the human PH20 polypeptide, is that also called the tertiary structure of the human PH20 polypeptide?
A. It's part of the tertiary structure because there are other elements in the protein besides that particular subdomain, if you will. But yes, it is part of the tertiary structure.

57. Dr. Petsko made clear that the "PH20 alpha-beta barrel fold" he is referring to is the one observed within the tertiary structure of the wild-type PH20 protein in its native state.[34]

58. Dr. Petsko's statements confirm that he views the "structurally homogeneous" forms of modified PH20 polypeptides to be enzymatically active.

---

deposition were directed towards the related U.S. Patent 11,952,600, which has a sequence identity parameter of 95% identity and claimed substitutions at position 320. I refer to the smaller 91% sequence identity parameter and the claimed substitutions at position 317 of the '758 Patent.

[33] EX1131 (Petsko Dep.), 68:6-69:5.

[34] EX1131 (Petsko Dep.), 68:16–69:5.

For example, he states that the "claims cover a homogeneous set of proteins, and the disclosed examples are representative of this homogeneous set, which ***share common structure-function relationships***."[35] The only PH20 function he could be referring to is hyaluronidase activity, and the only modified PH20 polypeptides that have enzymatic activity are those that Dr. Petsko describes as "properly folded." For example, Dr. Petsko said he uses enzymatic activity and the ability to be secreted as "proxies" for PH20 proteins that have "properly folded":[36]

> [12] When I say "enzymatically active" in this context, I am not referring to a specific level of enzymatic activity. Rather, I am using enzymatic activity as a proxy for a properly folded protein. Secretion is another proxy that could be used for a properly folded protein; proteins that do not fold will generally not be secreted.

59. While enzymatic activity confirms that an enzyme has folded, there would need to be additional evidence to justify a conclusion that the folded enzyme is sufficiently stable to retain its folded state. For example, a POSA would not believe an enzyme had "properly folded" if it did not maintain activity over a meaningful period or if it formed aggregates within a short period. Both of those

---

[35] EX2070 (Petsko Dec.), ¶ 58 (emphasis added); *see also* ¶ 101 ("[I]f the 91% identical modified PH20 folds, it will fold into the same alpha-beta barrel motif as the WT.").

[36] EX2070 (Petsko Dec.), ¶ 128, n.12; EX1131 (Petsko Dep.), 74:18-75:12, 86:17-87:14.

experimental results are indicative of a modified PH20 polypeptide being unstable, but not necessarily that the protein hasn't "properly folded."

B. The Amino Acid Sequences in the Claim Parameter Sequence Space Are Structurally Heterogeneous, Not Homogeneous

60. When POSAs consider the "sequence space" defined by the sequence identity claim parameters, what they see is an immense and structurally diverse set of amino acid sequences. POSAs would not assess the "structural" diversity of this sequence space by only considering a mathematical degree of sequence identity to a reference sequence. Instead, they would want to understand what the implications of the mutations are on the ability of the mutated sequence to fold and adopt the identical three-dimensional structure of the native state of the protein, on enzymatic activity, and even then whether the mutations affect the ability of the protein to maintain its native three-dimensional structure (the critical role of stability has been noted by Dr. Petsko[37]), among other questions.

61. I agree with Dr. Hecht that the common disclosure fails to "characterize the effects of any particular mutation on the structural features of

---

[37] *See, e.g.*, EX1152 (Ringe and Petsko 2009), 80.2 ("[M]ost metabolic disorders involve mutations that destabilize proteins rather than simply inactivating them.").

PH20."[38] There is no discussion on how any of the observed changes in activity of mutated PH20 proteins relates to the effects made to secondary structures or structural motifs found within the native protein. As Dr. Hecht notes, the common disclosure does not describe how the single amino acid substitutions affect PH20's native structure or affect any structure-function relationship, or the effect of multiple substitutions to the native protein.[39]

1. Amino Acid Sequences in the Claim Parameter Sequence Space Will Adopt Many Different Three-Dimensional Structures

62. A series of articles by Dr. Christopher Dobson at Cambridge University before 2011 provide a useful overview of the complexity and challenges of protein folding, and the diversity of conformations an amino acid sequence can adopt. Below are two of his figures. The first (Figure 5 from EX1136) illustrates the various events that can occur after synthesis of an amino acid sequence that can prevent a sequence from folding into the tertiary structure found in the native state of the wild-type protein (designated "native" in the figure).[40] It is perfectly possible

---

[38] EX1003 (Hecht Dec.), ¶ 69.

[39] EX1003 (Hecht Dec.), ¶¶ 69, 157.

[40] EX1136 (Dobson 2004a), 11 (Figure 5); *see also* EX1137 (Dobson 2003), 888 (Figure 4); EX1138 (Dobson 2004b), 11 (Figure 2); EX1131 (Petsko Dep.),

that proportions of a single sequence will end up as the different species as shown in Figure 5 (that is some folded, some aggregated, etc.).

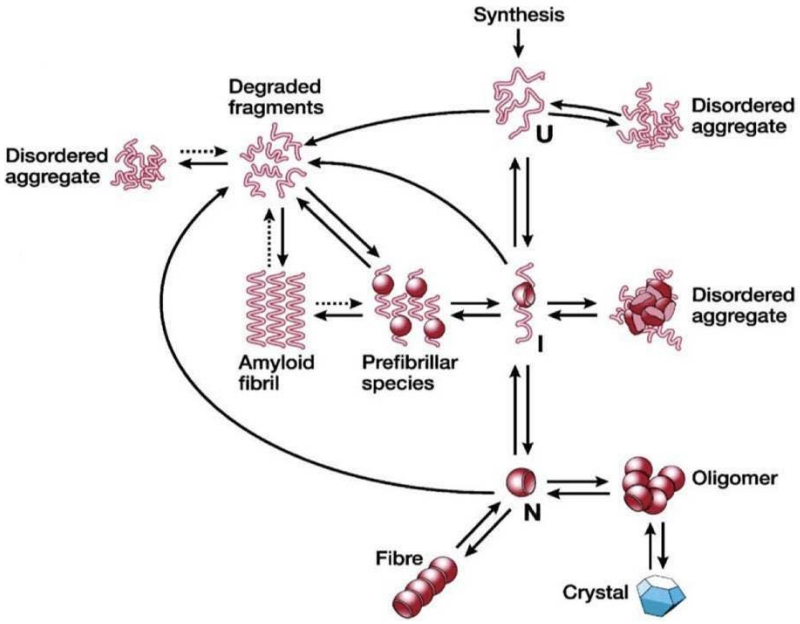


Fig. 5. Schematic representation of some of the states accessible to a polypeptide chain following its synthesis on a ribosome. The relative populations of the different states depend on the kinetics and thermodynamics of the various equilibria shown in the diagram. In living systems the fate of a given protein molecule is closely regulated by molecular chaperones and other quality control mechanisms, rather as metabolic pathways in cells are controlled by enzymes and associated molecules such as cofactors. From [36].

63. The second figure (Figure 2 of EX1139) identifies various conformational states that amino acid sequences can adopt, and how those conformation states can transition from one to the other. As the caption to Figure 2 explains, the figure provides a "…representation of some of the many

---

38:18–39:23 ("Q. Okay. So when a protein like PH20 folds, does it proceed through intermediate states or structures before it reaches its final folded tertiary state? A. It's thought to for all proteins that fold…").

conformational states that can be adopted by polypeptide chains and of the means by which they can be interconverted." [41]

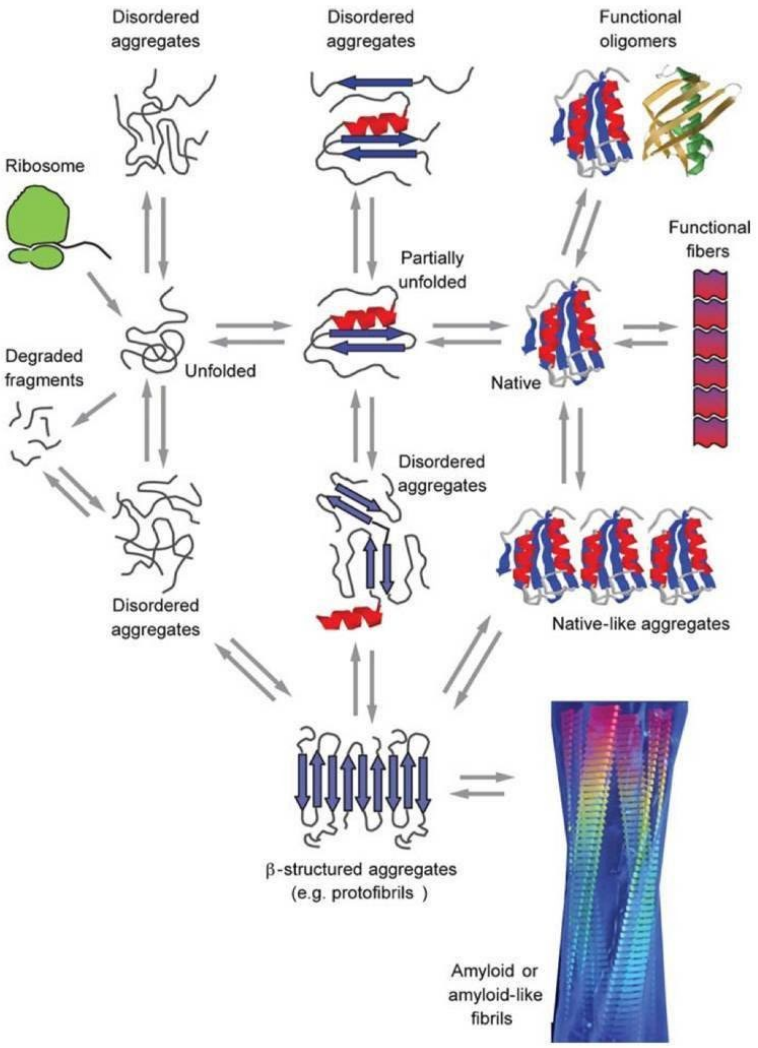


Figure 2

A schematic representation of some of the many conformational states that can be adopted by polypeptide chains and of the means by which they can be interconverted. The transition from β-structured aggregates to amyloid fibrils can occur by addition of either monomers or protofibrils (depending on protein) to preformed β-aggregates. All of these different conformational states and their interconversions are carefully regulated in the biological environment, much as enzymes regulate all the chemistry in cells, by using machinery such as molecular chaperones, degradatory systems, and quality control processes. Many of the various states of proteins are utilized functionally by biology, including unfolded proteins and amyloid fibrils, but conformational diseases will occur when such regulatory systems fail, just as metabolic diseases occur when the regulation of chemical processes becomes impaired.

41 EX1139 (Chiti & Dobson 2006), 351 (Figure 2).

64. As Dr. Dobson's figures illustrate, in order to fold stably into the native three-dimensional structure (tertiary), the naturally occurring amino acid sequence must navigate a complex pathway with various branches at key steps. Some of the key steps where the protein (or portions of it) can come off the track that leads to a native three-dimensional structure are:

a. The cell may not complete the synthesis of the encoded amino acid sequence.

b. The unfolded nascent polypeptide chain may form disordered aggregates or be degraded into fragments.

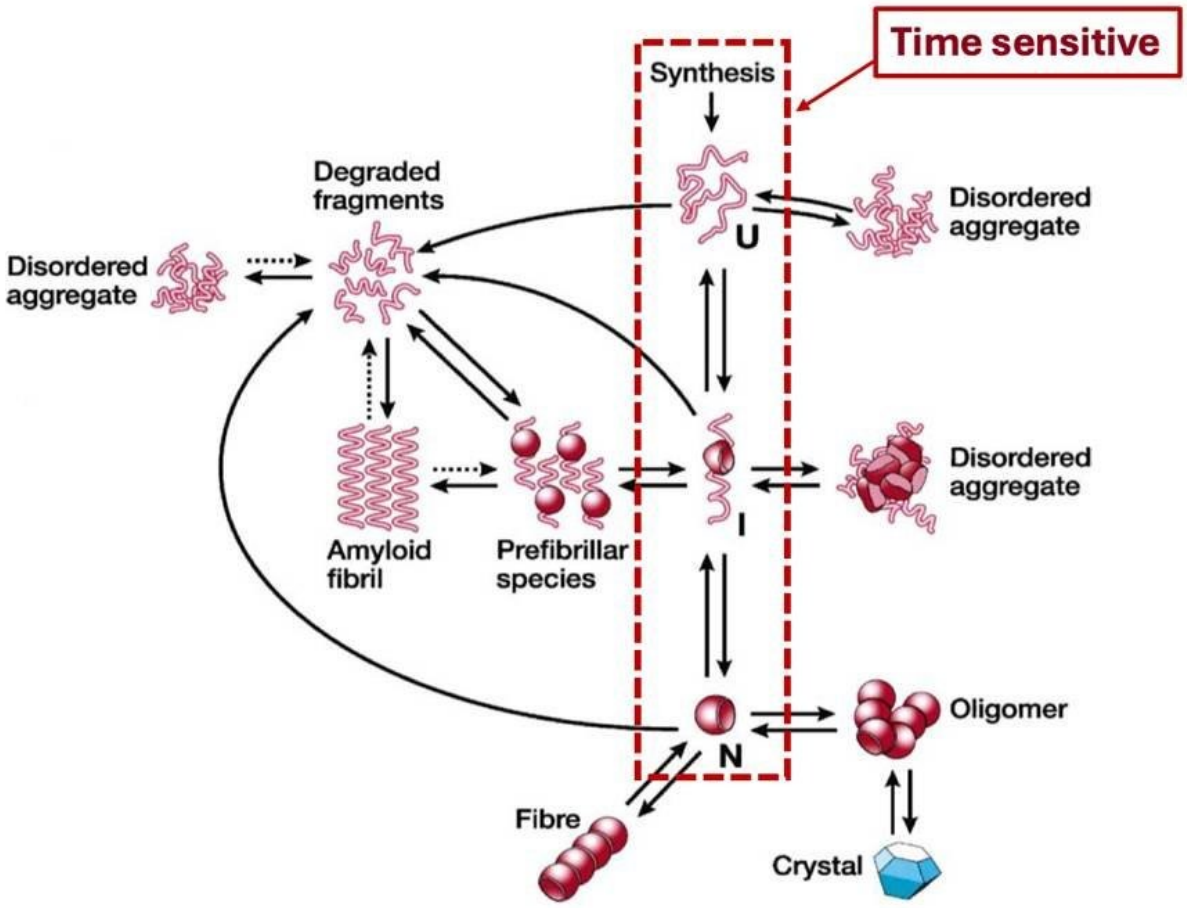
c. Partially folded forms of the sequence may lead to disordered or ordered aggregates.

d. The natively folded proteins may form various types of aggregates, or may unfold and be degraded within the extracellular environment.[42]

65. To successfully fold into the native folded state of the protein, an amino acid sequence must follow the path to the natively folded structure. In the annotated figure below, the dashed box shows the path an amino acid sequence

[42] EX1136 (Dobson 2004a), 11 (Figure 5); EX1139 (Chiti & Dobson 2006), 351 (Figure 2).

must follow to successfully fold into the natively folded, three-dimensional structure of the protein.



66. There is a time-dependence as well—the amino acid sequence must complete its folding pathway within a relatively short period of time. Otherwise, the amino acid sequence will be degraded. That degradation can occur within or

outside the cell. Dr. Petsko illustrated the time sensitive nature of folding during his deposition:[43]

> Proteins have to fold within a defined time period, otherwise a partially folded protein might accidentally get degraded while it hasn't got enough time to complete the folding. There are proteins, for example, that if they -- if they weren't protected from that in some fashion, would literally take minutes to hours to fold -- to be synthesized, much less fold. So that's an issue.

67. Dr. Petsko's point is well known. For example, the Branden & Tooze textbook provides an analogous explanation and illustrates that for a hypothetical 150 residue polypeptide, it would take 10^{48} years to stochastically search through all possible conformations:[44]

> Intuitively one might imagine that all protein molecules search through all possible conformations in a random fashion until they are frozen at the lowest energy in the conformation of native state. The biophysicist Cyrus Levinthal showed in 1968 by a simple calculation that this is impossible. Assume as a gross simplification that each peptide group has only three possible conformations, the allowed regions α, β and L in the Ramachandran diagram (see Figure 1.7), and that it converts one conformation into another in the shortest possible time, one picosecond (10^{-12} seconds). A polypeptide chain of 150 residues would then have $3^{150} = 10^{68}$ possible conformations. To search all these conformations would require 10^{48} years (10^{56} seconds)—an astronomical number compared with the actual folding time, which is between 0.1 and 1000 seconds both *in vivo* and *in vitro*. To occur on this short time scale, the folding process must be directed in some way through a kinetic pathway of unstable intermediates to escape sampling a large number of irrelevant conformations.

---

[43] EX1131 (Petsko Dep.), 43:15-23; *see also* 98:10–13 (explaining that an intermediate unfolded form can gain stability relative to the folded form).

[44] EX1014 (Branden & Tooze), 91.

68. A single amino acid sequence can adopt a tremendous number of possible conformations. The number of those conformations that a sequence can adopt will decrease during its folding process until it achieves a final native state, typically a single conformation. Sequences stochastically explore conformations during the folding process based on their ability to adopt those different conformations.

69. As Dobson 2004a explains, native (not mutant) sequences have been specifically selected by evolution to fold. Of particular note is that only these "special" evolutionarily selected amino acid sequences will induce the "correct (native-like) interactions between different residues [that] are on average more stable than the incorrect (nonnative) ones."[45] Mutated sequences do not benefit from this evolutionarily conferred advantage, which is why so many of them cannot find the pathway to the native state of the folded protein.

[45] EX1136 (Dobson 2004a), 4.

> It is now evident that the folding process does not involve a series of mandatory steps between specific partially folded states, but rather a stochastic search of the many conformations accessible to a polypeptide chain [5–8]. The conceptual basis of such a mechanism is shown in Fig. 1. In essence, the inherent fluctuations in the conformation of an incompletely folded polypeptide chain enable even residues at very different positions in the amino acid sequence to come into contact with one other. Because the correct (native-like) interactions between different residues are on average more stable than the incorrect (nonnative) ones, such a search mechanism is in principle able to find the lowest energy structure [9]. It is evident that this process is extremely efficient for those special sequences that have been selected during evolution to fold to globular structures, and indeed only a very small number of all possible conformations needs be sampled during the search process.

70. A "[s]chematic energy landscape for protein folding" is illustrated in Dr. Dobson's papers. For example, Figure 1 of Dobson 2004a models the energy landscape for a simple 27-residue protein.[46] Even this short sequence has 10^{16} possible starting configurations. The figure illustrates three possible folding pathways (red, yellow and green) within a landscape of possibilities. The polypeptide chain follows a particular sequence of conformations in each folding pathway. The sequence will adopt various conformations until it arrives at the conformation with the lowest free energy state that is the native state of the protein.

[46] EX1136 (Dobson 2004a), 4 (Figure 1).

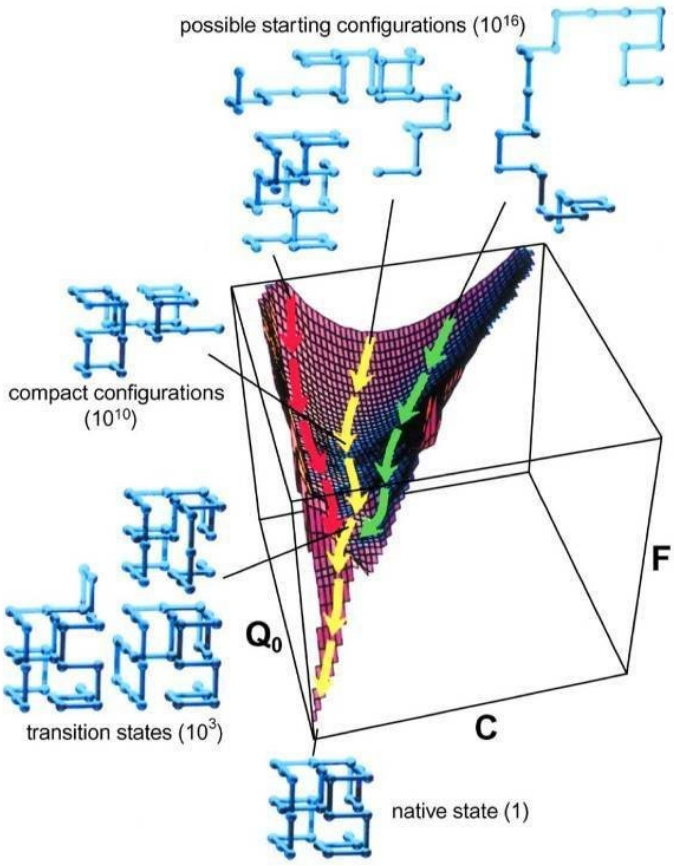


Fig. 1. Schematic energy landscape for protein folding. This surface is derived from a computer simulation of the folding of a highly simplified 27-residue protein. The free energy (F) of the system is shown as a function of the total number of contacts between residues (C) and the number of contacts that correspond to those of the most stable structure, denoted the native state (Q_0). The simulation shows that the surface serves to "funnel" the multitude of denatured conformations to the unique native state. From [9].

71. Figure 2 of Dobson 2004a illustrates a more complex energy landscape for a larger protein with two domains. This landscape has "hills" and "valleys" representing conformations of partially folded forms of the sequence that have higher or lower level energy barriers that must be correctly navigated for the sequence to adopt the native conformation. It also illustrates two folding pathways. The yellow path is a "fast track" in which the two domains fold concurrently. The red path involves sequential folding of the two domains and produces an intermediate that becomes trapped in an energy "valley." As it explains, "[f]urther

folding from this intermediate involves either a transition over a higher barrier, or partial unfolding to enable the remainder of the folding to take place along a fast track."[47]

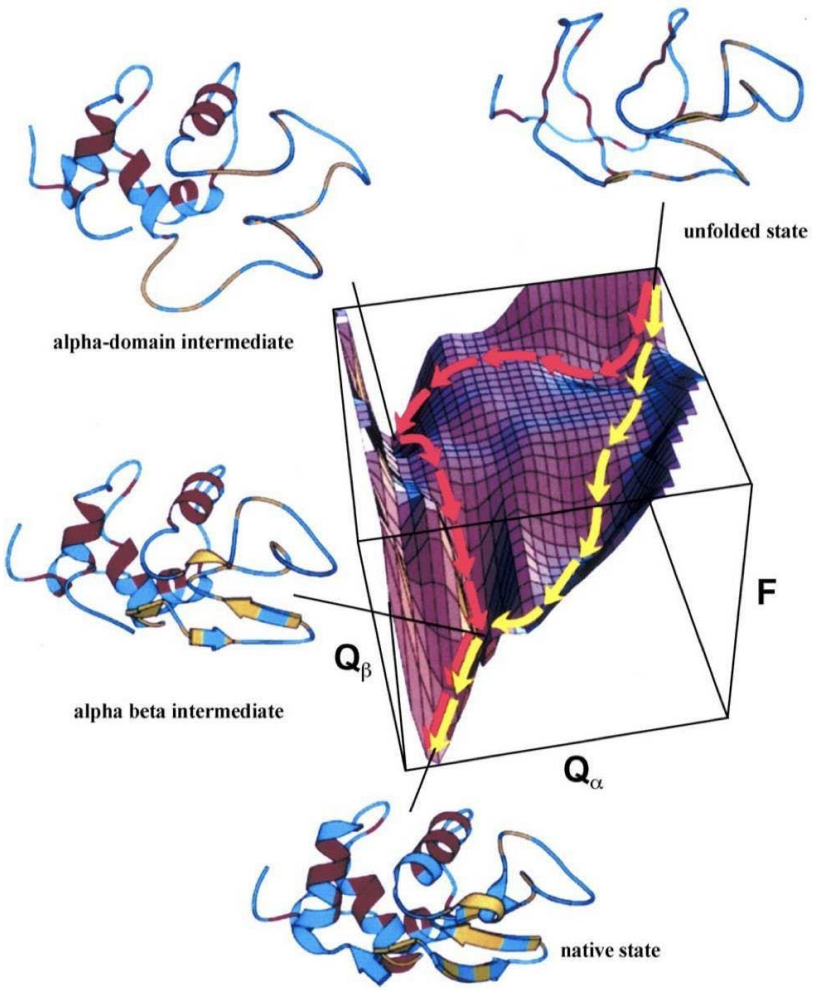


Fig. 2. Schematic energy landscape for the folding of lysozyme. This protein has 129 residues and its structure consists of two domains, denoted α and β. The free energy (F) is shown as a function of the number of native contacts (Q_α and Q_β) in each domain. The yellow trajectory represents a "fast track" in which the two domains fold concurrently and populate an intermediate (labelled alpha/beta) only briefly. The red trajectory represents a "slow track" in which the chain becomes trapped in a long-lived intermediate state with persistent structure only in the alpha domain. Further folding from this intermediate involves either a transition over a higher barrier, or partial unfolding to enable the remainder of the folding to take place along a fast track. The structures of the various species were generated by constraining molecular dynamics simulations with experimental data characterising each species, derived from NMR experiments. From [9].

[47] EX1136 (Dobson 2004a), 6 (Figure 2).

72. Dr. Dobson's illustration depicts deviations from the folding pathway as being reversible, but often they are, as a practical matter, irreversible. Aggregated proteins do not normally refold spontaneously in biology. A sequence may adopt an off-path intermediate conformation that is energetically more stable than the on-path intermediate conformation.[48] The protein in this incorrect but more energetically stable conformation cannot continue toward folding to the native state. The incorrect intermediate state does not need to persist for long given that, as noted by Dr. Petsko, the intermediate can degrade if it does not fold quickly enough.[49]

---

[48] EX1131 (Petsko Dep.), 98:4-20 ("The intrinsic -- … intrinsic stability of a protein is the difference in free energy between the folded form of the protein and the unfolded form of the protein. All right? So that can change not only because you can lose stability of the folded form, but you can -- ***that can change because you can gain stability of the unfolded form***.") (emphasis added).

[49] EX1131 (Petsko Dep.), 43:15-23 ("Proteins have to fold within a defined time period, otherwise a partially folded protein might accidentally get degraded while it hasn't got enough time to complete the folding.").

2. Mutations Increase Obstacles to an Amino Acid Sequence Successfully Folding to Adopt the Native State of a Protein

73. Amino acid sequences that differ from the evolutionarily selected wild-type sequence are at a greater risk of diversion from the folding pathway that must be followed to yield the natively folded protein structure. For example, the wild-type amino acid sequence of PH20 has resulted from evolutionary pressure that would have removed it from the human genome if it could not efficiently fold to adopt its native fold.[50] That reflects the ability of the wild-type sequence to be synthesized, to traverse the folding pathways to the folded native state of the PH20 protein fast enough to avoid degradation and to stably retain its native fold. Non-naturally occurring sequences, such as those with introduced mutations, have not been shaped by and survived the evolutionary pressure that removes sequences that cannot be synthesized by cells, cannot fold, or if they do fold, are not adequately stable.[51]

---

[50] EX1136 (Dobson 2004a), 4 (the folding "process is extremely efficient for those special sequences that have been selected during evolution to fold to globular structures, and indeed only a very small number of all possible conformations needs be sampled during the search process.").

[51] *See* EX1137 (Dobson 2003), 887 ("[M]any such disorders are familial because the probability of misfolding is often greater in mutational variants.").

74. As Dr. Hecht explained, the "net free energy of stabilization" is the difference in free energy between folded and unfolded states of a protein.[52] It represents the net difference in energy resulting from the decrease in enthalpy that occurs during the formation of hydrogen bonds and through van der Waals interactions in the protein's folded state.[53] The difference in energy required for a protein to transition from an unfolded state to a folded state is relatively small, usually no more than the energy contributions necessary to form a hydrogen bond. As a consequence, proteins are inherently unstable.[54] Dr. Petsko's textbook provides a similar explanation to Dr. Hecht's:[55]

> Finally, even though many hundreds of hydrogen bonds and van der Waals interactions occur in a folded protein, the net free energy of stabilization of most folded proteins—the difference in free energy between the folded and unfolded states—is actually rather small, about 21–42 kJ/mole, or only about 10 times the average thermal energy available at physiological temperature. Most folded proteins are marginally stable because the free energy released when hundreds of weak interactions form is almost exactly counterbalanced by the enormous loss of conformational flexibility (loss of entropy) that occurs when the unfolded chain folds into a compact, ordered structure. A folded protein is a thermodynamic compromise.

---

52 EX1003 (Hecht Dec.), ¶ 63.

53 EX2164 (Petsko Textbook), 26.

54 EX1003 (Hecht Dec.), ¶¶ 62–63.

55 EX2164 (Petsko Textbook), 26.

75. Mutations introduced into an amino acid sequence can (and usually will) alter the net free energy of stabilization of the protein. The change impacts the stability of the protein if it adopts the native fold or an alternative fold. If the changes decrease the net free energy of stabilization of the protein, the folded protein will be more prone to partial or complete unfolding. Mutations can thus prevent a modified PH20 protein from remaining stable at physiological conditions, which is necessary for it to exhibit its biological activity (*e.g.*, to bind the HA substrate and hydrolyze it).[56]

76. Dr. Petsko's textbook references Dr. Dobson's work and provides an illustration of the impact of even a single mutation to the structure and stability of a globular protein (*e.g.*, PH20).[57] In this example, the single mutation does not prevent secretion of the globular protein but disrupts its stability to the degree that the original folded protein partially or completely unfolds in the extracellular

---

[56] EX2164 (Petsko Textbook), 3 ("For a polypeptide to function as a protein, it must usually be able to form a ***stable*** tertiary structure (or *fold*) under physiological conditions.") (emphasis added); EX1131 (Petsko Dep.), 36:7-37:3 (for PH20 to "function as a protein, it must be able to form a stable tertiary structure under physiological conditions…").

[57] EX2164 (Petsko Textbook), 160 (Figure 4-53); *see also* EX1140 (Dobson 1999), 329 (Figure 1).

environment. After that occurs, the sequence can refold partially or completely. However, as Dr. Petsko's textbook explains, "the partially unfolded form is prone to aggregation, which results in the formation of fibrils [] and other aggregates that accumulate in the extracellular space."[58]

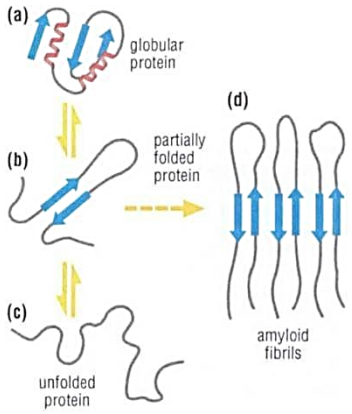


Figure 4-53 A possible mechanism for the formation of amyloid fibrils by a globular protein The correctly folded protein **(a)** is secreted from the cell. Under certain conditions, or because it contains a mutation, the protein unfolds partially **(b)** or completely **(c)**; the unfolded forms can also refold partially or completely. The partially unfolded form is prone to aggregation, which results in the formation of fibrils **(d)** and other aggregates that accumulate in the extracellular space.

Recent studies suggest that the ability to undergo a refolding leading to amyloid formation is not unique to these proteins, but can be observed in many other proteins under laboratory conditions such as low pH. One conclusion from such findings is that prions and other proteins that cause disease by this mechanism may differ from the vast array of "normal" cellular proteins only in having sequences that can undergo such refolding spontaneously under physiological conditions. Clearly, if this is the case, it is possible that a single point mutation may convert a harmless protein into one that can refold spontaneously; such mutations have been found associated with some of the amyloidoses.

3. Illustration of the Structural Diversity of the Claim Parameter Sequence Space

77. In the 2011-2012 timeframe, a POSA would not view the sequence space defined by the sequence identity claim parameters as being structurally homogeneous. That is because it contains many unfolded sequences that can adopt an immense number of conformational states, all of which are different from the native fold. Sequences can also become trapped in these intermediate conformational states. Dr. Petsko addressed these points in his deposition:[59]

---

[58] EX2164 (Petsko Textbook), 160 (Figure 4-53).

[59] EX1131 (Petsko Dep.), 98:4-25.

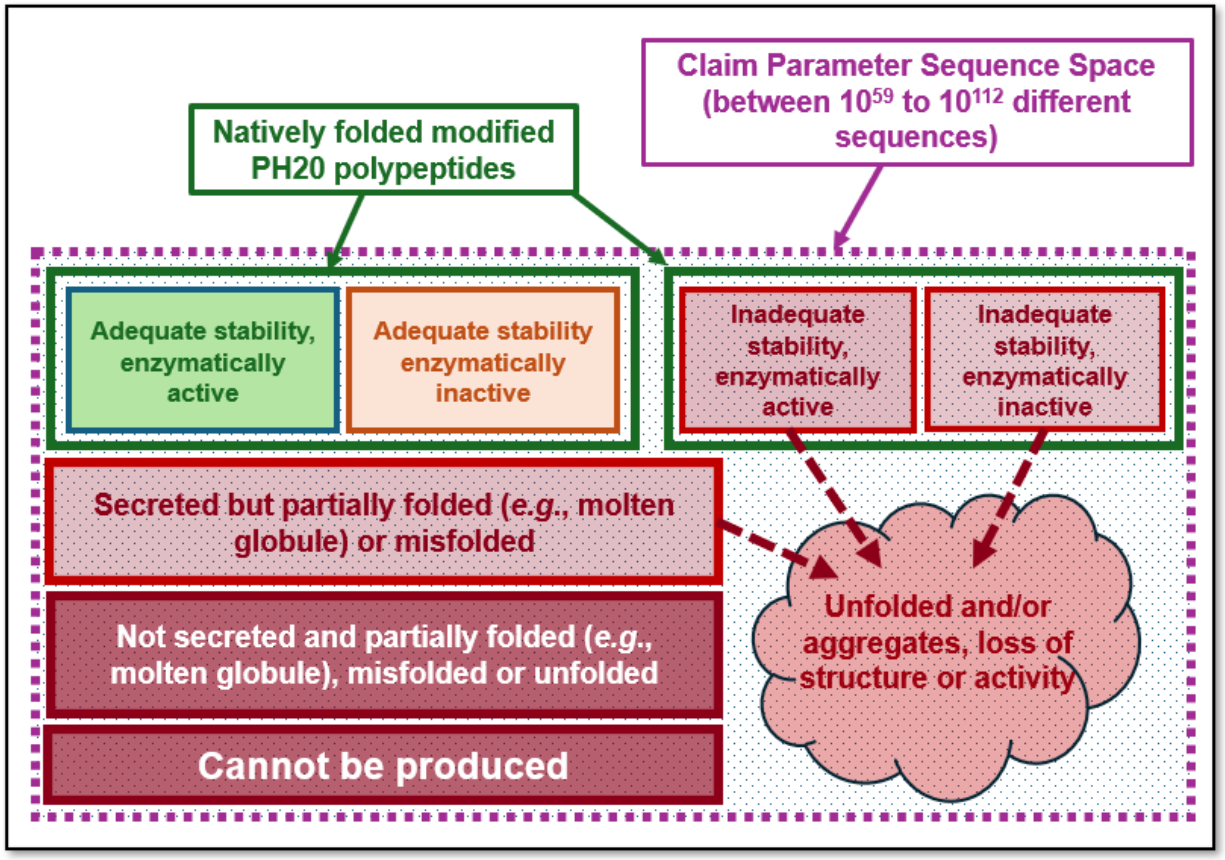
All right? The intrinsic -- if you want to define the term correctly, as I said, it gets used sloppy, intrinsic stability of a protein is the difference in free energy between the folded form of the protein and the unfolded form of the protein. All right?

So that can change not only because you can lose stability of the folded form, but you can -- that can change because you can gain stability of the unfolded form. And it matters which that is sometimes, and a careful scientist who studies protein folding and stability always has to keep that in mind, that you're dealing with a difference between a state that you can not only measure but in many cases visualize with crystallography and a state that isn't one state but that it's a distribution of unfolded forms.

And yet the stability -- the ability of that particular state to exist and how much of it exists in that form is critical to the stability of the protein, even though it's not -- I'm sorry. I'm -- I'm not going to ramble anymore.

78. Other portions of the sequence space may exist as partially folded or misfolded sequences, along with higher order aggregates (*e.g.*, disordered or ordered). These will include (1) proteins with an almost identical folded structure to the native which possess activity, (2) proteins with almost identical folded structure to the native which do not possess activity, (3) proteins with an almost identical folded structure to the native which possess activity but are unstable, causing them to unfold and/or aggregate resulting in loss of structure and activity and (4) proteins with an almost identical folded structure to the native which do not possess activity but are unstable, causing them to unfold and/or aggregate resulting in loss of structure and activity. A conceptual illustration of these distinct sets of

three-dimensional structures adopted by sequences within the claim parameter sequence space is provided in the figure below:



79. If the claims only require a polypeptide's amino acid sequence to meet the 91% overall sequence identity and position 317 claim parameters, regardless of whether the polypeptide will fold (or the nature of the fold it adopts), the claims would capture sets with between 10^{59} to 10^{112} different amino acid sequences.[60] As Dr. Hecht has explained, sets containing one molecule of each of those distinct

[60] *See* EX1004 (Park Dec.), ¶¶ 166-169.

amino acid sequences each would have an aggregate mass exceeding that of the Earth.[61] Large numbers of those amino acid sequences will exist in an unfolded state that exhibit myriad three-dimensional structures, while other sequences will adopt misfolded or partially folded (*e.g.*, molten globules) structures.[62] Many of these unfolded, misfolded or incompletely folded amino acid sequences also will exist as aggregates of individual molecules. The overall set will also include the various structurally distinct subsets of stably folded and enzymatically active proteins that Dr. Hecht described in his declaration.[63] That is the definition of heterogeneity, not homogeneity.

80. A POSA thus would not portray the claim parameter sequence space of 10^{59} to 10^{112} distinct amino acid sequences as being "structurally homogeneous." They instead would believe there is an incredible amount of structural diversity within it. The only relationships between these sequences are: (i) a mathematical one (*i.e.*, the numerical boundary of "at least 91% sequence identity"), and (ii) a

---

61 EX1003 (Hecht Dec.), ¶ 189.

62 EX1014 (Branden & Tooze), 91-92 ("The first observable event in the folding pathway of at least some proteins is a collapse of the flexible disordered unfolded polypeptide chain into a partly organized globular state, which is called the molten globule…"); *see also* 92 (Figure 6.2).

63 *See, e.g.*, EX1003 (Hecht Dec.), ¶¶ 155-169.

single amino acid (A, I, K, M, Q, or R at position 317 of the sequence). In the 2011-2012 timeframe, a POSA would have no reliable method of assigning sequences to the various categories I identify in my figure (*see* ¶ 78) except by extensive experimentation.

C. Sequence Identity Percentages Are Not a Structural Feature of a Modified PH20 Polypeptide

81. Drs. Simpson and Petsko both portray the "at least 91% sequence identity" claim requirement as a "structural" feature of a modified PH20 polypeptide.

a. Dr. Petsko states in ¶ 49 of his declaration that "[t]he ***91% sequence identity*** and substitution at position 317 are ***common structural features*** shared by the set of claimed modified PH20 polypeptides." (emphasis added). Similarly, in ¶ 424 he states that "a POSA would view the claims as directed to a narrow set of PH20 variants with common structural features including 91% sequence identity and a defined substitution at position 317."

b. Dr. Simpson similarly portrays the 91% sequence identity requirement as a "common structural feature" of the set of modified PH20 polypeptides being claimed. For example, Section X.B. of her Declaration identifies it as one of the "Common Structural Features of the Claimed PH20 Variants and Sub-Variants."

82. At his deposition, Dr. Petsko identified the 91% sequence identity claim requirement as one of the "structural features" of the set of modified PH20 polypeptides being claimed:[64]

> Q. So you're explaining in these two paragraphs that within -- or that claim 1 is directed to a set of modified PH20 polypeptides that are -- have particular structural features, right?
> A. They have structural features in common.
> Q. Structural features in common.
> Can you list for me what the common structural features are of that set?
> A. So every protein in that set, every protein in that set will contain an amino acid sequence that is at least 95 percent identical to the amino acid sequence of one of the reference proteins referred to in the claims. All right?

83. I do not agree that the 91% sequence identity requirement is a "structural feature" of a modified PH20 polypeptide. It is a claim requirement that is determined mathematically: one determines the number of changes between two amino acid sequences and then divides that number by the number of amino acids in the reference PH20 sequence (*e.g.*, 40 changes divided by 447 total residues =

[64] EX1131 (Petsko Dep.), 65:12-25.

.0895, which equates to .9105, or 91.05% identity).[65] It does not define any particular physical characteristic of a modified PH20 molecule. It functions solely to limit the size of the sequence space of the modified PH20 polypeptides.

84. Dr. Petsko also recognized that the 91% sequence identity requirement "limits" the sequence space the modified PH20 polypeptides can have. In other words, he recognized that the "91% sequence identity" claim requirement is defining a boundary of the claim parameter sequence space of the set of modified PH20 polypeptide sequences, and is not a common physical characteristic.[66]

> That's a common structural feature. It limits the space that those, that that sequence, that the proteins can have.

85. A POSA would consider "structural features" of a protein to be the physical characteristics of the protein molecule. For example, a particular amino acid located at a particular position in the amino acid sequence of the protein is a structural feature. Structural features of folded proteins are discrete three-dimensional structures within the tertiary structure of those proteins: secondary

---

[65] EX1026 ('731 App.), 66:32-34 ("As used herein, "sequence identity" refers to the number of identical or similar amino acids or nucleotide bases in a comparison between a test and a reference polypeptide or polynucleotide.").

[66] EX1131 (Petsko Dep.), 66:1–3.

structures, higher order motifs (*e.g.*, a catalytic site or substrate binding site), a domain, or the protein's overall tertiary structure. The phrase "structure of a protein" would, unless prefixed by "primary" or "secondary," be taken by a POSA to refer only to the three-dimensional structure of a protein.[67]

86. By definition, the primary structure of the modified PH20 polypeptides within any particular portion of the claim parameter sequence space (*i.e.*, the complete amino acid sequence of each of those polypeptides) cannot be a "common structural feature" of the claimed set of PH20 polypeptides. That is because every one of those modified PH20 polypeptides will have a ***unique*** amino acid sequence (primary structure), with between 1 and 42 differences relative to one of the reference wild-type PH20 sequences (*i.e.*, SEQ ID NOS: 3, 7 or 32-66).[68] *See* ¶ 51.

---

[67] EX1003 (Hecht Dec.), ¶ 36; EX2076 (Hecht Dep.), 240:1–20 ("When chemists think about structure, they're usually thinking about a three-dimensional structure, okay…").

[68] The "primary structure" does not describe the three-dimensional structure of a protein. It is only a description of the arrangement of amino acids along a linear polypeptide chain. *See, e.g.*, EX1014 (Branden & Tooze), 29 ("Primary structure is the amino acid sequence, or, in other words, the arrangement of amino acids along a linear polypeptide chain."); EX2164 (Petsko Textbook), 6 ("[P]rimary structure: the amino-acid sequence of a polypeptide chain.").

D. The "PH20" Requirement in the Claims Requires the Claimed Modified PH20 Polypeptides to be Folded Proteins

1. Modified PH20 Polypeptides are Proteins

87. The general introductions to protein structure in the declarations of Dr. Hecht and Dr. Petsko cover many well-established protein structure principles that would have been known to a POSA.[69] Dr. Petsko confirmed many of these at his deposition, including that POSAs would have known that:

a. An enzymatically active protein like PH20 exists in a stable tertiary structure at physiological conditions, which is a structure that results when the protein's amino acid sequence has properly folded.[70]

b. PH20 exists as a single compact active fold in its native state.[71]

---

[69] *See* EX1003 (Hecht Dec.), ¶¶ 34–46; EX2070 (Petsko Dec.), ¶¶ 59–82.

[70] EX1131 (Petsko Dep.), 36:22-25 ("Q. Is the human PH20 protein a protein that forms a stable tertiary structure under physiological conditions? A. Fortunately, yes.").

[71] EX1131 (Petsko Dep.), 37:4-11; EX1039 (Alberts), 130 ("As a result of all of these interactions, most proteins have a particular three-dimensional structure, which is determined by the order of the amino acids in its chain. The final folded structure, or conformation, of any polypeptide chain is generally the one that minimizes its free energy.").

c. A protein like PH20 must form a stable tertiary structure at physiological conditions to exhibit its ability to catalyze the hydrolysis of hyaluronic acid and to bind to the HA substrate.[72]

88. Dr. Petsko also confirmed that POSAs frequently use the terms "polypeptide" and "protein" interchangeably.[73] I note that Drs. Simpson and Petsko both interchangeably used the terms "polypeptide" and "protein" when they described the "modified PH20 polypeptides" being claimed by the patent. Some examples include the following (all emphases added):

- EX2068 (Simpson Dec.), ¶ 283: "[I]t is helpful to compare and contrast ***the claimed proteins*** from those that fall outside the scope of the claims";

---

[72] EX1131 (Petsko Dep.), 36:7-37:20; EX2164 (Petsko Textbook), 2 ("For a polypeptide to function as a protein, it must usually be able to form a stable tertiary structure (or *fold*) under physiological conditions."); EX1008 (Stern), 824 (discussing a possible mechanism of action for PH20).

[73] EX1131 (Petsko Dep.), 32:8-23 ("Q. Is there a difference between a polypeptide and a protein in the terminology you use? A. Sometimes. It's a sloppy terminology. …. So it is not a hard-and-fast terminology. Q. Do -- do the two terms -- are they used interchangeably? A. Frequently by people. Okay.").

- EX2068 (Simpson Dec.), ¶ 353: "POSAs would evaluate the scope of the claims by considering the variation across ***the set of claimed proteins***";

- EX2068 (Simpson Dec.), ¶ 376: "[P]ermit the POSA to make and use the full scope of ***the claimed proteins***";

- EX2070 (Petsko Dec.), ¶ 50: "POSAs would have understood that ***the set of claimed proteins*** share at least 91% of their primary structure, with less than 9% variation, and the recited modification at position 317"; and

- EX2070 (Petsko Dec.), ¶ 426: "To illustrate the structural similarities and differences within ***the set of claimed proteins***, I generated the figures below."

89. Considering all of the points above, I agree with the usage of Drs. Petsko and Simpson that the term "polypeptide" when used in this matter, conveys to a POSA a meaning that is interchangeable with "protein" (*i.e.*, a PH20 polypeptide means a PH20 protein).

2. Defined Terms Used in the Claims ("PH20" and "Modified PH20 Polypeptide")

90. The common disclosure expressly defines two of the terms used in the claims: "PH20" and "modified PH20 polypeptide."[74] Each definition is found

[74] EX1026 ('731 App.), 50:31-51:13, 54:5-14.

within a paragraph that provides additional explanations and context that helps to explain what each term means. The paragraphs containing the definitions of "PH20" and "modified PH20 polypeptide" are reproduced below:[75]

As used herein, PH20 refers to a type of hyaluronidase that occurs in sperm and is neutral-active. PH-20 occurs on the sperm surface, and in the lysosome-derived acrosome, where it is bound to the inner acrosomal membrane. PH20 includes those of any origin including, but not limited to, human, chimpanzee, Cynomolgus monkey, Rhesus monkey, murine, bovine, ovine, guinea pig, rabbit and rat origin. Exemplary PH20 polypeptides, including precursor and mature forms, include those from human (SEQ ID NO:6 and 7), chimpanzee (SEQ ID NO:8, 9, 10, 869 and 870), Rhesus monkey (SEQ ID NO:11 and 12), Cynomolgus monkey (SEQ ID NO:13 and 14), cow (*e.g.*, SEQ ID NOS:15-18); mouse (SEQ ID NO:19 and 20); rat (SEQ ID NO:21 and 22); rabbit (SEQ ID NO:23 and 24); sheep (SEQ ID NOS:25-27), guinea pig (SEQ ID NO:28 and 29); fox (SEQ ID NO: 30 and 31); Gibbon (SEQ ID NO:856 and 857), Marmoset (SEQ ID NO:858 and 859) and orangutan (SEQ ID NO:860 and 861) . Reference to PH20 includes precursor PH20 polypeptides and mature PH20 polypeptides (such as those in which a signal sequence has been removed), truncated forms thereof that have activity, and includes allelic variants and species variants, variants encoded by splice variants, and other variants, including polypeptides that have at least 40%, 45%, 50%, 55%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more sequence identity to the precursor polypeptides set forth in SEQ ID NO:7, or the mature forms thereof. PH20 polypeptides also include those that contain chemical or posttranslational modifications and those that do not contain chemical or posttranslational modifications. Such modifications include, but are not limited to, PEGylation, albumination,

As used herein, "modified PH20 polypeptide" or "variant PH20 polypeptide" refers to a PH20 polypeptide that contains at least one amino acid modification, such as at least one amino acid replacement as described herein, in its sequence of amino acids compared to a reference unmodified PH20 polypeptide. A modified PH20 polypeptide can have up to 150 amino acid replacements, so long as the resulting modified PH20 polypeptide exhibits hyaluronidase activity. Typically, a modified PH20 polypeptide contains 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 amino acid replacements. It is understood that a modified PH20 polypeptide also can include any one or more other modifications, in addition to at least one amino acid replacement as described herein.

---

[75] EX1026 ('731 App.), 50:31-51:13 (upper quote), 54:5-14 (lower quote).

91. Dr. Simpson testified at her deposition that the first sentence in each paragraph above (highlighted in yellow) is the definition of each term. Regarding the definition of "modified PH20 polypeptide," Dr. Simpson testified that the first sentence of the paragraph defined that term.[76] Dr. Simpson similarly referred to the first sentence of the paragraph defining the term "PH20" in the common disclosure.[77]

Q. Could you go to Page 50 of the common disclosure. It's 1026.
And at the bottom of the page, do you see a definition of the word "PH20" around Line 31?
A. Yes.
Q. And what is the definition of PH20 that the common disclosure provides?
A. (Reading:)
As used herein, PH20 refers to a type of hyaluronidase that occurs in sperm and is neutral-active.

[76] EX1130 (Simpson Dep.), 58:25-59:10 ("A. Yes, the phrase "modified PH20 polypeptide" is defined by the first sentence."), 60:9-13.

[77] EX1130 (Simpson Dep.), 60:14-25.

3. POSAs Would Understand the Term "PH20" Requires Modified PH20 Polypeptides to Be Folded Proteins

92. The common disclosure explicitly defines the word "PH20" to mean "a type of hyaluronidase that occurs in sperm and is neutral-active."[78] A POSA would know that proteins found on the surface of sperm and that are neutral-active are folded forms of PH20.[79] A POSA would thus be directed that the use of "PH20" to describe phrases like "modified PH20 polypeptide" requires that the polypeptides not only meet the "at least 91% sequence identity" and position 317 substitution requirements, but also to have stably folded to adopt the native state (the tertiary structure) of the PH20 protein.

---

[78] EX1026 ('731 App.), 50:31–32. The common disclosure defines "neutral active" as "the ability of a PH20 polypeptide to enzymatically catalyze the cleavage of hyaluronic acid at neutral pH, such as at a pH between or about between pH 6.0 to pH 7.8." EX1026 ('731 App.), 56:21-23.

[79] *See, e.g.*, EX2070 (Petsko Dec.), ¶ 53 ("Enzymatically active modified PH20s within the scope of the claims share a common mechanism of action in that they cleave hyaluronan ("HA") through a substrate-assisted hydrolysis reaction. All share the same active site."), ¶ 128 ("And enzymatically active 91% identical PH20 variants share a common tertiary structure (the PH20 alpha-beta barrel fold) as well as other common structural features…").

93. Dr. Petsko makes this point in his declaration. In ¶ 345 of his declaration, he states that "***POSAs think of PH20 as a folded protein***…" and not just "strings of text that happen to be 91% identical to one another." It is folding that gives rise to the spatial arrangement of amino acids linked by peptide bonds (the sequence). Dr. Petsko makes this point and echoes the definition of the PH20 ***protein*** that the patent uses:[80]

> 345. POSAs understand that PH20 variants with such a small number of substitutions will usually adopt the same general tertiary structure (*i.e.*, the alpha-beta barrel fold) as WT PH20, in the same way that PH20 has a tertiary structure homologous to Hyal-1. POSAs think of PH20 as a folded protein, and POSAs think of 91% identical variants as proteins typically with the same tertiary structure but with 9% variability in the residues in that tertiary structure. POSAs do not think of modified PH20s as 10^60 strings of text that happen to be 91% identical to one another.

I agree with Dr. Petsko that a POSA would understand the presence of the word "PH20" in the phrase "modified PH20 polypeptide" to be indicating that the modified PH20 polypeptides being claimed are folded proteins.

---

[80] EX2070 (Petsko Dec.), ¶ 345.

94. Dr. Simpson also agreed that the common disclosure defines the word "PH20" to be a folded and enzymatically active protein. As she testified: [81]

Q. So when the PH20 is on the sperm surface, is it a folded protein?
A. Yes.
Q. And it's enzymatically active. Correct?
A. In normal biology, surface-associated PH20 would be enzymatically active.
Q. And it's -- because it is properly folded, it is enzymatically active. Right?
MS. MARTIN: Object to form.
A. Properly folded PH20 should have enzymatic activity, yes.

Q. Okay. And this is how the common disclosure defines the term "PH20." Right?
A. PH20 is defined as a type of hyaluronidase that occurs in sperm and is neutral-active.
It explains further that it occurs on the sperm surface. In this particular paragraph, the folded structure is not specified.
Q. But it would just be the folded structure that occurs naturally with the PH20?
A. If PH20 is active on the sperm surface, it is assumed to be folded.

95. The common disclosure also uses the term "PH20" to refer to naturally occurring PH20 proteins from different species, which also would be understood by a POSA to be folded forms of PH20 polypeptides.[82] The common disclosure also uses the term "PH20" to refer to C-terminally truncated forms of the wild-type PH20 protein sequence that have retained activity. That again is using PH20 to indicate that the truncated wild-type sequence has folded.[83]

---

[81] EX1130 (Simpson Dep.), 61:13-62:10.

[82] EX1026 ('731 App.), 50:35-51:5 (listing "[e]xemplary PH20 polypeptides" from various species).

[83] EX1026 ('731 App.), 51:5-7 ("Reference to PH20 includes… truncated forms thereof that have activity…").

96. A modified PH20 polypeptide must preserve many aspects of the tertiary fold of the wild-type PH20 protein to exhibit hyaluronidase activity.[84] In Dr. Petsko's terminology, it must be "properly folded."[85] This includes not just the structure of the active site, but the spatial positioning of particular residues and structural motifs within it, as well as the structural relationships necessary for the protein to bind to and hydrolyze the HA substrate.[86] Only particular amino acid sequences may yield the structure-function relationship that is needed. From the literature before 2011, these include: (a) conserved acidic residues D111, E113 and E249 in the catalytic site, and (b) additional conserved residues including R176, R246 and R252.[87] Dr. Petsko states in his declaration that "[a]ll enzymatically

---

84 EX2070 (Petsko Dec.), ¶ 84 ("Specific binding occurs at sites on the protein that provide the complementarity for the ligand…. Such sites are formed as a consequence of the three-dimensional structure of the protein." (citing EX2164 (Petsko Textbook), 52)).

85 EX2070 (Petsko Dec.), ¶ 128, n.12.

86 EX2070 (Petsko Dec.), ¶¶ 58, 84 ("[A]n enzyme's three-dimensional structure directly affects its catalysis. These active sites create 'specialized microenvironments that result from protein tertiary structure' which 'contribute to catalysis.'" (citing EX2164 (Petsko Textbook), 52)).

87 EX1011 (Arming), 813 ("These three acidic amino acids [D111, E113, E249] lie within clusters of amino acids that are conserved between mammalian and

active modified PH20s within the scope of the claims will cleave HA through the same mechanism of action using the same chemistry carried out by the same active site residues oriented in about the same way."[88]

97. Not all modified PH20 polypeptides that fold will be "properly folded" as Dr. Petsko uses that terminology. A modified PH20 polypeptide may adopt a folded structure that is stable but has differences in its tertiary structure relative to the tertiary structure within the wild-type PH20 protein. As Dr. Hecht accurately explained in his declaration, making multiple changes to an amino acid sequence can alter secondary structures, the positioning of secondary structures, the formation of higher order structural motifs or domains and other features of the protein's three-dimensional shape.[89] Consequently, not all stably folded modified PH20 polypeptides will share the same structure, as the changes introduced may introduce clear and meaningful differences into the tertiary structure of the protein relative to the tertiary structure of the wild-type PH20 protein, including changes that impair the enzymatic activity or stability of the protein.

---

hymenopteran hyaluronidases…. [R252] is close to Glu249, which appears to be part of the active site…. [R176] is also important for substrate binding."); *see also* 813 (Figure 1) (labelling R246 as a "conserved" residue).

88 EX2070 (Petsko Dec.), ¶ 58.

89 EX1003 (Hecht Dec.), ¶¶ 55-61, 141-142.

98. Stable folding is what differentiates a "PH20 polypeptide" from being just a "polypeptide" (a sequence of amino acids of defined length that has not folded into a compact structure). The words "PH20" and "polypeptide" used together mean that the amino acid sequence of the polypeptide has folded to adopt a stable tertiary structure.

99. The definition of the phrase "unmodified PH20 polypeptide" in the common disclosure also supports reading "modified PH20 polypeptide" to mean a stably folded protein (below).[90] That explanation points out that a starting unmodified PH20 polypeptide can either be "a naturally-occurring, wild-type form of a polypeptide" or an "altered or mutated" PH20 that "differs from a native wildtype isoform" but has "been modified to have a desired increase or decrease in a particular activity or property."[91] Because both types of starting, unmodified

---

[90] EX1026 ('731 App.), 54:15-23.

[91] EX1026 ('731 App.), 54:15-27 ("Thus, existing proteins known in the art that have been modified to have a desired increase or decrease in a particular activity or property compared to an unmodified reference protein can be selected and used as the starting unmodified polypeptide. For example, a protein that has been modified from its native form by one or more single amino acid changes and possesses either an increase or decrease in a desired property, such as a change in an amino acid residue or residues to alter

PH20 polypeptides exhibit some hyaluronidase activity, a POSA would have believed both types were stably folded proteins.

> As used herein, an unmodified PH20 polypeptide refers to a starting PH20 polypeptide that is selected for modification as provided herein. The starting polypeptide can be a naturally-occurring, wild-type form of a polypeptide. In addition, the starting polypeptide can be altered or mutated, such that it differs from a native wild type isoform but is nonetheless referred to herein as a starting unmodified polypeptide relative to the subsequently modified polypeptides produced herein. Thus, existing proteins known in the art that have been modified to have a desired increase or decrease in a particular activity or property compared to an unmodified reference protein can be selected and used as the starting unmodified polypeptide. For example, a protein that has been modified from its native form

100. Reading the claims as referring to stably folded and active proteins is consistent with how the common disclosure describes producing modified PH20 polypeptides with multiple substitutions. After defining the phrase "modified PH20 polypeptide," the common disclosure explains that modified PH20 proteins "can have up to 150 amino acid replacements, so long as the resulting modified PH20 polypeptide exhibits hyaluronidase activity" (below).[92] A POSA would understand from the description that it is a natively folded, enzymatically active form of a PH20 polypeptide. After introducing a first modification into that starting PH20 polypeptide (*e.g.*, one of the position 317 substitutions required by the claims), up

---

glycosylation, can be selected for modification, and hence referred to herein as unmodified, for further modification.").

[92] EX1026 ('731 App.), 54:8-10.

to 150 additional changes can be made "so long as the resulting modified PH20 polypeptide exhibits hyaluronidase activity."[93] The POSA would understand that additional changes to the modified PH20 polypeptide are to be selected so as to not impair its ability to stably fold, which is a prerequisite to it retaining hyaluronidase activity. I agree with Dr. Hecht that this passage in the common disclosure supports reading the claims as requiring modified PH20 polypeptides that incorporate multiple substitutions to exhibit hyaluronidase activity.[94]

> As used herein, "modified PH20 polypeptide" or "variant PH20 polypeptide" refers to a PH20 polypeptide that contains at least one amino acid modification, such as at least one amino acid replacement as described herein, in its sequence of amino acids compared to a reference unmodified PH20 polypeptide. A modified PH20 polypeptide can have up to 150 amino acid replacements, so long as the resulting modified PH20 polypeptide exhibits hyaluronidase activity. Typically, a modified PH20 polypeptide contains 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 amino acid replacements. It is understood that a modified PH20 polypeptide also can include any one or more other modifications, in addition to at least one amino acid replacement as described herein.

101. From the definitions and discussions in the common disclosure, a POSA would have read the phrase "modified PH20 polypeptide" in the claims to mean a PH20 polypeptide that (i) has at least one modification relative to the wild-type PH20 sequence and (ii) has stably folded. The presence of the word "PH20"

---

93 EX1026 ('731 App.), 54:8-10.

94 EX1003 (Hecht Dec.), ¶¶ 126–129.

in the definition of "modified PH20 polypeptide" tells POSAs that the modified PH20 polypeptide is a stably folded protein.[95]

4. Dr. Simpson's Interpretation of "Modified PH20 Polypeptide" Effectively Removes the Term "PH20" from the Claims

102. Dr. Simpson testified that she did not believe the claim term "PH20" imposed any requirements on the modified PH20 polypeptides being claimed. As she explained:[96]

Q. I'm trying to understand whether PH20 is imposing any additional requirement beyond the features of the amino acid sequences that are recited in the claims.

MS. MARTIN: Object to form.

A. Because there is a specific sequence given, the term "PH20" doesn't impose any greater restraints. That is the sequence of PH20 in a modified polypeptide, as defined by the claims and by the patent.

Q. So if I'm visualizing claim 1 as defining a set of amino acid sequences -- is that term conceptually accessible to you, that the claim is defining a set of amino acid sequences that share these features? Is that -- you understand that?

A. Yes.

Q. If I were to remove the word "PH20" from the first line of the claim, would you see a different set of amino acid sequences relative to what is there in the claim?

MS. MARTIN: Object to form. Objection; scope.

A. The claim is on the basis of sequences that are identified and given. The sequences are absolute.

The nomenclature of "modified PH20 polypeptide" helps a POSA understand what the context is for those sequences and where they're found biologically, but the sequence structure itself doesn't necessarily require the term "PH20" to define it.

And the patent claim would not change if the word "PH20" wasn't there, but it is defined within the disclosure of the patent. So it's part of the specific phrase used to define this polypeptide sequence and refer to it throughout the patent disclosure.

103. In other words, Dr. Simpson is reading the phrase "modified PH20 polypeptide" to have the identical meaning as the phrase "modified polypeptide."

---

95 EX2070 (Petsko Dec.), ¶ 345 ("POSAs think of PH20 as a folded protein…").

96 EX1130 (Simpson Dep.), 55:23-56:13, 57:9-58:4.

This reading effectively strikes the word "PH20" out of the claims as illustrated below:

What is claimed:

1. A modified ~~PH20~~ polypeptide, comprising one or more amino acid modifications in an unmodified ~~PH20~~ polypeptide, wherein:

the unmodified ~~PH20~~ polypeptide consists of the amino acid sequence selected from the group consisting of SEQ ID NOs: 3, 7 and 32-66;

amino acid modifications are selected from the group consisting of amino acid replacements(s), deletion(s), and/or insertion(s);

the modified ~~PH20~~ polypeptide comprises an amino acid replacement at a position corresponding to residue 317, with reference to amino acid positions set forth in SEQ ID NO:3;

the replacement at the position corresponding to residue 317 is selected from the group consisting of A, I, K, M, Q, and R;

corresponding amino acid positions are identified by alignment of the ~~PH20~~ polypeptide with the polypeptide having the amino acid sequence of SEQ ID NO:3; and

the modified ~~PH20~~ polypeptide has at least 91% sequence identity to a polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NOs: 3, 7 and 32-66.

104. I have been informed that it is legally improper to read a patent claim in a way that makes words used in the claim superfluous or unnecessary. This seems particularly relevant here because "PH20" is an integral component of the phrase "modified PH20 polypeptide." PH20 is expressly defined in the common disclosure to have a form a POSA would understand is folded and is enzymatically active. I therefore disagree with Dr. Simpson that the claims cover every amino

acid sequence that meets the 91% sequence identity plus position 317 substitution requirements of the claims (*i.e.*, 10^{112} proteins), regardless of the folded state of those amino acid sequences. Instead, the defined meaning of PH20 would have led a POSA to conclude that "modified PH20 polypeptides" are those that (i) contain at least one amino acid modification and (ii) are properly folded proteins, as that is the form that the PH20 protein exists in on the surface of sperm cells.

E. The Claims as a Whole Define Enzymatically Active Modified PH20 Polypeptides

105. Dr. Simpson expressed her opinion that a POSA would not have read the claims of the patent to require the modified PH20 polypeptides to have enzymatic activity.[97] Her reasons include:

(i) Her belief that the claims require nothing more than an amino acid sequence that meets the 91% sequence identity and position 317 substitution requirements in the claims, and they therefore claim

[97] Dr. Petsko stated in his declaration that he was not asked to opine on claim construction and that he did not do so in his declaration. EX2070 (Petsko Dec.), ¶ 246.

between ~10^{59} and ~10^{112} distinct amino acid sequences (*see* ¶¶ 53-55 above).[98]

(ii) Her belief that the definition of "modified PH20 polypeptide" by itself does not require enzymatic activity.[99]

(iii) Her belief that Dr. Hecht incorrectly portrayed "active mutant" forms of modified PH20 polypeptides as having at least 40% hyaluronidase activity.[100]

(iv) Her belief that the common disclosure indicates that modified PH20 polypeptides can be "inactive mutants" and lists examples of single substitution mutants classified as "inactive mutants."[101]

(v) Her belief that it is possible to create inactive mutants by inserting single amino acid substitutions that yielded "active mutants" into truncated forms of the wild-type PH20 that are inactive.[102]

---

[98] Dr. Petsko views the 91% sequence identity and position 317 substitution requirements to be the only "common structural features" shared by all of the claimed modified PH20 polypeptides. EX1131 (Petsko Dep.), 65:12-66:17.

[99] EX2068 (Simpson Dec.), ¶¶ 259, 263-264

[100] EX2068 (Simpson Dec.), ¶¶ 256-258, 262.

[101] EX2068 (Simpson Dec.), ¶¶ 267-268, 273, 324-328.

[102] EX2068 (Simpson Dec.), ¶ 329.

106. I do not believe Dr. Simpson accurately portrayed Dr. Hecht's views and how different requirements in the claims work together to define a subset of modified PH20 polypeptides that is narrower than the definition of "modified PH20 polypeptide" considered in isolation. Most significantly, she did not account for the term "PH20" being used in the claim language.[103]

107. Dr. Simpson's analysis does not account for how a POSA would have evaluated the claims relative to the common disclosure. For example, she did not account for the common disclosure's classification of the mutants into one of two mutually exclusive types of modified PH20 polypeptides (*i.e.*, "active mutants" with greater than 40% hyaluronidase activity relative to unmodified PH20, and "inactive mutants" with less than 20% relative hyaluronidase activity).[104] She also did not address the requirement that the particular position 317 substitutions required by the claims (*i.e.* A, I, K, M, Q, R) are all substitutions that resulted in active mutants when incorporated into SEQ ID NO: 3.

108. I believe Dr. Hecht more accurately describes why a POSA would have read the claims as a whole, as being directed to modified PH20 polypeptides

---

[103] EX1130 (Simpson Dep.), 55:23-56:13, 57:9-58:4.

[104] *See* EX1003 (Hecht Dec.), ¶¶ 98-102; EX1026 ('731 App.), 251:1-20, 272:2-22.

that are enzymatically active.[105] I also understand his explanations to include the following observations.

a. The common disclosure depicts two, mutually exclusive types of modified PH20 polypeptides: (i) enzymatically active modified PH20 polypeptides (*i.e.*, "active mutants"), and (ii) enzymatically inactive modified PH20 polypeptides (*i.e.*, "inactive mutants").[106] It also differentiates these two types of modified PH20 polypeptides based on the relative hyaluronidase activity of the mutants compared to unmodified $PH20_{1-447}$ (*i.e.*, "active mutants" exhibit at least 40% relative hyaluronidase activity in at least one of two assays, "inactive mutants" exhibited less than 20% relative hyaluronidase activity).[107]

b. The claims include requirements beyond what is required by the definition of a modified PH20 polypeptide (*i.e.*, "…at least one amino acid modification").[108] Specifically, each claim requires the "modified PH20 polypeptide" to contain an amino acid

---

[105] *See* EX1003 (Hecht Dec.), ¶¶ 126-134.

[106] *See* EX1026 ('731 App.), 86:15-20, 130:30-131:3; EX1003 (Hecht Dec.), ¶¶ 98-102.

[107] *See* EX1003 (Hecht Dec.), ¶¶ 103-107; EX1026 ('731 App.), 251:1-20, 272:2-22.

[108] EX1026 ('731 App.), 54:5-8.

substitution at position 317 from leucine (L) to one of six alternative amino acids (A, I, K, M, Q, R).[109]

c. The common disclosure classifies modified PH20 polypeptides having the L317A, L317I, L317K, L317M, L317Q, or L317R substitutions as "active mutants" (*i.e.*, mutants that retained >40% enzymatic activity of unmodified $PH20_{1-447}$). Those six substitutions, for example, are listed only in Tables 3 and 9, which compile lists of "active mutants" (below).[110]

TABLE 9: ACTIVE MUTANTS

mutant	SEQ ID NO	AvgNorm Act.	mutant	SEQ ID NO	AvgNorm Act.	mutant	SEQ ID NO	AvgNorm Act.
						L317A	604	1.27
						L317D		0.61
						L317H		1.05
						L317I	605	1.76
						L317K	606	5.11
						L317M		1.20
						L317N	607	0.73
						L317Q	608	1.67
						L317R	609	2.41
						L317S	610	1.03
						L317T	611	0.93
						L317W	612	0.84

TABLE 3: Active Mutants

Corresponding Position	Replacement
317	A D H I K M N Q R S T W

d. The common disclosure explains that a modified PH20 polypeptide can have modifications in addition to a first one (here the position 317 substitution). It then explains when an active mutant is further modified to have additional changes,

[109] EX1001 ('758 Patent), 313:28-34.

[110] EX1026 ('731 App.), 95 (Table 3), 254 (Table 9).

those changes are selected to preserve hyaluronidase activity (*i.e.*, "[a] modified PH20 polypeptide can have up to 150 amino acid replacements, so long as the resulting modified PH20 polypeptide exhibits hyaluronidase activity."). [111] That is the case for the modified PH20 polypeptides being claimed—they all have at least a position 317 substitution that retained >40% hyaluronidase activity, which causes them to be classified as "active mutants," and each may have between zero and 1 or 41 additional modifications.

109. I recognize that the definition of "modified PH20 polypeptide" in the common disclosure considered in isolation does not explicitly state that each modified PH20 polypeptide is enzymatically active. I did not understand Dr. Hecht's opinion to be based solely on the words "modified PH20 polypeptide" in isolation, and I am not basing my opinion solely on those words. Instead, I believe that when a POSA considers all of the language used in the claims and does that in conjunction with the common disclosure, she would have concluded that the claims are directed to the "active mutant" alternative of "modified PH20 polypeptides," which are forms of modified PH20 polypeptides that exhibit hyaluronidase activity.

---

[111] EX1026 ('731 App.), 54:8-10.

110. I understand that determining what the claims define is a legal question, which is to be answered by judges. I am not a lawyer and do not have legal training. I understand the panel conducting this proceeding could conclude that the claims are not limited to modified PH20 polypeptides that have stably folded and possess hyaluronidase activity. That would mean the claims include additional forms of modified PH20 polypeptides (*e.g.*, unfolded, misfolded, and partially folded PH20 polypeptides, as well as higher order aggregates those types of sequences form). The claims would still include modified PH20 polypeptides that have stably folded to adopt the native state of the PH20 protein and thus possess hyaluronidase activity. That set of modified PH20 polypeptides will have distinct structural characteristics from the other sets of modified PH20 polypeptides I referenced above.

III. Response to Dr. Petsko's and Dr. Simpson's Opinions on the Common Disclosure's Description of the Claimed Modified PH20 Polypeptides

111. Drs. Petsko[112] and Simpson[113] suggest in their declarations that a POSA could use the mutagenesis data reported in the common disclosure to identify which of the ~10^{59} and ~10^{112} different amino acid sequences within the

---

[112] EX2070 (Petsko Dec.), ¶ 55 (stating the "patent provides the information a POSA requires to identify both single and multiple mutations that are likely to produce the desired activity"), ¶ 293 (stating the "patent renders modifying PH20 predictable by identifying these areas that are particularly intolerant to mutation, and those that are tolerant"), ¶ 294 (stating that with the patent, "the POSA can identify which regions are likely to tolerate change and which are not"); ¶ 297 (POSA "would have considered the activity data in the context of the PH20 homology model to help determine which positions are amenable to change and to identify additional variants"), ¶ 348 (stating the POSA will use the patent "to obtain modified PH20 polypeptides with the desired properties"), ¶ 423 (use of disclosure to identify substitutions that maintain/abrogate PH20 tertiary structure and activity).

[113] EX2068 (Simpson Dec.), ¶ 303 (patent teaches which residues "are necessary to enzymatic activity"), ¶ 374 ("Dr. Petsko provides numerous examples of how a POSA could have used the patent data to identify potential single and multiply substituted PH20s…"); ¶ 379 (noting that the patent teaches where a "rescue mutation" may be "introduced to restore activity"), ¶ 382 (use of data to "identify mutation targets"); ¶ 387 (use of data "to generate multiply modified PH20s with combinations of 5, 10, 15 or more substitutions").

claim parameter sequence space with combinations of more than a few changes will "properly fold" and exhibit enzymatic activity. They also both suggest that the common disclosure "teaches a correlation between the structure of modified PH20 and the hyaluronidase function of modified PH20."[114] They suggest that a set of 7 examples of sequences with more than one change in them is representative of the structural diversity of either the set of sequences that will fold and exhibit enzymatic activity, or the broader set of sequences within the claim parameter sequence space.[115] For the reasons I provide below, I disagree with those opinions.

A. It Was Impossible for POSAs to Determine from Information in the Common Disclosure Which Modified PH20 Polypeptides in the Claim Parameter Sequence Space Will Fold and/or Be Enzymatically Active, or Soluble or Able to Increase Delivery of a Therapeutic

112. Dr. Park's calculations showed (and Drs. Petsko and Simpson agreed) that there are between ~10^{59} and ~10^{112} different amino acid sequences within the sequence space defined by the sequence identity claim parameters.[116] As Dr. Hecht explained throughout his declaration, many of these sequences would include

---

[114] EX2070 (Petsko Dec.), ¶ 54; *see also* ¶¶ 58, 422; EX2068 (Simpson Dec.), ¶¶ 156, 159, 163, 305-315, 376, 384.

[115] EX2070 (Petsko Dec), ¶ 413–421; EX2068 (Simpson Dec.), ¶¶ 337–340.

[116] *See* ¶¶ 47-49.

changes that can lead to a large number of consequences that disrupt the protein's structure. The vast majority of this space is made up of sequences that have incorporated more than 5 changes.[117] Not all of those sequences will fold into a stable active native fold—some will misfold or form into aggregates. Those that do adopt the native fold may exhibit significantly reduced enzymatic activity due to subtle changes in the substrate binding or catalytic sites, and some that natively fold and are active may unfold or degrade to less than 40% of activity on relevant timescales. Some of the mutants that retain measurable enzymatic activity may not be soluble, or be properly secreted, or increase delivery of a therapeutic agent. It was not possible for a POSA to predict using the information available at the time and in the disclosure which of the mutants would or would not possess these functional properties.[118] Below I respond to comments made by Drs. Petsko and Simpson.

1. Dr. Petsko's 7 Examples Are Not Representative of the Diversity of the Claimed Modified PH20 Polypeptides

113. Dr. Hecht illustrated the structural diversity of the claim parameter sequence space that is made up of multiply modified PH20 polypeptides in his

---

[117] *See* ¶ 51.

[118] EX1003 (Hecht Dec.), ¶¶ 57–60, 141–142, 155–156.

declaration.[119] He identified many examples of combinations of substitutions, deletions and additions in sequences of different length that were very different than the modified PH20 polypeptides compiled in the common disclosure. I agree with Dr. Hecht's illustrations of this structural diversity of the claim parameter sequence space.

114. In response, Drs. Petsko and Simpson identified seven examples of PH20 polypeptides with more than one change:

a. Drs. Petsko and Simpson discuss four PH20 mutants in their declarations. These four mutants were not actually made or tested. They are hypothetical mutants made by inserting the single substitution (*i.e.*, L317Q) in a set of four alanine-substituted human PH20 mutants that were discussed in US-457.[120] The four alanine-substituted PH20 mutants included one with three alanine substitutions and three with two alanine substitutions)..[121] Inserting the position 317 single substitution

---

[119] EX1003 (Hecht Dec.), ¶¶ 116–125, 145.

[120] EX2068 (Simpson Dec.), ¶¶ 337-340 (citing Table 16 in US-457 as reporting three alanine-substituted double mutants (N82A/N166A, N82A/N254A, N166A/N254A) and one triple mutant (N82A/N166A/N254A) retained activity).

[121] EX2070 (Petsko Dec.), ¶¶ 214–217, 252, 419.

into this set of four alanine-substituted mutants yields four total mutants.

TABLE 16

Human PH20 Hyaluronidase Deglycosylation Mutants

Mutant	SEQ ID NO	Glycan Type	Secretion	Activity pH 5.5	Activity pH 7.4
PH20sHis (parent)	187	Both	YES	YES	YES
N82A	202	Complex	YES	YES	YES
N166A	203	Complex	YES	YES	YES
N235A	204	High Mannose (80%) Complex (20%)	YES	NO	NO
N254A	205	Complex	YES	YES	YES
N368A	188	High Mannose	NO	NO	NO
N393A	189	High Mannose	YES (WEAK)	NO	NO
N82A/N166A	206	Complex	YES	YES	YES
N82A/N254A	207	Complex	YES	YES	YES
N166A/N254A	208	Complex	YES	YES	YES
N82A/N166A/ N254A	209	Complex	YES	YES	YES
N368A/N393A	190	High Mannose	NO	NO	NO

b. Drs. Petsko and Simpson discuss three additional PH20 mutants in their declarations. These three mutants were not actually made or tested. They are hypothetical mutants derived by introducing the single substitution (*i.e.*, L317Q) into a set of three primate PH20 sequences—chimpanzee (5 differences), orangutan (13 differences), and gibbon (16 differences) relative to human PH20.[122]

[122] EX2070 (Petsko Dec.), ¶¶ 361–364.

115. These limited examples do not meaningfully probe the combinatorial sequence space covered by the claims. They are also not representative of the breadth of modified PH20 polypeptides within the scope of the claims. Dr. Hecht addresses these deficiencies in his declaration. I agree with his criticisms that:

a. The common disclosure identifies only six $PH20_{1–447}$ variants that fall within the scope of the claims—L317A, L317I, L317K, L317M, L317Q and L317R—and each of these is a single-substitution mutant, not a multiply modified PH20 polypeptide[123];

b. The common disclosure provides no examples of $PH20_{1–447}$ polypeptides containing the claimed L317 substitution in combination with ANY additional substitutions[124];

c. From the single-substitution mutants the common disclosure reports—including L317A, I, K, M, Q and R—it does not attempt to relate or characterize the effect on the PH20 structure[125];

d. Single-substitution $PH20_{1–447}$ mutants are not representative of multiply modified PH20 polypeptides because multiple substitutions can cumulatively disrupt secondary and tertiary

---

[123] EX1003 (Hecht Dec.), ¶¶ 140-143.

[124] EX1003 (Hecht Dec.), ¶¶ 140-143.

[125] EX1003 (Hecht Dec.), ¶¶ 55-56, 61, 69, 141-142.

structure through local and long-range effects not revealed by any single mutation and the common disclosure provides no multi-substitution PH20 variants that demonstrate structural preservation across the claimed ranges.[126]

e. Properties of a single protein were not reliably predictable from properties observed in proteins each having only one of the mutations in the relevant timeframe, because the combined effects of multiple mutations do not follow from the individual mutations in a simple or additive manner.

116. Dr. Petsko suggests a POSA would expect that if the L317Q substitution was introduced in three primate sequences, they would properly fold and be enzymatically active.[127] He suggests these would be three additional examples of a modified PH20 polypeptide with 6, 14 or 17 total changes. Dr. Petsko also bases his opinions on homology models of these sequences he created in Swiss-Model using the Hyal-1 structure as the template. He provides a cursory review of these models. For example, he shows images of the models and observes

[126] EX1003 (Hecht Dec.), ¶¶ 141-142, 144-145.

[127] EX2070 (Petsko Dec.), ¶¶ 361-365, 413-421 (generating models of primate PH20 proteins with the L317Q substitution introduced).

that each has an "alpha beta barrel fold" motif and that each preserves positioning of "key active region residues."[128]

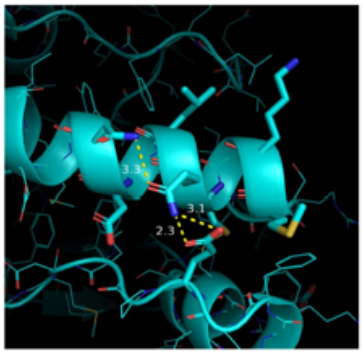
117. None of these proteins that Dr. Petsko described were made or tested. It is important to note that they represent special cases. Dr. Petsko's experiment of introducing the L317Q change in chimp, orangutan or gibbon PH20 is simply introducing a single mutation into a naturally occurring sequence that has been shaped by evolution to stably fold and be active. Dr. Petsko's hypothetical mutants represent an infinitesimally small portion of the patent's enclosed sequence space. There is no teaching in the patent or modelling software that would allow sets of 21 combined substitutions that differ by more than around five changes from a known stably folded active PH20 to be predictably active (or soluble or exhibit the ability to increase delivery of a therapeutic). As I explain in §§ IV.C and IV.E, homology models have limits in their accuracy and cannot predict the epistatic effects which were well known to confound predictions of the properties of multiply modified mutant proteins based on properties observed in different mutants having only a single substitution. *See* ¶¶ 67-72.

118. Dr. Petsko comments on an interaction that would occur with E31 and N321 in the human and three primate sequences if the leucine (L) at position 317 is

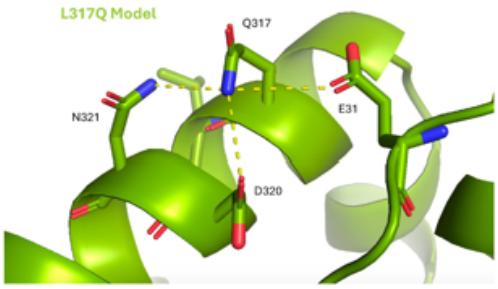
---

[128] EX2070 (Petsko Dec.), ¶¶ 417, 421.

changed to glutamine (Q).[129] All six PH20 proteins have L at 317, and glutamic acid (E) at position 31. Dr. Park determined in his analysis that an L to Q substitution at position 317 of PH20 would introduce a favorable stabilizing change to the PH20 structure, because the substituted glutamine (Q) could form strong hydrogen bonds with the glutamic acid (E) at position 31 and the asparagine (N) at position 321.[130] Dr. Petsko does not cite Dr. Park's analysis of this change, but their analyses agree.



Park



Petsko

119. I do not believe a POSA would consider the set of 7 hypothetical mutants Dr. Petsko discussed are at all representative of the diversity of enzymatically active modified PH20 polypeptides being claimed. That person also would not believe those 7 are in any way representative of the sequences within the

[129] EX2070 (Petsko Dec.), ¶ 404.

[130] EX1004 (Park Dec.), ¶¶ 112-119; EX2070 (Petsko Dec.), ¶ 404.

claim parameter sequence space. They do not meaningfully probe physicochemical changes within the PH20 protein structure or mutants with truncations or changes beyond position 447, for which no data is provided. Other than these 7 hypothetical mutants, neither Dr. Petsko nor Dr. Simpson identify or discuss any multiply modified PH20 polypeptide within the claim parameter sequence space in their declarations, particularly ones having 5, 10, 15 or 20 changes. Additionally, because the modified PH20 polypeptides can be based on SEQ ID NO: 7 (which contains the GPI anchor sequence), a POSA would expect them to be membrane bound.[131] A POSA would expect additional challenges and complexity in identifying, producing and assessing modified PH20 polypeptides. The common disclosure does not describe any examples of enzymatically active forms of modified PH20 polypeptides that include the GPI anchor sequence present in SEQ ID NO:7, and it provides no guidance about how a POSA would overcome these

---

[131] EX1013 (Frost) 429-430 ("The human PH20 enzyme was believed to be anchored to the plasma membrane using hydropathy plots, and the release of acid-active hyaluronidase activity from cells expressing full-length PH20 was achieved following phospholipase C treatment)"); EX1029 (Gmachl), 547-548 ("Like the guinea pig PH-20 protein, its human homologue is also bound to the membrane by a GPI-anchor."); EX2165 (US-457), ¶¶ 0109, 0120; EX1005 ('429 Patent), 2:56-57.

additional challenges and complexities when producing forms of modified PH20 polypeptides that are membrane bound (*e.g.*, assays that identify active mutants that are secreted would not identify membrane-bound forms of modified PH20s).

120. Drs. Petsko and Simpson suggest that a POSA could use Dr. Petsko's tolerance assessment, information about particular single substitution mutants in the patent, their knowledge and a PH20 homology model, to select and combine any combination of tested mutations, create a model of the mutated sequence with these multiple changes using Swiss-Model, assess the model and determine if each would yield a stably folded protein that is the same as the native PH20 protein. I do not think a POSA could have done that for the vast majority of combinations in the sequence space. It would require an impossible amount of time and effort to do that individualized assessment of the ~10^{59} to 10^{112} possible combinations of mutations in the claim parameter sequence space.

2. Dr. Petsko Confirms the Common Disclosure Does Not Identify the Enzymatically Active Multiply-Modified PH20 Polypeptides Being Claimed

121. At his deposition, Dr. Petsko confirmed that the claims include in their scope enzymatically active, properly folded modified PH20 polypeptides.[132] He was then asked what fraction of the 10^{59}+ different amino acid sequences within

[132] EX1131 (Petsko Dep.), 61:19-62:8.

the claim parameter sequence space would properly fold and be enzymatically active modified PH20 polypeptides.[133] He said that he did not know how many were nor was he able to even give an approximate estimate:[134]

> In the genus of polypeptides that are defined by claim 1, would you agree that there are enzymatically active, properly folded, modified PH20 polypeptides?
> A. Yes.
> Q. Okay. And so some number of those 1049-plus polypeptide sequences will be properly folded PH20 proteins that are enzymatically active?
> A. Yes.
> Q. Do you know how many?
> A. No.
> Q. Do you have an approximation that you can come up with of how many might be?
> A. Not an approximation, no.

122. Dr. Petsko was asked how a POSA could identify what portion of the set of 10^{59}+ modified PH20 polypeptides within in the scope of the claim would properly fold and be enzymatically active. He explained that a POSA would not attempt to do that because "there isn't a simple way to do that."[135] He then provided the reason why a POSA could not know what that set of modified PH20 polypeptides would be. It was not possible for a POSA to either calculate or

---

[133] Dr. Petsko's answers during his deposition were directed towards the related U.S. Patent 11,952,600, which has 10^{49} to 10^{66} possible combinations of mutations in its claim parameter sequence space. I refer to the larger 10^{59}+ sequence space within the claim parameter sequence space of the '758 Patent.

[134] EX1131 (Petsko Dep.), 61:19-62:8.

[135] EX1131 (Petsko Dep.), 63:19-64:8.

experimentally determine which of the 10^{59}+ different modified PH20 polypeptides meeting the sequence parameters would properly fold and be enzymatically active in the 2011-2012 timeframe:[136]

> If someone were asked in 2011 or 2012 and they were a person of ordinary skill in the art, and they were asked to determine the number of PH20 polypeptides in that genus of claim 1, that were properly folded and enzymatically active, how would they go about doing that?
>
> MS. MARTIN: Object to form.
>
> A. They wouldn't even go about doing that.
>
> BY MR. KUSHAN:
>
> Q. Can you -- again, my question is, I am asking that person to do that. And I am asking you to tell me how they would do that.
>
> A. What I'm saying is that would be their answer. They would say, I am not -- I won't do that because there isn't a simple way to do that.
>
> Q. Is it possible to calculate that number, in 2011, 2012?
>
> A. I don't believe so.
>
> Q. Is it possible to experimentally determine which ones are properly folded and enzymatically active in 2011 or 2012?
>
> MS. MARTIN: Object to form.
>
> A. Not from experimenting on the entire ensemble.

123. I agree with Dr. Petsko and Dr. Hecht that it would have been literally impossible in the 2011-2012 timeframe for a POSA to calculate or experimentally determine which of the 10^{59}+ modified PH20 polypeptides with more than about 5 modifications that meet the sequence identity claim parameters would properly fold and be enzymatically active or be soluble, or increase delivery of a therapeutic.[137] A POSA could not have even approximated the size of that fraction of the claim parameter sequence space using the information in the common disclosure in combination with their skills, knowledge and the tools available in the 2011-2012 timeframe. I do not believe a POSA in the 2011-2012 timeframe could

---

[136] EX1131 (Petsko Dep.), 63:19-64:17.

[137] EX1131 (Petsko Dep.), 63:19-64:17; EX1003 (Hecht Dec.), ¶¶ 158-159, 170.

have visualized from the common disclosure the set of modified PH20 polypeptides with more than about 5 changes within the claim parameter sequence space that will fold and be enzymatically active or be soluble, or increase delivery of a therapeutic. While evaluating combinations of two or three substitutions might have been considered a manageable and reasonable undertaking, by the time one reaches five substitutions, the combinatorial complexity and interdependence among residues make reliable prediction of which variants would retain enzymatic activity—and thus increase delivery—effectively impossible. I also do not believe it is possible to read the common disclosure in a way that demonstrates the inventors were in possession of that subset of properly folded, enzymatically active modified PH20 polypeptides within the claim parameter sequence space when the applications were filed in 2011 and 2012. Since Dr. Petsko cannot approximate what portion of the claim parameter sequence space will fold and be enzymatically active or be soluble, or increase delivery of a therapeutic, no POSA could visualize and have possession of all of those modified PH20 polypeptides.

124. Dr. Petsko's practical answer that a POSA would not have even tried to make modified PH20 polypeptides with more than about 5 modifications reflects the practical limitations of experimental and computational capabilities in the 2011-2012 timeframe for producing multiply modified PH20 polypeptides. I agree. For example, even the largest scale mutagenesis and screening techniques available

in that timeframe could not screen more than around ~10^{12} mutants, which represent an extremely small fraction of the sets of sequences captured by the 91% sequence identity parameters of the claims (roughly in proportion to one grain of sand to the planet earth, ~10^{30}).[138] A June 2013 review of large scale screening technologies used with enzymes placed an upper limit of about 10^9 sequences:[139]

Table 1. Characteristics of general *in vivo* screening and selection strategies. It is the product of the enzymatic reaction or the enzymatic conversion itself that needs to be detected, making these strategies reaction dependent (in contrast to most reporter-based strategies). Requirements encompass everything except basic microbial techniques/reagents and substrate. m, metagenomic library; ev, enzyme variant library. Adapted from [3].

Strategy	Library type	Library size	Requirements	Advantages	Disadvantages
Selection	m/ev	~ 10^9	None	Only desired variants	Low activity variants might be missed
				Simple to operate	Difficult to quantify
					Higher false positive risk
Agar plate screening	m/ev	~ 10^5	Reagents[a]	Simple to operate	Low dynamic range
			Detection device[a]		Difficult to quantify
Microtiter plate screening	m/ev	~ 10^4	Detection device	Multiple analytical detection methods possible	
				Increased sensitivity by cell lysis	
				Large dynamic range	
Cell-in-droplets (screening)	m/ev	~ 10^9	Detection device	Pico/femto-litre scale (not much reagent needed)	Product should stay in the droplet
			Cell sorter		Technically challenging
Cells as micro-reactors (screening)	m/ev	~ 10^9	Detection device	Femto-litre scale (not much reagent needed)	Product should stay in the cell
			Cell sorter	No compartmentalization needed	Technically challenging
Cell surface display (screening)	ev	~ 10^9	Detection device	Substrate can stay extracellular	Product needs to attach to cell surface
			Cell sorter		Technically challenging
					Careful selection of host and anchoring motif needed
					Fusion can alter enzyme structure/activity

[a] In some cases.

[138] EX1141 (Goldsmith), 409 (describing limit of "ultrahigh throughput" screening techniques being 10^6 to 10^{12} mutants).

[139] EX1147 (Van Rossum), 2991 (Table 1).

125. There also were significant challenges in designing and carrying out large scale screening and selection experiments in the 2011-2012 timeframe. These challenges were summarized concisely in a 2012 review by Goldsmith (EX1141):[140]

> **To screen or select?**
> The main dilemma facing those designing directed enzyme evolution experiments is whether to apply a high, or even ultrahigh throughput selection methodology (10^6 to $>10^{12}$ variants processed in parallel), or a screen that might be of much lower throughput (10^2–10^4 variants tested individually). Ultra or high throughput screens or selections become advantageous with libraries of variants generated by extensive random mutagenesis or by simultaneous saturation mutagenesis of many positions. Such libraries need to be applied when: firstly, the available phylogenetic and/or structural information is insufficient for the design of small and focused libraries. Secondly, the identity and position of the mutations required to induce the desired property are hard to predict. Thirdly, the target change in activity depends on highly epistatic mutations that demand the simultaneous exploration of more than one position. In these, and possibly in other cases, large libraries might be the method of choice.

126. One challenge Goldsmith notes is that a "target change in activity depends on highly epistatic mutations that demand the simultaneous exploration of more than one position."[141] That challenge reveals the shortcomings of the limited guidance in the common disclosure about possible designs for an iterative

---

[140] EX1141 (Goldsmith), 409.

[141] EX1141 (Goldsmith), 409.

mutagenesis strategy. For example, the common disclosure proposes to target positions identified in a first library that were identified as being deleterious in a first library (below).[142] But because of epistasis, combinations generated in a second or subsequent library may alter the effect of the first mutation.

> The further modifications can be targeted near particular regions (*e.g.*, particular amino acid residues) associated with activity and/or stability of the molecule. For example, residues that are associated with activity and/or stability of the molecule generally are critical residues that are involved in the structural folding or other activities of the molecule. Hence, such residues are required for activity, generally under any condition. Critical residues can be identified because, when mutated, a normal activity of the protein is ablated or reduced. For example, critical residues can be identified that, when mutated in a hyaluronan-degrading enzyme, exhibit reduced or ablated hyaluronidase activity under a normal or control assay condition. A further library of modified proteins can be generated with amino acid mutations targeted at or near to the identified critical amino acid residues, such as adjacent to the identified critical amino acid residues. In some examples, the mutations can be amino acid replacement to any other of up to 19 other amino acid residues. The secondary library can be tested using the assays and methods described herein.

127. The illustration below shows the growth of the size of the smallest sequence space defined by the sequence identity claim parameters (*i.e.*, reference sequence SEQ ID NO:32 (1-430), 1 choice at position 317, and 21 total changes). The orange row identifies the upper limits of screening capabilities (~4 to 5 mutations per PH20). This illustration does not use the parameters of this claim set,

[142] EX1026 ('731 App.), 150:14-26.

but is used to demonstrate the scale of possibilities relative to the capability of the library to screen the possibilities.

Total Change	Total Sequences	#
1	1	1.00E+00
2	8,152	8.15E+03
3	33,150,118	3.32E+07
4	89,660,073,504	8.97E+10
5	181,449,739,545,075	1.81E+14
6	293,077,978,086,132,000	2.93E+17
7	393,555,483,371,437,000,000	3.94E+20
8	451,914,977,218,566,000,000,000	4.52E+23
9	452,989,259,911,317,000,000,000,000	4.53E+26
10	402,658,125,051,895,000,000,000,000,000	4.03E+29
11	321,362,356,487,055,000,000,000,000,000,000	3.21E+32
12	232,608,649,507,388,000,000,000,000,000,000,000	2.33E+35
13	153,968,078,195,545,000,000,000,000,000,000,000,000	1.54E+38
14	93,849,823,863,830,800,000,000,000,000,000,000,000,000	9.38E+40
15	52,991,852,975,982,400,000,000,000,000,000,000,000,000,000	5.30E+43
16	27,859,708,843,273,000,000,000,000,000,000,000,000,000,000,000	2.79E+46
17	13,698,336,949,198,000,000,000,000,000,000,000,000,000,000,000,000	1.37E+49
18	6,323,829,801,952,950,000,000,000,000,000,000,000,000,000,000,000,000	6.32E+51
19	2,750,529,891,360,250,000,000,000,000,000,000,000,000,000,000,000,000,000	2.75E+54
20	1,130,619,221,191,820,000,000,000,000,000,000,000,000,000,000,000,000,000,000	1.13E+57
21	440,435,474,482,722,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000	4.40E+59

128. Computational tools available in the 2011-2012 timeframe also were not capable of identifying and predicting the myriad types of subtle disruptions to a protein structure that collectively determine if a particular mutated sequence will properly fold and remain stably folded.[143]

[143] *See* EX1003 (Hecht Dec.), ¶¶ 49–50; EX1004 (Park Dec.), ¶ 159 ("[A] skilled artisan in 2011 could not have used SWISS-MODEL to reliably predict the effects on the structure of a modified PH20 polypeptide of many substitutions (*e.g.*, 5 or more) in the PH20 sequence, if each position is allowed to vary independently to one of several amino acids.").

129. Dr. Petsko confirmed that it would not have been possible for a POSA to predict the change to a modified PH20 polypeptide's free energy of stabilization that would be associated with 10, 15 or 20 mutations.[144] This is another reason why a POSA could not have predicted which of the 10^{59}+ sequences with 10, 15 or 20 mutations within the claim parameter sequence space would fold to adopt the tertiary structure of the native state of the PH20 protein. If a POSA cannot accurately predict the effect of a set of mutations to a protein's free energy of stabilization (and particularly the overall magnitude of those effects), that person cannot predict whether that set of mutations will collectively increase, have no net effect on or decrease the stability of the folded protein (and if it folds). As Dr. Petsko's textbook explains, "the net free energy of stabilization of most folded proteins —the difference in free energy between the folded and unfolded states — is actually rather small, about 21-42 kJ/mole, or only about 10 times the average thermal energy available at physiological temperature."[145] Changes that cause even modest reductions in the net free energy of stabilization of a modified PH20

---

[144] EX1131 (Petsko Dep.), 100:10–101:5 ("[T]here are programs that will allow you to do that, but the numbers are very imprecise, so I would normally not do that.").

[145] EX2164 (Petsko Textbook), 26.

polypeptide can slow or prevent folding or result in rapid unfolding of the protein after it has folded and/or been secreted.

130. Dr. Petsko suggested a POSA would not attempt to make such mutants because they could achieve their objectives by making mutants with fewer changes.[146] A more accurate answer would have been that a POSA could not have used the information in the common disclosure, what was known in the 2011-2012 timeframe, and a PH20 homology model to calculate or experimentally determine what portion of the set of modified PH20 polypeptides in the protein space would stably fold, much less the portion that will stably fold into the same tertiary structure of the native state of wild-type PH20 protein and be active or soluble, or increase delivery of a therapeutic agent to a patient.

131. Drs. Petsko and Simpson propose a methodology for selecting and assessing combinations of mutations to incorporate into a PH20 polypeptide using the data in the patent with a PH20 homology model. Their methodology is not described anywhere in the common disclosure. For example, Dr. Simpson said a POSA could use the information in the patent "as a starting point,"[147] look at structural models, consider alignments of sequences, consider conserved features

---

[146] EX1131 (Petsko Dep.), 207:15–208:19.

[147] EX1130 (Simpson Dep.), 93:20–24.

in the protein family, perform homology modeling of the five human homologs, and then "coupled with the data from the patent" choose and assess sets of mutations:[148]

Q. So the information in the patent is kind of like a starting point for them to then devise the sequences that will be enzymatically active PH20 polypeptides?

A. The information in the patent, yes, is a starting point.

However, the POSA would also need their ability to look at the structure of HYAL1, develop a model of PH20 structurally overlaid on HYAL1.

I'd look at the bee venom hyaluronidase structure, in which the tetrasaccharide substrate is docked, compare where the substrate binding domain is.

And they would have the information from alignments, sequence alignments among the hyaluronidases; their knowledge of the structural features conserved within the family of hyaluronidase proteins, possibly by doing an -- homology modeling of all five human hyaluronidases to increase their confidence in their model.

Then, coupled with the data from the patent on activity, the POSA would have strong certainty about their likelihood of success in creating a 95% identical mutant polypeptide -- PH20 polypeptide soluble that would be enzymatically active.

Dr. Petsko described a similar methodology. In ¶ 359 of his declaration (below), he provides an example where the POSA inspects positions that tolerated between 60% and 100% of the tested changes using a homology model and selects sets of 22 substitutions from those positions. Dr. Petsko also suggests that "[t]he POSA seeking to make multiple mutations could generate a model using, for example, Swiss Model to evaluate the results of the anticipated modifications."[149] He also frequently mentions targeting surface residues. A POSA also would need to use a

[148] EX1130 (Simpson Dep.), 101:21–102:22.

[149] EX2070 (Petsko Dec.), ¶ 360.

homology model to determine whether residues are surface residues or partially buried.[150] He outlines his proposed methodology in several other paragraphs.

> 359. A POSA could select 23 residues to change and reasonably expect to obtain an active PH20 as those changes were based on the data in the patent. For example, there are 40 positions that tolerated all tested substitutions, 87 that tolerated 90% or more and 158 that tolerated 60% or more. A POSA could select sets of 22 residues within this set based on the highlighted homology model that I described above and reasonably anticipate that most of the resulting proteins will be active.

132. A POSA could not have performed this homology model-guided selection and positional assessment methodology for even the tiniest fraction of the claim parameter sequence space. Assume that one worked with the shortest reference sequence (430 residues), allowed 1 change at position 317, avoided changing the 57 essential residues identified in the common disclosure (*see* ¶¶ 304-305) and only made changes at Dr. Petsko's " ≥60% tolerated positions". Also assume that only 5 additional changes were made, and the options were restricted to the tested active mutants at those ≥60% positions (an average of 11.6 alternatives). For that "simple" investigation of only one change at position 317 and 5 additional substitutions per polypeptide, a POSA would have to individually

[150] EX2070 (Petsko Dec.), ¶¶ 298-303, 351.

model and assess 85,480,506,952,353 distinct sequences. Spending only 1 minute per assessment would translate into 59 billion days of working 24 hours a day. That is plainly impossible. As Dr. Hecht explained, in the 2011-2012 timeframe, computational tools for modeling protein structure were not capable of accurately predicting the effect of 5 or more mutations on a native protein sequence and its corresponding structure.[151]

133. The 2024 Chemistry Nobel Price Background paper that Dr. Hecht references in his declaration provides an accurate historical perspective on the capabilities of computational tools to accurately predict a protein's tertiary structure from its amino acid sequence. As that paper shows, those computational tools did not reach predictive accuracy levels comparable to experimental determination of structure until around 2019 (below).[152] While experimental techniques were known by 2011-2012 that could investigate the preservation of secondary structures and generally characterize the folded state of proteins, it would not have been possible for a POSA to use them to assess even a tiny fraction of the sequence space defined by the claim parameters that include modified PH20 polypeptides with 5, 10, 15 or 20 changes.

---

[151] EX1003 (Hecht Dec.), ¶¶ 49–52, 158–159.

[152] EX1027 (Nobel Chemistry Background), 8 (Figure 6).

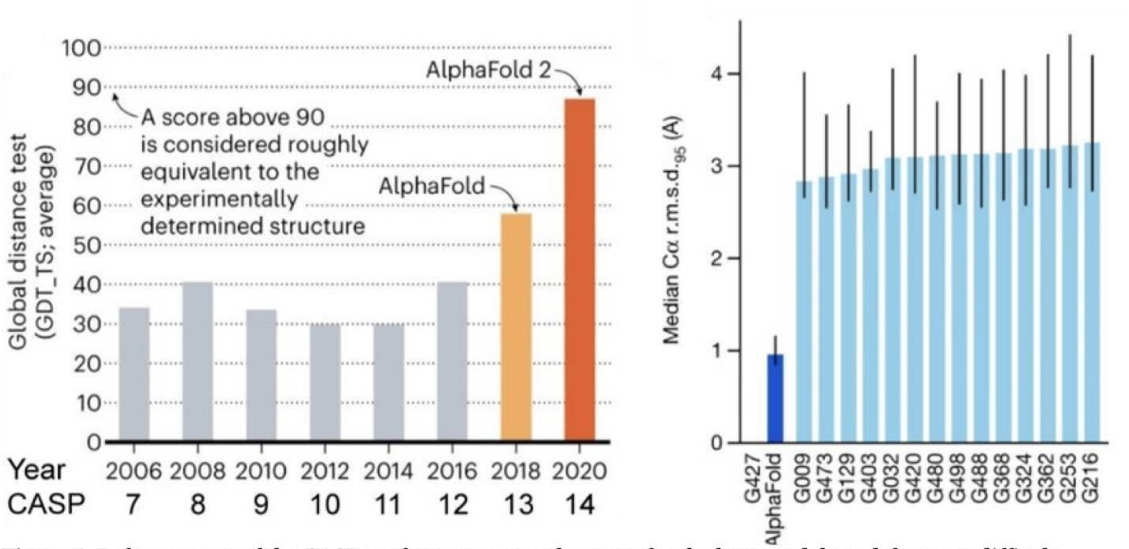


Figure 6. Left: progress of the CASP performance over the years for the best models and the most difficult targets.[38] Right: performance of AlphaFold2 relative to the top 15 entries by other groups in CASP14. Data are the median coordinate error and the 95% confidence interval of the median, estimated from 10 000 bootstrap samples.[41]

B. "PH20 Alpha-Beta Barrel" Does Not Describe Which Amino Acid Sequences in the Claim Parameter Sequence Space Will Produce It

134. Dr. Petsko observes that the wild-type PH20 protein has within it a motif referred to as an α/β barrel.[153] He suggests that "properly folded" and "enzymatically active" modified PH20 polypeptides will "adopt the same tertiary structure-an alpha-beta barrel fold," and that "all folded PH20s that meet the claims will have the same alpha-beta barrel fold."[154] He then portrays the "alpha-

---

[153] EX2070 (Petsko Dec.), ¶¶ 79, 126, 292.

[154] EX2070 (Petsko Dec.), ¶ 51 ("The POSA would have understood that properly folded PH20s adopt the same tertiary structure-an alpha-beta barrel fold. The claimed modified PH20s (that fold) thus share this fundamental common

beta barrel" motif in the PH20 protein as a "common structural feature" of the claimed modified PH20 polypeptides that have properly folded.[155]

135. Dr. Petsko is correct that an α/β barrel motif can only exist within a protein that has folded. It is a motif that is composed of secondary structures arranged in a particular manner, and forms part of the protein's tertiary structure. Dr. Petsko, however, uses the α/β fold terminology in a varying and imprecise manner in his declaration (as does Dr. Simpson) (*see* § III.B.3). The term is also incapable of accurately or completely describing the structural features shared by enzymatically active modified PH20 polypeptides.

136. In the 2011-2012 timeframe, a POSA could have identified an α/β barrel motif within a valid structural model of a folded, modified PH20 polypeptide. That person's ability to identify that motif after it has formed cannot be equated to the ability to predict from inspection of a mutated PH20 amino acid

---

structural feature (the PH20 alpha-beta barrel fold)."); ¶ 58 ("All folded PH20s that meet the claims will have the same alpha-beta barrel fold."); ¶ 128 ("And enzymatically active 91% identical PH20 variants share a common tertiary structure (the PH20 alpha-beta barrel fold) as well as other common structural features I will describe in detail below."); ¶ 396 ("And enzymatically active PH20s will adopt the same alpha-beta barrel fold.").

[155] EX2070 (Petsko Dec.), ¶ 128.

sequence whether it will adopt that same α/β barrel structural motif within the native state of the PH20 protein. As I explain below (*see* § IV), in the 2011-2012 timeframe, a POSA could not have done that.

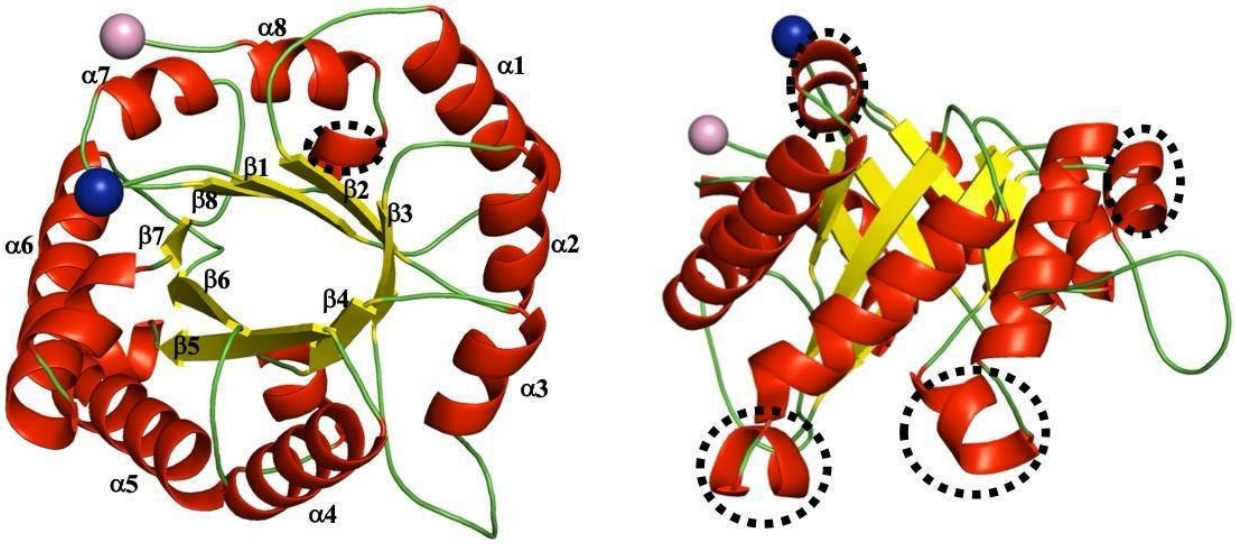
1. Alpha-Beta Barrel Motifs Are Found in Many Enzymes

137. The α/β barrel motif (also called a TIM barrel after its first observation in triose-phosphate isomerase) is a well-known protein structural unit. The canonical α/β barrel consists of 8 parallel β-strands that form a barrel (with each strand acting as stave). In the canonical arrangement this inner barrel is surrounded by 8 α-helices on the outside.[156] The canonical arrangement of the protein sequence follows the pattern β1-α1-β2-α2-β3-α3-β4-α4-β5-α5-β6-α6-β6-α7-β8-α8, with loops connecting the secondary structural elements. Chicken muscle triose-phosphate isomerase was the first experimental description of the α/β barrel (hence its other name, the TIM barrel). Phe 6 is the N-terminal residue of β1 and the C-terminal residue of α8 is Ala 246. The C-terminal residue is His248.

---

[156] EX1142 (Farber & Petsko), 228 ("The α/β enzymes have a domain with eight parallel β-strands surrounded by seven or eight α-helices. Each inner β-strand is connected to an outer α-helix, so to a first approximation the enzyme can be considered to be an $(\alpha/\beta)_8$ structure.").

138. Using PyMol 3.1.0, I prepared two orthogonal cartoon images of PDB 1TIM (chicken triose-phosphate isomerase) to show the TIM barrel. I used the DSS command within PyMol to assign secondary structure elements in the cartoon and "smooth loops" to simplify the representation (common in publications). Helices are colored red, strands yellow, loops green, and the N-terminus is blue. There are five additional helices in the tertiary structure that are not strictly part of the α/β barrel motif, which are enclosed by black dotted lines.



139. Many α/β barrel proteins are enzymes. In his textbook, Dr. Petsko estimated that ~10% of all enzymes share this protein structural fold classification.[157] As Dr. Petsko explained in his 1990 paper, "[e]nzymes with an α/β

[157] EX2164 (Petsko Textbook), 38; EX1143 (Nagano), 761 (α/β barrel fold enzymes "are ubiquitous in all the kingdoms of life, and are amongst the most frequent folds observed in all genomes.").

barrel domain often have other structures (α-helices, β-strands or entire domains) that precede, interrupt, or follow the barrel."[158] A 2002 review of α/β barrel proteins by Nagano also reported that barrel domain elements are often mixed and matched with other domains, which frequently have functional importance.[159] The actual structural fold of each protein containing an α/β barrel motif may have considerably more structural size, diversity and complexity beyond what the term "alpha-beta barrel" might suggest. Those structural differences also translate into functional diversity. As Dr. Petsko has observed in his publications, α/β barrel enzymes "all have one domain with approximately the same structural motif, but they have very different functions."[160]

2. Publications Prior to 2011 Identified the Alpha-Beta Barrel Motif in Hyaluronidases as Having an Unusual Structure

140. Dr. Petsko references publications before 2011 that discuss attributes of the tertiary structure of human hyaluronidases, including PH20. His review omits several significant insights that were reported in those papers about the

---

[158] EX1142 (Farber & Petsko), 228.

[159] EX1143 (Nagano), Abstract ("In many structures, the barrels are 'mixed and matched' with other domains generating additional variety.").

[160] EX1142 (Farber & Petsko), 228.

nature of the α/β barrel motif in hyaluronidases, most significantly that it is an unusual and distorted motif relative to the canonical α/β barrel motif.

141. Dr. Petsko extensively discussed the bee venom hyaluronidase structure reported in Markovic-Housley (EX1033).[161] He did not comment on an important observation in it, which was that the most visually striking feature of the bee venom hyaluronidase structure (PDB 1FCV) is that it does not contain a regular β-barrel and has other unusual features:[162]

> others three residues each; Figure 3). The overall fold (Figures 3 and 4) resembles that of a classical $(\beta/\alpha)_8$ TIM barrel [27], which is common to many hydrolases [28]. However, in Hya only seven strands (β1–β7) form the barrel that is surrounded by α helices A–J. There is a lack of closure between strands β1 and β2, which do not interact through hydrogen bonding, leaving a wide gap (7–8 Å) that could accommodate an additional strand (Figure 3). Loop 34–49 partly fills this gap with two polar mainchain atoms of Leu39 engaged in parallel β sheet hydrogen-bonding interactions with the mainchain atoms of Val13 and Trp15 of β1. Other unusual features, that deviate from the regular $(\beta/\alpha)_8$ barrel, are the presence of two α helices between strands 3 and 4 (helices C and D) and 4 and 5 (helices E and F) and two β hairpins comprising residues 63–75 and 269–279 at the C-terminal end of β2 and β6, respectively. Hya contains two disulfide bridges (Figure 3b). The first, Cys189–Cys201, stabilizes the base of a long loop at the C-terminal end of β4, whereas the Cys22–Cys313 bridge joins distant secondary structure elements, a 3_{10} helix near the N terminus and α helix I near the C terminus. The Hya sequence contains

[161] EX2070 (Petsko Dec.), ¶¶ 115, 135-136, 140, 196-206.

[162] EX1033 (Markovic-Housley), 1027.

142. Chao (EX1006) also reported that the human Hyal-1 structure (2PE4) had a distorted barrel fold and that the beta barrel is open despite the presence of all eight β-strands (below). The bee venom hyaluronidase lacks the β2 strand.[163] In addition, the α-helix following the β2 strand is missing in human Hyal-1 (but not in the bee venom hyaluronidase). To a POSA these represent important and meaningful structural distinctions from the canonical α/β barrel motif, and between the bee venom and these mammalian hyaluronidases.

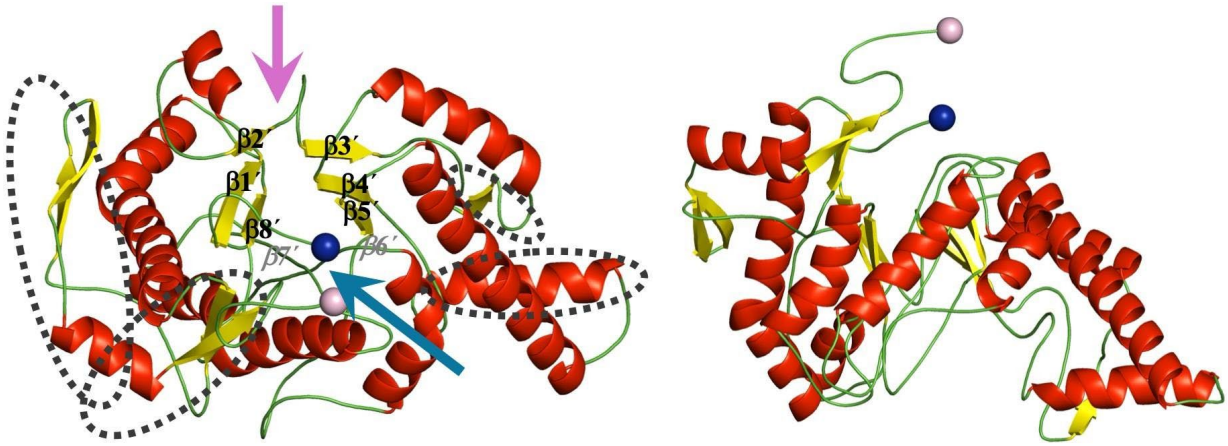
> The hHyal-1 molecule comprises two closely associated domains: the N-terminal catalytic domain (Phe22−Thr352) and the smaller C-terminal domain (Ser353−Trp435) (Figure 2A and B). The catalytic domain adapts a distorted $(\beta/\alpha)_8$-barrel fold similar to that of the bvHyal structure (*22*) but with some variations (Figure 2C): The hHyal-1 β2-strand is missing in bvHyal. However, this β2 does not form β-sheet interactions with β3, hence the β-barrel is open despite the presence of all eight strands. The barrel imperfection extends to the α-helix following β2, which is absent in hHyal-1. A number of secondary structure units are inserted in loops on the C-terminal side of the β-barrel: an α-helix (labeled α4′ following the numbering of the primary β/α motifs), two 3_{10} helices (η4′ and η8′), and a β-hairpin (β3′-3′′) (Figures 2B and 3). Two of the five hHyal-1 disulfide bonds (Cys43−Cys333 and Cys207−Cys221) reside in the catalytic domain and are also present in bvHyal (Figure 2C).

143. In the figures below, I show two orthogonal views of the PH20 model produced by Dr. Petsko (helixes are colored in red, strands in yellow, smoothed

[163] Chao (EX1006), 6913.

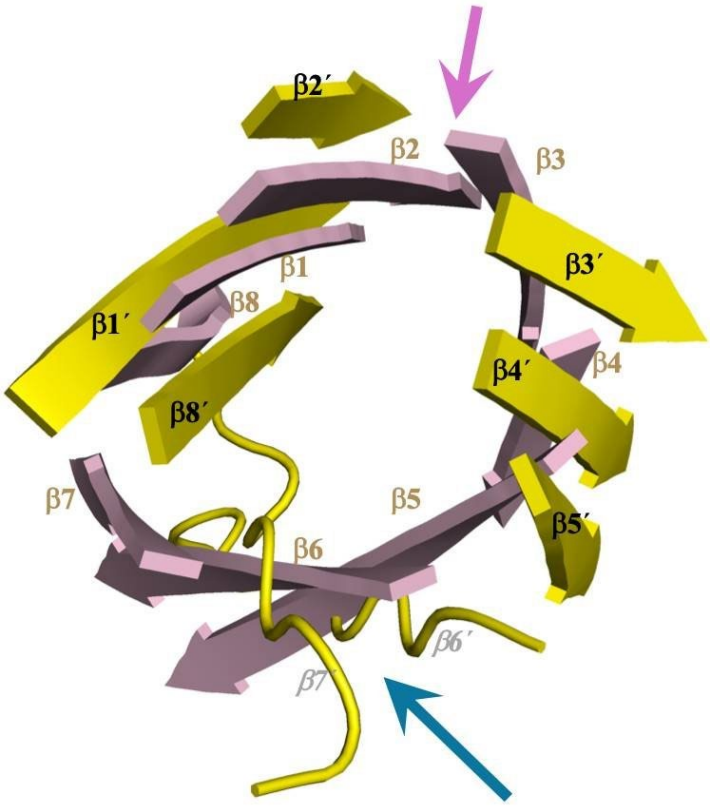
loops in green, N-terminus blue sphere and C-terminus pale red sphere). The β strands are labelled. A comparison with the TIM structure shown in ¶ 138 demonstrates the PH20 fold (left figure) is larger and more complex. In the 2011-2012 timeframe, the spatial arrangement of the secondary structural elements within the α/β barrel motif of the hyaluronidases had not been publicly reported for any proteins other than human Hyal-1 and bee venom hyaluronidase.[164] Some of the additional secondary structural elements (to the α/β barrel motif) are circled in the figures below. A pink arrow points to the gap in the barrel, and a blue arrow highlights the irregular nature of β6′ and β7′, neither of which fall within the geometric definition of PyMol DSS assignment for a β strand. This is also seen in 2PE4 (human Hyal-1) and 1FCV (bee venom hyaluronidase). Different secondary structure assignment algorithms, such as PROMOTIF (EBI), would identify a regular two-residue β strand (Tyr286–Val287) on β7 in the PH20 model.

[164] Searches using secondary-structure matching algorithm hosted by the European Biotechnology Institute (SSM at EBI) reported no matches prior to 2013.



144. Not only does the barrel in PH20 have a large gap and two irregular strands, but its shape is also distorted. This can be visualized by superimposing the β strands of 1TIM and PH20 (essentially the same image is obtained by using 2PE4 or 1FCV). In the image, the β strands of PH20 are colored yellow and labeled in black / gray whilst for TIM, the β strands are colored in pale pink and labeled in light brown. The gap is highlighted by the pink arrow and the irregular strands by a blue arrow. Given these significant structural deviations of the α/β barrel motif, a POSA could not have used knowledge of the canonical α/β barrel motif to predict the effect of mutations in the PH20 sequence on the ability of a mutated sequence to adopt and maintain this atypical α/β barrel motif structure and the overall tertiary structure of the protein. Since this secondary structural arrangement was not seen elsewhere (other than hyaluronidases) a POSA could not find guidance from other studies of the arrangement. Use of "α/β barrel motif" or

"α/β barrel fold" to describe the tertiary structure of the PH20 protein is also imprecise and incomplete.



145. Dr. Petsko references the Stern review (EX1008) several times.[165] For example, he cites it as describing the mechanism of action of catalysis as well the nature of the active site in hyaluronidases.[166] He also describes the active site in PH20, Hyal-1 and bee venom hyaluronidases as "a long groove" with "the same critical residues (oriented in about the same position)" that "catalyze the reaction in

---

[165] EX2070 (Petsko Dec.), ¶¶ 114, 127, 135, 137-138, 140, 147 n.14.

[166] EX2070 (Petsko Dec.), ¶ 135.

all three proteins."[167] He then states that "Stern teaches that 'HA-degrading enzymes' share a 'common structural feature,' which is an elongated cleft for accommodating substrate."[168] He also cites Markovic-Housley here as reporting that "the 'dominant feature of the [bee venom hyaluronidase] protein surface is a large groove' that forms the active site and 'would be large enough to accommodate a hexasaccharide' of HA."[169]

146. Dr. Petsko does not discuss several other observations in Stern that would be highly relevant to modification of hyaluronidase proteins like PH20. For example, he does not mention Stern's comments characterizing the α/β barrel motif in hyaluronidases as being an unusual one that was "only loosely associated with the TIM barrel" (below).[170] Stern here is conveying that the hyaluronidase α/β barrel motif structure has significant structural deviations from the canonical α/β barrel motif.

---

[167] EX2070 (Petsko Dec.), ¶ 139.

[168] EX2070 (Petsko Dec.), ¶ 140.

[169] EX2070 (Petsko Dec.), ¶ 140.

[170] EX1008 (Stern), 833.

> and vertebrate-like). Similar arguments can be made for the extreme N-terminus of the catalytic domain. That domain is only loosely associated with the TIM barrel. Its proteolytic removal should not affect in vitro specific activity. However, in vivo activity of such N- or C-terminal truncated enzymes in tissues would likely be affected, in their individual physiological requirements.

147. Stern makes an observation about the unusual character of the α/β barrel motif in hyaluronidases. Stern points out that in the "distorted" α/β barrel structure the "central β-sheets deviate the most from the standard TIM barrel fold, as three of the β-sheets appear to be missing from the BVHyal enzyme complex (below)."[171] Dr. Petsko does not acknowledge this observation at any point in his declaration.

> The structural properties of BVHyal, as well as those of bovine PH-20 (BPH-20), encompass a large cleft (see above) that allows for substrate binding, followed by catalysis. As these two structures demonstrate, vertebrate-like Hyals are globular proteins that encompass a large groove or cleft that traverses the entire enzyme (Figure 2). These two structures display an enzyme with a classical distorted $(\beta/\alpha)_8$ triose phosphate isomerase (TIM) barrel fold.[87] The central β-sheets deviate the most from the standard TIM barrel fold, as three of the β-sheets appear to be missing from the BVHyal enzyme complex (PDB code: 1FCV). But only one is missing in the high resolution native BVHyal structure (PDB code: 1FCQ), suggesting that the substrate induces an allosteric modification of the secondary structure of the enzyme. The cleft that transverses the wider end of the TIM

[171] EX1008 (Stern), 824.

148. In the same paragraph, Stern also observes that allosteric effects may be induced by substrate binding, suggesting the protein's structure is not rigid. A POSA would expect that if allosteric effects can be induced by substrate binding, the structure could also be sensitive to epistasis, including effects from changes at a distance.

149. Another important structural insight reported in Stern is that the residues involved in catalysis in hyaluronidases are located ***outside*** the α/β barrel core, rather than being integral to the α/β barrel motif structure (below).[172] While enzymatically active hyaluronidases do preserve the spatial positioning of many residues involved in substrate binding and catalysis, Stern explicitly excludes attributing those conserved structural features to the α/β barrel motif.

> octasaccharide segment of HA. The extreme N-terminal segment of the catalytic domain, ranging in length from 28 aa for human Hyal-1 to 41 aa for HPH-20, always assumes a helical conformation. It takes on a structure of a single helix for Hyal-1 and Hyal-2 (the Hyals with the shortest segment at their N-terminus) or a set of two helices connected by a short loop for Hyal-3, Hyal-4, and HPH-20. Hyal-4 and HPH-20 have the longest aa segments at their N-termini (Figure 3). In all five cases, however, the helices are located on the outside of the TIM barrel core structure, are not incorporated into the distorted TIM barrel fold, and are primarily associated with the catalytic domain only; the interactions with the C-terminal domain are minimal, if any (Figures 4 and 5). The sequence similarity in this region of all human Hyals is very low, compared to the remaining portion of these Hyals, and therefore allows for some variations in structure.

[172] EX1008 (Stern), 829.

150. From these observations in the scientific literature before 2011, a POSA would have inferred that while there is a distorted α/β barrel motif in hyaluronidases like PH20, that motif alone does not define the catalytic site or associated substrate binding sites of the proteins. Knowledge of the presence of that motif would provide at best an incomplete picture of the structural features shared by enzymatically active PH20 proteins.

151. None of the unusual characteristics of the α/β barrel motif in the PH20 protein that had been reported in the scientific literature are identified, described or discussed in the common disclosure. Dr. Petsko and Dr. Simpson also do not mention these unusual characteristics of the PH20 α/β barrel motif in their declarations. A POSA could not have identified or evaluated these unusual characteristics of the PH20 α/β barrel motif by inspecting only the amino acid sequences or single mutations listed in the common disclosure. Other than reporting that mutants at positions previously reported to be essential in hyaluronidases were inactive, the common disclosure does not explain how any of the experimental results reported in the common disclosure relate to any of the structural observations reported in the Stern, Markovic-Housley, Chao or Zhang publications. It also does not tell a POSA how to how to use those findings in combination with the mutagenesis data when designing modified PH20

polypeptides to preserve the unique α/β barrel motif or other structural features in the tertiary structure of the wild-type PH20 protein.

152. I have been informed that an adequate written description for a set (or genus) of chemical compounds requires "a precise description, such as by structure, formula, chemical name, physical properties, or other properties of species falling within the genus sufficient to distinguish the genus from other materials." I do not believe the phrase "PH20 alpha-beta barrel fold" as it is being used by Drs. Petsko and Simpson does or is capable of doing that.

a. First, the use of "fold" with "PH20 alpha-beta barrel" is scientifically inaccurate. In PH20, the α/β barrel is a motif, not the fold.

b. Second, the common disclosure does not associate a particular sequence within PH20 to the α/β barrel motif or explain how the active site or substrate binding sites relate to α/β barrel motif.

c. Third, Drs. Petsko and Simpson are using the phrase "alpha-beta barrel" motif at a level of abstraction that cannot differentiate the PH20 protein from other human hyaluronidases, or non-mammalian hyaluronidases. The structures in all these proteins differ (*e.g.*, the bee venom omits the significant portion at the C-terminus of PH20).

3. Dr. Petsko's Description of the PH20 Alpha-Beta Barrel Fold Varies and Is Not an Objective or Precise Description of a Common Structure of the Claimed PH20 Polypeptides

153. Dr. Petsko's use of the "alpha-beta barrel fold" terminology in his declaration where it discusses PH20 is imprecise. This is a problem because he is attempting to use this term to define a particular structural feature in the tertiary structure of a modified PH20 polypeptide. To do that, the POSA would refer to the "alpha-beta barrel" as a motif or sub-domain within the tertiary structure of the protein, which makes clear there are other structural features in the tertiary structure. The POSA also would describe how the α-helix and β-strand elements of the α/β barrel relate to other elements of the protein structure. The phrase α/β barrel ***fold***, by contrast, conveys that the α/β barrel motif ***is*** the tertiary structure of the protein.[173] That is not the case for PH20.

154. In several paragraphs of his declaration, Dr. Petsko appears to be using the phrase to refer to the entire tertiary structure of PH20.

a. In ¶ 79, he states: "As I explain below, PH20 adopts an ***alpha-beta barrel fold***."

b. Similarly, in ¶ 424, he appears to equate the "tertiary structure" of PH20 (*i.e.*, the structure of the entire protein) with the PH20 α/β barrel fold ("The claims encompass enzymatically active

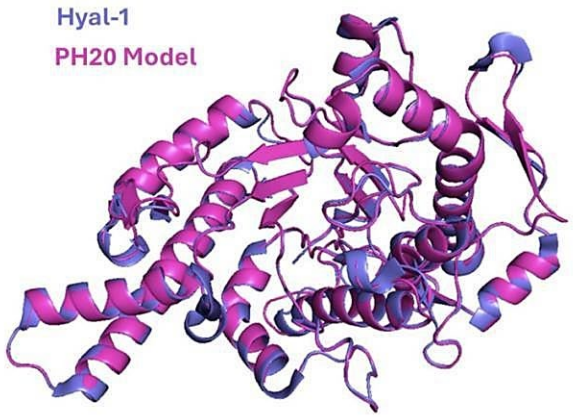
[173] Few proteins with an α/β barrel exist as an α/β barrel fold.

PH20s with the same tertiary structure—the PH20 alpha-beta barrel fold.").

c. In ¶ 130, he states "Hyal-1 has a tertiary structure that the POSA would characterize as an alpha-beta barrel fold."

d. Also in ¶ 132, he compares Hyal-1 and PH20 and states that "it is evident that both proteins adopt essentially the same alpha-beta barrel fold."[174] The Hyal-1 structure shown in that image is the entire Hyal-1 protein (2PE4),[175] while the PH20 model only models positions 3 to 415 of the PH20 sequence.

132. When the PH20 model is superimposed on the Hyal-1 structure in PyMOL, as in the figure below, it is evident that both proteins adopt essentially the same alpha-beta barrel fold.



[174] EX2070 (Petsko Dec.), ¶ 132.

[175] EX1006 (Chao), 6913 ("Structure refinement at 2.0 Å resolution… yielded a high quality electron density map… covering the entire coding region of the mature hHyal-1 (Phe22-Trp435).").

155. Elsewhere Dr. Petsko refers to the α/β barrel ***fold*** as a ***part*** of the PH20 protein. As he states: "[E]nzymatically active 91% identical PH20 variants share a common tertiary structure (the ***PH20 alpha-beta barrel fold***) as ***well as other common structural features***…."[176]

156. In other paragraphs, Dr. Petsko accurately refers to the α/β barrel as a structural "motif" in his declaration (*i.e.*, a portion of the larger tertiary structure of the protein).[177] At his deposition, Dr. Petsko was also accurate when referring to the α/β barrel motif as a "subdomain" within the PH20 tertiary structure.[178]

> 101. As I will explain in detail below, based on these principles, a POSA would have expected that folded, secreted proteins that are 90-91% identical in sequence adopt the same fold. In other words, with respect to PH20, a POSA would have expected that a folded modified PH20 that is 91% identical in sequence to wild type ("WT") would adopt the same alpha-beta barrel fold as WT PH20. The POSA would not expect the 91% identical protein to adopt a completely different protein fold. Put another way, if the 91% identical modified PH20 folds, it will fold into the same alpha-beta barrel motif as the WT.

---

[176] EX2070 (Petsko Dec.), ¶ 128 (emphasis added).

[177] EX2070 (Petsko Dec.), ¶ 101 ("[I]t will fold into the same alpha-beta barrel motif as the WT.").

[178] EX1131 (Petsko Dep.), 68:24-69:23.

Is the alpha-beta barrel fold of the human PH20 polypeptide, is that also called the tertiary structure of the human PH20 polypeptide?

A. It's part of the tertiary structure because there are other elements in the protein besides that particular subdomain, if you will. But yes, it is part of the tertiary structure.

Q. So is it fair to say that -- that a portion of the amino acid sequence of the human PH20 polypeptide is responsible for creating the alpha-beta barrel fold in the PH20 protein?

A. I wouldn't use that terminology, no.

Q. How would you explain the relationship between the part of the sequence in the part that causes the folded protein?

A. You used the phrase "responsible for creating." That's an active verb.

Q. Right.

A. Okay? I would say that there was a particular sequence in the complete sequence of the protein that is -- what's the exact word I want to use? -- that is contained within the alpha-beta barrel fold. Okay?

The way you phrased it, it implied a mechanism, and I don't think we want to do that.

157. Dr. Simpson makes similarly imprecise use of the "alpha beta barrel fold" terminology throughout her declaration.[179]

158. Dr. Petsko clearly knows that a protein's tertiary structure is the structure of the entire protein.[180] For example, when he was asked what a protein's tertiary structure at his deposition, he explained that it referred to the structure of the entire protein:[181]

---

[179] EX2068 (Simpson Dec.), ¶¶ 105–106, 122, 135, 289, 295, 305, 319, 368.

[180] EX1131 (Petsko Dep.), 38:12-17 ("Q. Okay. So if I say -- if a scientist uses the word "fold" generally in connection with the protein, that's when they would be referring to the entire tertiary structure? A. Normally, that would be the default interpretation."); *see also* EX2164 (Petsko Textbook), 3 ("For a polypeptide to function as a protein, it must usually be able to form a stable tertiary structure (or *fold*) under physiological conditions.").

[181] EX1131 (Petsko Dep.), 35:5-16.

Q. And then just the last one is: What's the -- what is a protein's tertiary structure?
A. So these individual secondary structure elements can be visualized as independent entities. The tertiary structure involves how the protein folds up to bring all of those elements together to make a compact structure. It isn't always globular, but it's more compact than when it's a linear sequence.
And the interactions between those elements are what hold the whole globular protein together.

159. Dr. Petsko correctly restricts his comments about the "PH20 alpha-beta barrel motif" to properly folded proteins (below).[182] That is appropriate because the only modified PH20 polypeptides that could contain the PH20 α/β barrel motif are properly folded forms of those proteins.

101. As I will explain in detail below, based on these principles, a POSA would have expected that folded, secreted proteins that are 90-91% identical in sequence adopt the same fold. In other words, with respect to PH20, a POSA would have expected that a folded modified PH20 that is 91% identical in sequence to wild type ("WT") would adopt the same alpha-beta barrel fold as WT PH20. The POSA would not expect the 91% identical protein to adopt a completely different protein fold. Put another way, if the 91% identical modified PH20 folds, it will fold into the same alpha-beta barrel motif as the WT.

[182] EX2070 (Petsko Dec.), ¶ 101.

4. The Phrase "PH20 Alpha-Beta Barrel Fold" Does Not Identify the Structural Features Shared by Enzymatically Active Forms of PH20

160. At the level of abstraction being used in Dr. Petsko's and Dr. Simpson's declarations, the phrase "the PH20 alpha-beta barrel" does not identify to a POSA what common structural features are present in and shared by enzymatically active forms of the PH20 protein. Dr. Petsko's and Dr. Simpson's use of the term is simply confirming the presence of a characteristic pattern of secondary structural elements associated with the motif that was first observed in the TIM structure. *See* ¶¶ 137-138. Dr. Petsko's declaration also does not provide an objective or complete description of the α/β barrel motif within PH20, such as by identifying the amino acid sequence that he believes to correspond to this structural motif. Dr. Petsko chose to not annotate any of the images of PH20 structures he produced to show where the components of the α/β barrel fold start and end, or the motif's relationship to other structural elements in the tertiary structure. The common disclosure also does not direct POSAs how to determine what the "common structural features" are that Dr. Petsko has referred to. As a result, a POSA cannot determine which modified PH20 polypeptides might or might not possess them from Dr. Petsko's varying use of the α/β barrel terminology.

161. Only one paragraph (¶ 249) in Dr. Petsko's declaration references a particular amino acid sequence range in PH20 (below). Dr. Petsko quotes this from the common disclosure, which refers to it as the "common core hyaluronidase domain" that spans positions 3-339 of the mature PH20 sequence. Neither the disclosure nor Dr. Petsko describe how this identified sequence within PH20 relates to the α/β barrel motif.

> 249. The '758 patent recognizes that "sequence and structure of PH20 polypeptides is highly conserved," and specifically "[s]equence identity between and among PH20 proteins from various species is about 50% to 90%." EX1001, 69:60-62. There is a "common core hyaluronidase domain region of about 340 amino acids in length that corresponds to amino acid residues 38-374 of the precursor human PH20 sequence set forth in SEQ ID NO:6." *Id.*, 69:65-70:2. The patent cites several articles[16] for the proposition that "[w]ithin the common hyaluronidase domain region, at least 57 amino acids are conserved between and among species." *Id.*, 70:12-17-.

162. If Dr. Petsko is suggesting that positions 3 to 339 correspond to the α/β barrel motif in PH20, that would be misleading. That is because several loops in PH20 are much more extensive and contain additional and elongated secondary structural elements when compared to the canonical α/β barrel (which comprises 248 residues in 1TIM). Some examples are the strand loop strand (Tyr 64 – Ser 84), a helix (Lys126-Gln140), a sixteen residue longer helix (Leu146-Leu175), a longer loop (Pro187-Asn205), and a twenty residue longer helix (Met313 –

Leu345). Similar observations were made by Chao (EX1006) in the paragraph discussed in ¶ 142.[183]

> The hHyal-1 molecule comprises two closely associated domains: the N-terminal catalytic domain (Phe22−Thr352) and the smaller C-terminal domain (Ser353−Trp435) (Figure 2A and B). The catalytic domain adapts a distorted $(\beta/\alpha)_8$-barrel fold similar to that of the bvHyal structure (*22*) but with some variations (Figure 2C): The hHyal-1 β2-strand is missing in bvHyal. However, this β2 does not form β-sheet interactions with β3, hence the β-barrel is open despite the presence of all eight strands. The barrel imperfection extends to the α-helix following β2, which is absent in hHyal-1. A number of secondary structure units are inserted in loops on the C-terminal side of the β-barrel: an α-helix (labeled α4′ following the numbering of the primary β/α motifs), two 3_{10} helices (η4′ and η8′), and a β-hairpin (β3′-3″) (Figures 2B and 3). Two of the five hHyal-1 disulfide bonds (Cys43−Cys333 and Cys207−Cys221) reside in the catalytic domain and are also present in bvHyal (Figure 2C).

163. The "common core hyaluronidase domain" of PH20 referred to in the common disclosure (quoted by Dr. Petsko), when produced, has no enzymatic activity as Dr. Hecht noted.[184] The model of PH20 also shows that position 339 is part of an elongated alpha helix. The "common core hyaluronidase domain" also excludes the EGF domain, an indispensable feature of the structure-function relationship of PH20.

---

[183] EX1006 (Chao), 6913.

[184] EX1003 (Hecht Dec.), ¶¶ 92-93.

5. The Common Disclosure Does Not Portray the Alpha-Beta Barrel Fold as a Common Structural Feature of Modified PH20 Polypeptides and Does Not Describe Representative Examples of Multiply-Modified PH20s

164. Based on the legal requirements that have been explained to me, I agree with Dr. Hecht that the common disclosure does not describe either (i) the common structural features shared by all members of the claimed genus of modified PH20 polypeptides,[185] or (ii) examples of multiply modified PH20 polypeptides that are representative of the structural diversity of the claimed modified PH20 polypeptides.[186] *See* ¶¶ 29-30.

165. The common disclosure does not include or reference any experimentally determined hyaluronidase structures published before 2012, or any homology models depicting the tertiary structure of PH20.

a. The common disclosure does not reference or discuss the findings in Markovic-Housley (EX1033) about the unusual structure of the α/β barrel motif within the bee venom hyaluronidase protein. It also does not discuss or explain the structural significance of this highly distorted α/β barrel fold, or any of the insights that were flagged in that publication.

---

[185] EX1003 (Hecht Dec.), ¶¶ 113, 145, 157.

[186] EX1003 (Hecht Dec.), ¶¶ 141-143, 157-159.

b. The common disclosure does not reference or discuss the published reports of the human Hyal-1 structure reported in Chao (EX1006) and does not reference or discuss the experimental studies of the human Hyal-1 protein reported in Zhang (EX1010). These omissions are particularly significant because Chao and Zhang reported the structure of a homologous human hyaluronidase.

c. While the common disclosure references earlier work with C-terminally truncated forms of wild-type PH20 and the presence of a glycosyl phosphatidylinositol (GPI) anchor sequence within it, it does not correlate any of those findings to the Hyal-EGF domain structure or the characteristic pattern of residues associated with it that was published in 2007 in Chao (EX1006) and discussed in Zhang (EX1010).

166. The common disclosure does not use the phrase "alpha-beta barrel motif" or "alpha-beta barrel fold" (or α/β) and does not identify a particular amino acid sequence that a POSA would understand to correspond to the "alpha-beta barrel" motif in PH20.

167. The common disclosure does not include or reference a structural model of PH20 protein.[187] It also does not identify a particular three-dimensional

[187] The structure of human PH20 was not experimentally determined until 2024. EX1126 (Im), 1067.

structural feature within the tertiary structure that corresponds to the α/β barrel motif in PH20, or portray that as being a common structural feature of all modified PH20 polypeptides that exhibit enzymatic activity. While it does reference the active site residues and substrate binding residues, it does not describe their structural context.[188] It also cites three publications: (i) Arming (EX1011), (ii) Ten Have (EX1149) and (iii) Chowpongpang (EX1150). None of those publications contain an experimentally determined structure or a homology model of a hyaluronidase enzyme. None of these publications discuss a PH20 α/β barrel fold structure or its capacity to be mutated without consequence to folding or activity of the PH20 protein.

168. The common disclosure does not associate any particular amino acid sequence(s) within PH20 to an α/β barrel motif in the PH20 protein. It also does not identify or discuss other structural elements present within the tertiary structure of the PH20 protein, including within or outside the α/β barrel region, and does not portray them as being "common structural features" of enzymatically active modified PH20 polypeptides (*e.g.*, the Hyal-EGF domain reported in Chao (EX1006)).

---

[188] EX1026 ('731 App.), 80:11-18.

169. The common disclosure does not identify any examples of modified PH20 polypeptides with more than 3 amino acid changes relative to the PH20 wild type sequence, much less demonstrate that multiply-modified PH20 polypeptides with 5, 10, 15 or 20 changes (including deletions) will properly fold.[189] It reports only six examples of enzymatically active modified PH20 polypeptides that meet the claim requirements that could contain an α/β barrel motif (*i.e.*, the L317A, L317I, L317K, L317M, L317Q, L317R single substitution PH20 mutants that exhibited hyaluronidase activity).[190] These examples are not representative of the structural diversity of the immense sets of multiply-modified PH20 polypeptides being claimed.

170. The PH20 α/β barrel is not a "common structural feature" of the entire set of 10^{59} to 10^{112} polypeptides within the claim parameter sequence space that Dr. Simpson describes.[191] Within that claimed sequence space will be many sequences that cannot be produced, will not properly fold, or will unfold because they are

---

[189] EX1003 (Hecht Dec.), ¶ 131 ("[T]here is no meaningful guidance in the common disclosure on what happens when different kinds of mutations are combined, or different mutations in a region of the PH20 are combined.").

[190] EX1026 ('731 App.), 254-255 (Table 9).

[191] Dr. Petsko stated that he was not asked "to opine regarding claim construction and [did] not do so in this declaration." EX2070 (Petsko Dec.), ¶ 246.

unstable. There also will be secreted forms that have misfolded or are incompletely folded. Because none of those sequences will possess a stable tertiary structure, they will not contain the "PH20 α/β barrel" motif present in the tertiary structure of the PH20 protein. This structural motif cannot therefore be a "common structural feature" of that set of polypeptides.

171. Dr. Petsko also did not address the discussion in the common disclosure that distinguishes the "common core hyaluronidase domain" sequence from a larger portion of the PH20 amino acid sequence that the common disclosure describes as being "the minimal sequence required for hyaluronidase activity." The common disclosure identifies the latter sequence as running from positions 1 to 429 of the mature PH20 sequence (below).[192]

among PH20 hyaluronidase polypeptides. PH20 hyaluronidases contain a common core hyaluronidase domain region of about 340 amino acids in length that corresponds to amino acid residues 38-374 of the precursor human PH20 sequence set forth in SEQ ID NO:6. A mature PH20 polypeptide lacking the signal sequence and containing a contiguous sequence of amino acids having a C-terminal amino acid residue corresponding to amino acid residue 464 of SEQ ID NO:6 (*e.g.*, amino acid residues corresponding to positions 36-464 of the amino acid sequence set forth in SEQ ID NO:6) is the minimal sequence required for hyaluronidase activity (see *e.g.*, U.S. Patent Application No. 10/795,095,which is issued as U.S. Patent No. 7,767,429; see also U.S. Publication No. US20100143457).

172. Dr. Simpson also did not identify where the α/β barrel sub-domain is within the PH20 sequence or identify it within illustrations of the PH20 protein

[192] EX1026 ('731 App.), 79:23-31.

structure she uses in her declaration. I understand she also did not inspect any PH20 homology models using visualization software (*e.g.*, PyMol) but only evaluated the images of the PH20 structure in Dr. Petsko's declaration.[193]

C. Multiple Mutations to PH20 Are No More or Less Predictable than in Other Proteins

173. In ¶ 292 of his declaration, Dr. Petsko suggests that PH20 is "likely to be relatively robust in structure" because (in part) they contain "an alpha-beta barrel fold," which he suggests is a "robust fold that is tolerant to substitution." Dr. Petsko cites no scientific support for this proposition—either generally or critically for PH20, which contains a unique α/β barrel motif. However, as Dr. Petsko notes, the data in the common disclosure actually contradicts this proposition.[194]

174. The data (summarized below at ¶ 236) shows that most single amino acid substitutions in PH20 were deleterious, with another 13% uncharacterized (pie chart below). Even among the "active mutants" many substitutions clustered at the boundary of the 40% cutoff (histogram below, Table 9), indicating substantial functional impact from these substitutions. There is a substantial impact of single amino acid substitutions. Only approximately 7% of mutants retained near-native

---

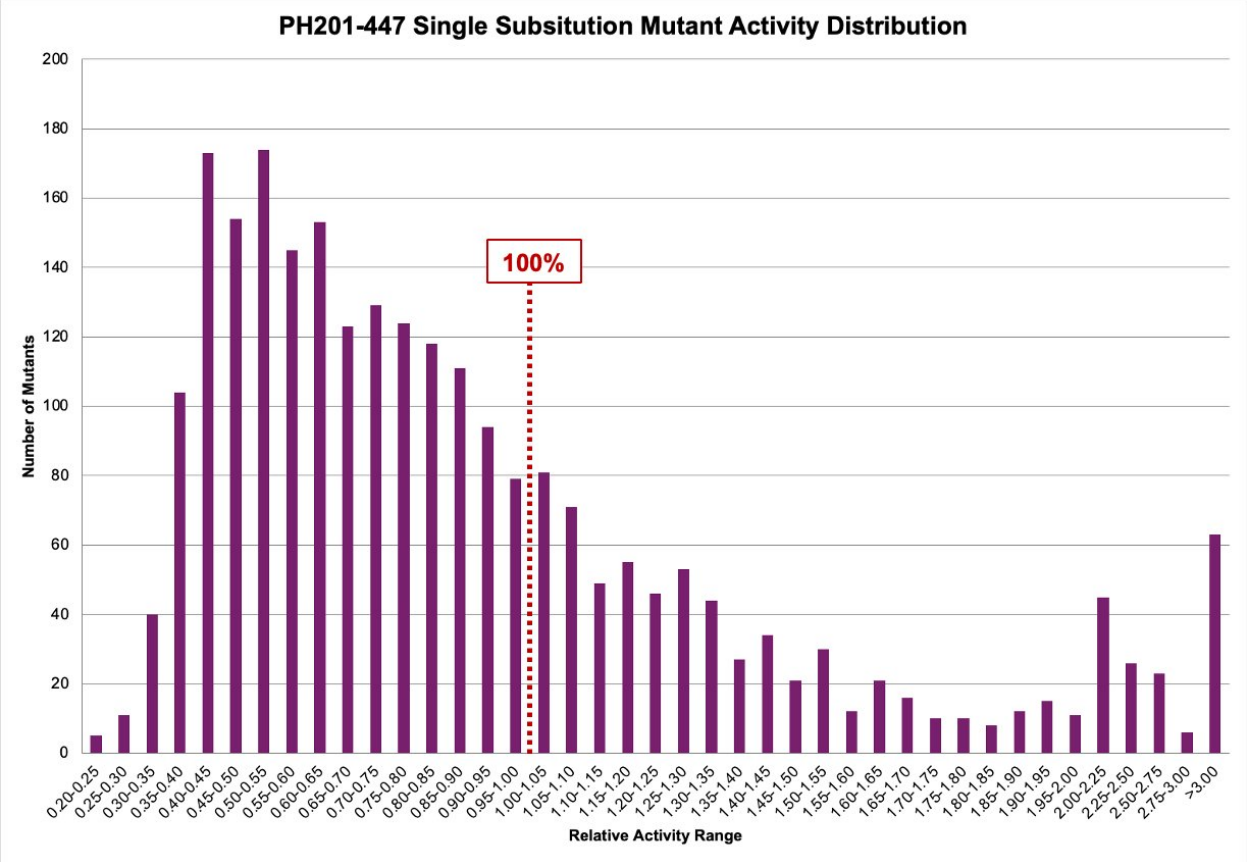
[193] EX1130 (Simpson Dep.), 35:6-20.

[194] EX2070 (Petsko Dec.), ¶ 292 ("But the data in the '758 patent shows that it is not.").

activity (*i.e.*, between 95-105% activity), thus the actual data show that most mutations perturb in a measurable way (improving or degrading) the structure-function relationship. The common disclosure does not address this observation, nor do the expert reports of Dr. Petsko or Dr. Simpson. Both Drs. Simpson and Petsko contend that the single-mutant activity data disclosed in the patent establish a structure-function relationship that extends across the claimed PH20 variants. That conclusion does not follow from the data. A POSA would conclude from the disclosure itself that even single mutants classified as active but distant from the active site alter the structure-function relationship. Such changes in the structure are likely to be so subtle that, for most, if not all variants, they would have been impossible to identify using tools in 2011-2012. Even this limited set of experimental data (the 28% of possible mutants that are classified as active) demonstrated high degree of sensitivity of the structure-activity relationship for single mutants, which in turn highlights the difficulty of combining multiple mutations in a predictable manner. The disclosure does not analyze or report the physical reasons why certain mutants are classified as inactive, and therefore provides no insight into the structure-function relationship.

175. As the histogram below shows, there is a significant portion of mutants below the 40% activity cutoff (histogram below, Table 9). A POSA reviewing the data in the patent that the PH20 would agree with Dr. Petsko, PH20

is not tolerant to single mutations.[195] Given the protocol (at least one sample > 40% activity), average values can be based on quite different individual measurements. For example, for an average of 0.35 the smallest difference between the two measurements is 0.12 equivalent to ~ 30 % (0.41 and 0.29). The common disclosure does not comment on whether such large discrepancies between duplicates were common.



[195] *See* Appendix B-5 (Histogram of $PH20_{1\text{-}447}$ Activity Data).

176. Most proteins with the α/β barrel motif include many other structural features and are also a very heterogenous class of proteins (and that is the case for PH20).[196] By 2002, 21 different superfamilies of these proteins were known, with varying numbers and combinations of domains and motifs.[197]

177. Experimental evidence published before 2011 shows there are cases in addition to PH20 (see ¶175) that do not follow Dr. Petsko's general assumption of robust tolerance to substitution. In 2009, Ochoa-Levya et al (EX1144) reported on experiments that investigated the effects of replacing loops within phosphoribosylanthranilate isomerase (PRAI) enzymes, an enzyme with a canonical α-β barrel fold. Importantly, the study investigated substitutions of 4 to 19 amino acids in a single protein and their effects on protein folding, directly by addressing the tolerance of α-β structures to substantial sequence changes.

178. The study examined a highly constrained set of changes: substitutions were limited to loop sequences already explored and optimized by natural evolution and known to be tolerated in naturally occurring α-β barrel enzymes with

---

[196] EX1008 (Stern), 825, 829.

[197] EX1143 (Nagano), 744 (Figure 1) (showing distribution of single vs. multiple domain α/β barrel motif proteins).

the same fold, increasing the likelihood that the resulting proteins would remain folded and screenable for function.[198]

Table 1. Characteristics of the exchanged loops in the PRAI-LoxP scaffold

Enzyme (EC number)	Position of the exchanged loop[a]	Sequence of the loop	Lengths
MR (EC 5.1.2.2)	β/α loop 2	GYPAL	5
FBPA (EC 4.1.2.13)	β/α loop 2	SNGGASFIAGKGVKSDVPQ	19
Ure (EC 3.5.1.5)	β/α loop 2	GGTGPAAGTHATTCTPG	17
PRAI WT (5.3.1.24)	β/α loop 2	VATSPRCVN[b]	9
DHDPS (EC 4.2.1.52)	β/α loop 4	PYYNRPS	7
TPS (EC 2.5.1.3)	β/α loop 4	LGQEDLH	7
MR (EC 5.1.2.2)	β/α loop 4	EPTLEHD	7
FBPA (EC 4.1.2.13)	β/α loop 4	DLSEES	6
αTS (EC 4.2.1.20)	β/α loop 4	DVPVQQS	7
Ure (EC 3.5.1.5)	β/α loop 4	EDWGAT	6
ADA (EC 3.5.4.4)	β/α loop 4	GDELGFPGSLF	11
PBGS (EC 4.2.1.24)	β/α loop 4	AAMDG	5
PRAI WT (5.3.1.24)	β/α loop 4	GNEE[b]	4
DHDPS (EC 4.2.1.52)	β/α loop 6	TGNL	4
PBGS (EC 4.2.1.24)	β/α loop 6	PAGAY	5
αTS (EC 4.2.1.20)	β/α loop 6	SRAGVTGAENRAALP	15
PRAI WT (5.3.1.24)	β/α loop 6	NGQGGSGQRFD[b]	11

[a] Structural position in which the WT loop was replaced for the new loop in the PRAI-LoxP scaffold.
[b] Sequence corresponding to the WT loops from PRAI-LoxP.

179. The locations for modification were further constrained to three loops "where there is higher participation of their residues in substrate binding and catalysis."[199] Figure 1 annotates these loop replacement sites within the α-β barrel sequence.

---

198 EX1144 (Ochoa-Levya), 950 ("The exchanged loops belong to eight different proteins sharing the same fold but have diverse functions.").

199 EX1144 (Ochoa-Levya), 950.

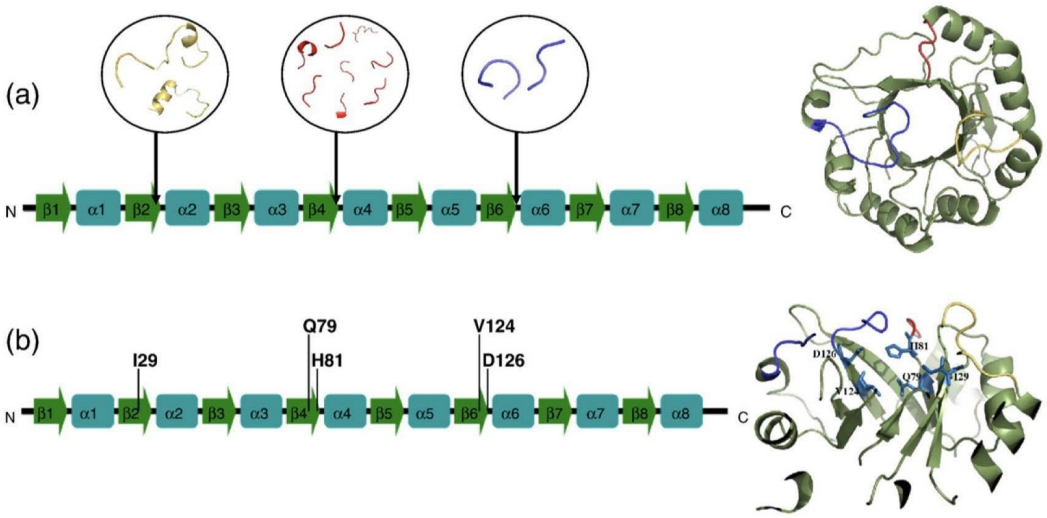
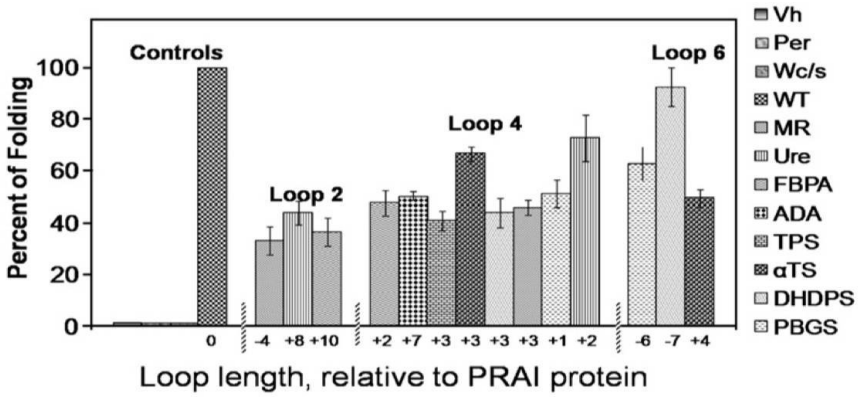


Fig. 1. Secondary-structure elements and tertiary structure (PDB code 1pii) of WT PRAI (ribbon diagram showing a view from the top of the central β-barrel). (a) The loops 2, 4, and 6 are shown in yellow, red, and blue, respectively. The circular insets show the loops to exchange in the PRAI-LoxP scaffold at the three structural positions (β/α 2, β/α 4, and β/α 6). (b) The mutated positions are indicated and a side-view slice of the TIM barrel structure of the PRAI scaffold is shown on the right.

180. The paper reported that 545 folded mutants were identified using a folding reporter gene, with frequencies varying by loop: approximately 30-40% for loop 2, 40-70% for loop 4 and 50 to 90% for loop 6 (below, and Figure 4). [200]

> The percentage of folded proteins for the three different libraries generated in loop 2 had a very similar value, around 30–40%, and in fact they were the lowest values found in the whole set of constructions. Exchanges at loop 4, where more libraries were generated, afforded a higher viability, from 40% to approximately 70% of folded peptides, depending on the incoming guest loop. Finally, in the case of loop 6, for which only three libraries were constructed, we observed the highest viability, ranging from 50% to 90%.

[200] EX1144 (Ochoa-Levya), 953-954.



181. The authors found it "somewhat surprising" and observed that "considering the drastic changes introduced in the PRAI-LoxP sequence, the relatively high percentages of folded proteins that were obtained according to the Cm resistance are somewhat surprising."[201] They found that the insertion location mattered more than the inserted sequence, consistent with the active-site end of the TIM barrel, which was more tolerant to variability. They also observed that selective pressure used in the experiments influenced these successful outcomes.[202]

---

[201] EX1144 (Ochoa-Levya), 955-956.

[202] EX1144 (Ochoa-Leyva), 958 ("It is clear from the data that the introduction of variability at the hinges connecting the loops with the scaffold has a noticeable effect for obtaining folded proteins. The distribution of amino acids found in randomized positions without selective pressure of the analyzed samples shows only the normal bias inherent to the combinations NNG/C and NNT in the oligonucleotides used for mutagenesis. Thus, the amino acid preferences

182. The Ochoa-Levya paper illustrates the challenges to folding of proteins containing a set of multiple (even highly constrained) changes in a canonical α/β barrel enzyme. When 4 to 19 changes were introduced, up to 70% of the mutations did not fold. The disclosure's claims place no constraints on the location or nature of the changes made to modified PH20 polypeptides (*see* § II.A.2), and those changes would made in an enzyme with a non-canonical and unusual α/β barrel motif. *See* §§ III.B.1-III.B.2. The results illustrate another point. When one changes the length of the loops in an α/β barrel motif, it also can have surprising effects on the ability of the protein to fold. For example, the set of loops being exchanged introduce varying length loops into the test molecules across the set of mutants tested. The Ochoa-Levya paper reinforces that even constrained modifications to canonical α-β barrel enzymes can have unpredictable effects on protein folding.

D. Hyaluronidases Share a "PH20 Alpha-Beta Barrel" Motif But Exhibit Distinct Functional Properties

183. The human hyaluronidase family provides a good illustration of the limits of using an α/β barrel motif to describe a set of related proteins. By 2011-

---

observed with the introduction of each particular loop are indicative of the importance of the position adjacent to the loop for conformational fitness.").

2012, it had been demonstrated that the five human hyaluronidase enzymes exhibit distinct profiles of biological activities and expression patterns (below).[203]

on chromosome 7q31.3.[21] Human Hyal-1 and Hyal-2 are the two major hyaluronidases for the degradation of HA in somatic tissues. Hyal-2 of human origin degrades high molecular mass HA to an approximately 20 kDa product (~50 saccharide units), whereas Hyal-1 can degrade high molecular weight HA to small oligomers, primarily to tetrasaccharides. Human Hyal-3 is an acid-active enzyme (B. Triggs-Raine and R. Hammond, personal communication) as is Hyal-1. Hyal-2 has a broader pH optimum. A product of the human *hyal-4* gene, Hyal-4, based on preliminary studies, is also a chondroitinase with a predominant activity toward Ch and ChS. The detailed properties of neither Hyal-3 nor Hyal-4 have been documented to date. Details of Hyal-3 catalysis are currently being investigated and will be forthcoming (B. Triggs-Raine and R. Hammond, personal communication).

Table 2. Sequence Identity (%) among Human Hyaluronidases Hyal-1−4 and PH-20[a]

enzyme	Hyal-1	Hyal-2	Hyal-3	Hyal-4	HPH-20
Hyal-1	100	38.3	38.3	38.5	35.1
Hyal-2		100	38.0	36.4	34.5
Hyal-3			100	33.1	33.7
Hyal-4				100	41.7
HPH-20					100

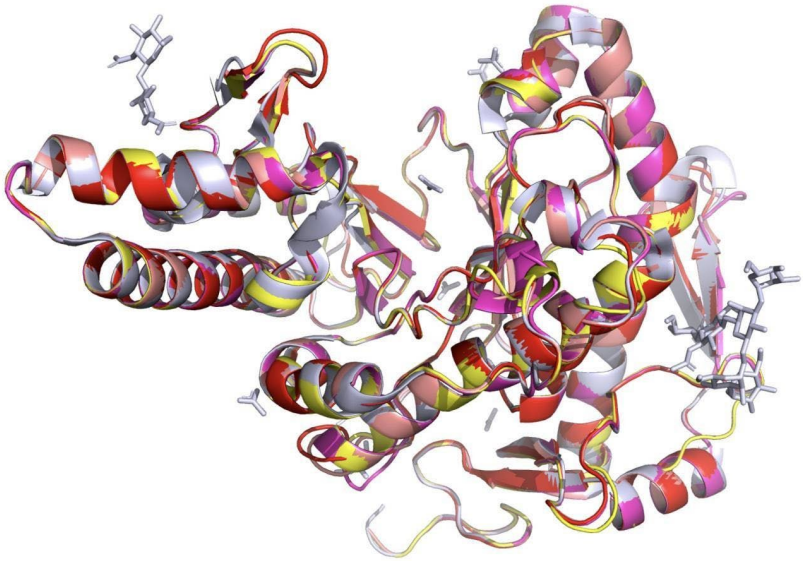
[a] The data are based on Clastal W 1.82[101] as reported by Jedrzejas and Stern.[12] The alignments are as shown in Figure 3.

184. When modeled on the human Hyal-1 structure as the template, the overlaid structures have only subtle visual differences (below).[204] This is a similar

[203] EX1008 (Stern), 825.

[204] 2PE4 structure was used for Hyal-1, the PH20 homology model was from Dr. Petsko, and homology models for Hyal-2, Hyal-3 and Hyal-4 were produced using Swiss-Model with the Hyal-1 2PE4 structure as the template.

overlay of the homology models Dr. Petsko used in his declaration.[205] Using Dr. Petsko's high level perspective, all appear have the same "alpha-beta barrel" motif. Despite this similarity, the five proteins exhibit subtle but meaningfully different substrate preferences and activity.



E. The Hyal-EGF Domain Was Known to Be a Key Part of the PH20 Structure But Is Not Discussed in the Common Disclosure

185. Drs. Petsko and Simpson identify the "Hyal-EGF" domain as a "common structural feature" of Hyal-1, PH20, and some of the PH20 variants,[206]

---

[205] EX2070 (Petsko Dec.), ¶ 421 (overlay of L317Q homology models of human, chimp, gibbon, orangutan PH20, plus N47A/N131A/L371Q, N47A/N219AL317Q, N131A/N219A/L317Q and N47A/N131A/N219A/L317Q variant PH20 sequence models).

[206] EX2070 (Petsko Dec.), ¶ 207.

but they offer conflicting characterizations that diverge from pre-2011 reports (Chao (EX1006) and Zhang (EX1010)). Most significantly, the common disclosure does not discuss this structural motif at all.

186. Dr. Hecht explained that Chao was the first publication to experimentally characterize the Hyal-EGF structure, and identify the residue patterns responsible it,[207] an advance Dr. Simpson described as "a huge breakthrough."[208] I agree that Chao provides important structural insights into hyaluronidases, including showing that earlier *ab initio* models of the mammalian C-terminal region was inaccurate:[209]

> The HyalEGF-like fold does not resemble the Hyal-1 C-terminal domain fold predicted by ab initio approaches (*33*). The close interaction between the two domains is also very different from the earlier prediction that proposed a flexible linker and lack of interactions between the two domains. Moreover, the modeling of the five human hyaluronidases C-terminal domains yielded a different fold in each case, whereas we expect that these domains adopt the HyalEGF-like fold because the amino acid sequences contain the characteristic disulfide bond pattern.

---

[207] EX1006 (Chao), 6912, 6916-17; EX1003 (Hecht Dec.), ¶¶ 84-86; EX1004 (Park Dec.), ¶¶ 97-98.

[208] EX1130 (Simpson Dep.), 115:10-116:9 ("That is among the things that this paper showed for us. Yes, it was a huge breakthrough in the field.").

[209] EX1006 (Chao), 6913.

187. Dr. Petsko portrays the Hyal-EGF domain as a "common structural feature" of PH20 and initially states that it "is required for activity," [210] but he then cites Zhang as concluding that the domain "is not absolutely required for the catalytic mechanism."[211] This mischaracterizes Zhang, which reports that the deletion of the Hyal-EGF domain abolishes Hyal-1 activity ("Deletion of the EGF-like Domain ***Abolishes*** Hyal1 Activity Independently of Asparagine 350.").[212] Zhang also reported that "loss of this domain is deleterious to both secretion" with any residual activity described as "indirect."[213] I note this level of "residual" activity of these constructs, and their lack of secretion, would classify them as "inactive mutants" in the disclosure.

188. If Dr. Petsko is suggesting the catalytic domain and the Hyal-EGF are independent, that is contrary to what the Zhang paper reports. In fact, Zhang did not investigate that question. What Zhang found was that the Hyal-1 structure has a

---

[210] EX2070 (Petsko Dec.), ¶ 207 ("[T]he EGF region is required for enzymatic activity in [Hyal-1 and PH20] proteins.").

[211] EX2070 (Petsko Dec.), ¶ 209.

[212] EX1010 (Zhang), 9439, 9433 ("Deletion of the C-terminal epidermal growth factor like domain in which Asn^{350} is located, also eliminated activity irrespective of glycosylation.").

[213] EX1010 (Zhang), 9439.

"***single overall folded domain***" with extensive contacts between the EGF-like region … and the remainder of the protein"[214] and that "because the inactivity of the truncated mutants and their inability to be secreted imply that this region ***does not function as an independently folded domain***."[215] Dr. Simpson confirms that Zhang reported the Hyal-EGF motif "is not a separate domain but instead "adopts a folded structure that packs tightly against the rest of the protein as a single domain."[216] Zhang would have led a POSA to believe that the truncation disrupted wild-type protein's native structural arrangement required for substrate binding and catalysis.

189. Drs. Hecht and Park discussed the characterization of the Hyal-EGF domain in Chao and Zhang,[217] and I agree with their opinions, for example, that POSAs would have understood this domain to be necessary for proper folding and enzymatic activity of PH20.

190. The common disclosure does not cite Chao or Zhang, or discuss any of their findings. For example, it does not identify the specific portion of the C-

---

[214] EX1010 (Zhang), 9439.

[215] EX1010 (Zhang), 9441.

[216] EX2068 (Simpson Dec.), ¶ 143.

[217] EX1003 (Hecht Dec.), ¶¶ 84-88; EX1004 (Park Dec.), ¶¶ 94-99.

terminal sequence of PH20 corresponding to the Hyal-EGF domain as being associated with a particular domain within PH20 or other human hyaluronidase proteins. It does not explain the relationship of the Hyal-EGF domain with the catalytic domain, which are reported in Chao and Zhang.

191. Chao et al. (2PE4) defined the Hyal-EGF fold as Ser353 to Trp435 (which included part of the last helix arranged around the barrel). Using this definition, the EGF fold in PH20 comprises Leu336 to Asp416.[218] However, as Dr. Park explained, "the HYAL1 template structure cannot provide structural information for PH20 amino acids past position 403, because the sequence is no longer homologous after position 421 in HYAL."[219] I agree that in 2011-2012, a POSA would not have relied on a PH20 homology model to assess the structural effects of mutations beyond 403, a region that includes a portion of the Hyal-EGF domain in PH20 (*i.e.*, either 404-416 or 404-409).

---

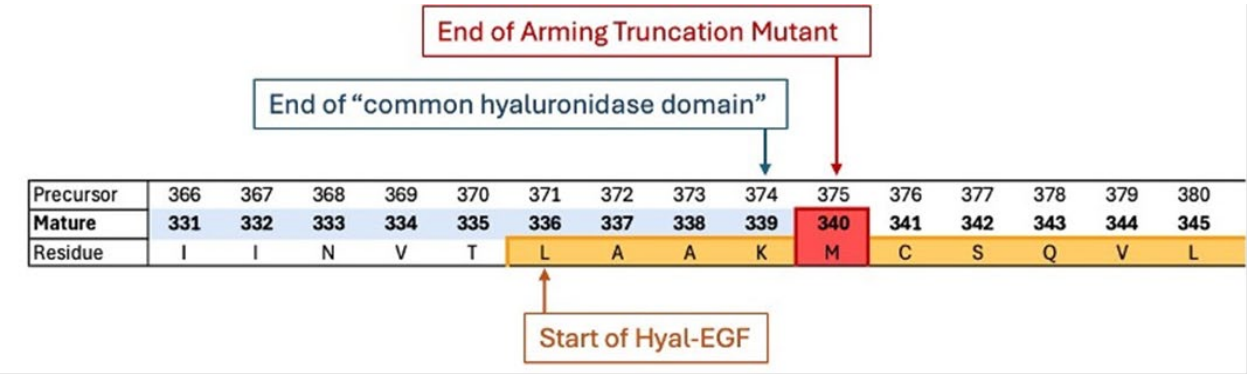
[218] EX1004 (Park Dec.), ¶ 98. Dr. Park identified the location in PH20 as being at 337-409. The terminus of the sequence can be plausibly located at either position 409 or position 416 in human PH20.

[219] EX1004 (Park Dec.), ¶ 149.

192. The Arming paper reported two C-terminal truncation mutants of PH20:[220] one was truncated at position 438, which failed to secrete but did exhibit some activity when recovered from a cell lysate, and another truncated at position 341 which was neither secreted nor active. Neither truncation corresponds to the Hyal-EGF domain in PH20 that is indicated by Chao (*i.e.*, positions 336-416 or 337-409). Arming therefore would not lead a POSA to conclude that the 336-416 or 337-409 portions of the PH20 sequence were associated with an EGF domain motif in PH20.

193. Appendix B-10 shows the terminus of the "common core hyaluronidase domain" reported in the common disclosure (position 339) (EX1026), the location of the C-terminal (340) truncation mutant reported in Arming (EX1011), and the start of the Hyal-EGF domain (EX1006). As can be seen, the common disclosure and Arming selected truncation positions that show no awareness of the Hyal-EGF domain and its location in the PH20 sequence.

[220] EX1011 (Arming), 813-814.



Precursor	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380
Mature	**331**	**332**	**333**	**334**	**335**	**336**	**337**	**338**	**339**	**340**	**341**	**342**	**343**	**344**	**345**
Residue	I	I	N	V	T	L	A	A	K	M	C	S	Q	V	L

194. Dr. Simpson, a co-author of Zhang,[221] explained that her group refined their assays to enable detect low levels of Hyal-1 mutant activity, showing that "the S245A substitution impairs activity at lower concentrations of HA but does not abrogate intrinsic catalytic potential of the enzyme."[222]

195. Zhang highlights the limits of homology models in predicting mutation effects. Dr. Simpson confirmed that her group used a Hyal-1 structural model to predict catalytic residues, then tested those predictions experimentally., with mixed results—for example, mutating Glu131 abolished activity, while mutating Asp129 did not.[223] This reflects the understanding of a POSA in 2011-2012: that one could make a limited number of reasonable predictions using a

---

[221] EX2068 (Simpson Dec.), ¶¶ 143-144, 147-148.

[222] EX2068 (Simpson Dec.), ¶ 151.

[223] EX1130 (Simpson Dep.), 130:9-16; EX1010 (Zhang), 9436-9437.

homology or structural model but experimental confirmations are needed to test those predictions.

F. Drs. Petsko and Simpson Rely on Information Not Referenced or Contained in the Common Disclosure to Support their Opinions on What the Common Disclosure Would Have Conveyed to a POSA

196. Drs. Petsko and Simpson extensively rely on a PH20 homology model built using the Hyal-1 structure reported in Chao (EX1006) to interpret the common disclosure and explain what and how a POSA would purportedly predict mutation effects in 2011-2012.[224] For example, Dr. Petsko uses his annotated versions of this model, incorporating his positional tolerance analysis, to support his view of what the common disclosure would have conveyed to a POSA, including what it informed the POSA about:

- which positions relatively tolerant to change (¶¶ 317-324);
- which positions intolerant to change (¶¶ 325-333);
- which surface residues intolerant to change (¶¶ 334-342);
- which positions can be modified to retain activity (¶¶ 355-365);
- which positions can be modified to increase activity (¶¶ 366-373);

---

[224] EX2070 (Petsko Dec.), ¶¶ 281, 297-304; *see also* 118, 120-121; EX2068 (Simpson Dec.), ¶¶ 312-313, 380-383.

- how to identify "rescue mutations" (¶¶ 374-383).

197. Dr. Petsko also compared homology models of primate and alanine-substituted PH20 proteins, overlaying them to show similar positioning of key residues despite sequence differences.[225] As I explained in ¶ 282, however, Swiss-Model homology models cannot reliably detect 1-2 Å structural deviations. Such undetectable shifts affect the positioning of critical catalytic atoms of an enzymes' substrate recognition and catalytic site, where much less than 1 Å can translate into significantly different biological effects, meaning Dr. Petsko's overlays cannot rule out subtle but functionally meaningful structural distinctions.

198. The common disclosure does not suggest using a PH20 homology model together with the mutagenesis results to identify combinations of substitutions. The common disclosure also makes no mention of the Hyal-1 experimentally determined structure that was reported in Chao in 2007, or any of the insights Chao provided using that structure, such as structural insights about the Hyal-EGF domain. The common disclosure also does not reference or discuss any of the findings in the Markovic-Housley paper that published the experimentally determined structure of the bee venom hyaluronidase protein.

---

[225] *See* EX2070 (Petsko Dec.), ¶¶ 395-410, 426-429.

G. Dr. Petsko's and Dr. Simpson's Opinions Contradict the Guidance in the Common Disclosure

199. Several opinions expressed by Drs. Petsko and Simpson are contrary to the guidance and observations in the common disclosure.

200. Drs. Simpson and Petsko suggest that a POSA would use the data in the common disclosure—together with homology modeling—to estimate the tolerance of positions in PH20 to change and then use those assessments to make PH20 polypeptides with multiple mutations. For example, in ¶ 348 of his declaration, Dr. Petsko states that a POSA could identify tolerated and non-tolerated positions from common disclosure and use that information to make and test PH20 variants to confirm they have desired properties.

> 348. With the '758 patent in hand, the POSA knows what positions tolerate substitution and which do not. The POSA would use that information to obtain modified PH20 polypeptides with the desired properties. The POSA would then make and test the desired polypeptide. With the desired protein in hand, the POSA would test it to characterize the protein and confirm desired properties. That is routine; it is what POSAs do. Making a desired protein and testing it to confirm activity is not a "research program." It is routine experimentation.

201. Dr. Simpson expressed similar opinions, referencing Dr. Petsko's positional tolerance assessment.

> 314. Accordingly, in my opinion, the '758 patent provides a POSA the information needed to predictably identify both single and multiple substitutions capable of producing hyaluronidase activity, and distinguish them from inactive variants. Considered as a whole, and consistent with Dr. Petsko's analysis, the patent's data allow a POSA to evaluate each position in the sequence to determine the degree of tolerance to substitution, the types of residues that are most likely to preserve activity, and the positions that could be selected to maintain or enhance enzymatic function.

202. In reaching these conclusions, Drs. Petsko and Simpson rely on predictive analyses and residue-selection strategies that are neither taught by nor consistent with the guidance provided in the common disclosure.

1. Dr. Petsko's Opinion that POSAs Would Not Use Directed Evolution Techniques to Produce Multiply Modified PH20 Polypeptides Contradicts the Common Disclosure

203. In ¶ 347 of his declaration, Dr. Petsko asserts that a POSA would not generate modified PH20 polypeptides with multiple mutations using directed evolution techniques, particularly by iteratively generating and screening multiple libraries of mutants because doing so would yield "billions of sequences that the POSA expects are not going to fold into PH20 proteins…" (below). That accurately reflects expectations of a POSA in the 2011-2012 timeframe about the frequency with which combinations of mutations will prevent folding.

> 347. A POSA seeking to make a modified PH20 with the '758 patent in hand would not make a randomly generated combinatorial library of billions of sequences that happen to be at least 91% identical to SEQ ID NO: 3 and then screen it for activity, because it makes no sense to do so. It is difficult to imagine why anyone would do such a thing. If the sequences truly are random, then it is likely that a substantial number would include substantial modifications to residues known to be invariant because they are involved in catalysis, the active site, or in protein structure. A POSA would not make a library with billions of sequences that the POSA expects are not going to fold into PH20 proteins because it makes no sense to do that, and thanks to the '758 patent, the POSA has a much more sensible strategy.

204. Dr. Petsko's view of what a POSA would do directly contradicts the instructions of the common disclosure. The ***only*** description in the common disclosure for producing PH20 polypeptides with multiple substitutions is an iterative, directed evolution technique. As Dr. Hecht described in his declaration,[226] it involves generating a first library of mutants (each having one substitution), screening that library for activity, then using it to generate a second library that yields mutants with two substitutions in each mutant, screening those, and so on.[227]

---

[226] EX1003 (Hecht. Dec.), ¶¶ 173-178.

[227] EX1026 ('731 App.), 141:30-34 ("The method can be performed a plurality of times, whereby the steps of the method are repeated I, 2, 3, 4, or 5 times. The method provided herein also is iterative."); 150:1-26 ("A secondary library can be created by introducing additional modifications in a first identified modified

One aspect of this example also uses the mutagenesis results (*e.g.*, to target "critical residues").[228] In other words, the common disclosure instructs POSAs to do precisely what Dr. Petsko claims a POSA would not do. This technique is also incapable of introducing more than a few substitutions in a single polypeptide, as I explained in ¶¶ 124-127.

205. I agree with Dr. Hecht's criticism of the common disclosure's description of how to identify and produce modified PH20 polypeptides with desired properties.[229] Although the disclosure provides an iterative directed evolution approach, it provides no guidance on which residues or regions to target or how to meaningfully narrow the search space.[230] As a result, the common

---

hyaluronan-degrading enzyme. For example, modifications that were identified as conferring stability, such as increasing stability, can be combined to generate a combinatorial library. The secondary library can be tested using the assays and methods described herein.").

[228] EX1026 ('731 App.), 150:14-26.

[229] EX1003 (Hecht Dec.), ¶¶ 178-185, 187-192.

[230] EX1026 ('731 App.), 141:30-34, 150:1-26; EX1003 (Hecht Dec.), ¶ 180 ("In other words, under the logic of the common disclosure's research plan, the POSA should "target" mutations to positions "near" 90% of the protein sequence. Thus, the research plan provides little to no guidance on how to design and/or choose mutations in this protein.").

disclosure leaves POSAs no practical alternative to the iterative mutagenesis approach that Dr. Petsko claims a POSA would avoid, notwithstanding that it necessarily entails generating and screening enormous numbers of largely nonfunctional variants.[231]

2. The Common Disclosure Does Not Describe Using Tolerance Ratings to Design Modified PH20 Polypeptides

206. The common disclosure does not propose using a tolerability rating system to select combinations of mutations to incorporate into multiply modified PH20 polypeptides. It does not describe merging the "active" and "inactive" mutant data sets to score positional tolerance. It also does not provide criteria corresponding to Dr. Petsko's cutoffs (*i.e.*, ≥ 60% tolerated, ≤ 30% tolerated, and 30-60% tolerated), or an underlying scientific rational for these cutoffs. There is also no description in the common disclosure of using a homology model to guide selection or assessment of sets of mutations in a modified PH20 polypeptide that one might select out of the 10^{59} to 10^{112} combinations possible.

207. The common disclosure uses variants of the word "toler*" only four times, and only one instance addresses how mutations should be selected for

[231] EX1003 (Hecht Dec.), ¶¶ 181–185.

combination in modified PH20 polypeptides.[232] That guidance instructs POSAs to exclude certain types of mutations when seeking enzymatically active variants: specifically, to not make ***any*** change at a list of certain positions, and to not include single amino acid substitutions led to the mutant being classified as an "inactive mutant" (*i.e.*, included in Tables 5 and 10).[233]

> To retain hyaluronidase activity, modifications typically are not made at those positions that are less tolerant to change or required for hyaluronidase activity. For example, generally modifications are not made at a position corresponding to position 7, 16, 17, 18, 19,

3. Dr. Simpson Contradicts the Common Disclosure's Instructions to Not Modify Certain Non-Tolerated Positions in Enzymatically Active Mutants

208. Dr. Simpson testified that a POSA seeking to retain hyaluronidase activity would rely on Dr. Petsko's positional tolerance analysis[234] and would consider modifying positions including 431, 282, 303 and 299 to make an

---

[232] EX1026 ('731 App.), 185:26 ("…the tolerance of the subject to a particular administration route…"); 308:32-33 ("These results confirm that F204P tolerates EPB level of preservative…"); 148:2-6 ("…some reduction of enzyme activity, as a result of denaturation, can be tolerated in various applications, … (*e.g.*, therapeutic activity).").

[233] EX1026 ('731 App.), 91:30-92:22.

[234] EX1130 (Simpson Dep.), 186:10-191:4, 192:12-193:3; EX2068 (Simpson Dec.), ¶¶ 243-246.

enzymatically active modified PH20 polypeptide.[235] She also stated that a POSA would consider making changes at positions 303 and 299 despite Dr. Petsko's analysis marking those positions as poorly or non-tolerant to substitution.

209. Dr. Simpson's reading of the common disclosure, and Dr. Petsko's tolerance scheme that he built from the information in the disclosure, is directly contrary to the guidance in the common disclosure. In the excerpt below, it instructs POSAS to not modify certain positions (including 282, 299, 303 and 431) in modified PH20 polypeptides that are intended to retain hyaluronidase activity."[236] This illustrates how Dr. Petsko's tolerance findings could lead POSA's to deviate substantially from the actual guidance in the common disclosure.

To retain hyaluronidase activity, modifications typically are not made at those positions that are less tolerant to change or required for hyaluronidase activity. For example, generally modifications are not made at a position corresponding to position 7, 16, 17, 18, 19, 21, 25, 53, 55, 56, 57, 62, 64, 76, 78, 80, 88, 95, 100, 101, 109, 111, 112, 113, 115, 116, 121, 123, 126, 129, 185, 187, 188, 189, 190, 191, 194, 199, 201, 203, 207, 210, 223, 225, 227, 228, 229, 241, 243, 244, 246, 249, 250, 252, 254, 262, 268, 282, 295, 296, 298, 299, 303, 319, 322, 329, 330, 332, 333, 336, 337, 340, 341, 344, 345, 346, 350, 352, 354, 355, 362, 363, 364, 365, 366, 370, 372, 382, 384, 386, 390, 400, 402, 408, 423, 424, 429, 430, 431, with reference to

---

[235] EX1130 (Simpson Dep.), 191:5-16 (431), 191:17-192:7 (282), 192:8-194:6 (303), 194:7-195:25 (299).

[236] EX1026 ('731 App.), 91:30-92:1.

4. Dr. Petsko's Rescue-Mutation Theory is Not Supported by the Common Disclosure or the Literature

210. In ¶¶ 374–383, Dr. Petsko suggests that a POSA could use the data in the common disclosure and a PH20 homology model to identify "rescue" mutations that restore activity lost from deleterious substitutions, and that examining structural interactions would allow a POSA to predictably design such activity restoring combinations. This is not described in the common disclosure. There also is no technique or passage in the common disclosure that could provide a basis for Dr. Petsko's theory that a POSA could reliably identify rescue mutations to restore activity in inactive modified PH20 mutants. Instead, it only proposes methods that discover multiply-modified PH20 variants after performing iterative rounds of mutagenesis and experimental testing, in which libraries of PH20 variants are generated, assayed for activity, and used to guide further mutation.[237] This is not a technique that attempts to predict rescue mutations in advance.

211. Dr. Petsko's rescue mutation proposal would not have yielded predictable results based on what was known as of 2011. Although the general concept of "rescue mutations" was recognized, there was no reliable way to predict in advance which, if any, additional mutations would restore activity in a given

[237] EX1003 (Hecht Dec.), ¶¶ 175-179.

protein, or PH20. Available approaches—including (i) mutagenesis-and-screening methodologies, (ii) computational or in silico modeling, and (iii) rational or semi-rational design—were inherently uncertain. Dr. Petsko's focus on local structural interactions would likely miss many—if not most—potential rescue mutations for a given "inactive" mutant, as such mutations are often distributed throughout a protein and act through indirect effects on stability or dynamics, rather than through obvious local structural corrections. The rational or semi-rational approaches that Dr. Petsko proposes also were known to succeed only in a fraction of cases and still required experimental testing to confirm restoration of function.[238]

5. Dr. Simpson Contradicts the Data and Conclusions in the '457 Publication's Data and the Common Disclosure Regarding PH20 Truncation Mutants

212. Dr. Simpson portrays C-terminal truncation mutants of the human wild-type PH20 protein terminating at positions 430 and 431 as being enzymatically inactive, based on her analysis of the common disclosure and

[238] EX1159 (Brachmann), 1847 ("This strategy of rational suppressor mutation design, although promising, could generate false predictions and will almost certainly be incomplete in scope"); *see also* EX1003 (Hecht Dec.), ¶¶ 47-53.

published US patent application 2010/0143457 (US-457) (EX2165).[239] As I explain below, her opinions are inconsistent with both the descriptions and the experimental data reported in the common disclosure and US-457.[240]

213. Dr. Simpson cites to the common disclosure's explanation that soluble, secreted PH20 requires a C-terminal truncation.[241] She then discusses hyaluronidase activity data for various PH20 truncation variants reported in US-457 (EX2165). Based on her interpretation of those data, Dr. Simpson created a summary table classifying specific truncation mutants with the labels "inactive," "little activity/inactive," "reduced activity," or "active."[242]

---

[239] EX2068 (Simpson Dec.), ¶¶ 188 (including table), 330 (including table).

[240] I understand that the US-457 publication is incorporated by reference into the common disclosure, and its contents form part of the common disclosure and may be relied upon as express teachings of the disclosure.

[241] EX2068 (Simpson Dec.), ¶¶ 171, 332.

[242] EX2068 (Simpson Dec.), ¶ 188; *see also* ¶ 330.

188. From the data in the table above, I created the table below showing what C-terminal truncation variants a POSA would have understood to be inactive, active, or have reduced or limited activity. To be clear, when I say "C-terminal position" in the chart below, I am referring to the position of the last amino acid in the sequence (*i.e.,* C-terminal position 430 in the mature sequence has 430 amino acids).

C-terminal Position (Precursor sequence)	**C-terminal Position (Mature sequence)**	**Expected Activity**
465	430	Inactive
466	431	Inactive
467	432	Reduced activity
468	433	Reduced activity
469	434	Active

214. I disagree with Dr. Simpson's conclusions that a POSA would have concluded that the common disclosure portrays the $PH20_{1-430}$ and $PH20_{1-431}$ truncation mutants as "inactive" or provides experimental evidence that supports that conclusion. US-457 instead reports test results showing both mutants are enzymatically active, and both the common disclosure and US-457 characterize both mutants as being enzymatically active.

215. US-457 reports activity data for PH20 C-terminal truncation mutants terminating between positions 415 to 472.[243] Mutants terminating before position

[243] EX2165 (US-457), ¶ 0362 (Tables 6, 6A).

430 (*i.e.*, 415-429) show little to or no activity (either zero or <15% of the activity of the $PH20_{1-430}$ mutant). By contrast, US-457 reports that the $PH20_{1-430}$ mutant exhibited significant enzymatic activity at both pH values tested (Table 6, below). These reported values correspond to approximately 80% and 99% of the activities reported for the $PH20_{1-432}$ mutant, which Dr. Simpson classified having "reduced activity." The $PH20_{1-431}$ mutant shows no detectable activity at pH 7.4, but exhibits measurable activity at pH 5.5 (0.2115 U/ml).[244]

TABLE 6

	Hyaluronidase Activity			
Deletion Mutant	Precursor (AA)	Mature (AA)	pH 7.4 Activity (Units/ml)	pH 5.5 Activity (Units/ml)
SPAM1-AFLK (SEQ ID NO: 82)	470	435	3.204	2.925
SPAM1-DAFL (SEQ ID NO: 83)	469	434	2.3895	2.2365
SPAM1-IDAF (SEQ ID NO: 84)	468	433	0.5625	0.62775
SPAM1-CIDA (SEQ ID NO: 85)	467	432	0.5535	0.4725
SPAM1-VCID (SEQ ID NO: 86)	466	431	0	0.2115
SPAM1-GVCI (SEQ ID NO: 87)	465	430	0.441	0.468

[244] EX2165 (US-457), ¶ 0362 (Tables 6, 6A).

216. The $PH20_{1-430}$ and $PH20_{1-431}$ mutants are described as being enzymatically active in US-457. Specifically, ¶ 0362 states that truncation mutants ending ***before*** position 430 (*i.e.*, $PH20_{1-415}$ through $PH20_{1-429}$) are inactive,[245] and separately notes that mutants terminating between positions 463 and 467 "have little detectable activity due to low expression level." US-457 then states that "***[a]ll other mutants*** (SEQ ID NO:55-57 and **60-87**) ***are active hyaluronidases*** at both pH 7.4 and pH 5.5." (below).[246] In US-457, SEQ ID NO:87 corresponds to $PH20_{1-430}$, while SEQ ID NO:86 corresponds to $PH20_{1-431}$.[247]

> **[0362]** The results are shown in Tables 6 and 6A, below. Human PH20 hyaluronidase carboxy-terminus deletion mutants shorter than I430, corresponding to SPAM1-GDVC to SPAM1-ADVK (SEQ ID NOS:88-95), are inactive. Human PH20 hyaluronidase carboxy-terminus deletion mutants ending at 1498 (SEQ ID NO:101), L499 (SEQ ID NO:59), F500 (SEQ ID NO:100), L501 (SEQ ID NO:58) and I502 (SEQ ID NO:99) have little detectable activity due to low expression level. All other human PH20 hyaluronidase carboxy-terminus deletion mutants (SEQ ID NOS:55-57 and 60-87) are active hyaluronidases at both pH 7.4 and pH 5.5.

217. Dr. Simpson's conclusion that $PH20_{1-430}$ and $PH20_{1-431}$ would have been considered inactive mutants by a POSA appears to be based on her assertion

---

[245] EX2165 (US-457), ¶ 0362 ("Human PH20 hyaluronidase carboxy-terminus deletion mutants shorter than I430 . . . are inactive.").

[246] EX2165 (US-457), ¶ 0362.

[247] EX2165 (US-457), ¶ 0362.

that a POSA would discount individual numerical differences in the reported activity data. Dr. Simpson states that, because variability is expected in hyaluronidase assays, a POSA would "look for larger numerical differences and consistent trends among truncation positions to identify significant changes in activity," and that POSAs "would not consider values of 2 and 3 to be meaningfully different, particularly between adjacent residues if not maintained as a trend," while considering values of "0.05 and 2" to be meaningfully different.[248]

218. To evaluate that assertion and her conclusions, I normalized the reported hyaluronidase activity of each truncation mutant as a percentage of the activity for $PH20_{1\text{-}447}$ and compared those values with the classification assigned by the US-457 publication and the classification assigned by Dr. Simpson:[249] As shown, Dr. Simpson reports classifications that are internally inconsistent and cannot be reconciled with the reported activity data.

---

[248] EX2068 (Simpson Dec.), ¶¶ 186–87.

[249] *See* Appendix B-6 (Activity Data Plot of C-terminal Truncation PH20 Mutants from US-457 and the '429 Patent); *see also* Appendix B-7 Comparison Table of Dr. Simpson's Activity Classification of C-terminal Truncation PH20 Mutants vs. US-457).

a. The report labeled $PH20_{1-430}$ "inactive" but labeled 5 mutants with comparable or less activity "reduced activity" (*i.e.*, $PH20_{1-432}$, $PH20_{1-456}$, $PH20_{1-458}$, $PH20_{1-460}$, $PH20_{1-462}$);

b. The report labeled 3 mutants with comparable or less activity than $PH20_{1-430}$ as having "inactive/little activity" (*i.e.*, $PH20_{1-463}$, $PH20_{1-464}$, $PH20_{1-465}$);

c. The report labeled one mutant ($PH20_{1-437}$) "active" even though it had only 5% more activity than $PH20_{1-430}$; and

d. The report labeled three mutants ($PH20_{1-457}$, $PH20_{1-459}$, $PH20_{1-461}$) as having "reduced activity" even though no activity data for those mutants are reported in US-457.

Terminus (precursor)	Terminus (mature)	pH 7.4	pH 5.5	Normalized Activity (pH 7.4)	Normalized Activity (pH 5.5)	Simpson Classification	US-457 Classification	Table
482	**447**	**2.03175**	**1.647**	**100%**	**100%**	**Active**	**Active**	**6**
465	**430**	0.441	0.468	22%	28%	Inactive	Active	6
466	**431**	0.000	0.212	0%	13%	Inactive	Active	6
467	**432**	0.545	0.900	27%	55%	Reduced Activity	Active	6
468	**433**	0.860	0.916	42%	56%	Reduced Activity	Active	6
472	**437**	0.545	0.900	27%	55%	Active	Active	6
490	**455**	0.819	1.238	40%	75%	Reduced Activity	Active	6
491	**456**	0.365	0.563	18%	34%	Reduced Activity	Active	6
491	**456**	0.851	0.729	42%	44%	Reduced Activity	Active	6A
492	**457**	n/t	n/t	n/a	n/a	Reduced Activity	-	none
493	**458**	0.689	0.851	34%	52%	Reduced Activity	Active	6
494	**459**	n/t	n/t	n/a	n/a	Reduced Activity	-	none
495	**460**	0.689	0.770	34%	47%	Reduced Activity	Active	6
496	**461**	n/t	n/t	n/a	n/a	Reduced Activity	-	none
497	**462**	0.176	0.216	9%	13%	Reduced Activity	Active	6
497	**462**	0.365	0.162	18%	10%	Reduced Activity	Active	6A
498	**463**	0.527	0.203	26%	12%	Inactive/Little Activity	Low expression	6A
499	**464**	0.020	0.045	1%	3%	Inactive/Little Activity	Low expression	6
499	**464**	0.486	0.122	24%	7%	Inactive/Little Activity	Low expression	6A
500	**465**	0.365	0.162	18%	10%	Inactive/Little Activity	Low expression	6A
501	**466**	0.446	0.122	22%	7%	None	Active	6
502	**467**	0.608	0.324	30%	20%	None	Active	6
503	**468**	1.094	0.648	54%	39%	None	Active	6

219. The conclusions of Dr. Simpson's report are also inconsistent with the characterization of C-terminal truncation mutants in the common disclosure. It states that a "mature PH20 polypeptide lacking the signal sequence and containing a contiguous sequence of amino acids having a C-terminal amino acid residue corresponding to amino acid 464 of SEQ ID NO:6 is the minimal sequence

required for hyaluronidase activity."[250] From this, a POSA would have understood that wild-type PH20 sequences terminating after position 429 (*e.g.*, including the 1-430 and 1-431 PH20 truncation mutants)[251] would be enzymatically active, as both retain the contiguous 1-429 region identified as sufficient for activity.

220. Dr. Petsko appears to have reproduced a portion of Dr. Simpson's table in his declaration (¶¶ 384-388), but offers no explanation for his classifications. For the same reasons discussed with respect to Dr. Simpson's opinions, I disagree with his opinions.

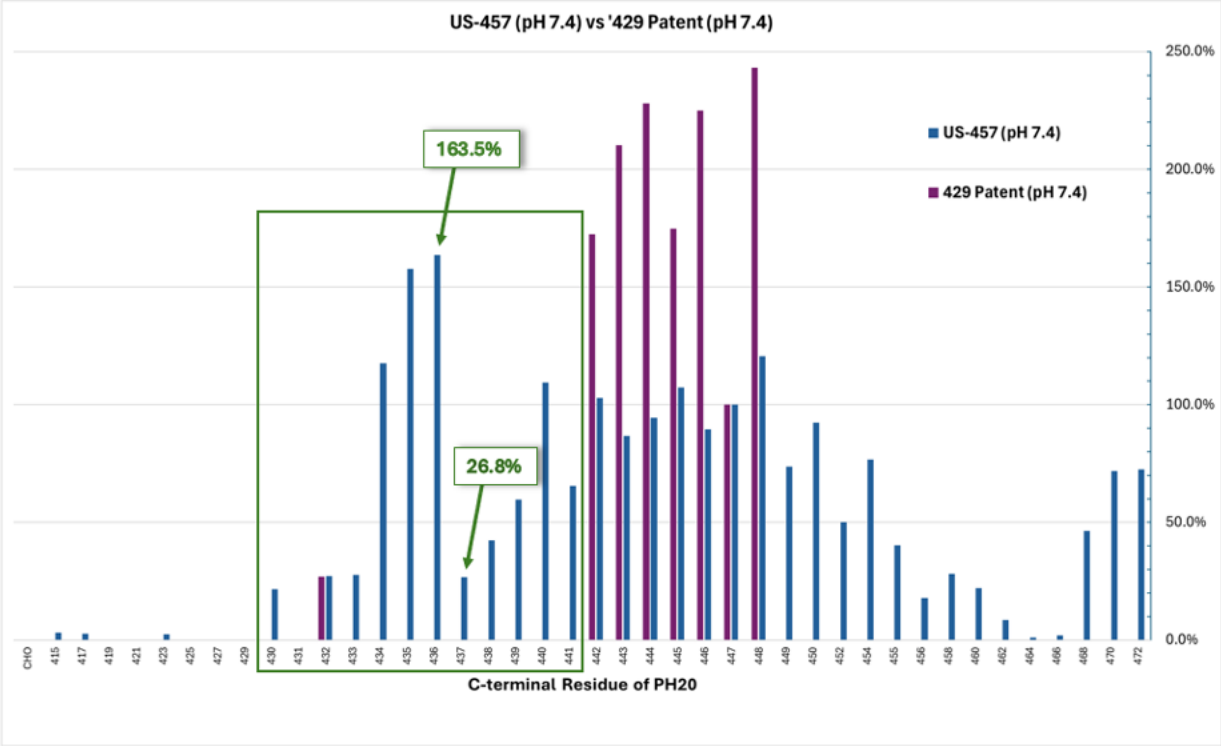
221. In ¶ 385 of her declaration, Dr. Simpson disputes Dr. Hecht's observations about the unpredictable activity of PH20 C-terminal truncation mutants. For example, Dr. Hecht noted that mutants differing by a single residue can show more than twofold differences in activity (¶ 91). Dr. Simpson responds by citing the C-terminal truncation data reported in US-457. However, the US-457 publication validates Dr. Hecht's point about the unpredictability of C-terminal truncations in adjacent C-terminally truncated PH20 proteins. While fluctuations between truncation mutants terminating between 446 and 448 were less pronounced than those reported in the '429 Patent (EX1005), US-457 shows

---

[250] EX1026 ('731 App.), 79:25-31.

[251] Position 464 in SEQ ID NO:6 corresponds to position 429 in SEQ ID NO:3.

similarly unpredictable activity elsewhere—for example, the $PH20_{1-437}$ mutant exhibited over 500% of the activity of the neighboring $PH20_{1-436}$ mutant, despite differing by only a single residue.



222. Mutations in this region are further complicated by epistasis, which can produce unpredictable and non-local effects. Combining multiple mutations in the C-terminus makes the aggregate impact increasingly more difficult to predict. I agree with Dr. Hecht[252] that, given the variability among the C-terminal truncation

[252] EX1003 (Hecht Dec.), ¶¶ 55-58, 91, 152-154.

mutants and the epistasis effect of making additional mutations, POSAs would be unable to reliably predict the outcome of mutating residues in this region.

6. Dr. Petsko's Theories About Substrate Binding Sites Is Contrary to the Guidance in the Common Disclosure

223. In ¶¶ 334-341, Dr. Petsko provides opinions about possible HA substrate binding sites on the PH20 protein based on his use of a computational tool called AutoDock Vina.[253] AutoDock Vina predicts and ranks ligand binding conformations using estimated binding energies, producing docking files that include PDB structures, binding scores, and predicted RMSD values for comparing and evaluating the docked conformations.[254]

224. The common disclosure discusses HA binding sites on PH20 in two paragraphs (below), based on published reports before 2011.[255] It indicates that the HA binding site is located in the Peptide 2 region (positions 170-200) of the PH20 protein:

---

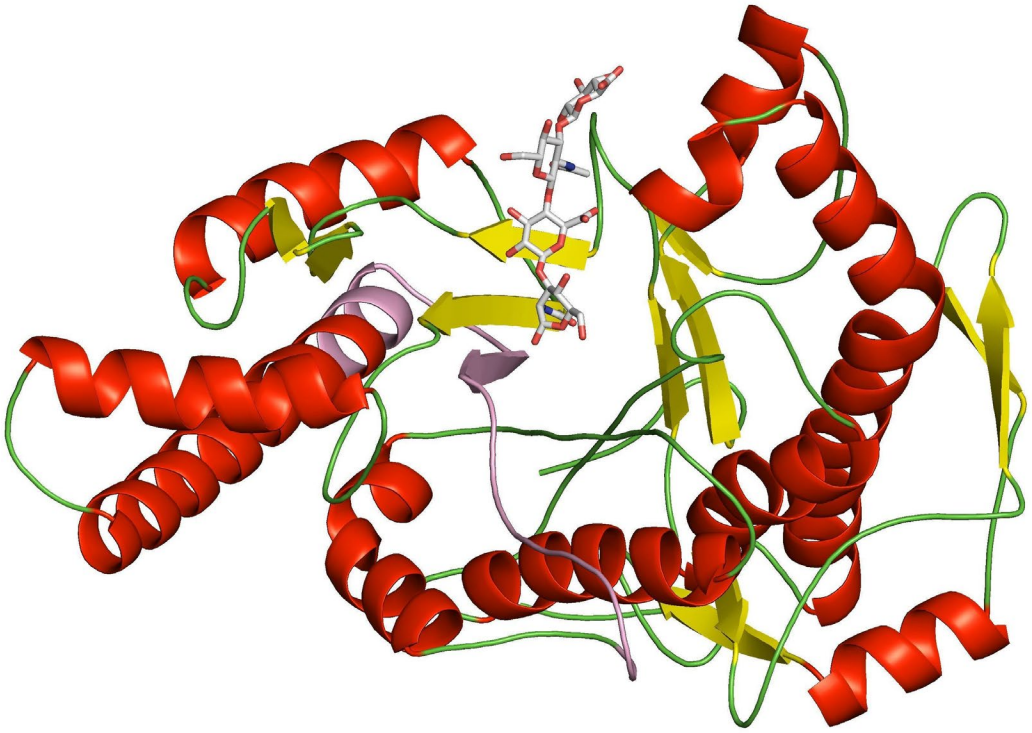
[253] EX2070 (Petsko Dec.), ¶¶ 334-341.

[254] EX1145 (AutoDock Vina Tutorial 2011), 7.

[255] EX1026 ('731 App.), 80:13-18, 81:7-14.

> Amino acid residues corresponding to amino acid residue D111, E113 and E249 of the sequence of amino acids set forth in SEQ ID NO: 3 or 7 are acidic residues part of the enzyme active site and are conserved between and among PH20 species. Amino acid residues R176, R246, R252 of the sequence of amino acids set forth in SEQ ID NO: 3 or 7 are also conserved between and among species and contribute to substrate binding and/or hyaluronidase activity. Amino acid mutations D111N, E113Q, R176G, E249N and R252T result in enzymes that have no detectable enzymatic activity or residual enzymatic activity (see *e.g.*, Arming *et al.* (1997) *Eur. J. Biochem.*, 247:810-814).
>
> In addition to the catalytic sites, PH20 also contains a hyaluronan-binding site. This site is located in the Peptide 2 region, which corresponds to amino acid positions 205-235 of the precursor polypeptide set forth in SEQ ID NO:6 and positions 170-200 of the mature polypeptide set forth in SEQ ID NO:3 or 7. This region is highly conserved among hyaluronidases and is similar to the heparin binding motif. Mutation of the arginine residue at position 176 (corresponding to the mature PH20 polypeptide set forth in SEQ ID NO:3 or 7) to a glycine results in a polypeptide with only about 1% of the hyaluronidase activity of the wild type polypeptide (Arming *et al.*, (1997) *Eur. J. Biochem.* 247:810-814).

225. Prior to 2011, Markovic-Housley et al (EX1033) described the crystal structure of a complex that forms between a hyaluronic acid tetramer and bee venom hyaluronidase. That structure would have allowed a POSA to approximate the positioning of hyaluronic acid on a PH20 structural model (as shown by Park). To illustrate this, I overlaid the bee venom structure (1FCV) on Dr. Petsko's PH20 homology model to show the predicted positioning of the bound tetrasaccharide in PH20.



226. Applying this structural overlay identifies predicted hyaluronic acid contact regions on PH20, including portions of the loop near Pro44, the loop near Tyr57, Asp111, the loop near Asp275, Trp304, and part of the helix near Leu307. Because productive substrate binding requires at least a hexamer of hyaluronic acid, the tetramer-bound structure reveals at most approximately two-thirds of the full substrate-binding site. The modeled substrate occupies the gap in the barrel with the catalytic Glu113 located on the loop connecting the additional α-helix spanning Lys126–Gln140. Marković-Housley et al. further predicted that the remaining portion of the binding site is formed by the loop following the β5′ strand (Pro187–Asn205), together with β6′ and β7′ on one side and the Glu113–Asn122

loop on the other.[256] The predicted PH20 tetrasaccharide complex does show residues within the 170-200 region interacting with the ligand, but it does not provide a comprehensive account of the interaction and includes many protein residues that seem almost certain to have no direct binding role.

227. Dr. Petsko also presents images of PH20-HA complexes generated with AutoDock Vina. The images are not annotated to show the HA positions or PH20 residue numbers involved in the purported interactions, and that information is also not provided in his declaration. Although Dr. Petsko states that his analysis relied on PDB files from AutoDock Vina,[257] and shows two-dimensional images depicting a single structure in complex with six residues,[258] he does not explain how those structures were selected or superimposed, or how interactions were evaluated.[259] The materials provided by Halozyme to Merck also do not identify the residues Dr. Petsko determined to interact with HA. As a result, based on the information in his declaration and the materials produced, his docking analysis cannot be independently reproduced or meaningfully evaluated.

---

256 EX1033 (Marković-Housley), 1028-1030.

257 EX1131 (Petsko Dep.), 258:13–259:8, 262:13–19, 259:9–24.

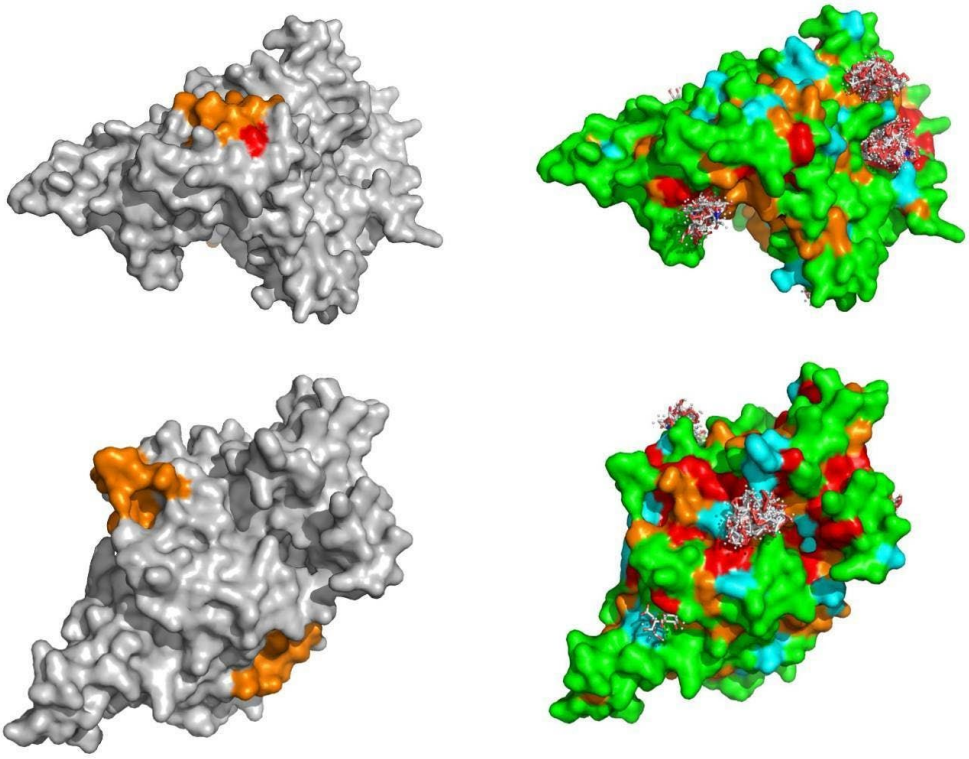
258 EX2070 (Petsko Dec.), ¶¶ 336–342.

259 EX1131 (Petsko Dep.), 263:7–264:2, 265:5–20.

228. The common disclosure discusses that residues R176, R246 and R252 "contribute to substrate binding and/or hyaluronidase activity."[260] It also reports that mutating the arginine residue at position 176 to a glycine reduces activity to 1% of wild-type. I note that Table 9 lists one mutant (R176K) that retained 60% relative hyaluronidase activity, and that R176G is listed as an inactive mutant in Table 10.

229. In the images below, I provide two other perspectives of the PH20 structure showing positions 170-200 in orange. I show adjacent to these the images from Dr. Petsko's report with the PH20 protein in the same orientation. These comparisons indicate that the additional HA-binding sites proposed by Dr. Petsko based on AutoDock Vina do not involve residues in the Peptide 2 region (positions 170–200).

[260] EX1026 ('731 App.), 80:13-18, 81:7-14.



230. This part of Dr. Petsko's analysis relies on potential additional HA binding sites that differ from those identified in the common disclosure and on information not disclosed in any published report prior to 2011-2012. I find, as of the date of this report, no experimental identification of such additional extended HA sites on PH20. The common disclosure does not reference the AutoDock Vina tool or a comparable computational tool for identifying HA binding sites. The hypothesis he offers is one for future experimental work, not guidance a POSA would have derived from the common disclosure.

IV. Response to Opinions of Drs. Petsko and Simpson Regarding the Ability of a POSA to Make and Use the Claimed Modified PH20 Polypeptides

231. Drs. Petsko and Simpson both characterize the common disclosure as providing a "comprehensive and systematic" analysis of the tolerance of each position in PH20 to amino acid substitutions. They also portray it as providing POSAs with "a functional map of PH20" that can be used to identify "which types of residues at each position are likely to be successful" in maintaining activity, can be targeted to increase activity, or can "inactivate" PH20. For example, Dr. Petsko stated:[261]

> 44. The '758 patent provides a comprehensive and systematic analysis of each position on PH20, achieved through large-scale mutagenesis and enzymatic activity testing of thousands of variants, with up to 18 mutations per position. These data provide a functional map of PH20 that identifies for the POSA the extent to which each position in the sequence tolerates substitutions and which do not, which types of residues at each position are likely to be successful in terms of maintaining activity, and which positions in particular could be targeted to maintain or improve activity. The patent also teaches which positions to change to inactivate PH20.

Dr. Simpson similarly states:[262]

---

[261] EX2070 (Petsko Dec.), ¶ 44; *see also* ¶¶ 54, 233-234.

[262] EX2068 (Simpson Dec.), ¶ 246; *see also* ¶¶ 45, 158, 161, 193, 212, 348, 375-376, 384.

> 246. The '758 patent provides a detailed functional map of PH20 through its comprehensive and systematic analysis of PH20, achieved through large-scale mutagenesis and enzymatic activity testing of thousands of variants, with up to 18 mutations per position. The results show positions that a POSA would understand are critical and tolerate minimal modification or no modification, as well as residues that do tolerate substitution, and particular substitutions that substantially increase activity.

232. These statements are based on experimental work discussed in the common disclosure that investigated the effects of single substitutions within one PH20 sequence ($PH20_{1\text{-}447}$).[263] I do not share these opinions of Dr. Petsko and Dr. Simpson. As I explain further (below), several of the opinions provided by Drs. Petsko and Simpson also contradict the guidance in the common disclosure. I discussed these in § III.G, above.

a. The well-established phenomenon of non-additivity/epistasis observed when proteins incorporate multiple changes precluded POSAs from being able to predict the effects of multiple changes to the PH20 sequence in the 2011-2012 timeframe.

---

[263] EX1026 ('731 App.), 219:28-220:28, 247:2-248:6 (Example 2, including Table 8), 248:11-250:31 (Example 3), 250:32-251:20 (Example 4- Active Mutants, including Table 9), 272:2-22 (Example 4- Inactive Mutants, including Table 10).

b. Information on which positions are generally known to inactivate hyaluronidase proteins was known before 2011 but could not provide insights into the effects of making multiple changes outside of those positions.

c. The common disclosure did not provide a "comprehensive and systematic" functional map establishing the tolerance of each position in PH20 to change. The common disclosure did not present any analysis of the data generated in its mutagenesis study, other than a classification of the mutants into "active mutant" and "inactive mutant" categories.

d. The experimental inputs for Dr. Petsko's positional tolerance analysis have methodological and reporting deficiencies that would have caused a POSA to doubt the scientific credibility, precision, and reproducibility of his tolerability ratings. A POSA would have viewed that data as providing only limited insights into the effects of substitutions at certain positions on the folding and activity of PH20 proteins.

e. Homology models and techniques of using them to assess changes to a protein's structure were incapable of predicting the effects of more than a few concurrent changes. They are incapable of predicting effects of long-range interactions and other phenomena that can prevent folding or retention of the protein's tertiary structure.

233. I therefore agree with Dr. Hecht's opinions that a POSA could not have used the information in the common disclosure with what was known by

2011-2012 to predict which multiply-modified PH20 sequences within the immense sequence space defined by the claim parameters could be produced, much less would properly fold and be enzymatically active or be soluble, or increase delivery of a therapeutic.[264] Determining which multiply-modified PH20 polypeptides do, would require individualized assessments of changes and experimentation, that would demand an impossible scale of effort.[265]

A. POSAs Could Not Identify or Predict Which Multiply Mutated PH20 Polypeptides in the Claimed Sequence Space Can Be Produced or Will Stably Fold

234. The common disclosure classifies single-substitution mutants made in a single "background" (*i.e.*, the 1-447 residue wild-type sequence of PH20) ("$PH20_{1-447}$") based on a biotinylated hyaluronic acid activity assay.[266] Of the ~ 6,753 mutants reportedly generated, ~ 5,857 were classified as either "active" or "inactive" based on measured hyaluronidase activity.

235. Hyaluronidase activity assessed from two supernatant samples per mutant. Mutants were classified as "Active" if at least one measurement was ≥

---

[264] EX1003 (Hecht Dec.), ¶¶ 61, 140-142, 156-159, 167-168.

[265] *See* EX1003 (Hecht Dec.), ¶ 181 ("…the major challenge in using directed evolution techniques in protein engineering is scale.").

[266] EX1026 ('731 App.), 249:4-251:20 (Examples 3-4).

40% of the wild-type $PH20_{1-447}$ activity and "inactive" if at least one measurement was < 20%. "Active mutants" are listed in Tables 3 and 9, with averaged activity values, while the "inactive mutants" are listed in Tables 5 and 10, without any reported activity values. I discuss shortcomings in these experimental procedures and the reporting of data further below (*see* § IV.C).

236. I reviewed the mutant classifications reported in the common disclosure and in Dr. Petsko's analysis as well as Dr. Hecht's survey of the mutants.[267] The chart below shows the number of mutants in each category.[268]

a. Approximately 2,357 supernatants exhibited at least 40% relative activity and were classified as "active mutants";

b. Approximately 3,324 supernatants were classified as being "inactive mutants";

c. Approximately 126 of the mutants had measured activities between 20 and 40%;

d. There are approximately 50 errors in the classifications of the mutants (*e.g.*, due to double counting, errors in counts, etc.)[269];

---

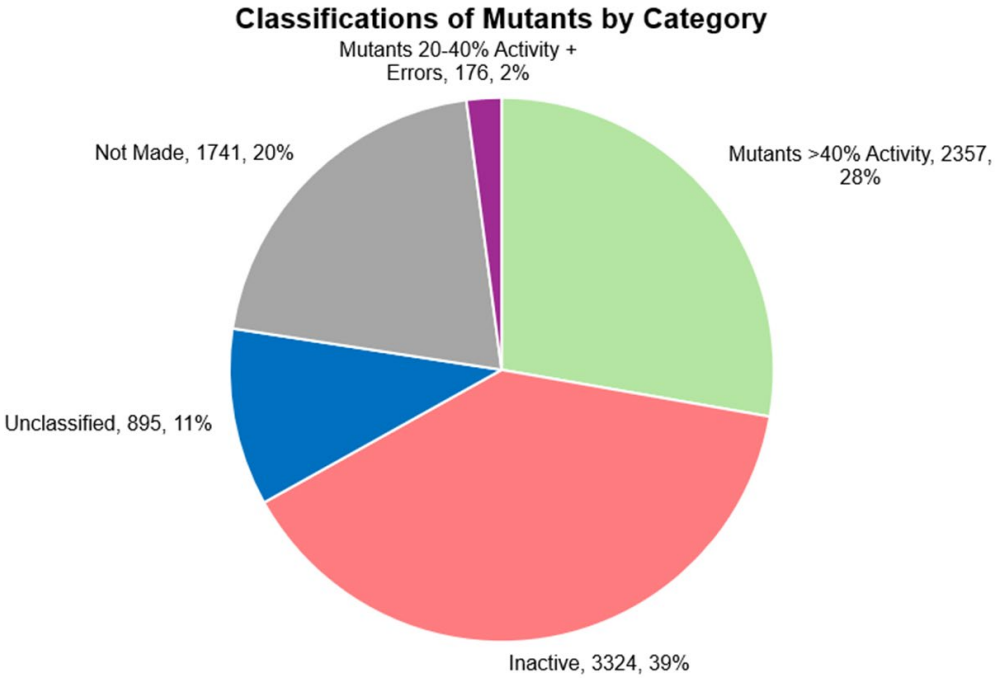
[267] EX1003 (Hecht Dec.), ¶¶ 103-107.

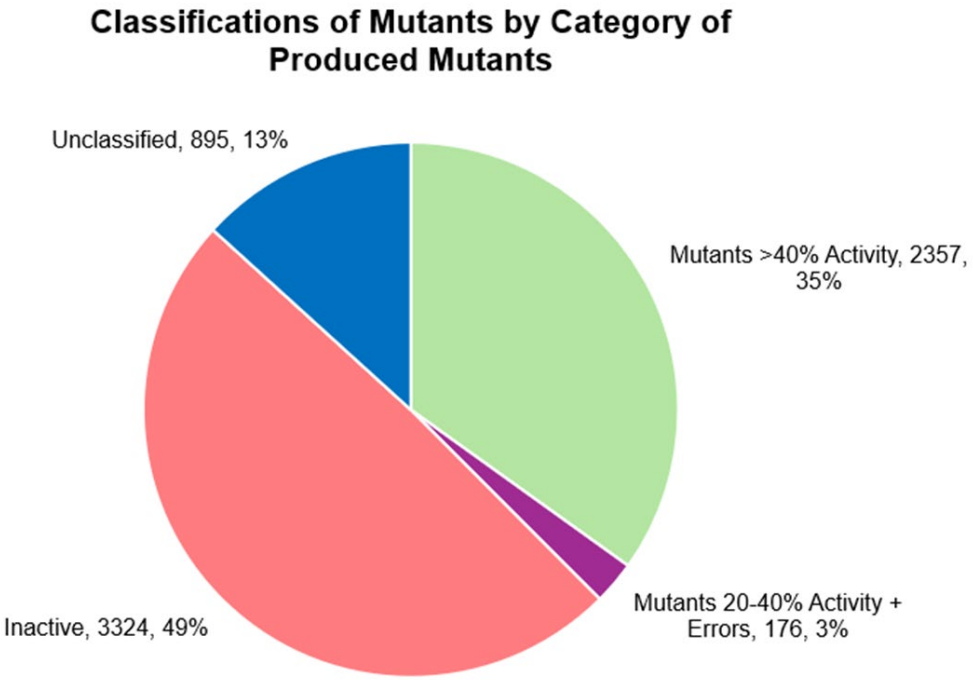
[268] *See* Appendix B-1 (Counts and Classifications of $PH20_{1-447}$ Mutants).

[269] *See* § IV.B.

e. Approximately 895 of the mutants made (Table 8) were not classified as an "active" or "inactive" mutant. No reason is provided for the lack of classification of these mutants; and

f. Approximately 1,741 substitutions were not made.





237. Drs. Petsko and Simpson both agree that the claim parameter sequence space is immense, encompassing from 10^{59} to 10^{112} possible amino acid sequences. As Dr. Petsko explains (¶ 347), a POSA would expect that billions of these sequences would not fold into PH20 proteins because they alter residues critical to protein structure. While some such residues can be predicted (*e.g.*, highly conserved positions), many others cannot without experimentation, as Dr. Park determined (¶ 261), including residues and patterns that play unknown but critical roles in folding and stability. The limited testing of approximately 5,857 single-amino-acid substitutions does not support broader conclusions across this immense sequence space.

B. Double-Counting and Other Errors in Dr. Petsko's Positional Tolerance Analysis

238. In ¶¶ 263-274 of his declaration, Dr. Petsko explained how he prepared his positional tolerance table on pages 127-128. As he explains, he used the data compiled in Tables 3 and 9 for "active mutants" and in Tables 5 and 10 for "inactive mutants." He also explains in footnote 18 (page 125) that if there were discrepancies for "active mutants" between Tables 3 and 9, he used the

classification in Table 9, and if there were discrepancies for "inactive mutants" in Tables 3 and 10, he used the classification in Table 10.[270]

239. Several individual mutants are classified as both "active" and as "inactive" mutants in the common disclosure (*e.g.*, mutants listed in both Tables 9 and 10). Dr. Petsko did not address these "double counted" mutants in his declaration. At his deposition, Dr. Petsko said that "he looked for that a little bit, but I don't know that I was completely systematic."[271] He also agreed that a POSA would need to be cautious about using any of those double counted mutants in his tolerance assessment table.[272]

240. Dr. Petsko's response suggests that he did not himself conduct a systematic review for these internally inconsistent classifications of mutants in the common disclosure.[273] Had Dr. Petsko performed such a review, it would have

---

[270] EX1131 (Petsko Dep.), 213:18-24 ("Q. So basically, if you had an inconsistency in the active mutant pair, you went with the Table 9 classification. And then with the inactive 5 and 10, you went with the Table 10 classification? A. Yes.").

[271] EX1131 (Petsko Dep.), 213:25-214:8.

[272] EX1131 (Petsko Dep.), 214:9-215:5.

[273] Dr. Petsko confirmed that he did not prepare the spreadsheet that was used to prepare the tolerance classifications in his declaration. *See* EX1131 (Petsko

revealed that a substantial number of mutants (50) were double counted as both active and inactive (below), including: L061F, P065R, I067R, D068G, D068L, D068P, S069T, G072H, Q081P, K082Y, I083K, S084Y, Q086P, D087P, L089W, D090G, A092H, L105C, P117D, W119P, Q143C, L165C, P177V, W181K, Y183E, K195S, G198W, N211W, W218V, P226W, D275L, D275V, K279A, V287N, T289S, G291C, I303D, W304G, W304I, M310F, D320V, I331C, I331E, V351C, F398L, A412H, D413K, D413H, V414K and A425Y.

241. These double counted mutants create an ambiguity that cannot be resolved using the information only in the common disclosure. The common disclosure, for example, does not report measured hyaluronidase activity values for the inactive mutants. A POSA would have found the classifications of these mutants to be unreliable and would not have used them in analyzing the tolerability of a position to change.[274] There are tabulation errors in the spreadsheet provided to Dr. Petsko (*e.g.*, position 149 and 205).

Dep.), 13:16-21 ("Q. … Did you prepare this spreadsheet? A. Myself directly? No. I don't work with Excel. Q. Okay. So you were given an Excel worksheet, and then you reviewed it? A. Yes.").

[274] Tables 9 and 10 also contain other errors, for example: E249V has no reported activity value but is listed in Table 9 (EX1026 ('731 App.), 267); Table 9 lists mutant K297R even though G297 is wildtype (EX1026 ('731 App.), 262);

242. Appendix B-1 is a spreadsheet, made with the assistance of counsel, that corrects the various errors made in Dr. Petsko's spreadsheet and identifies positions where the errors justify reclassification of his original tolerance ratings. It also shows positions where there were fewer mutants that were reliably investigated than what is shown in Dr. Petsko's spreadsheet (EX2166). To identify these errors and tabulate different perspectives of the information provided by Dr. Petsko and the common disclosure, I relied upon spreadsheets made with the assistance of counsel.[275] Additional information from the spreadsheet is included as part of Appendix B:

a. Appendix B-2 is a composite spreadsheet of the activity data reported in the common disclosure and from the work of Dr. Park;

b. Appendix B-3 compares the assessment of which residues are "tolerant" to modification and offers lists of positions that meet various criteria (*e.g.*, adequate physicochemical testing);

c. Appendix B-4 reports solvent assessable surface area data reported by Dr. Park (EX1004);

R311G is listed twice in Table 9 with two different activities (EX1026 ('731 App.), 253); and Table 10 lists mutants G378D, G378E, G378F, and G378I twice (EX1026 ('731 App.), 272-273, 282).

[275] EX1135 (Dr. Naismith Analysis Spreadsheet).

d. Appendix B-5 contains a histogram distribution plot of activity data of $PH20_{1-447}$ mutants reported in the common disclosure;

e. Appendix B-6 compiles activity data of C-terminally truncated PH20 mutants from US-457 (EX2165) and the ’429 Patent;

f. Appendix B-7 compares Dr. Simpson’s analysis against the data reported by US-457;

g. Appendix B-8 contains alternate amino acids considered by Dr. Park as part of his MSA analysis;

h. Appendix B-9 illustrates the breadth of Dr. Park’s analysis when evaluating alternate amino acids across positions in PH20;

i. Appendix B-10 contains a PH20 positional template showing recurring domains, features, and residues identified in the prior art; and

j. Appendix B-11 contains Dr. Petsko’s position tolerance analysis vs. residues identified as essential in $PH20_{1-447}$.

C. Dr. Petsko’s Positional Tolerance Analysis Rests on Unreliable Assumptions and Data

243. Drs. Petsko and Simpson base many of their opinions on Dr. Petsko’s positional tolerance analysis, using it to identify residues they deem critical to

PH20 folding and activity, assess positional sensitivity to mutation, and guide selection of amino acid substitutions.[276]

244. In his positional tolerance analysis, Dr. Petsko rates a position's "tolerance" based on whether a mutation retains its ability to properly fold and adopt its native tertiary structure, which is necessary for hyaluronidase activity. His use of "tolerance" in a manner that is consistent with Dr. Park's definition[277] and is derived solely from classifying individual mutants as "active" or "inactive" based on measured hyaluronidase activity—the only experimental readout reported in the common disclosure. Those classifications, in turn, are based on the measured

---

[276] *See* EX2068 (Simpson Dec.), ¶¶ 243-247, 310-314, 380-382; EX2070 (Petsko Dec.), ¶¶ 54, 276-282, 291, 317-324, 353-360, 366-373; EX1130 (Simpson Dep.), 35:6-20, 186:10-22, 191:11-16, 194:10-195:11; EX1131 (Petsko Dep.), 74:18–75:23, 76:6–12, 148:21-150:5, 152:18-24, 214:9–25.

[277] EX1131 (Petsko Dep.), 119:4–22 ("Q. So is his usage of 'tolerance' consistent with your usage of tolerating in your opinion? A. It's a subset of my usage of tolerate, yes…. Q. So if he, for example, is saying that -- or if a POSA was saying, 'This amino acid substitution at this position was tolerated,' so that you were kind of using the word 'tolerated' in connection with a particular position in the sequence, when your interp- -- your usages would align with Dr. Park's? A. In that particular situation.").

hyaluronidase activity of each mutant, which is the only experimental readout reported in the common disclosure.

245. "Active mutants" (≥ 40% of wild-type PH20 activity) represent a "positive" experimental result that may be informative as to proper folding. For example, Dr. Petsko indicated that he assumed that a mutant had "properly folded" if it exhibited enzymatic activity. [278] The common disclosure, however, did not investigate independently whether mutants were secreted.

246. Dr. Hecht expressed concerns over use of the mutants classified as inactive to draw conclusions about PH20 folding,[279] and I share those concerns. The primary reason for this conclusion is that the common disclosure does not provide information that permits a POSA to determine ***why*** any particular single amino acid substitution in the $PH20_{1-447}$ wild type sequence exhibited reduced (<

---

[278] *See* EX2070 (Petsko Dec.), ¶ 128, n.12 ("When I say "enzymatically active" in this context, I am not referring to a specific level of enzymatic activity. Rather, I am using enzymatic activity as a proxy for a properly folded protein. Secretion is another proxy that could be used for a properly folded protein; proteins that do not fold will generally not be secreted.").

[279] EX1003 (Hecht Dec.), ¶ 106 ("The relative activities reported are difficult to assess because the specification does not provide measured values for the activity of the unmodified $PH20_{1-447}$ used as the reference for percentages reported in Table 9."); *see also* ¶¶ 113, 142, 151, 169.

20 %) hyaluronidase activity. A POSA cannot tell from what is reported in the common disclosure whether an inactive mutant was classified as such because (i) it was not successfully synthesized by the cell, (ii) it was not secreted, (iii) it did not fold within the cell, (iv) it properly folded and was secreted but unfolded after secretion due to instability, (v) it was secreted in a misfolded or partially folded state, or (vi) it was properly folded but exhibited impaired or reduced enzymatic activity.

247. The possible outcomes of a single substitution in $PH20_{1-447}$ wild-type protein (*see* ¶ 48) have distinct implications for folding and positional tolerance. Outcomes (i)–(v) reflect defects in expression, secretion, correct folding, stability, or aggregation. It is not possible from the data provided in the common disclosure to ascertain which of these reasons (or experimental errors) caused the mutant to be reported having activity below the 20% threshold. Grouping these disparate outcomes under a single "inactive" classification obscures the cause of inactivity and prevents a POSA from drawing a meaningful structure-function inference about positional tolerance in PH20.

248. Among the flaws I identified in the hyaluronidase assay design and reporting of results for inactive mutants were the following:

a. no steps taken to confirm that each inactive mutant PH20 was successfully produced or results from those investigations;

b. no steps taken to confirm secretion or quantify levels of inactive mutants in the supernatants, or report any results;

c. no steps taken to detect aggregation of inactive mutants in supernatants, or results from such investigations;[280]

d. no steps taken (*e.g.*, circular dichroism) to detect whether any inactive mutants had folded, or results from such investigations;

e. reports that inactive mutants were measured at two time points[281] but does not report any of the measured hyaluronidase activity values for inactive mutants (*i.e.*, only listed in Tables 5 and 10);

f. no evidence of controls regarding cryopreservation methods of −20°C freezing followed by freeze/thaw prior to activity screening. This protocol could plausibly promote uncontrolled variability in the measured sample activity;

g. no turbidity measurements or other assessments of sample clarity to detect aggregation;

---

[280] EX1131 (Petsko Dep.), 239:12-23 ("Q. It could be that the protein was secreted but then aggregated to the point where it had no activity? A. Less likely. You'd probably pick that up as turbidity. Q. Right. There is no measurement of turbidity in these - A. You just notice it, if you were careful. Q. Just from what we have in the patent, we don't know if it is one possible outcome is -- it is an aggregated outcome, right? A. Um-hmm.").

[281] EX1026 ('731 App.), 248:1-5, 272:3-19.

h. no disclosure of the standard curve, positive-control data, or statistical analysis needed to determine whether "inactive" classifications reflect true loss of activity or assay variability; and

i. no protein-level evidence that PH20 was expressed or secreted for mutants classified as "inactive," and no validation that SEAP normalization controlled for construct-to-construct variation in PH20 expression or secretion.

249. Dr. Petsko confirmed that a POSA could not determine which outcome led to an "inactive" classification or whether such a mutant had properly folded, agreeing that an "inactive" label does not differentiate a mutant that was produced and secreted from a cell from one that never exited the cell:[282]

Q. Sure. In -- in the table, you wouldn't be able to differentiate that inactive mutant that had secreted and folded from an inactive mutant that never got out of the cell?
A. In the table --
Q. Yes.
A. -- that would probably be true.
Q. Okay. So do you -- the mutants in -- listed in Table 10 give you information about the folded state of each of those mutants?
A. Not from just the list of the mutants per se.

[282] EX1131 (Petsko Dep.), 242:24-243:10.

250. Dr. Petsko confirmed that because the common disclosure does not report measured hyaluronidase activity values for inactive mutants, a POSA cannot tell whether an "inactive" mutant had near zero activity, activity near the 20% threshold, or inconsistent replicate results. As Dr. Petsko explained, a level of activity near the 20% threshold (provided it was above the error in the assay) would suggest the inactive mutant had folded and secreted.[283]

Q. If you had activity values for the inactive mutants that were above zero, that might give you some indication that the protein had folded?
A. It would depend on how much above zero they were.
Q. If you had, let's say, 18 percent activity?
A. Oh, I would think that was a folded protein.
Q. If you had 18 percent activity for an inactive mutant, you also would be confident that it had been secreted, right?
A. Yes. Some of it anyway.

251. Dr. Petsko confirmed that activity above assay error at different time points would suggest stable folding, while the loss of activity over time could indicate unstable folding, and that assay noise precludes drawing conclusions

[283] EX1131 (Petsko Dep.), 241:3-16; EX2070 (Petsko Dec.), ¶ 128 n.12.

where measurements fall below the error level. I agree that conclusions about folding require measurements above the error level (or noise) in the assay.[284]

So if I -- you said earlier that the fact that the first mutant with 18 percent activity at the first reading and the second reading, that protein has folded and has stayed folded at least for another 12 hours.

A. If your assay permits you to assign a value of 18 percent as a measurement of activity above error -- the error level, yes. But if it doesn't, then no.

Q. Okay. Assuming that it is above the error level, then that --

A. If the error levels say 10 percent, then you have 8 percent above that, you might be satisfied at calling that active.

Q. Okay. And -- and that means it's also folded?

A. It would -- mean that it's folded, yes.

Q. For the other mutant which drops to 0 percent at 12 hours, that is an unstable protein because it's got no residual activity at --

A. Could be, or you could have -- the two measurements could just have that much variation.

Q. All right.

A. You would have to know that.

Q. So the variation of the assay in terms of the activity variation is something that would influence these classifications that are made in the patent?

A. It would -- it would certainly -- it would certainly need to be considered in the way you looked at the data. Okay?

Q. And that's not possible for the dataset of the inactive mutants, which are classified only by falling below the threshold?

A. Yes, that's correct.

D. POSAs Could Not Have Predicted Which Combinations of Substitutions Tolerated Individually Would Yield, Enzymatically Active, Soluble Modified PH20 Polypeptides

1. Epistasis Associated with Multiple Changes in a PH20 Polypeptide Is Unpredictable

252. Neither Dr. Simpson nor Dr. Petsko directly address the well-known phenomenon of epistasis in discussing the predictability of making PH20 polypeptides with multiple amino acid changes. Epistasis was widely recognized by POSAs in 2011-2012 and has profound implications for introducing multiple

[284] EX1131 (Petsko Dep.), 247:1-248:11.

mutations into a polypeptide. Dr. Petsko made observations in his declaration and at his deposition reflecting his awareness of the phenomenon of epistasis.

a. In ¶¶ 323-324, he discusses making combinations of changes using surface residues that are tolerant to change. He contrasts making substitutions "in different parts of the protein" where "they would be independent of each other" from "making several changes close together on the protein." In the latter situation, he says the POSA would need "an assessment of the local environment of the residues" and would need to determine if "the desired changes are likely to preserve the interactions formed by the WT residues" or if they "are likely to have an adverse impact on, for example, packing." His use of the term "independent" reflects an awareness that combinations of mutations can interact unpredictably through epistatic effects. However, in the 2012 timeframe, a POSA could not reliably predict whether mutations would behave independently; that could only be determined experimentally.

b. At his deposition he referred to a mutation involving "action at a distance," and described it as common, complicated and subtle (two or more distant mutations affecting the active site would raise concerns about epistatic effects):[285]

[285] EX1131 (Petsko Dep.), 79:15-22.

But the most common things are more complicated and more subtle. They involve, for example, action at a distance. You could have a mutation that changes the electrostatic potential of the protein at the active site even though it's not in the active site, for example. But it would still have exactly the same fold. We see that sort of thing often.

c. He referenced "long-range" effects, where mutations could cause effects even though two residues were not directly interacting with each other (below).[286] For example, he identified "long range interactions" as one of the types of effects that a mutation could have that would result in loss of activity in a secreted mutant.

[286] EX1131 (Petsko Dep.), 89:3-90:24; 270:9-17 ("Q. So are you suggesting those are some of the residues that interact with hyaluronic acid? A. They might be. But the patent doesn't form that conclusion. They say they could be. All right? They could contribute in other ways. These are charged residues. All right? And charged residues can exert long-range electrostatic effects, for example, that have nothing to do with direct binding.").

Q. So if I -- if I understand you, there is one kind of mutation that you can rationalize as explaining the lack of activity of a secreted, mutated PH20 polypeptide, and that is one which affects the -- one of those residues involved in catalysis or hydrolysis?

A. That's one of the mutations I can envision. There are others.

Q. And the -- the other, I think, we talked about it earlier, mutations that were made to the invariant residues of the protein that are found across the homologous protein set?

A. That's another.

Q. Okay. The other category, I think, were residues that you indicated a person wouldn't understand or associated with the alpha-beta barrel fold structure?

A. That's third. But there are others.

Q. What are -- what are other ones?

A. For example, I mentioned earlier long-range interactions. Usually of the electrostatic type. Those are often -- it is often possible to determine that that's likely what happened by the nature of a substitution in its location.

Q. And the disruption of that long-range electrostatic interaction would cause the protein to no longer adopt the PH20 alpha-beta barrel fold?

A. Quite the opposite.

Q. Quite the opposite.

A. We'd still have the alpha-beta barrel fold, but it would have an effect at the active site that would perturb it's catalytic activity without directly modifying the amino acids in the active site.

Q. So that scenario is one where the spatial positioning of the atoms that are involved in hydrolysis in the catalysis activity of the protein have shifted somehow so --

A. No. That's yet another category, which we can talk about. No, it's actually more a question of the chemical nature of the residue. So for example, there is a glutamic acid residue in the active site as of hyaluronic base. It must be able to donate a proton to the substrate in the catalytic mechanism.

253. Research has demonstrated that additive effects of individual mutations are frequently not observed when they are combined. Proteins are complex, non-linear molecular systems, and deviations from the additive effects of the mutations is epistasis. Generally, epistasis refers to a non-additive interaction of an additional mutation, where the functional effect of the functional mutation is context-dependent.

254. Epistasis is commonly categorized into three types.[287] *Magnitude epistasis* occurs when the combined effect of two mutations differs from the effect

[287] EX1160 (Poelwijk), 383-385.

predicted by simple additivity, even though the direction (beneficial or deleterious) of each mutation remains the same. *Sign epistasis* occurs when a mutation that is beneficial in one sequence background becomes deleterious in the presence of another mutation. *Reciprocal sign epistasis* occurs when two mutations that are each individually beneficial become jointly deleterious when combined, yielding a protein that is less active than the native form. These principles apply not only to enzymatic activity, but also to protein stability and folding, reflecting the fundamentally non-linear effects of combining mutations on protein properties. Quantifying epistatic interactions is highly complex, and in the 2011–2012 timeframe such effects could not be predicted in a general or reliable manner.

255. The observed impact of an additional mutation—whether beneficial, neutral, or deleterious—depends on the other mutations already present in the protein sequence. This context-dependence is the central reason why single-mutant data, which measures each mutation only in the wild-type context, has limited predictive power for combining mutations in a single polypeptide. Epistasis progressively limits the ability of POSAs to predict the functional impact of additional mutations and strongly hinders rational protein design strategies where multiple mutations are introduced.

256. Throughout his declaration, Dr. Hecht explains why POSAs could not predict the effects of combining single substitutions in modified PH20 based on

single mutant data. For example, he notes that the introduction of combinations of changes can alter patterns of amino acids needed to maintain secondary structures or introduce new steric and/or chemical interactions, which may not occur with a single substitution on its own.[288] Dr. Park discussed the challenges of predicting the effects of combining mutations, reflecting the unpredictable nature of epistasis.[289]

257. Drs. Hecht and Park both observed that POSAs could not reliably predict the effect of more than a few concurrent changes within a protein.[290] This was due in part to the limits of computational tools available in 2011-2012 (*see* § IV.E) to model epistasis. A single mutation to a protein will create a different set of interactions with its neighboring residues than the wild-type residue, which leads to small changes in the positions of these neighboring residues which may alter the dynamics, stability, activity or secondary structure (or any combinations

---

[288] *See, e.g.*, EX1003 (Hecht Dec.), ¶¶ 54, 56, 60, 61, 158, 159, 184.

[289] EX1004 (Park Dec.), ¶¶ 157-159.

[290] EX1003 (Hecht Dec.), ¶¶ 50, 58-59, 158-159; EX1004 (Park Dec.), ¶ 159 (explaining that POSAs could not reliably predict effects of modeling more than about 5 concurrent changes in 2011-2012 timeframe).

of these) of the protein. And, as Dr. Petsko noted, effects can occur at a distance.[291] When a second mutation is introduced into the protein having the first mutation, the protein being mutated is subtly different than the native, and there is therefore the potential that the changes introduced by the second mutation interact with the changes introduced by the first in a way that is different than how the second mutation would interact in the native protein. As each additional mutation is introduced, the potential for epistasis grows

258. I agree with Dr. Hecht and Dr. Park that after about 5 mutations, a POSA could no longer predict enzyme behavior by simple linear combinations of single site mutations.[292] Such deviations from linearity are likely to occur with fewer changes where the mutations are concentrated in a particular region of the protein, as noted by Dr. Hecht. Similar perspectives are reflected in the pre-2011 scientific literature.[293] Dr. Petsko also seems to share this perspective. *See* ¶¶ 129-

---

[291] EX1131 (Petsko Dep.), 79:15-22.

[292] EX1003 (Hecht Dec.), ¶ 158 (noting that homology models cannot reliably assess the effect of more than 5 substitutions); *see also* ¶ 54-61; EX1004 (Park Dec.), ¶ 159 (stating that POSAs could not reliably predict the effect of 5 or more substitutions to a sequence using Swiss-Model).

[293] EX1146 (Pal), 22381 ("Interestingly, the addition of a fourth mutation (D171S) did not improve affinity further suggesting that additivity principles

131. That is also the practical limit for screening enzyme libraries in the 2012 timeframe. Even using the most sophisticated screening technologies, the largest library that could be screened was around ~10^9 proteins. *See* ¶¶ 124-127.

259. Deleterious epistatic interactions can arise when individually tolerated mutations are combined, disrupting proper folding. Such efforts may be local and long-range, reflecting interactions between otherwise independent changes.[294] Consequences of these deleterious epistatic interactions include:

a. Disruption of successful folding of a nascent polypeptide chain within the cell by disrupting co-translational folding kinetics or creating aggregation-prone folding intermediates. For example, the combination may alter local or long-range intramolecular interactions not caused by any single substitution;

b. Unpredictable changes in protein stability, where independently neutral substitutions may collectively disrupt packing interactions, hydrogen-bonding networks, electrostatic

---

begin to deteriorate when too many mutations are incorporated into a proximal area.").

[294] EX1131 (Petsko Dep.), 9-17 ("And charged residues can exert long-range electrostatic effects, for example, that have nothing to do with direct binding."); EX1003 (Hecht Dec.), ¶ 58 ("Introducing nine more substitutions… may collectively destabilize that region of the protein structure…."); *see also* ¶¶ 59-61.

complementarity, or dynamic motions that are only critical when multiple sites are altered simultaneously. Impacts on stability can also occur via long-range electrostatic or allosteric effects associated with the individual changes; and

c. Disruption of different functional effects in combination that are not observed as individual mutations (*e.g.*, altering the spatial positioning of catalytic residues, impairing substrate binding, causing long-range electrostatic effects on either).

2. The Data in the Common Disclosure and a PH20 Homology Model Cannot Be Used to Predict Which Sequences in the Claim Parameter Sequence Space Will Fold or Exhibit Hyaluronidase Activity, or Solubility or Increased Delivery of a Therapeutic Agent

260. In the 2011-2012 timeframe, epistatic interactions were not predictable from single-mutation analyses, and POSAs could not reasonably infer from the common disclosure or PH20 homology models which combinations of 5, 10, 15 or 20 changes would result in a PH20 variant that would fold or exhibit activity, or increase delivery of a therapeutic agent.

261. Dr. Park explained that SWISS-MODEL cannot reliably predict the effects of multiple mutations on a PH20 sequence, which become "intractable" as mutations accumulate due to model limitations and compounding uncertainty of

each additional substitution (this is epistasis).[295] For example, the visual analysis of protein structure can suggest sites of modification, but further investigation is needed to navigate the intricate networks of amino acid interactions, as substitutions may lead to rearrangements in protein cores and binding interfaces that depend on three-dimensional, geometric packing of native residues.[296]

262. In ¶ 292, Dr. Petsko states that POSAs would have expected PH20 to be "relatively robust in structure" because "it adopts an alpha-beta barrel fold … that is known to be a generally robust fold that is tolerant to substitution." As I explained in § III.C, the data show as Dr. Petsko notes this is not true for PH20. Even if mutations each individually yield ≥ 100% relative activity, a POSA could not predict the effect of combining 5, 10, 15, or 20 such changes in a single PH20 polypeptide due to epistasis. As Dr. Hecht explained, a single amino acid substitution tolerated by PH20 does not lead to the conclusion that this same substitution will be tolerated when additional substitutions are introduced because of the collective impact these mutations may have on the PH20 structure.[297]

---

[295] EX1004 (Park Dec.), ¶¶ 156–158; EX2077 (Park Dep.), 204:7–206:3.

[296] EX1017 (Green), 234.

[297] *See* EX1003 (Hecht Dec.), ¶ 61.

263. Dr. Petsko notes that the data "shows that [PH20] is not" tolerant to substitution, contrary to what a POSA might have expected.[298] That observation reinforces that predictions about multiple mutations would be unreliable because most single substitutions were shown to exert effects on the structure–function relationship. Dr. Petsko, however, suggests that the patent data "are extremely valuable to the POSA seeking to make variant PH20 polypeptides because the data show the POSA which positions are likely to tolerate substitutions (and what types of substitutions) and which are not."[299] Where data exist, this is correct for single substitution changes in PH20. However, the data on single site mutants do not allow a POSA to reliably predict the effect of making 5, 10, 15 or 20 concurrent changes.

3. The Active Mutant Experimental Evidence In the Common Disclosure Suggests Epistasis is Likely to Occur in Modified PH20 Polypeptides

264. Drs. Simpson and Petsko emphasize the experimental data in the common disclosure as showing that certain PH20 positions are 100% tolerant to

---

[298] EX2070 (Petsko Dec.), ¶¶ 52, 292.

[299] EX2070 (Petsko Dec.), ¶ 292.

amino acid substitutions.[300] Dr. Petsko highlights these positions in dark green, in his positional tolerance table, indicating that at those positions, 100% of the tested mutants retained enzymatic activity, and he relies on these highly tolerated positions when discussing making multiple changes in the PH20 protein. Both experts characterize these findings as "extremely valuable" and "useful" to POSAs seeking to engineer new forms of multiply modified PH20 polypeptides. I show some examples of these positions in Dr. Petsko's positional tolerance table below.

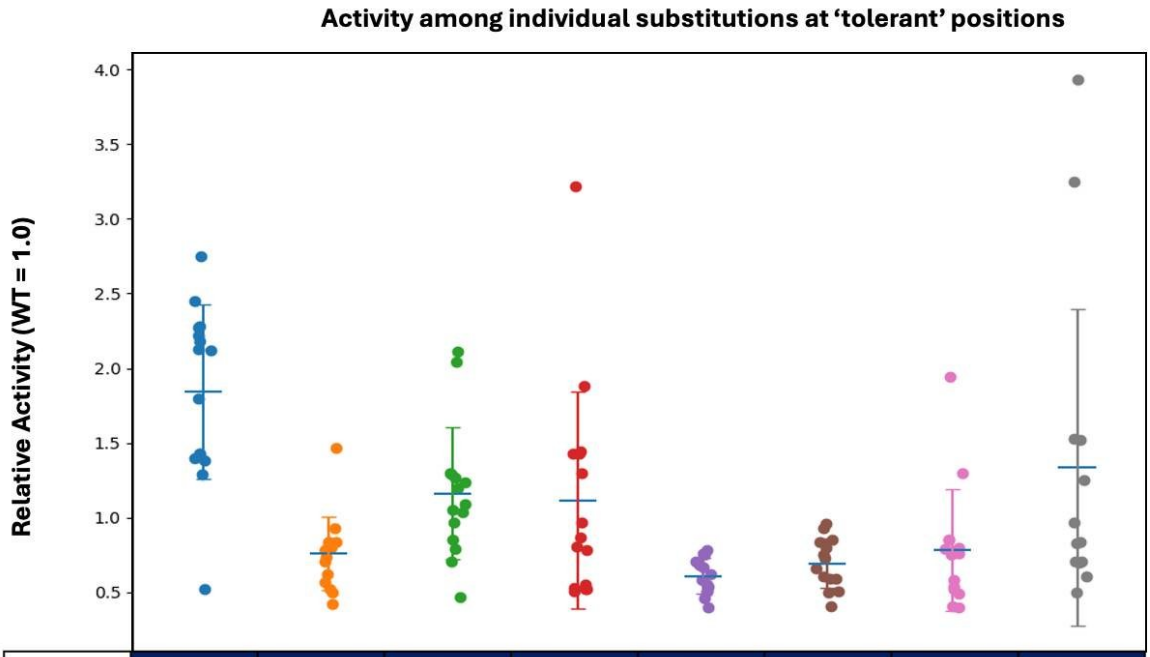
[300] EX2070 (Petsko Dec.), ¶¶ 264, 291-292, 319, 322, 333, 355, 359; EX2068 (Simpson Dec.), ¶¶ 244, 246, 311, 381-382.



Position	WT	30	D	61	L	93	K	123	W	154	K	185	L
		31	E	62	G			124	K	155	Q	186	F
1	L	32	P	63	Y			125	P	156	E	187	P
2	N	33	L	64	Y	95	D	126	K	157	F	188	D
3	F	34	D	65	P	96	I	127	D	158	E	189	C
4	R	35	M	66	Y	97	T	128	V	159	K	190	Y
5	A	36	S	67	I	98	F	129	Y	160	A	191	N
6	P	37	L	68	D	99	Y	130	K	161	G	192	H
7	P	38	F	69	S	100	M	131	N	162	K	193	H
8	V	39	S	70	I	101	P	132	R	163	D	194	Y
9	I	40	F	71	T	102	V	133	S	164	F	195	K
10	P	41	I	72	G	103	D	134	I	165	L	196	K
11	N	42	G			104	N	135	E	166	V	197	P
12	V	43		74	T	105	L	136	L	167	E	198	G
13	P	44	P			106	G	137	V	168	T	199	Y
14	F	45	R	76	N	107	M	138	Q	169	I	200	N
15	L	46	I	77	G	108	A			170	K	201	G
16	W	47	N	78	G	109	V	140	Q	171	L	202	S
17	A	48	A	79	I	110	I			172	G	203	C
18	W	49	T	80	P	111	D	142	V	173	K		
19	N	50	G	81	Q	112	W	143	Q	174	L	205	N
20	A	51	Q	82	K	113	E	144	L	175	L		
21	P	52	G	83	I	114	E	145	S	176	R	207	E
22	S	53	V	84	S	115	W	146	L	177	P	208	I
23	E	54	T	85	L	116	R	147	T	178	N	209	K
24	F	55	I	86	Q	117	P	148	E	179	H	210	R
		56	F	87	D	118	T	149	A	180	L	211	N
26	L	57	Y	88	H	119	W	150	T	181	W	212	D
		58	V	89	L	120	A	151	E	182	G	213	D
28	K	59	D	90	D	121	R	152	K	183	Y	214	L
		60	R	91	K	122	N	153	A	184	Y	215	S

265. A POSA relying on Dr. Petsko's "100% tolerated" classification might mistakenly assume that any amino acid can be substituted at one of these positions without perturbing the structure and activity of the protein. This, however, is contrary to a POSA's understanding of how substitutions affect protein structure. More critically, it also ignores the experimental data reported in the common disclosure for the mutants with changes at these positions. Below, I plot the range of hyaluronidase activities observed for the active mutants that were tested at each of positions T74, L146, T147, L26, K28, Q140, N205 and K93. Each data point represents the measured activity of one mutant. The mean and standard deviation

values indicated for the set of mutants at that position and summary statistics reported in the accompanying table give some quantification of the range of variation.



	T74	L146	T147	L26	K28	Q140	N205	K93
Mean activity	1.84	0.76	1.16	1.12	0.61	0.70	0.78	1.33
Std. dev.	0.58	0.25	0.44	0.73	0.12	0.17	0.41	1.06
CV	32%	33%	38%	65%	20%	24%	52%	80%
Tested	15	15	15	15	14	14	14	13
Active (>40%)	15	15	15	15	14	14	14	13

266. Mutants at these positions show wide activity ranges, with coefficients of variation from ~24% to 80%, demonstrating the existence of substitution- and position-specific differences in effects on the structure-function relationship of the PH20 protein. For example, T74M roughly halves activity, whilst T74F increases it by roughly half. T74 is on the surface of the protein and points away from the

substrate binding and catalytic sites, yet different changes at this position have significant and entirely opposite effects on the structure-function relationship. Thus, even at positions that Dr. Petsko assigns as "100 % tolerated" based on single-substitution testing, there is no consistent functional outcome that could be mapped to a defined structural role, and therefore no reliable structure–function correlation can be drawn for such positions. Since these effects were not predictable in 2012 and in light of the data showing such effects were extensive in PH20, a POSA would have no confidence they could reliably predict the combination of more than around 5 mutations.

267. The experimental data associated with "active mutant" classifications would alert a POSA to the high potential for epistasis when multiple mutations are introduced. Even at positions Dr. Petsko labels as "100% tolerated", PH20 activity of each mutant varies, reflecting each mutant's unique profile of perturbations to local and long-range interaction networks and, thus, to the protein's structure-function relationship. A position with no influence on PH20 activity would show no activity change across substitutions. The data instead show that different mutations at the same position—even far from the active site—produce different activity outcomes, demonstrating sensitivity to distal perturbations. When multiple such mutations are combined particularly outside the active site, a POSA in 2011-

2012 could not predict how their "effects at a distance" would interact to alter activity.

268. The bar chart in ¶ 175 shows that only ~160 substitutions retain near-native activity (0.95 to 1.05), and in no position do all changes result in native activity. Seeing such a complex picture of structure-function effects for even single mutants would only increase a POSA's concerns about epistasis. Neither the common disclosure nor PH20 homology models provide a basis for predicting or managing such long-range effects.

269. Since even a single amino acid substitution can induce long-range effects on packing, electrostatics, dynamics, or folding pathways, substitutions deemed "tolerated" at a given position do not place the protein into the single native functional state, but rather into a range of states. This dispersion underscores the limits of predicting function from single-mutant tolerance data: neither the magnitude nor the direction of functional effects for even untested single substitutions can be reliably inferred from the behavior of the tested subset, even when all tested variants exceed an activity cutoff. What is true for single substitutions is only more so true for combinations of mutants.

4. Many Effects that Influence Folding, Activity and Stability of a Protein are Imperceptible in a Homology Model

270. Drs. Petsko and Simpson suggest that POSAs in 2011-2012 could rely on homology models to manage the unpredictability of making multiple changes to

the PH20 polypeptide. This misrepresents the capabilities of homology models in 2011-2012 in important ways, which I also address in § IV.E.

271. The human Hyal-1 structure was determined by X-ray crystallography to a resolution of 2.0 Å.[301] This is considered medium high resolution and accurately positions atoms, although some side chains in some loops are less well defined by the experimental electron density. The PH20 models were generated by Dr. Petsko and Dr. Park using SWISS-MODEL. In 2011 and 2012, Swiss-Model would generate models that would be very closely based on the 2PE4 structure, with an overall RMSD deviation of 0.34 Å across 410 aligned residues used in the superposition of the two structures (positions 404-447 in PH20). The non-aligned regions are at the N and C-termini and the loop at 376. Although the models would be broadly correct, a POSA would have expected an experimental structure of PH20 determined to equivalent resolution in 2011-2012 to show differences in loops and other subtle changes.

272. PH20 and Hyal-1 both catalyze hyaluronic acid (although there are some subtle differences in substrate profile). Despite the large number of changes in sequence between the two proteins, the spatial arrangement of the atoms responsible for catalysis in them is structurally conserved. This structural

---

[301] EX1006 (Chao), 6913 (Table 1) (reporting 2PE4 structure).

conservation is maintained through conserved positions and compensating evolutionary mutations, such that similar constraints apply to their substrate-binding sites. The situation is very different when there are deliberate introductions of multiple changes into the modeled protein. As more of these are introduced, the protein increasingly differs from a known active protein. *See* ¶ 182. This is the critical difference: substrate recognition and enzyme catalysis depend on precise positioning of atoms in the enzyme. Catalysis, in particular, is sensitive to very small (< 0.3 Å) changes in position that can be induced by residue changes which are remote from the active site.

273. Much of the missing substitution data in the common disclosure cannot be reliably predicted using a homology model. This is because the disclosure reported different effects depending on sequence position, even where there are barely perceptible spatial differences in atoms. For example, Val and Thr are similarly sized residues; the difference is in Thr there is an OH in the side chain rather than a CH3. When compared to Ser which also has an OH, Thr has an additional CH3. As reflected in the table below, substitutions among these closely related residues at L175, L165, and L33—each located in hydrophobic environments—produce inconsistent and position-specific effects on activity, revealing no clear structural or chemical pattern that a POSA could have reliably predicted based on residue size or chemistry alone. The common disclosure does

not explain this pattern of results, which a POSA in 2011-2012 could not have predicted from homology modeling or residue size and chemistry alone. (Grey denotes not reported, purple inactive, green active).

Position	WT	V	S	T	W
33	L	-	0.48	0.45	1.58
165	L	1.22	1.31	-	1.14
175	L	0.94	-	1.43	

274. Other examples are readily identified in the data set. They illustrate a variety of structural contexts and show activity is influenced by the nature of the substitution in complex ways. As reflected in the table below, substitutions at G291 (within a helix), S315 (on a surface-exposed helix), and N47 (on the protein surface) show a pattern of outcomes. These patterns reveal no simple structural or chemical rule that would allow a POSA to predict whether an untested single substitution would be classified as “active” or “inactive.” Accordingly, substitutions that were not made or not tested represent genuine gaps in the data. And because these results arise from single substitutions, the unpredictability would only increase when multiple changes are combined as physicochemically similar residues causing opposite functional outcomes.

Position	WT	V	A	L	S	T	F
47	N	-	0.48	ND	0.85	1.49	1.32
291	G	1.63	-	ND	0.45	-	-
315	S	-	0.85	0.42		0.97	ND

5. Dr. Petsko's Positional Tolerance Table Does Not Enable Prediction of Multiply Modified PH20 Polypeptides

275. Drs. Petsko and Simpson maintain that a POSA could have used Dr. Petsko's table classifying the tolerance of individual positions in PH20 to amino acid substitution either alone or in combination with other information publicly available in 2011-2012 to reasonably predict the effects of introducing combinations of approximately 10, 15 or 20 modifications into the PH20 amino acid sequence, including whether (i) such multiply modified sequences would properly fold, (ii) retain hyaluronidase activity, and (iii) be soluble or increase delivery of a therapeutic agent. In my opinion, it would not.

276. Dr. Petsko's positional tolerance table is derived from a limited, unevenly sampled single-substitution dataset. The active/inactive classifications that are the foundation of Dr. Petsko's tolerance assessments do not differentiate reasons for loss of activity. For example, they do not differentiate between mutations that cause loss of activity between those that impair expression or secretion, induce aggregation, impair folding, or disrupt positioning of residues

involved in catalysis, substrate binding or unknown biological functions. They also do not evaluate how the effects of individual substitutions change when combined—there is no investigation in the common disclosure of epistatic effects which manifest only when multiple residues are changed in one polypeptide. Narrowing the analysis to a smaller subset of positions that appear tolerant to single substitutions does not resolve these limitations, because the data still provides no reliable basis for predicting how multiple substitutions distributed across the protein will interact to affect folding, stability, solubility or activity. Thus, an immense sequence space remains for which the dataset provides no predictive insight, meaning the individual combinations would need to be produced and experimentally investigated.

277. I agree with Dr. Hecht's conclusion that the common disclosure does not provide a reliable basis for predicting the effects of multiply substituted PH20 polypeptides. As Dr. Hecht explained, combinations of substitutions were not meaningfully investigated in the mutagenesis study, and a POSA could not have realistically predicted how multiple substitutions will interact to affect folding, solubility or enzymatic function. This limitation persists even when the single-substitution data are corrected or restricted to positions that were classified as tolerant to single substitutions in the $PH20_{1\text{-}447}$ protein.

278. The common disclosure does not discuss nor explain how to predict or avoid any non-additive (epistatic effects) among mutations, which were known by 2011–2012 to confound the predictability of effects of making multiple concurrent changes in proteins. Dr. Petsko's positional tolerance table cannot provide a POSA with a reliable foundation for predicting the structural or functional behavior of modified PH20 polypeptides that incorporate 5, 10, 15 or 20 substitutions, deletions or additions (or combinations of such changes).

E. Homology Models Have Limits in Predicting Effects of Mutations

279. Drs. Petsko and Simpson emphasize the importance of using a homology model of the PH20 protein with the data in the common disclosure from testing single substitutions in the $PH20_{1-447}$ protein to evaluate potential mutations in the PH20 protein.[302] For example, Dr. Petsko explains that a "POSA analyzing the specification would have considered the activity data in the context of the PH20 homology model to help determine which positions are amenable to change

[302] EX2070 (Petsko Dec.), ¶¶ 264, 291-292, 319, 322, 333, 355, 359; EX2068 (Simpson Dec.), ¶¶ 244, 246, 311, 381-382.

and to identify additional variants."[303] His comments make clear this is especially important when considering multiple mutations in the protein.

280. In ¶¶ 323-324, Dr. Petsko discusses two scenarios of making modified PH20 polypeptides with more than one mutation and requires the use of a PH20 homology model that is reliable at the location(s) of the mutations.[304]

a. The first scenario involves a set of substitutions at positions that "are very tolerant of substitution on the surface of PH20" and "in different parts of the protein" such that "they would be independent of each other."[305] A homology model would be needed here to determine the spatial positioning of the residues, however, it cannot determine if two (or more) residues are independent. For example, epistasis can occur between two spatially separated positions (action at a distance). *See* ¶ 257.

---

[303] EX2070 (Petsko Dec.), ¶ 351 ("The POSA would, as I have explained, use the patent data in conjunction with the PH20 homology model to evaluate potential substitutions and make variants with the desired properties.").

[304] EX2070 (Petsko Dec.), ¶ 360 ("The POSA seeking to make multiple mutations could generate a model using, for example, Swiss Model to evaluate the results of the anticipated modifications.").

[305] EX2070 (Petsko Dec.), ¶ 322; *see also* ¶ 355 (addressing multiple substitutions made to "residues that tolerate all tested substitutions [that] are scattered across the surface of the protein.").

b. The second involves changes in spatial proximity to each other or which are not limited to surface residues.[306] For these types of combinations, Dr. Petsko explains a POSA would need to perform "an assessment of the local environment of the residues in question in the protein." He also suggests the POSA "would be able to evaluate whether the desire[d] changes are likely to preserve the interactions formed by the WT residues or whether the desired substitutions are likely to have an adverse impact on, for example, packing." All of these steps require use of a reliable structural model of the PH20 protein.

281. Dr. Petsko does not engage with the experimental data in the patent which shows that, at every position where mutant(s) were made, at least one mutation was reported to alter the structure-activity relationship, and that only ~ 160 single mutations result in a mutated protein with between 0.95 and 1.05 the activity of $PH20_{1-447}$. Nor does his approach account for the limitations of homology models known in the 2011-2012 timeframe would limit a POSA's efforts to predict effects of combinations of mutations. First, as Dr. Park explained, homology models are reliable for modeling a single amino acid change at a location in a wild-type protein where there is a sufficiently reliable QMEAN score and the model is supported by the aligned portion of the template structure (PH20

[306] EX2070 (Petsko Dec.), ¶ 324.

residues).[307] That is because the sequence being mutated is a naturally occurring protein (which properly folds) and the position of the mutation is based on an experimentally-determined structure. Multiple changes to a modeled sequence erode the presumption of structural accuracy because it is no longer the wild-type sequence and the additional changes cause the model to deviate more from the experimentally determined structure. Each change can influence the positioning of many local neighbors of the changed residue and will cause more subtle long-distance effects on other residues. A second change causes not only its set of local and long-distance effects, but can influence the effects caused by the first mutation. By the time 5 changes are made to the wild-type sequence, the challenge of identifying and evaluating all the possible individual and interactive effects becomes intractable and exceeded the predictive capabilities of computational tools in 2011-2012 (¶¶ 260-261).

282. Second, depending on the degree of sequence identity, the accuracy of a model generated by Swiss-Model can vary significantly from reality. If the modeled protein has at least 50% sequence identity to the protein used as the template, the positioning of the backbone α-carbon atoms typically deviate by ~1 Å RMSD from an experimentally determined structure, with side chains and loop

---

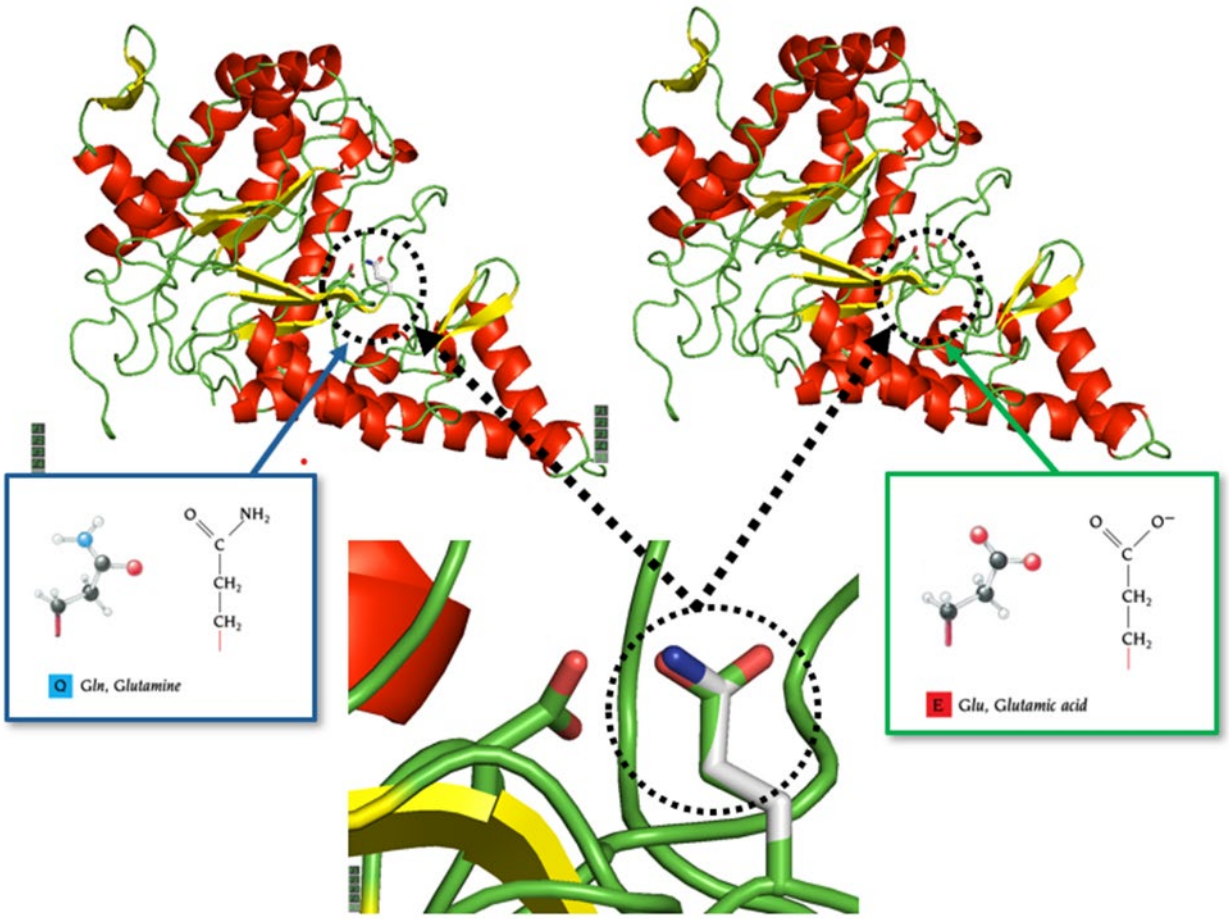
[307] EX1004 (Park Dec.), ¶¶ 157-159; *see also* ¶¶ 150-153.

structures deviating more.[308] PH20 and Hyal-1 have approximately ~36% overall sequence identity, with about 41% sequence identity in the overlapping portion of the PH20 sequence. This suggests that structural deviations less than around 1-2 Å will not be reliably detectable in the homology model. Epistasis (in multiple mutants) and changes in amino acids can alter the positioning of atoms by less than 1-2 Å. Yet enzyme catalysis and substrate recognition are exquisitely sensitive to much smaller positional and electrostatic differences. Indeed, as shown in the literature, the conservative Glu131→Gln substitution abolishes activity.[309] This is because the Glu residue is essential for catalysis, emphasizing that it is the precise structure-function relationship that is key to enzyme activity; not simply the overall fold of the protein.

---

[308] EX1012 (Bordoli), 2.

[309] EX1010 (Zhang), Abstract, 9437 (Figure 2).



283. Third, portions of the PH20 sequence that lie outside the Hyal-1 template (*i.e.*, positions 1-3 and 404-447) are not based on a structure and are not reliable.[310] In addition regions with unreliable local QMEAN values would not be

[310] EX1131 (Petsko Dep.), 203:5-21 ("Q. … another basic question about SWISS-MODEL. When I have two proteins, a template protein and a sequence that's being modeled, the SWISS-MODEL can only reliably give you a model for the overlapping portion of the two sequences if the -- like, for example, if the -- if the sequence you're sending in is longer than the structure, the sequence

reliable for modeling. A local QMEAN score above 0.6 is considered necessary to provide adequate reliability. Generally, several contiguous residues below 0.6 must be observed to classify them as having insufficient reliability in the model. In PH20, positions 3-5, 27-35, 44-48, 104-105, 234-237, 256-262, 274-278, 357-358, 405-406 have unreliable local QMEAN values, and modeling at those positions would be unreliable.[311]

284. Dr. Petsko's analysis did not account for any of these limitations on the ability to reliably use a PH20 homology model to assess and predict the effects of a potential mutation in the 2011-2012 timeframe. For example, he does not account for the lack of a reliable PH20 model at positions 1-5, 27-35, 44-48, 104-105, 234-237, 256-262, 274-278, 357-358, 404-447 of his PH20 homology model in his positional tolerance table (¶ 274, pages 127-128). A POSA would have done so.

---

associated with the 4 structure can't model the part that extends past the boundary of the other protein, right? A. In 2010, it could not do that. Q. Okay. So when you're looking at the models coming back from SWISS-MODEL in 2010, you're going to restrict your analyses to the overlapping part that's based on a template? A. Those are the reliable parts.").

311 *See* Appendix B-2 ($PH20_{1-447}$ Physiochemical and Activity Data).

F. Dr. Petsko's Positional Tolerance Table Is Unreliable and Cannot Be Used to Predict the Effects of Multiple Changes on a PH20 Polypeptide's Ability to Fold and/or Exhibit Activity

285. In his declaration, Dr. Petsko rates each position in the 1-447 form of PH20 protein as to its tolerance to amino acid substitutions. The results are compiled in a table after ¶ 274 of his declaration (pages 127-128) (reproduced below). Dr. Petsko explains that in his table, dark green entries represent positions where 100% of the tested substitutions were tolerated, light green are positions that tolerated >60% of the tested substitutions, uncolored positions tolerated 30% to 60% of the tested substitutions, orange positions tolerated <30% (but more than 0%) of the tested substitutions, and red positions tolerated none of the tested substitutions.

		30	D	61	L	92	A	123	W	154	K	185	L	216	W	247	V	278	L	309	I	340	M	371	A	402	C	433	F
Position	WT	31	E	62	G	93	K	124	K	155	Q	186	F	217	L	248	R	279	K	310	M	341	C	372	I	403	Y	434	L
1	L	32	P	63	Y	94	K	125	P	156	E	187	P	218	W	249	E	280	F	311	R	342	S	373	Q	404	S	435	K
2	N	33	L	64	Y	95	D	126	K	157	F	188	D	219	N	250	A	281	L	312	S	343	Q	374	L	405	T	436	P
3	F	34	D	65	P	96	I	127	D	158	E	189	C	220	E	251	I	282	S	313	M	344	V	375	E	406	L	437	P
4	R	35	M	66	Y	97	T	128	V	159	K	190	Y	221	S	252	R	283	Q	314	K	345	L	376	K	407	S	438	M
5	A	36	S	67	I	98	F	129	Y	160	A	191	N	222	T	253	V	284	D	315	S	346	C	377	G	408	C	439	E
6	P	37	L	68	D	99	Y	130	K	161	G	192	H	223	A	254	S	285	E	316	C	347	Q	378	G	409	K	440	T
7	P	38	F	69	S	100	M	131	N	162	K	193	H	224	L	255	K	286	L	317	L	348	E	379	K	410	E	441	E
8	V	39	S	70	I	101	P	132	R	163	D	194	Y	225	Y	256	I	287	V	318	L	349	Q	380	F	411	K	442	E
9	I	40	F	71	T	102	V	133	S	164	F	195	K	226	P	257	P	288	Y	319	L	350	G	381	T	412	A	443	P
10	P	41	I	72	G	103	D	134	I	165	L	196	K	227	S	258	D	289	T	320	D	351	V	382	V	413	D	444	Q
11	N	42	G	73	V	104	N	135	E	166	V	197	P	228	I	259	A	290	F	321	N	352	C	383	R	414	V	445	I
12	V	43	S	74	T	105	L	136	L	167	E	198	G	229	Y	260	K	291	G	322	Y	353	I	384	G	415	K	446	F
13	P	44	P	75	V	106	G	137	V	168	T	199	Y	230	L	261	S	292	E	323	M	354	R	385	K	416	D	447	Y
14	F	45	R	76	N	107	M	138	Q	169	I	200	N	231	N	262	P	293	T	324	E	355	K	386	P	417	T		
15	L	46	I	77	G	108	A	139	Q	170	K	201	G	232	T	263	L	294	V	325	T	356	N	387	T	418	D		
16	W	47	N	78	G	109	V	140	Q	171	L	202	S	233	Q	264	P	295	A	326	I	357	W	388	L	419	A		
17	A	48	A	79	I	110	I	141	N	172	G	203	C	234	Q	265	V	296	L	327	L	358	N	389	E	420	V		
18	W	49	T	80	P	111	D	142	V	173	K	204	F	235	S	266	F	297	G	328	N	359	S	390	D	421	D		
19	N	50	G	81	Q	112	W	143	Q	174	L	205	N	236	P	267	A	298	A	329	P	360	S	391	L	422	V		
20	A	51	Q	82	K	113	E	144	L	175	L	206	V	237	V	268	Y	299	S	330	Y	361	D	392	E	423	C		
21	P	52	G	83	I	114	E	145	S	176	R	207	E	238	A	269	T	300	G	331	I	362	Y	393	Q	424	I		
22	S	53	V	84	S	115	W	146	L	177	P	208	I	239	A	270	R	301	I	332	I	363	L	394	F	425	A		
23	E	54	T	85	L	116	R	147	T	178	N	209	K	240	T	271	I	302	V	333	N	364	H	395	S	426	D		
24	F	55	I	86	Q	117	P	148	E	179	H	210	R	241	L	272	V	303	I	334	V	365	L	396	E	427	G		
25	C	56	F	87	D	118	T	149	A	180	L	211	N	242	Y	273	F	304	W	335	T	366	N	397	K	428	V		
26	L	57	Y	88	H	119	W	150	T	181	W	212	D	243	V	274	T	305	G	336	L	367	P	398	F	429	C		
27	G	58	V	89	L	120	A	151	E	182	G	213	D	244	R	275	D	306	T	337	A	368	D	399	Y	430	I		
28	K	59	D	90	D	121	R	152	K	183	Y	214	L	245	N	276	Q	307	L	338	A	369	N	400	C	431	D		
29	F	60	R	91	K	122	N	153	A	184	Y	215	S	246	R	277	V	308	S	339	K	370	F	401	S	432	A		

286. Dr. Petsko explained that Examples 3 and 4 of the common disclosure described how individual mutants were tested and classified.[312] As he observes, the common disclosure classified tested mutants as an "active mutant" if the measured relative hyaluronidase activity of one of two measured samples was at least 40%.[313] The activities being reported for these mutants is an average of the two measured hyaluronidase activity values.[314] Conversely, if the measured hyaluronidase activity of the sample was below 20%, it was classified as an "inactive mutant."[315] The active mutants classified in this manner were compiled in Tables 3 and 9, while the inactive mutants were compiled in Tables 5 and 10.

---

[312] Each mutant is based on the human PH20 wild-type sequence that spans positions 1-447 of the mature protein numbering scheme.

[313] EX2070 (Petsko Dec.), ¶ 258; EX1026 ('731 App.), 251:2-4 ("Active mutants were selected whereby at least one duplicate sample exhibited greater than 40% of wildtype activity when normalized to SEAP activity. The identified active mutants are set forth in Table 9.").

[314] EX1026 ('731 App.), 251:6-13 ("The Table also sets forth the average hyaluronidase activity of tested duplicates[.]").

[315] EX2070 (Petsko Dec.), ¶ 260; EX1026 ('731 App.), 272:3-5 ("The other mutants that exhibited less than 20% hyaluronidase activity of wildtype PH20, in at least one of the duplicates, were rescreened to confirm that the dead mutants are inactive.").

1. Drs. Petsko and Simpson Relied on the Common Disclosure's Active/Inactive Mutants Classifications to Form Their Opinions on Positional Tolerance in PH20

287. Dr. Petsko emphasized the importance to his opinions of not only the measured activity data that was generated by testing the mutants, but also the common disclosure's classification of the mutants using the "active mutant" / "inactive mutant" classification scheme described in the common disclosure. As he states in ¶ 262:

> 262. These data are extremely useful because, as I illustrate below, they show the tolerability of each position to substitution. They also show which amino acids (and which types of amino acids) are tolerated at each position. With the '758 patent data in hand, the POSA seeking to generate PH20 variants now knows which positions, if changed, are likely to abrogate activity and which are not. As I explain below, the POSA could use this information to generate single and multiple mutants with desired properties.

288. At his deposition, Dr. Petsko confirmed the importance of the classification of mutants in the common disclosure to his opinions on the tolerance of individual positions in PH20 to single or multiple changes.[316]

[316] EX1131 (Petsko Dep.), 146:14-148:8.

This is the part of your declaration where you are introducing the analysis you did of mutagenesis data in the patent, right?
A. Yes.
Q. Okay. And you relied on the data from that mutagenesis experiment which is reported in Tables 9 and 10 of the patent, right?
A. Yes.
Q. You said that the data are extremely useful because they show the tolerability of each position in PH20 to substitution, right?
A. Yes.
Q. And you said the data shows which amino acids and which types of amino acids were tolerated at each position of PH20, right?
A. Yes.
Q. All right. And you also say that the data provides a POSA seeking to make PH20 variants with certain insights about making modifications to the PH20 protein, right?
A. I specifically said "are likely to abrogate activity," on one hand, and which do not.

In your opinion, the data in the patent enables a POSA to generate single and multiple mutants with desired properties?
A. Yes.
Q. So is it fair to say that you could not have formed those opinions about the tolerability of individual positions in PH20 to modification without the mutagenesis data in the common disclosure?
A. I could not have formed that opinion for most of the residues in the protein without the mutagenesis data.

289. Dr. Simpson relied on Dr. Petsko's positional tolerance analysis for many of her opinions in her declaration.[317]

290. Neither Dr. Petsko nor Dr. Simpson disputed Dr. Hecht's conclusion that the common disclosure portrays "active mutants" and "inactive mutants" as mutually exclusive classifications of modified PH20 polypeptides.[318]

[317] EX2068 (Simpson Dec.), ¶¶ 243-247, 310-314, 379-382.

[318] EX1003 (Hecht Dec.), ¶¶ 98-102, 107.

2. Only Mutants with Activity Above 40% Are Reliable and Provide Insights into Folded PH20 Proteins

291. The common disclosure does not report intra-assay or inter-assay variability data for the activity measurements in Examples 2 through 4. Example 5, however, reports 17 measurements of the activity of the same wild-type $PH20_{1-447}$ protein used as the positive control. It also includes one set of measured values at the same temperature (37°C) used in the primary screen of the mutant library. Example 5 also reports each measured replicate hyaluronidase activity value for these positive controls in Table 11.[319]

292. Dr. Hecht reviewed the positive control values in Example 5 and indicated that the variability in them "raises serious doubts about how probative or instructive the values for individual tested mutants are that fall within the range of variability observed for the control."[320] I agree.

293. I used the reported wild-type control values in Example 5 to assess the inter-plate (between plates) and intra-plate (between duplicates) variability of the wild-type $PH20_{1-447}$ positive control in the three assay conditions tested. I believe that provides an approximate indicator of the level of variability of the biotinylated

---

[319] EX1026 ('731 App.), 284:26–36, 293 (Table 11), 302-303 (Table 12).

[320] EX1003 (Hecht Dec.), ¶ 72; *see also* ¶¶ 70, 73, 76, 106.

hyaluronan substrate assay used in Examples 2 through 4.[321] To estimate inter-plate variability using the Example 5 data, I analyzed the individual wild-type PH20 activity values reported for each of the three assay conditions. As noted in the table below, many of the duplicate measurements showed large intra-plate variability (below):[322]

---

[321] Example 5 reports activity values for 17 wild-type PH20 control measurements under each condition but does not explicitly state whether those measurements were obtained on separate plates or in separate assay runs. I conservatively assume that these measurements reflect results obtained across multiple plates or runs because the values are presented individually rather than as averaged replicates, and because Example 5 describes rescreening and repeated testing under distinct conditions. Even if some measurements were obtained on the same plate, the observed dispersion among wild-type values would still reflect assay-level variability relevant to the interpretation of high-throughput screening results.

[322] EX1026 (’731 App.), 293 (Table 11).

TABLE 11: Absolute Hyaluronidase Activity						
Mutant	No incubation (4 °C)		37 °C no cresol (37 °C)		37 °C with m-cresol (37 °C plus m-cresol)	
positive control (OHO)	2.919	2.173	2.773	2.105	0.145	0.178
	3.984	4.463	4.215	4.823	0.189	0.253
	3	2.725	3	3.325	0.1	0.125
	2.501	2.883	2.370	3.158	0.452	0.522
	7.629	2.989	10.835	3.914	0.485	0.219
	5.783	5.356	2.609	3.643	0.542	0.402
	5.279	5.422	2.815	4.026	0.618	0.401
	4.775	4.385	2.845	3.327	0.718	0.540
	3.617	4.264	3.322	3.427	0.633	0.479
	5.881	4.511	5.518	4.359	0.743	0.848
	6.754	4.932	3.902	4.120	0.665	0.724
	3.911	3.494	3.911	5.179	0.726	0.841
	5.406	7.559	4.018	4.620	0.735	0.429
	4.015	3.887	3.9400	3.4080	0.3340	0.3410
	2.604	2.339	2.4430	2.3910	0.2350	0.2330
	3.736	3.473	3.6210	3.0560	0.3100	0.2770
	3.759	3.509	3.6330	3.0490	0.3600	0.3030

Positive Control #	Absolute Hylarudonidase Activity														
	No Incubation (4°C)					37°C (no cresol)					37°C (with m-cresol)				
	Dup. 1	Dup. 2	Mean	STDEV	CV	Dup. 1	Dup. 2	Mean	STDEV	CV	Dup. 1	Dup. 2	Mean	STDEV	CV
1	2.919	2.173	2.546	0.527502	20.7%	2.773	2.105	2.439	0.472347	19.4%	0.145	0.178	0.1615	0.023335	14.4%
2	3.984	4.463	4.2235	0.338704	8.0%	4.215	4.823	4.519	0.429921	9.5%	0.189	0.253	0.221	0.045255	20.5%
3	3	2.725	2.8625	0.194454	6.8%	3	3.325	3.1625	0.22981	7.3%	0.1	0.125	0.1125	0.017678	15.7%
4	2.501	2.883	2.692	0.270115	10.0%	2.37	3.158	2.764	0.5572	20.2%	0.452	0.522	0.487	0.049497	10.2%
5	7.629	2.989	5.309	3.280975	61.8%	10.835	3.914	7.3745	4.893886	66.4%	0.485	0.219	0.352	0.18809	53.4%
6	5.783	5.356	5.5695	0.301935	5.4%	2.609	3.643	3.126	0.731148	23.4%	0.542	0.402	0.472	0.098995	21.0%
7	5.279	5.422	5.3505	0.101116	1.9%	2.815	4.026	3.4205	0.856306	25.0%	0.618	0.401	0.5095	0.153442	30.1%
8	4.775	4.385	4.58	0.275772	6.0%	2.845	3.327	3.086	0.340825	11.0%	0.718	0.54	0.629	0.125865	20.0%
9	3.617	4.264	3.9405	0.457498	11.6%	3.322	3.427	3.3745	0.074246	2.2%	0.633	0.479	0.556	0.108894	19.6%
10	5.881	4.511	5.196	0.968736	18.6%	5.518	4.359	4.9385	0.819537	16.6%	0.743	0.848	0.7955	0.074246	9.3%
11	6.754	4.932	5.843	1.288349	22.0%	3.902	4.12	4.011	0.154149	3.8%	0.665	0.724	0.6945	0.041719	6.0%
12	3.911	3.494	3.7025	0.294864	8.0%	3.911	5.179	4.545	0.896611	19.7%	0.726	0.841	0.7835	0.081317	10.4%
13	5.406	7.559	6.4825	1.522401	23.5%	4.018	4.62	4.319	0.425678	9.9%	0.735	0.429	0.582	0.216375	37.2%
14	4.015	3.887	3.951	0.09051	2.3%	3.94	3.408	3.674	0.376181	10.2%	0.334	0.341	0.3375	0.00495	1.5%
15	2.604	2.339	2.4715	0.187383	7.6%	2.443	2.391	2.417	0.03677	1.5%	0.235	0.233	0.234	0.001414	0.6%
16	3.736	3.473	3.6045	0.185969	5.2%	3.621	3.056	3.3385	0.399515	12.0%	0.31	0.277	0.2935	0.023335	8.0%
17	3.759	3.509	3.634	0.176777	4.9%	3.633	3.049	3.341	0.41295	12.4%	0.36	0.303	0.3315	0.040305	12.2%
				Std Dev	13.9%				Std Dev	14.4%				Std Dev	12.9%

294. Inter-plate variability for the WT replicates also varied significantly: between 33% and 49% for some conditions below. For each condition, I treated the reported wild-type measurements as independent observations and calculated basic

descriptive statistics, including the mean activity (μ), standard deviation (σ), and coefficient of variation (CV)[323]:

$$CV = \frac{\sigma}{\mu}$$

The results are reported in the table below:

Assay Condition	Number of WT Measurements (n)	Mean WT Activity (μ)	Standard Deviation (σ)	Coefficient of Variation (CV = σ/μ)
37 °C, No Preservative	17	3.756	1.488	40%
37 °C, + Preservative (m-cresol)	17	0.444	0.222	49%
Control Condition (4 °C baseline)	17	4.233	1.407	33%

295. Dr. Simpson did not address Dr. Hecht's analysis of the variability of the positive control data. She also did not report any statistical assessment of these values. Despite that, she expresses her opinion that the variability observed in the

[323] EX1148 (Reed), 1235 ("In many laboratories, the variability of the ELISA and other methods of chemical assay that produce continuous-type values is summarized not by the standard deviation (SD) but by the CV, which is defined as the SD divided by the mean, with the result often reported as a percentage.").

Example 5 positive controls was "expected in hyaluronidase activity ELISA assays like this"[324] and "in line with what a POSA would expect."[325]

296. I disagree. The variability observed for the wild-type positive controls in Example 5 is substantially higher than what a POSA would consider acceptable for reliable quantitative interpretation of a high-throughput screening assay. As shown in ¶ 294, the estimated coefficients of variation for the 37° C assay run without m-cresol was ~ 40%, and there is evidence of large differences (≥ 30 %) between some duplicates. This is substantially greater than the approximately 10–15% inter-assay variability that Dr. Petsko testified would be an acceptable upper limit of variation for assays of this type.[326]

---

[324] Dr. Simpson incorrectly refers to the assays as "ELISA" assays used for antigen–antibody detection. They are microtiter-plate activity assays that use a biotinylated hyaluronan substrate and a colorimetric readout to quantify residual substrate after enzymatic digestion.

[325] EX2068 (Simpson Dec.), ¶ 182.

[326] EX1131 (Petsko Dep.), 217:16–24.

Q. What would -- how big does a variability in the positive control have to be for you to have a concern of the values that you are evaluating at the --

A. So it depends on the nature of the assay and so forth. For the data I looked at, okay, I was happy if the variability was less than 10. 10, 15 -- 10 to 12 percent, something like that. That didn't surprise me too much.

297. The variability is also significantly higher than the inter-assay variability benchmarks reported for properly characterized biotinylated hyaluronan microtiter assays in Halozyme's US-457 publication, which reported inter-assay variation of the biotinylated hyaluronan microtiter assay was less than 10%.[327]

298. Variability of the magnitude observed in Example 5 indicates that measured activity values are highly sensitive to assay conditions and/or handling. Reported differences between mutants and native on the order of tens of percent could therefore reflect assay noise rather than true differences in enzymatic function. The decision to omit variability data for Examples 2 through 4 of the common disclosure, when combined with assay variability problems which are clearly visible elsewhere in the common disclosure (IV.F.2), undermines the

[327] EX2165 (US-457), ¶ 0271 ("The sensitivity permits rapid measurement of hyaluronidase activity from cultured cells and biological samples with an inter-assay variation of less than 10%").

confidence a POSA would have in the assignment of mutants as "active" or "inactive". Dr. Petsko's tolerance table did not comment on or correct for these issues.

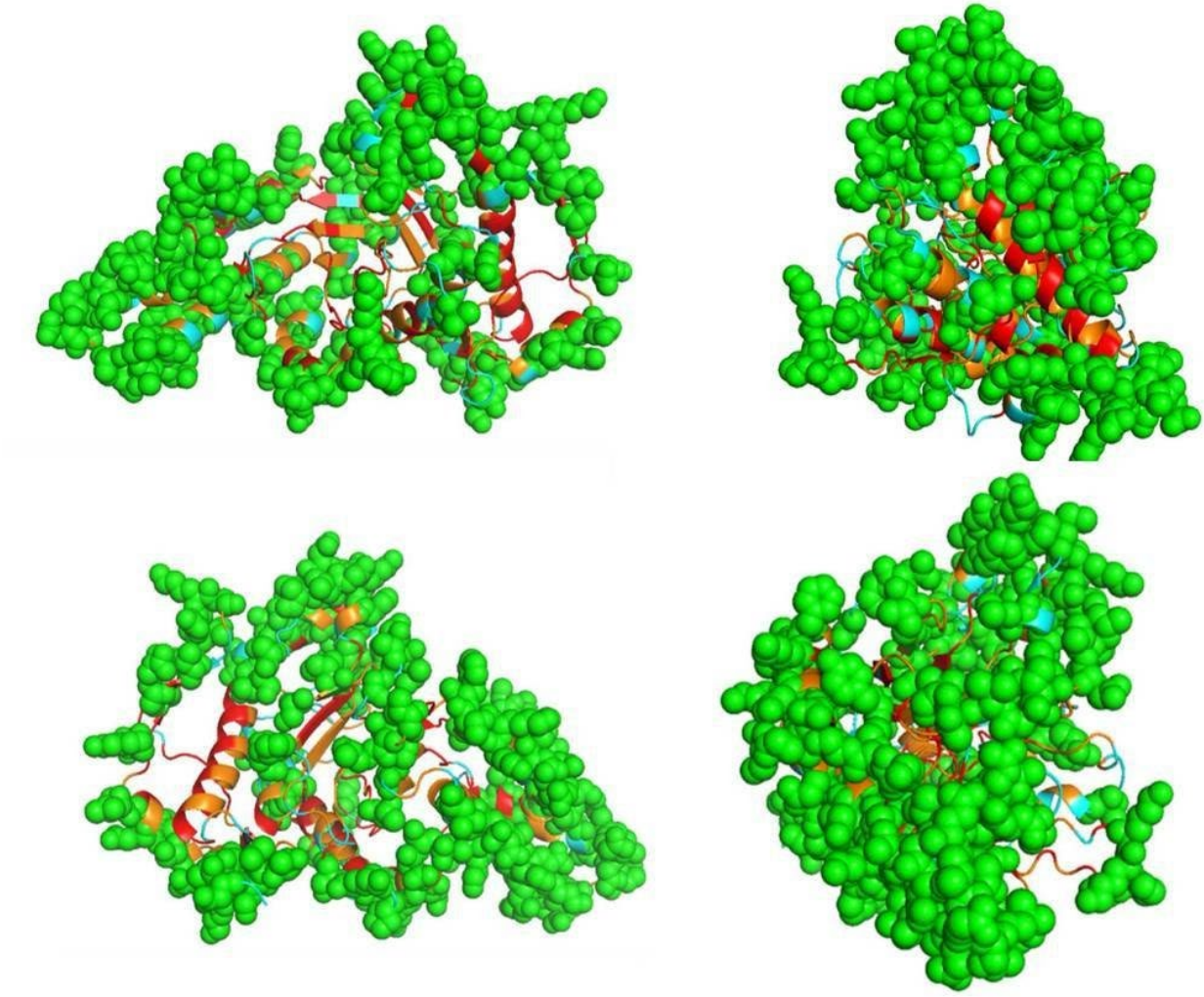
299. A POSA could not rely on the precise 20% and 40% thresholds to classify mutants as "active" or "inactive" because of the variability observed in the same biotinylated-HAS assay is run under comparable conditions (*i.e.*, 37°C, no added m-cresol). I believe a POSA would only have confidence in data for "active mutants" with reported activities above 40%, and I have used that set of data in my analyses.

G. Dr. Petsko's Positional Tolerance Analysis Overstates What the Data Support

300. Dr. Petsko presented his conclusions about the tolerance of each position of the $PH20_{1-447}$ protein to mutation using: (i) a table with coloring to show tolerance ratings, and (ii) annotated images of a PH20 homology model with residues colored based on the coloring scheme in his positional tolerance table. Dr. Petsko's table and annotated PH20 images overstate the conclusions that can be drawn from experimental data in the common disclosure.

301. For example, in ¶ 317, Dr. Petsko presents annotated images that suggest (incorrectly) that nearly all of the surface residues on the PH20 protein will

be highly tolerant to change.[328] The PyMol image visually emphasizes residues classified as "tolerant" by rendering them as green spheres, which thus appear much larger relative to other residues shown only as main chain cartoon.



302. The classifications made by Dr. Petsko are not informative of positions that are "intolerant" to substitution. As I explained in § IV.C, the portion

[328] EX2070 (Petsko Dec.), ¶ 317.

of the data associated with inactive mutants does not provide reliable insights into whether single amino acid substitutions at those positions simply reduce activity below 20% or perturb the folding, secretion, or stability of the folded PH20 structure. Accordingly, classifications of residues as "intolerant" to change that are based on "inactive mutant" data lack scientific foundation. Likewise, positions for which Dr. Petsko found only 30% to 60% of the tested substitutions retained activity do not provide meaningful information about "tolerance", because this binary classification (based on a 40% activity threshold) does not distinguish between variants with marginal activity and those with substantially higher activity, and thus such limited outcomes cannot be extrapolated to the full set of possible substitutions. Finally, even among mutants classified as "active," deviations in activity between duplicates in the activity assay warrants excluding variants with average hyaluronidase activity between 20% and 40% (since one measurement must be below the 40% threshold) and provides an additional reason not to rely on the common disclosure's inactive-mutant classifications (*see* § IV.F.2).

1. Positions Identified in the Prior Art as Essential to Folding, Secretion or Activity

303. As Dr. Simpson acknowledged in her deposition, a POSA would require independent functional and structural analysis demonstrating that

substitutions at a residue disrupt PH20 activity or compromise the folded structure before concluding that the position is essential:[329]

> A. If my objective was to make a modified PH20 polypeptide, I would need more information than is given in the alignment.
>
> The alignment also is supported by actual previously published functional data that demonstrated which residues of PH20 are inactive residues or are key catalytic residues. So that's why they're colored separately.
>
> To retain enzymatic activity, certainly you would not mutagenize those, because it's previously been confirmed that mutagenesis at those sites eliminates hyaluronidase activity.
>
> A. It is my opinion that a POSA would use the information about this structure, use the sequence alignment, and model PH20 according to this new structural information, and evaluate whether the blue invariant residues were located in a position of critical -- of a critical nature to the folded structure of PH20, and evaluate them individually, potentially, for their location in the structure, if they wanted absolute certainty about their role in maintaining the structure.
>
> It's also my opinion, though, that if I have this information about invariant residues in an alignment and about functional data that implicates certain of those invariant residues in maintaining enzymatic activity of the enzyme, that a POSA would avoid residues that are invariant if they wish to maintain activity of the polypeptide.

304. The common disclosure references publications that identified 57 conserved residues in PH20 proteins.[330] It also reports that those publications had identified a number of those positions as playing important structural or functional roles, such as cysteine residues that form disulfide bonds, acidic residues within the active site, and residues that contribute to substrate binding and/or hyaluronidase activity. It observes that mutagenesis of several positions results in

---

[329] EX1130 (Simpson Dep.), 119:25–120:11; 121:8–24.

[330] EX1026 (’731 App.), 79:32-80:18 (citing EX1011 (Arming 1997), EX1149 (Ten Have) and EX1150 (Chowpongpang)).

enzymes that have no detectable or residual enzymatic activity.[331] From those studies, a number of positions were identified that, if changed, abolished enzymatic activity or prevented secretion. The common disclosure then provides a list of those previously characterized positions where "the results herein confirm the requirement of" those residues "for hyaluronidase activity, since mutagenesis of these residues results in an enzyme that is not active (*e.g.*, it is not expressed or is inactive when expressed, *see e.g.*, Tables 5 and 10),"[332] which testing showed comparable results (below).[333]

> **The results herein confirm the requirement of PH20 amino acid residues corresponding to positions 25, 111, 113, 176, 189, 203, 246, 249, 252, 316, 341, 346, 352, 400, 402, 408, 423 and 429 of the sequence of amino acids set forth in a mature PH20 lacking the signal sequence such as set forth in SEQ ID NO: 3 or 7 for hyaluronidase activity, since mutagenesis of these residues results in an enzyme that is not active (*e.g.*, it is not expressed or is inactive when expressed, see *e.g.*, Tables 5 and 10). The exception is that amino acid replacement corresponding to R176K and C316D resulted in mutants that generated some residual hyaluronidase activity.**

305. Because these residues are supported by clear functional and structural evidence not the inactive-mutant classifications relied upon by Dr. Petsko, a POSA would have viewed them as being positions that are essential to

---

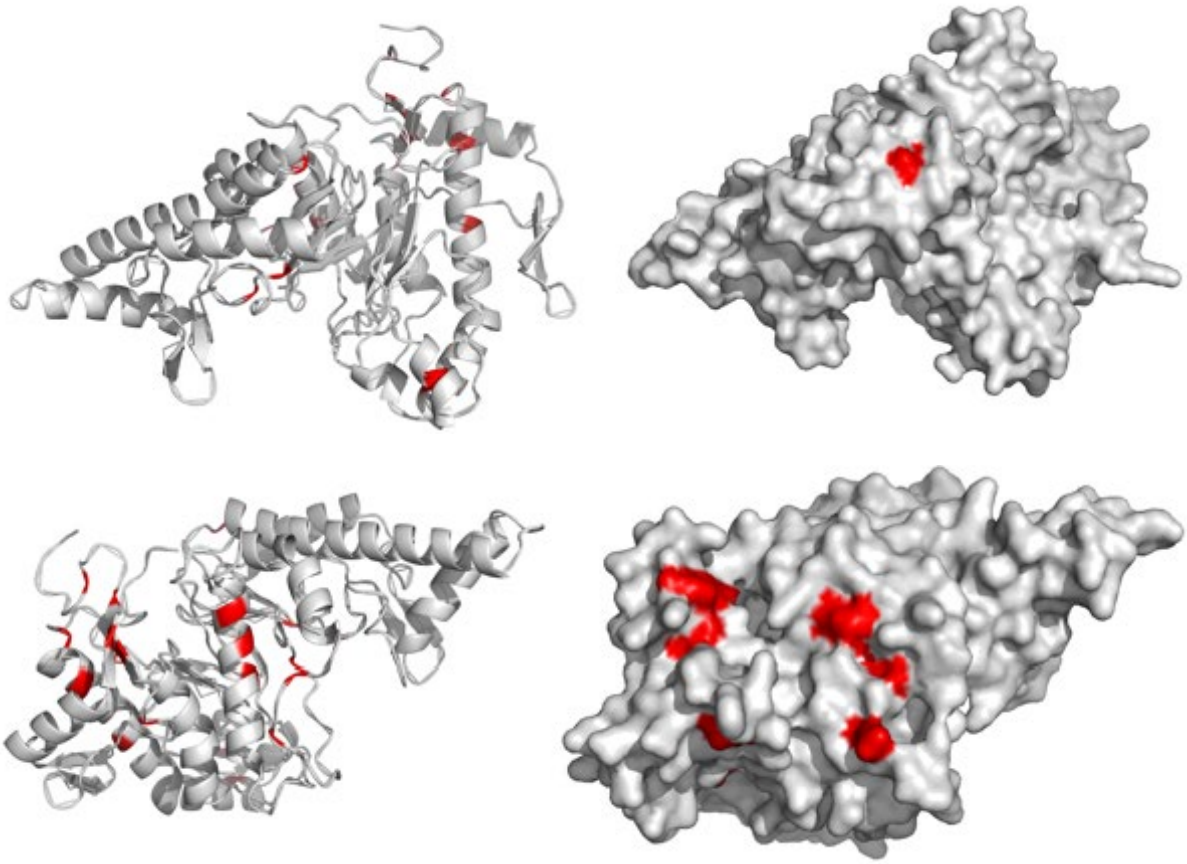
[331] EX1026 ('731 App.), 80:16-18.

[332] EX1026 ('731 App.), 80:19-24.

[333] EX1026 ('731 App.), 80:19-26.

PH20 folding, secretion, or enzymatic activity and, when substituted, abolish function. I use this set of experimentally validated essential residues in place of Dr. Petsko's "red" classifications and present a corresponding structural model below to provide a more accurate depiction of positions that a POSA would have avoided modifying.

WT	WT	Position	WT	Position	WT
25	C	246	R	352	C
111	D	249	E	358	N
113	E	252	R	400	C
176	R	316	C	402	C
189	C	333	N	408	C
200	N	341	C	423	C
203	C	346	C	429	C



2. Positions that Retain Activity Despite Meaningful Interrogation

306. Dr. Petsko characterizes many positions as being "generally tolerant of substitutions" because they were found to "tolerate" "60% or more" of the substitutions that were tested.[334] He and Dr. Simpson rely on these classifications not only to suggest those individual amino acid substitutions would be "tolerated", but that those positions in the sequence are generally "tolerant" to change.[335]

307. Dr. Petsko acknowledged at his deposition that his tolerance classifications of ≥60% or 100% tolerated in his table reflect only the percentage of tested mutants that were classified as "active mutants," rather than the percentage of mutants out of the 19 alternative amino acids substitutions that were possible at each position.[336] His ratings also do not reflect any structural context of the

---

[334] EX2070 (Petsko Dec.), ¶¶ 267, 303; *see also* EX2068 (Simpson Dec.), ¶¶ 310-311.

[335] EX2070 (Petsko Dec.), ¶¶ 267, 303, 323, 333; *see also* EX2068 (Simpson Dec.), ¶¶ 45, 310-312, 322, 348, 382.

[336] EX1131 (Petsko Dep.), 150:11-16 ("Q. So that -- so that color is going to be true -- that dark green color is going to be true if you had seven mutants tested and all seven were active, as well as if you had 15 mutants 15 tested and all 15 were active? A. Correct."); 153: 6-12 ("…A. No, they don't reflect the absolute number of mutations at each position.").

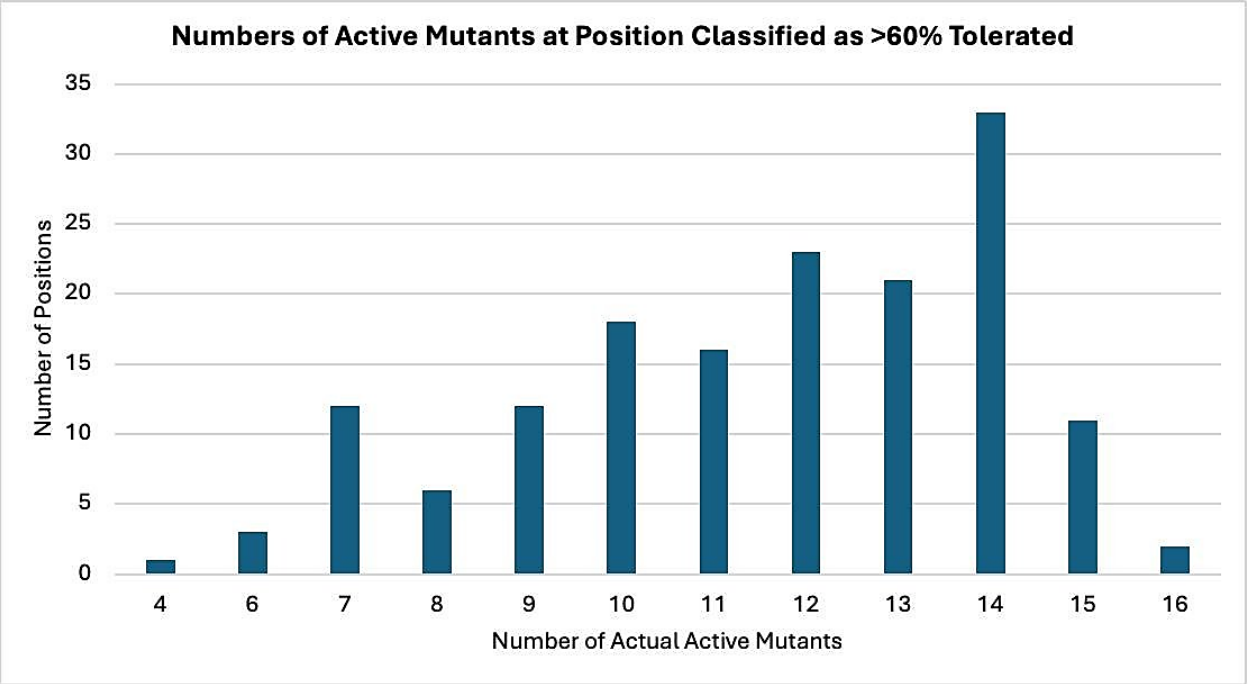
position within the protein, or the degree of physicochemical diversity of the substitutions that were tested and retained activity.[337]

308. For example, Dr. Petsko's table and illustrations designated a position with 7 active mutants out of 11 tested substitutions, the same as a position with 11 active mutants out of 18 substitutions tested. A POSA would find no objective reason to give two residues an identical tolerance rating when they reflect materially different levels of experimental investigation.[338] As shown below, many of the positions in Dr. Petsko's table are designated as tolerating "60% or more" substitutions based on as few as 4 to 10 active mutants.

---

[337] EX1131 (Petsko Dep.), 153:13-25 ("Q. There is also no indicator or factor that's built into your coloring scheme that reflects the character of the change of the amino acids at these positions that were -- that were substituted.… A. In this table, no. But the data exists in the patent to get that information very easily.").

[338] *Compare* EX2166 (Petsko Spreadsheet), [position 310] where 12 substitutions were made and 8 were found active, *with* [position 72] where 18 substitutions were made and 11 were found active.



309. Because there are 19 possible amino acid substitutions at any given position, a properly applied 60% tolerance threshold would credit only those positions at which 11 or more active mutants were observed.[339]

310. A POSA would consider whether the active mutants observed at each position adequately reflect the physicochemical diversity of the 20 amino acids. Amino acids fall into distinct physicochemical classes—including (i) charged residues (positively or negatively charged), (ii) polar uncharged residues, (iii)

---

[339] *See* Appendix B-3 (Count Comparison of $PH20_{1-447}$ Residues). The identified residues had 60% tolerability of the tested substitutions under Dr. Petsko's analysis after correcting for errors such as substitutions listed in both Table 9 and 10.

hydrophobic residues, and (iv) large hydrophobic, aromatic residues—that impose different structural and energetic constraints on a protein. Because substitutions from different classes probe different structural constraints, it is not a valid assumption to find that a position is highly tolerant to mutation if the only mutants that retained activity were those that were exposed to a subset of the different physicochemical attributes that are present in the 19 amino acids. Dr. Petsko's positional tolerability analysis does not account for whether active mutants were observed that survived interrogation with all of the physicochemical categories of the 20 amino acids.[340] A POSA would not have classified these residues as "generally tolerant."

311. The structural context of a position within the protein is relevant to how a POSA would view mutations at that position. As Dr. Petsko notes in his declaration, a POSA needs to understand the structural context of a position being

---

[340] EX1135 (Naismith Spreadsheets), (physicochemical classes used were: small hydrophobes (G, A, P, V, L, I, M), polar uncharged (C, T, S, N, Q), polar negative (D, E), polar positive (H, K, R), large hydrophobes (Y, F, W). A "0" in any one physicochemical group required exclusion. Positions were found to not be sufficiently interrogated when no active mutants with any residues in a category were observed.); *see also* Appendix B-2 ($PH20_{1\text{-}447}$ Physiochemical and Activity Data).

mutated when attempting to characterize its tolerance to change.[341] For example, one important structural context is whether the position is a buried residue or is a surface residue. Residues that are buried within the protein core or that form part of a binding interface are more likely subject to substantially tighter structural and energetic constraints than solvent-exposed surface residues.

> 297. A POSA analyzing the specification would have considered the activity data in the context of the PH20 homology model to help determine which positions are amenable to change and to identify additional variants.

312. Dr. Petsko relies on this point when classifying the tolerance of residues in PH20. He states, "[t]ypically, surface residues not involved in binding or activity are expected to be generally tolerant to mutation."[342] Consequently, a

---

[341] EX2070 (Petsko Dec.), ¶ 297; *see also* ¶ 10 (explaining use of PyMol to "understand the structural context of individual residues"); ¶ 18 (agreeing with Dr. Park that when assessing a mutation it was common and expected for a POSA to generate a model and consider the substitution in that structural context"); ¶ 352 (agreeing with Dr. Hecht why it is important to consider the context of a "position's physical chemical environment" when considering a substitution").

[342] EX2070 (Petsko Dec.), ¶ 03; *see also* ¶¶ 294, 317 ("The residues that are relatively tolerant to change are generally on the surface of PH20."); ¶ 318 ("The figures above show that there are residues that are reasonably tolerant to

POSA would ordinarily require a substantially more stringent and comprehensive interrogation of buried or functionally constrained residues than solvent-exposed surface residues.[343] However, Dr. Petsko's extrapolation from single-mutation data to combinations of mutations does not turn on "tolerance" of individual mutations alone, but instead it also requires independence among the multiple substitutions. The experimental data in the patent show that the vast majority of single substitutions measurably perturb the structure–activity relationship of PH20, indicating that multiple substitutions will not be independent.

---

change across the surface of PH20."); ¶ 22 ("there are many positions that are very tolerant of substitution on the surface of PH20 that can be targeted by the POSA seeking to generate a modified PH20 that retains hyaluronidase activity"); ¶ 55 (explaining that a POSA considering making multiple modifications would focus on surface residues, stating "[a] POSA choosing this route has many potential substitution options to consider; as shown below, residues that tolerate all tested substitutions are scattered across the surface of the protein.").

[343] EX1146 (Pal), 22379, 22381 (showing that mutational tolerance depends on "different packing environments" and on "local dielectric and packing environments," and that evaluating substitutions across "the full spectrum of chemical and conformational diversity within the 20 natural amino acids" is fundamental to understanding the energetic consequences of mutation at structurally constrained positions).

313. A POSA can classify residues as buried or surface residues using information in a PH20 homology model file ("pdb" file) that measures the solvent-accessible surface area ("SASA") for each modeled residue. For PH20, that information is only available for positions 4-403 because that is the portion of the PH20 protein that is supported by the Hyal-1 template structure. A POSA would not rely on portions of a PH20 homology model that fall outside of the template structure used to generate it, or for which insufficient quality in that part of the model exists. A POSA thus would not have included residues 1-3 and 404-447 in the set of positions Dr. Petsko labeled "generally tolerant" because there is no structural context available for those positions. Likewise, the POSA would have considered positions 3–5, 27–35, 44–48, 104–105, 234–237, 256–262, 274–278, 357–358, 405–406 to have unreliable local QMEAN values.

314. Because buried or partially buried residues are subject to especially restrictive packing, stability, and energetic constraints, a POSA would regard anything short of near-complete tolerance as not being meaningfully predictive of mutations. For residues with a fractional solvent-accessible surface area (fSASA) below approximately 0.65, tolerance conclusions would require a stringent and comprehensive showing—typically approaching tolerance of nearly all substitutions tested—rather than a threshold such as 60%, which would not be

considered indicative of general tolerability for such structurally constrained positions.[344]

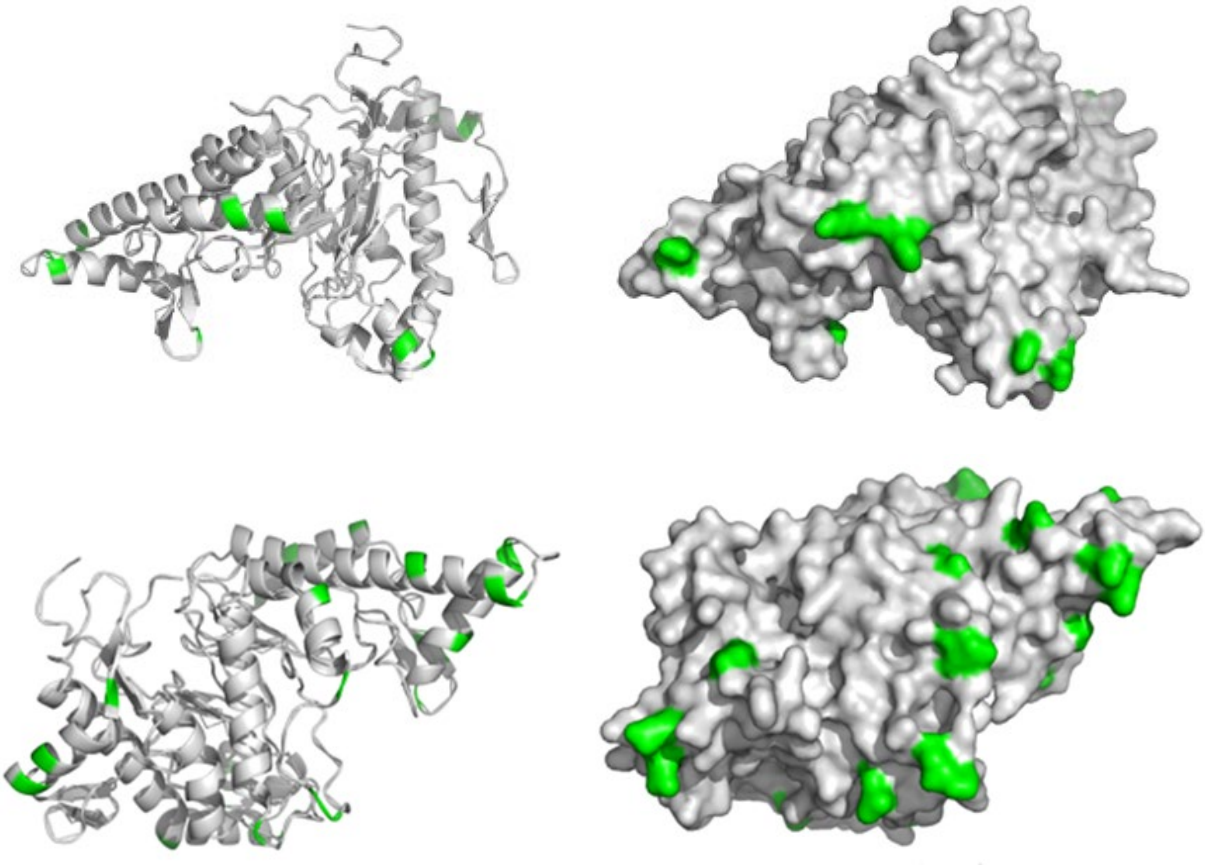
315. After excluding positions for which fewer than 11 active mutants were produced, excluding positions that did not have active mutants that survived interrogation across all physicochemical classes, and excluding positions with fSASA values below 0.65 (i.e., structurally buried residues), 12 positions remain. Even with respect to this limited subset, a POSA would not have considered combinations of changes that included these positions to be predictable in PH20 polypeptides containing 10, 15, or 20 total substitutions, given the well-recognized phenomenon of epistasis.

316. I have compiled these positions in the table below and illustrate their position using Dr. Petsko's PH20 homology model below.[345]

Position	WT	Position	WT	Position	WT
27	G	98	F	197	P
30	D	140	Q	237	V
69	S	146	L	389	E
94	K	147	T	392	E

[344] *See* Appendix B-3 (Count Comparison of $PH20_{1-447}$ Residues).

[345] Data is compiled in EX1135 (Naismith Spreadsheets); *see also* Appendix B-3 (Count Comparison $PH20_{1-447}$ Residues).



317. A POSA would not have viewed single substitution mutations in PH20 that retained activity to have been predictive of making sets of 5, 10, 15 or 20 of those individually tolerated mutations in a single PH20 protein. Even using Dr. Petsko's perspective on combining substitutions (but correcting for errors in his compilations of residues and the factors I discuss above), there are only a small number of residues that could be considered "generally tolerant" to change. Just exploring this set of 12 "highly tolerated" positions with 19 alternatives in the 430-length PH20 sequence translates into ~10^{15} possibilities, which is still a tiny fraction of the claim parameter sequence space (*e.g.*, ~10^{59}). Beyond the 12 "highly tolerated" positions, the remaining portion of that multiply-modified sequence

space would be unpredictable given that even at these small numbers, the data show different substitutions do perturb activity differently, meaning there are complex effects on structure-function relation. Combinations of around 5 or more substitutions would give rise to unpredictable effects on activity in my opinion because there is no way to quantify or identify which combinations of mutations are "independent" of each other thus avoiding the well-known issue of epistasis. The common disclosure does not address whether mutations are "independent" of one another, nor does it provide any guidance on how to combine mutations. Dr. Petsko's analysis and proposed approach assumes independence but does not articulate how such independence would be determined or even estimate how many such "independent" combinations might exist. And even if the number of "tolerated" positions was expanded, and substitutions were made at all residues not identified as essential by Dr. Park, the prior art, and "confirmed" by the common disclosure,[346] the sequence space would still contain ~10^{57} possibilities (below).

Seq ID No.	Residue Length	Total # of Essential Residues Reported by Prior Art, Confirmed by Common Disclosure, and Identified by Dr. Park	Remaining Non-essential Residues	Required Change	# of Distinct PH20 Polypeptides
32	430	75	354	L317Q	8.57×10^{57}

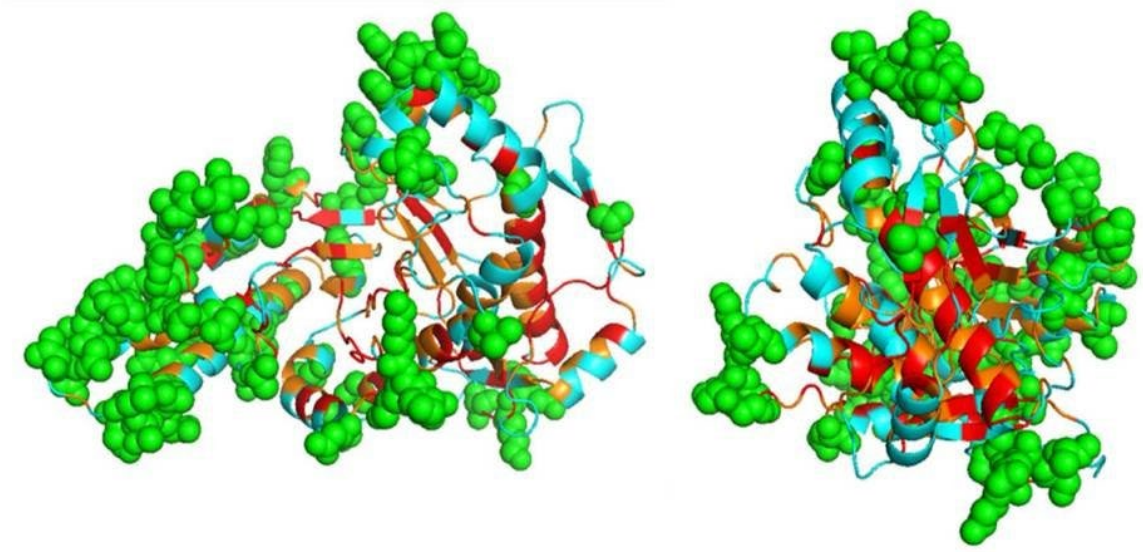
[346] *See* Appendix B-3 (Count Comparison $PH20_{1-447}$ Residues).

H. The Tolerance Table Does Not Enable Prediction of Multiply Modified PH20 Variants

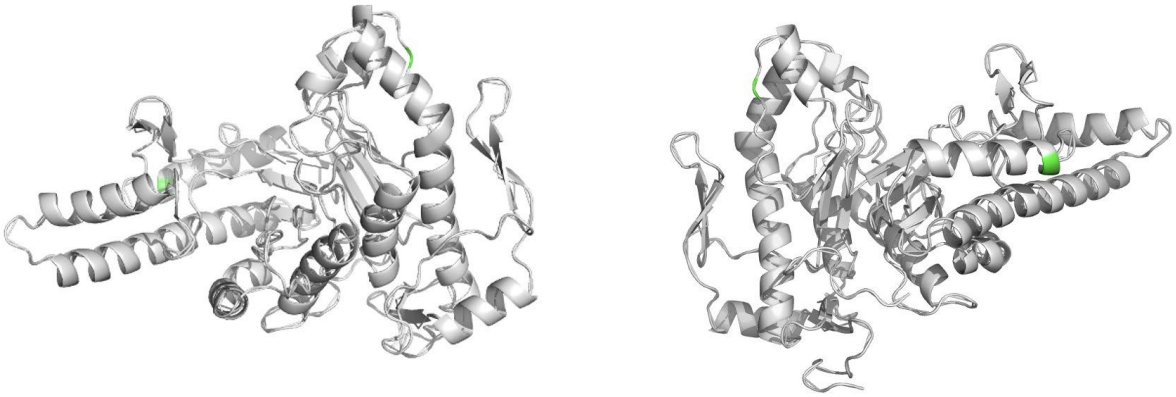
318. In ¶ 355 of his declaration, Dr. Petsko opines that a POSA wishing to make multiple modifications to PH20 while retaining activity could "select from those positions that best tolerate substitution," and, in particular, could choose positions that tolerated all or nearly all tested substitutions. He further asserts that a POSA would reasonably expect that introducing multiple substitutions at such positions would not abrogate activity. In ¶ 356, Dr. Petsko extends his reasoning to positions where supposedly 90% of substitutions retain activity and provides an image suggesting most of the surface is tolerant to any change. He then suggests that in that situation that "a POSA would have a very large number of potential mutants to choose from, and the POSA would expect those variants to maintain activity."[347]

> 355. Second, the POSA wishing to make multiple modifications but retain activity could select from those positions that best tolerate substitution. The POSA seeking to be conservative in this regard could select positions that tolerated all tested substitutions, as shown in the figures below. A POSA choosing this route has many potential substitution options to consider; as shown below, residues that tolerate all tested substitutions are scattered across the surface of the protein. A POSA would reasonably expect that introducing multiple substitutions at these positions would also not abrogate activity.

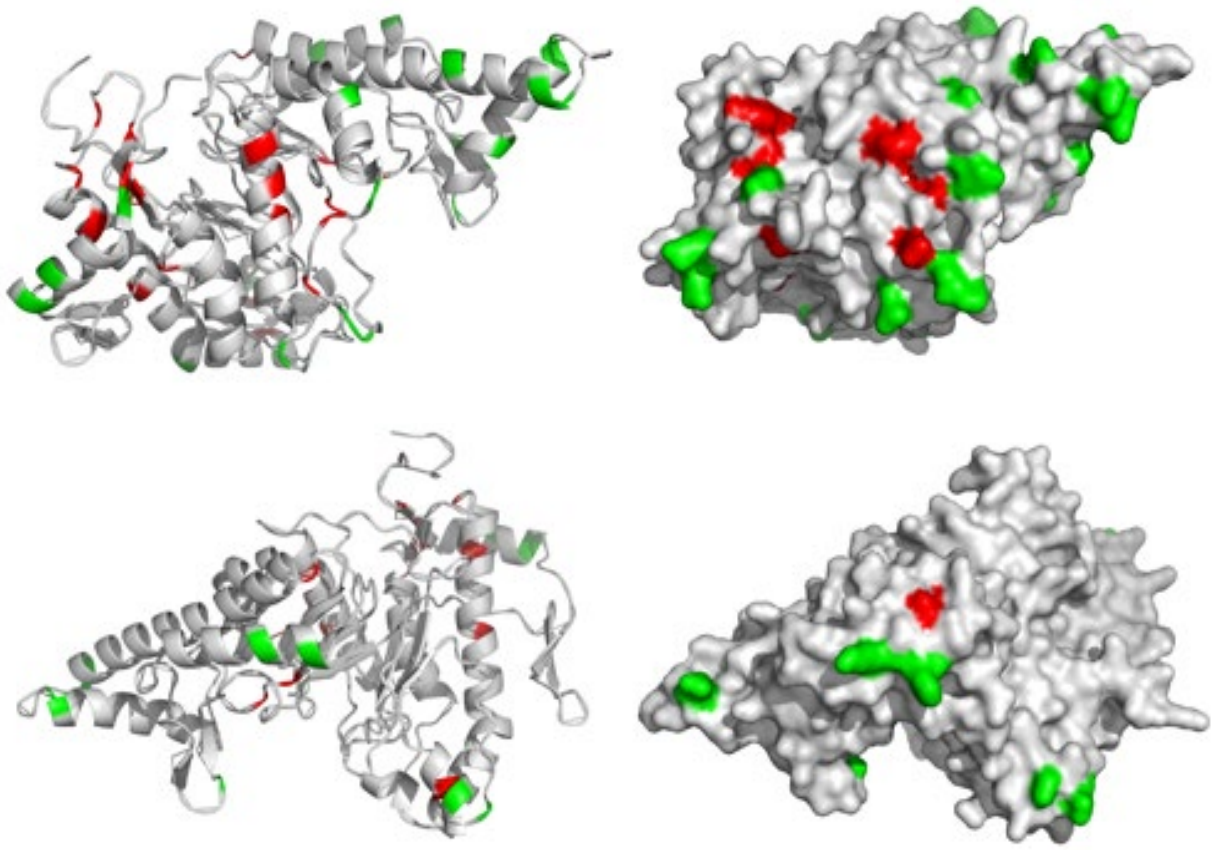
[347] EX2070 (Petsko Dec.), ¶ 357.



319. Imposing the threshold of 90% of 19 possible mutations results in only two positions: (i) Q86 which has 17 tested as active (89%) and (ii) P32 has 16 mutants which tested as active (85%). The remainder of the protein has 15 or less mutants tested as active (< 80 %). One of these is in an α-helix (below), while the other is in a loop.



320. The more realistic picture that a POSA would have seen from the data in the common disclosure and known before 2011-2012 is a protein with the vast majority of its positions presenting no insights as to general effects of mutations, and no insights beyond combinations of up to 2 or 3 substitutions in extremely constrained settings (below).



321. I agree with Dr. Petsko's statement on making combinations only to the extent it is limited to the introduction of only two or three substitutions at surface residues where each mutation was classified as an active single amino acid substitution, consistent with observations that Drs. Hecht and Park made in their

declarations.[348] However, Dr. Petsko's broader propositions about predicting effects of multiple changes (*i.e.*, more than about 5) are not supported by the underlying patent data or by the state of the art in 2011–2012. His reasoning rests on the assumption that tolerance to isolated single substitutions can be extrapolated to predict the behavior of proteins containing many combined substitutions in a myriad of structural contexts. That assumption is not valid, given the known non-additive effects of mutations on protein folding, stability, and activity (¶¶ 254-256). In my opinion, this approach does not provide a reliable basis for designing PH20 variants containing ten, fifteen, or twenty combined substitutions.

322. As Dr. Hecht explained in his opening report—and I agree—the effects of amino-acid substitutions on protein folding, stability, solubility, and activity, or the ability of the PH20 to increase delivery of a therapeutic are inherently context-dependent and frequently non-additive.[349] Dr. Hecht described that while a single substitution may be tolerated—or even beneficial—when evaluated in isolation, that observation provides no reliable basis for predicting the

---

[348] EX1003 (Hecht Dec.), ¶¶ 50, 58-59, 158-159, 229; EX1004 (Park Dec.), ¶¶ 158-159.

[349] EX1003 (Hecht Dec.), ¶¶ 54–61.

behavior of proteins containing multiple additional substitutions.[350] He also explained that each substitution alters the local and long-range network of noncovalent interactions that collectively stabilize protein structure, including hydrogen bonds, ionic interactions, van der Waals contacts, hydrophobic packing, and aromatic interactions.[351] Introducing multiple substitutions simultaneously can disrupt these interaction networks in ways that cannot be inferred from single-mutant behavior, such as by destabilizing secondary structure, altering folding pathways, interfering with the formation of structural motifs, or negating the stabilizing effect of earlier substitutions.[352]

323. Dr. Hecht further explained that the number of possible substitution combinations grows combinatorially as additional substitutions are introduced, and that even substitutions that are individually tolerated can collectively render a protein unstable or inactive.[353] These principles, which were well understood in the field by 2011–2012, directly contradict Dr. Petsko's assertion that positional

---

[350] EX1003 (Hecht Dec.), ¶¶ 55–56, 61.

[351] EX1003 (Hecht Dec.), ¶¶ 56–57, 142.

[352] EX1003 (Hecht Dec.), ¶¶ 58–60, 142.

[353] EX1003 (Hecht Dec.), ¶¶ 58–59, 61.

tolerance inferred from isolated single substitutions can be extrapolated to predict the behavior of PH20 proteins containing many combined substitutions.

324. Dr. Hecht's explanations are reinforced by systematic scanning mutagenesis studies published before and around the relevant time period, which directly examined the effects of large numbers of single, double, and higher-order substitutions. These studies demonstrated that positional tolerance cannot be inferred from a small or "representative" subset of substitutions, and that increasing numbers of combined mutations rapidly lead to deleterious outcomes. For example, Pal and colleagues performed comprehensive combinatorial scanning of a protein–protein interaction interface and showed that mutational tolerance is highly context-dependent and cannot be predicted from limited sampling or from chemically "conservative" substitutions alone.[354] Similarly, Fowler and colleagues applied deep mutational scanning to map protein sequence–function relationships at high resolution and explicitly quantified the effects of single, double, and triple mutants across a large mutational space. Although many individual single substitutions were tolerated, the overwhelming majority of higher-order variants were deleterious relative to wild type. In that study, while hundreds of single substitutions retained activity, approximately 97.2% of the variants observed—

---

[354] *See, e.g.*, EX1146 (Pal), 22381–82.

including the vast majority of double and triple mutants—were deleterious.[355] These results illustrate that tolerance observed for individual substitutions does not translate into predictability for combinations of substitutions.

325. By 2011–2012, persons of ordinary skill in the art understood from such studies that reliance on incomplete or sparsely sampled mutagenesis data could not support reliable predictions about the behavior of proteins containing many combined substitutions. Even when a small fraction of double or triple mutants was explored, experimental results showed that most combinations disrupted folding, stability, or activity, often in ways that could not be anticipated from single-mutant data alone. As Dr. Hecht explained, the number of possible combinations grows combinatorially as substitutions are added, and individually tolerated changes can collectively destabilize a protein or negate one another's effects. *See* EX1003 (Hecht Dec.) ¶¶ 58–61. Accordingly, these studies confirmed that increasing mutational complexity leads to unpredictability and frequent loss of function, directly contradicting the assumption that positional tolerance inferred from limited single-mutant data can be extrapolated to proteins containing many simultaneous substitutions.

---

[355] EX1151 (Fowler), 743-744.

I. Dr. Petsko Does Not Account for Protein Aggregation in Multiply Modified PH20 Polypeptides

326. Dr. Petsko's conclusions regarding how a POSA could use the tolerability table to predict the behavior of multiply modified PH20 proteins do not account for the well-recognized risk of protein aggregation. His analysis assumes that individually "tolerated" substitutions can be combined without introducing new and unpredictable consequences for aggregation that are not captured by single-substitution mutant data.

327. At his deposition, Dr. Petsko agreed—and I agree—that aggregation is a common phenomenon that can occur when substitutions are made to a protein:[356]

Q. Can aggregation occur during the process of protein folding if there has been a mutation to the protein?
A. It can occur even if there hasn't been a mutation in the protein depending on the external conditions. Okay. But usually, that's concentration dependent. And it typically requires nonphysiological concentrations.
Q. Can aggregation occur within the cell during transcription/translation?
A. Yes. It can. And often, that's the result of overexpression that occurs because of, say, a promoter mutation.
Q. So those -- those interactions, those are hydrophobic patches that become exposed on the surface of the protein?
A. Or that are created on the surface, yes. Those are the most common ways that you would see that kind of aggregation, would be hydrophobic.
Q. And when those interactions occur, those are basically random interactions between the locations on the proteins that are interacting with each other, right?
A. No. It can be, but in the case of sickle cell hemoglobin it is an extremely precise interaction.

328. I understand, however, that Dr. Petsko has opined that in the context of modified PH20 proteins, aggregation would be uncommon or unlikely:[357]

---

[356] EX1131 (Petsko Dep.), 42:20–43:7; 47:7–19.

[357] EX1131 (Petsko Dep.), 48:23–49:13.

A. Yes. That can happen. It is not -- it's not a common consequence.
Q. When you say "common," give me some context for that.
A. If I make, typically, 100 mutations in a protein, okay, one or two might do that. None of them might do that. But it's -- I have never seen a case where 20 or 30 do that.
Q. So you have not seen a case where one or two, maybe just say three, mutations made to the amino acid sequence caused a protein to aggregate?
A. Maybe once.

329. I disagree. Protein aggregation is a well-established phenomenon and recognized in the scientific literature as a "fundamental problem in biotechnology."[358] For example, "the *de novo* design of novel proteins or the rational modification of existing proteins, is frequently frustrated by the fact that the polypeptide chains aggregate into large assemblies."[359] The ability to aggregate is not "restricted to a relatively few amino acid sequences associated with specific diseases but is a generic phenomenon of polypeptide changes."[360] As the Branden & Tooze textbook explains, aggregation is a phenomenon that occurs commonly

[358] EX1155 (Calloni), 10607.

[359] EX1155 (Calloni), 10607.

[360] EX1156 (Chiti 2002), 16426.

when a protein is not folded into its native state, which can occur inside or outside the cell, and is one of "[t]he most common obstacles to correct folding"[361] (below).

> even if such calculations could be made with great accuracy. One important question therefore is how a living cell can prevent the folding pathway from becoming blocked at an intermediate stage. The most common obstacles to correct folding seem to be (1) aggregation of the intermediates through exposed hydrophobic groups, (2) formation of incorrect disulfide bonds, and (3) isomerization of proline residues. To circumvent these three obstacles cells produce special proteins that assist the folding process, as we shall discuss later in this chapter.

> Before protein molecules attain their native folded state they may expose hydrophobic patches to the solvent. Isolated purified proteins will aggregate during folding even at relatively low protein concentrations. Inside cells, where there are high concentrations of many different proteins, aggregation could therefore occur during the folding process. This is prevented by **molecular chaperones**, ubiquitous and abundant families of proteins that assist the folding of both nascent polypeptides still attached to ribosomes and released completed polypeptide chains. Several chaperones are induced

> The cytoplasm of all cells contains folded proteins and folding polypeptides at high concentrations. Unfolded proteins with exposed hydrophobic patches aggregate easily by non-specific hydrophobic interactions. To circumvent this problem a class of proteins called chaperones have evolved to sequester unfolded polypeptides. The complex structure of one class of multimeric chaperones, the chaperonins GroEL and GroES, has been elucidated and has shed light on how chaperones function. These chaperonins are short cylinders which, because they have hydrophobic residues in the interior, can bind to any unfolded polypeptide that has large exposed hydrophobic patches, regardless of its amino acid sequence. Once the polypeptide chain is shielded inside the chaperonins it is protected from aggregation with other protein molecules. During folding some polypeptide chains go through many cycles of binding and release from the chaperonins.

330. Individual mutations can directly impact the likelihood of protein aggregation by, for example, increasing "the hydrophobicity of the polypeptide chain", increasing "its propensity to convert from α-helical to β-sheet structure" or

[361] EX1014 (Branden & Tooze), 91, 99-100, 118.

decreasing "the overall net charge on the protein molecule."[362] Critically, the effects of multiple mutations on aggregation can be cumulative, arising from combined perturbations to fundamental physicochemical properties of the polypeptide chain.[363] As a result, substitutions that do not individually produce detectable aggregation can, when combined, increase the rate or likelihood of aggregation sufficiently to yield an aggregated protein.[364]

331. Protein aggregation is a "key challenge in the development of protein biotherapeutics."[365] Beyond its impact on product quality—such as loss of biological activity and undesirable drug-product appearance—protein aggregation has been shown to cause cytotoxicity and to increase the immunogenicity of protein products.[366]

---

[362] EX1155 (Calloni), 10607.

[363] EX1157 (Chiti 2003), 808; *see also, e.g.*, EX1003 (Hecht Dec.), ¶¶ 55-56.

[364] EX1157 (Chiti 2003), 808.

[365] EX1158 (Wang 2012), 1.

[366] EX1158 (Wang 2012), Abstract, 1.

332. Consistent with the literature, the common disclosure acknowledges that aggregation can have negative effects on stability and on protein activity.[367]

333. Dr. Petsko has written extensively about the relationship between protein misfolding, aggregation, and immunogenicity. For example, he correctly observes that "[p]roteins are only marginally stable under physiological conditions," that misfolded proteins can aggregate with "deleterious consequences," (top, below) [368] and that immunogenicity in biopharmaceuticals may arise "largely because misfolded proteins break tolerance."[369]

> The term chaperone is borrowed from the name of a class of proteins that function in living cells [1]. Protein molecules are usually only marginally stable under physiological conditions, so some percent of them are often unfolded or misfolded. Such molecules can aggregate with one another, or with properly functioning proteins, with deleterious consequences to the cell. Protein chaperones prevent these unwanted associations by sequestering unfolded and misfolded proteins and providing them with an environment in which they have the opportunity to

[367] EX1026 ('731 App.), 118:15-119:6 ("In particular, increasing or high amounts of preservatives can negatively impact the stability of a protein, including effects on physical stability (aggregation or precipitation) that can impact protein activity.").

[368] EX1152 (Ringe and Petsko), 80.1.

[369] EX1152 (Ringe and Petsko), 80.3; EX1153 (Petsko 2011), 2.

> (iii) ***Understanding why most biopharmaceuticals are immunogenic, even though they are human (or humanized) proteins.*** My own guess is that it is largely because misfolded proteins break tolerance (*10*); if that is true, research on improving the folding and stability of expressed proteins would be worthwhile.

334. There is no scientific basis to treat PH20 variants containing numerous concurrent substitutions as exempt from the principles governing aggregation that is observed. While aggregation may be uncommon for the wild-type PH20, that observation does not permit reliable prediction across the broad genus of modified PH20 polypeptides claimed here. A POSA would have recognized aggregation as being a meaningful and unpredictable risk, especially for multiply modified PH20 variants. As Dr. Petsko has opined, the presence of such aggregates is suspected to be a problem in biopharmaceuticals.

V. Dr. Park Performed a Rigorous, Unbiased and Reliable Assessment of Single-Amino Acid Substitutions in PH20 Polypeptides as a POSA Would Have in 2011

335. Dr. Simpson criticizes Dr. Park's analysis of single amino acid substitutions of PH20 in non-essential regions that a POSA would have expected to retain some hyaluronidase activity.[370] She said his analysis does not explain why a

[370] EX1004 (Park Dec.), ¶ 20 Dr. Park explained how he was using "tolerance" in his declaration ("I use tolerate here to mean that the presence of a single amino

POSA would have picked out any of the 800+ single amino acid that Dr. Park concluded would have been tolerated as single amino acid substitutions in PH20.[371] I understand she believes that Dr. Park did not perform an objective assessment to identify which single amino acid substitutions at non-essential regions of PH20 a POSA would have identified, but was instead instructed by counsel to only consider particular substitutions, such as the substitution of glutamine (Q) for the wild-type leucine (L) at position 317.[372]

336. At her deposition, Dr. Simpson did not appear to be entirely familiar with the details of Dr. Park's analysis of substitutions or his testimony about how

---

acid at a particular position of the protein's amino acid sequence that is different than the naturally occurring ("wild-type") amino acid at that position of the protein does not materially alter the local structure around that position in the protein and thus does not meaningfully alter the biological activity of the protein.").

371 EX2068 (Simpson Dec.), ¶¶ 411-414.

372 EX2068 (Simpson Dec.), ¶ 411, n.18 ("I further note that Dr. Park admitted that his focus on position 317 was not based on an independent scientific analysis, but that he was 'asked by counsel to report [his] conclusions with respect to position 317.' EX1004 ¶ 103. I understand that during his deposition, Dr. Park likewise acknowledged that instead of relying on his own scientific analysis to select position 317, counsel 'asked [him] to elaborate on certain positions….' EX2078 (Park II Tr.), 320:14-15.").

he conducted his work. For example, when she was asked about the spreadsheet[373] that Dr. Park used to compile the results of his analysis, she did not seem to appreciate that it contained Dr. Park's actual assessments of many hundreds of individual amino acid substitutions in PH20 (below).[374]

Q. And do you understand this spreadsheet basically to be compiling his conclusions on particular single amino acid substitutions at positions throughout the PH20 sequence?
MS. MARTIN: Object to form.
(Witness reading.)
A. Dr. Park was compiling his conclusions about substitution at certain residues with respect to their -- the frequency with which an alternative residue appeared in his alignment models, his homology alignments, and with respect to its position within the PH20 structure modeled -- the PH20 model modeled on the structure of HYAL1.

337. When she was asked if she knew if Dr. Park formed his opinions about which substitutions would have been tolerated before he was asked by counsel about particular substitutions, such as the L to Q substitution at position 317, she said "I don't know that."[375]

---

[373] EX2176 (Park Analysis Spreadsheet).

[374] EX1130 (Simpson Dep.), 184:6-18.

[375] EX1130 (Simpson Dep.), 185:14-21.

Q. Correct. So I just want to understand whether you appreciate that Dr. Park created this spreadsheet and captured his views from his assessments before he was asked about these positions, and including the one at position 320.
MS. MARTIN: Object to form, scope.
A. I don't know that.

338. It is therefore unclear to me how Dr. Simpson's report concluded that Dr. Park's analysis was not an independent and objective scientific analysis and was instead limited to substitutions selected by counsel. My review of Dr. Park's description of how he designed and carried out his analysis demonstrates it was a well-designed study that was comprehensive, objective and unbiased. I also believe it accurately reflects what a POSA would have done to answer the questions he was asked (*i.e.*, could a POSA have determined which single amino acid substitutions in $PH20_{1-447}$ would retain biological activity, and if so, what would those single amino acid substitutions be?).[376]

[376] EX1004 (Park Dec.), ¶¶ 15-17; EX2077 (Park Dep.), 67:22-68:14 ("The task that I was given was to identify non-essential positions in PH20 and identify substitutions at those positions. So my task came down to two things: One to identify what those non-essential positions may be; and second, evaluate various substitutions at those positions to see if there are substitutions at non-essential positions that would be tolerated. And that is motivated by patent

339. Dr. Park explained that he devised his methodology to be how a POSA would have conducted the analysis using tools and techniques known by 2011. He outlined the primary steps in ¶ 22 of his declaration (below). The papers Dr. Park cited in ¶ 23 of his declaration describe similar workflows using the same tools and techniques Dr. Park used for his analysis.[377] The analysis he performed was described in sufficient detail that it could be replicated by Dr. Simpson or Dr. Petsko if there were doubts about its correctness.

'427[sic], which I believe is quoted in Paragraph 32. We use the term non-essential because that's what is in the patent, and I was trying to see -- I was tasked to identify the substitutions at non-essential positions as described in the patent to see where they are and what they may be.").

[377] EX1004 (Park Dec.), ¶¶ 2-23 (citing EX1017 (Green), 223-230, 236 (discussing rational design techniques) and EX1016 (Steipe), 181-186).

> 22. To identify single amino acid substitutions that would be tolerated in human $PH20_{1-447}$, I used procedures that were widely used in rational design methods of protein engineering before December 2011. Generally, these include the following steps:
>
> (a) identify homologous sequences based on sequence comparison;
>
> (b) perform a multiple sequence alignment of retrieved sequences;
>
> (c) identify the frequency of occurrence ("profile") of amino acids at each position of the protein across the set of sequences;
>
> (d) generate a protein model using SWISS-MODEL;
>
> (e) use the model to visually assess individual substitutions within the local environment of the protein at a particular position.[3]

340. Dr. Park explained that he investigated many individual substitutions at the non-essential positions of PH20 within the portion of the PH20 sequence that overlapped with the Hyal-1 template used to create his PH20 homology model, and for which the model had reliable local QMEAN values.[378] To assess his methodology and the output of his work, I reviewed a compilation of the numbers of positions and substitutions that he assessed using the data in EX2176.[379]

---

[378] EX1004 (Park Dec.), ¶ 149 (used portion of Hyal-1 template); ¶¶ 148, 150, 157, n.83, n.88, n.94, n.98, n.119 (confirmed adequate local QMEAN values).

[379] *See* Appendix B-8 (Frequency Chart of Alternative Residues Identified in MSA and Number of MSA Alternatives Tested); Appendix B-9 (PH20

Position	WT	Position	WT	Position	WT	Position	WT	Position	WT	Position	WT	Position	WT	Position	WT	Position	WT	Position	WT	Position	WT
4	R	42	G	80	P	118	T	156	E	194	Y	232	T	270	R	308	S	346	C	384	G
5	A	43	S	81	Q	119	W	157	F	195	K	233	Q	271	I	309	I	347	Q	385	K
6	P	44	P	82	K	120	A	158	E	196	K	234	Q	272	V	310	M	348	E	386	P
7	P	45	R	83	I	121	R	159	K	197	P	235	S	273	F	311	R	349	Q	387	T
8	V	46	I	84	S	122	N	160	A	198	G	236	P	274	T	312	S	350	G	388	L
9	I	47	N	85	L	123	W	161	G	199	Y	237	V	275	D	313	M	351	V	389	E
10	P	48	A	86	Q	124	K	162	K	200	N	238	A	276	Q	314	K	352	C	390	D
11	N	49	T	87	D	125	P	163	D	201	G	239	A	277	V	315	S	353	I	391	L
12	V	50	G	88	H	126	K	164	F	202	S	240	T	278	L	316	C	354	R	392	E
13	P	51	Q	89	L	127	D	165	L	203	C	241	L	279	K	317	L	355	K	393	Q
14	F	52	G	90	D	128	V	166	V	204	F	242	Y	280	F	318	L	356	N	394	F
15	L	53	V	91	K	129	Y	167	E	205	N	243	V	281	L	319	L	357	W	395	S
16	W	54	T	92	A	130	K	168	T	206	V	244	R	282	S	320	D	358	N	396	E
17	A	55	I	93	K	131	N	169	I	207	E	245	N	283	Q	321	N	359	S	397	K
18	W	56	F	94	K	132	R	170	K	208	I	246	R	284	D	322	Y	360	S	398	F
19	N	57	Y	95	D	133	S	171	L	209	K	247	V	285	E	323	M	361	D	399	Y
20	A	58	V	96	I	134	I	172	G	210	R	248	R	286	L	324	E	362	Y	400	C
21	P	59	D	97	T	135	E	173	K	211	N	249	E	287	V	325	T	363	L	401	S
22	S	60	R	98	F	136	L	174	L	212	D	250	A	288	Y	326	I	364	H	402	C
23	E	61	L	99	Y	137	V	175	L	213	D	251	I	289	T	327	L	365	L	403	Y
24	F	62	G	100	M	138	Q	176	R	214	L	252	R	290	F	328	N	366	N		
25	C	63	Y	101	P	139	Q	177	P	215	S	253	V	291	G	329	P	367	P		
26	L	64	Y	102	V	140	Q	178	N	216	W	254	S	292	E	330	Y	368	D		
27	G	65	P	103	D	141	N	179	H	217	L	255	K	293	T	331	I	369	N		
28	K	66	Y	104	N	142	V	180	L	218	W	256	I	294	V	332	I	370	F		
29	F	67	I	105	L	143	Q	181	W	219	N	257	P	295	A	333	N	371	A		
30	D	68	D	106	G	144	L	182	G	220	E	258	D	296	L	334	V	372	I		
31	E	69	S	107	M	145	S	183	Y	221	S	259	A	297	G	335	T	373	Q		
32	P	70	I	108	A	146	L	184	Y	222	T	260	K	298	A	336	L	374	L		
33	L	71	T	109	V	147	T	185	L	223	A	261	S	299	S	337	A	375	E		
34	D	72	G	110	I	148	E	186	F	224	L	262	P	300	G	338	A	376	K		
35	M	73	V	111	D	149	A	187	P	225	Y	263	L	301	I	339	K	377	G		
36	S	74	T	112	W	150	T	188	D	226	P	264	P	302	V	340	M	378	G		
37	L	75	V	113	E	151	E	189	C	227	S	265	V	303	I	341	C	379	K		
38	F	76	N	114	E	152	K	190	Y	228	I	266	F	304	W	342	S	380	F		
39	S	77	G	115	W	153	A	191	N	229	Y	267	A	305	G	343	Q	381	T		
40	F	78	G	116	R	154	K	192	H	230	L	268	Y	306	T	344	V	382	V		
41	I	79	I	117	P	155	Q	193	H	231	N	269	T	307	L	345	L	383	R		

Note:
Red - essential residues identified by Dr. Park
Dark Green - positions where Dr. Park evaluated ≥ 66% of alternate residues identified in the MSA
Green - positions where Dr. Park evaluated between 33% and 66% of alternate residues identified in the MSA
Light Green - positions where Dr. Park evaluated < 33% of alternate residues identified in the MSA
Unshaded - positions where Dr. Park did not evaluate alternate substitutions

341. It is evident from a simple inspection of the results of Dr. Park's analysis that it was not focused on any particular position or particular substitution. Given the scale of time and effort required to perform these assessments, I believe he performed a robust assessment.

a. Dr. Park identified 379 non-essential positions across the PH20 sequence based on his sequence alignment of 88 proteins

---

Positional Alignment: Tested MSA Alternatives); EX2176 (Park Analysis Spreadsheet), cols. D-G, pp. 1-50.

homologous to PH20 known in 2011 (*i.e.*, the positions in between highly conserved positions in the homologous proteins).[380]

b. Dr. Park recorded assessments of alternative amino acids at 351 of those non-essential positions (*i.e.*, ~93%).

c. All of the candidate amino acid substitutions he evaluated were included in the amino acid frequency table he compiled using his MSA analysis. In that analysis, at each position, the amino acids that occur at the corresponding position in homologous proteins are listed with most frequent first and then in decreasing frequency.

d. The particular amino acids Dr. Park assessed at the non-essential positions were the ones the MSA identified as being the most frequently occurring ones at the corresponding position in PH20 being assessed.

---

[380] EX1004 (Park Dec.), ¶¶ 25-32 (describing the identification of non-essential residues and essential residues of PH20), ¶ 31 ("The 379 positions in $PH20_{1-447}$ other than the 68 essential residues are positions at which there is evolutionary variation among the 88 protein sequences that I analyzed."), ¶¶ 42-44, Appendix D-2 (compiling the 379 non-essential residues); EX2077 (Park Dep.), 77:12-17 (Q. Okay. So when - - when you did your alignment for PH20, did you include all of the hyaluronidase sequences that were available in 2011, or did you use some other cut-off? A. Yes. I used the sequences that were available as of the end of year 2011.").

e. Dr. Park assessed approximately 830 single–amino-acid substitutions across those 351 non-essential positions. That is ~30% (830/2813) of the entire set of amino acids that occur in the set of homologous proteins he used to produce his MSA analysis.

342. The records of Dr. Park's analysis are consistent with the description of his work that he provided in his declaration and at his depositions. For example, he said that he evaluated the amino acids that occur most frequently at the corresponding positions in the homologous proteins to PH20 of each position in PH20 he assessed.[381] That is consistent with what I observed in Dr. Park's spreadsheet. The most frequently occurring amino acids in the MSA analysis are the ones that were assessed at nearly each position. I note that in many positions,

---

[381] EX2077 (Park Dep.), 176:6-177:1 ("Q. So you said that you looked at about – I think you said you looked at about a quarter of the 3,000 and some; is that right? A. About a quarter, yes. 820 or 830. A. I don't know exactly. Somewhere around there, slightly over 800 substitutions. Q. How did you pick those 800-some -- how did you pick the ones that you looked at? A. I started with substitutions that appeared with high frequency. Ultimately I looked at all substitutions that involved amino acids that appeared with a frequency of 5 or greater percentage. So I started at the top. After doing 10 percent, I think I had about 450. I said: Okay. 450's not bad but I can do better. So let's do additional 5 percent. That added another 400.").

the most frequently occurring amino acid is the wild-type amino acid in PH20, so that was not assessed. There are a few other positions where all or nearly all of the MSA-listed alternatives were assessed.

343. Dr. Park said that he did not consider every amino acid in the MSA at every position for two reasons. One appears to be practical – it would require a lot more time to review every single amino acid at every position.[382] The other appears to reflect his scientific judgment to prioritize the most frequently occurring amino acid changes in the MSA that were present at greater than 5% of the assessed homologous sequences.[383] This is understandable, as amino acids that are found in many homologous proteins to PH20 at a particular position would be viewed as more likely to retain some activity than ones that occur in only 1 or 2 proteins. He also did not assess alternatives at residues 4-7 and around 398-403, because of the uncertainty in the structural model in those regions.[384] I believe these decisions are and would be seen by any POSA as reasonable.

---

[382] EX1004 (Park Dec.), ¶ 42; EX2077 (Park Dep.), 169:10-16; EX2078 (Park Dep. II), 322:10-323:18.

[383] EX1004 (Park Dec.), ¶ 42; EX2077 (Park Dep.), 169:20-170:12, 176:6-20; EX2078 (Park Dep. II), 272:2-17.

[384] EX1004 (Park Dec.), ¶¶ 149-150.

344. Dr. Park assigned each amino acid substitution he evaluated a score of 1, 2, or 3. A score of 1 indicated his judgment that the substitution was predicted to be destabilizing and not tolerated, 2 indicated the substitution was predicted to be tolerated with neutral or minor effects, and a 3 indicated the substitution was predicted to be stabilizing.[385] Dr. Park found that 43 substitutions were likely to be stabilizing.[386] He also identified 83 substitutions that were likely to be destabilizing.[387] The remaining 706 substitutions were rated "2" meaning they would likely be tolerated with neutral or modestly positive or negative effects on stability, including L317Q.[388] He did not attempt to assess the effect of substitutions on activity, other than an assumption that a single mutant that does not perturb or that stabilizes the structure would likely retain some activity, as indicated by the '429 Patent.[389] I agree with Dr. Hecht that one cannot predict a

---

[385] EX1004 (Park Dec.), ¶¶ 85-87 (describing Dr. Park's three-point tolerability scale); *see generally* EX2077 (Park Dep.), 128:14-133:5 (describing classification system).

[386] EX1004 (Park Dec.), ¶¶ 137; EX2176 (Park Analysis Spreadsheet).

[387] EX2176 (Park Analysis Spreadsheet).

[388] EX2176 (Park Analysis Spreadsheet).

[389] EX1004 (Park Dec.), ¶¶ 20, 32; EX1005 ('429 Patent), 16:14-22.

particular magnitude of change to the enzymatic activity based on this analysis, but a POSA would expect a minimum level of enzymatic activity for a neutral or stabilizing substitution.[390]

345. Dr. Park explained that his tolerance ratings reflected his assessment of the collective effect of interactions that he identified for the wild-type amino acid and its neighbors, and then between each substituted amino acid at that position with the neighbors of that substituted amino acid.[391] The types of interactions he considered and effects of those interactions on protein stability were well-known before 2011.[392] Dr. Park also stated that his impressions reflected judgment calls he had to make about the nature of interactions (and the collective

---

[390] EX1003 (Hecht Dec.), ¶¶ 222, 230.

[391] EX1004 (Park Dec.), ¶¶ 102-103; EX2077 (Park Dep.), 146:20-147:22, 170:4-12; EX2078 (Park Dep. II), 293:3-15, 302:12-18 ("Absolutely all substitutions were evaluated based on a combination of the intrinsic properties of the amino acid, its compatibility with the secondary structure, as well as its potential interactions with neighbors.").

[392] EX1004 (Park Dec.), ¶¶ 48-50; EX1031 (Yue), 462, 470-471; EX1032 (Wang 2001), 266.

set of them) that he observed.[393] This was a commonly understood approach in 2011-2012[394] and analogous to what other groups had done before 2011.[395] Drs. Petsko and Simpson did not contest the methodology and offered no alternative approach.

346. I believe Dr. Park's assessment process was an appropriate way to assess the potential effect of a single amino acid substitution on the local environment of the protein at the site of the substitution. This approach allows a POSA to make a reasonable prediction as to the general effect of a single amino acid substitution in terms of destabilizing, neutral or stabilizing to the structure.

---

[393] EX1004 (Park Dec.), ¶¶ 102-103; EX2078, (Park Dep. II), 288:7-10; 290:13-18.

[394] EX1004 (Park Dec.), ¶¶ 33-35; EX1017 (Green), 228-29, 234 ("Visual analysis of protein structure, either with or without an energetic guide such as the residual potential, can suggest sites of modification; chemical intuition can then motivate particular amino-acid substitution. However, additional analysis is often needed due to the intricate networks of interactions typically found in proteins."); EX1031 (Yue), 459.

[395] EX1031 (Yue), 464 (describing classifications of "mutants that stabilize or mildly destabilize" the protein and those with a "high level of destabilization"); EX1032 (Wang 2001), 266-68; EX1004 (Park Dec.), ¶ 51.

347. Halozyme's experts used the same technique of inspecting a residue's interactions with its neighbors to assess the role of the wild-type or to assess the likely consequence of a substitution. For example, Dr. Simpson performed many of the techniques Dr. Park used in his analysis when her team investigated structural issues in Hyal-1, which was published in the Zhang publication in 2009 (EX1010). She explained that she performed sequence alignments to identify fully conserved residues, and generated and used homology models in SWISS-MODEL to evaluate amino acid substitutions.[396] The Zhang publication captures the same type of assessment of interactions that Dr. Park performed.[397]

[396] EX2068 (Simpson Dec.), ¶¶ 126-131.

[397] EX1010 (Zhang), 9436.

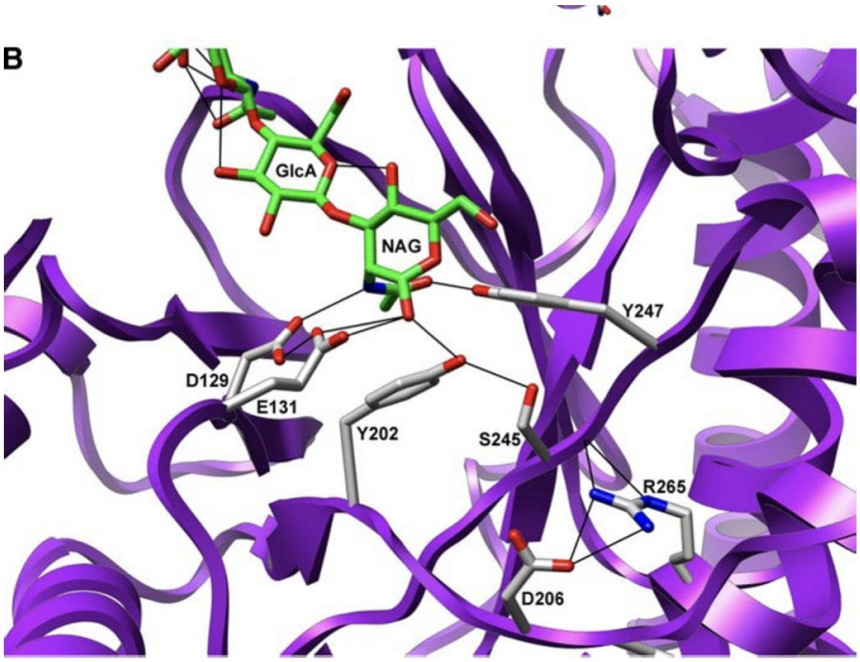


FIGURE 1. **Ribbon representation of the human Hyal1 crystal structure with bound HA tetrasaccharide.** Crystal coordinates from the HA substrate crystallized in the bee venom hyaluronidase structure (PDB code 1FCV) were used to place an HA tetrasaccharide in stick representation at the active site groove of human Hyal1 (PDB code 2PE4). *A*, side chains of residues examined in this study for potential catalytic importance are labeled and shown in stick representations. One active site proximal disulfide is indicated (Cys^{207}–Cys^{221}), as well as the three asparagines that are sites for *N*-glycosylation (Asn^{99}, Asn^{216}, and Asn^{350}), shown modified with short carbohydrate chains as determined in the crystal structure. The molecular interactions occurring between bound HA and putative catalytic residues at the active site are depicted in an active site magnification (*B*), with hydrogen bonds indicated by *thin black lines*. Carbon atoms of the protein residues and modifications are represented in *gray*, whereas those of the substrate are in *green*. In both cases, nitrogen atoms are *blue*, oxygen are *red*, and sulfurs are *yellow*. *GlcA*, glucuronic acid; *NAG*, *N*-acetylglucosamine.

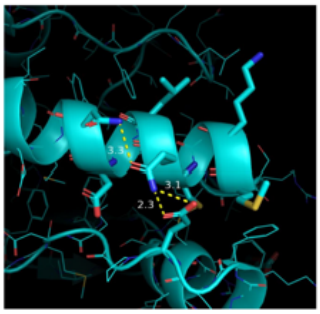
348. Dr. Petsko also replicated Dr. Park's steps to generate a PH20 homology model (*i.e.*, using Hyal-1 as a template in SWISS-MODEL).[398] He then performed the same types of assessments of interactions of the wild-type or substituted amino acid with its neighbors, using the PH20 homology model.[399]

---

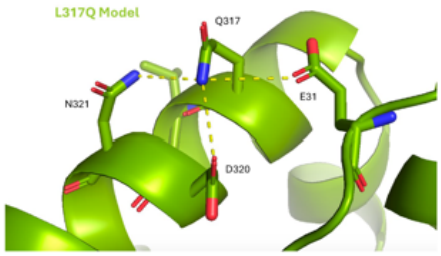
[398] EX2070 (Petsko Dec.), ¶¶ 105-107, 116-124; EX1004 (Park Dec.), ¶¶ 146-148.

[399] *See, e.g.*, EX2070 (Petsko Dec.), ¶¶ 325-328 (examining interactions between wild-type PH20 residues using a homology model to explain why substitutions at certain of those positions were not tolerated).

349. Dr. Petsko also validated Dr. Park's findings regarding the L317Q substitution. Dr. Park explained in his observations in his declaration that he observed that when glutamine (Q) is substituted for leucine (L) at position 317, it can form hydrogen bonds with E31 and N321 that enhance the stability of the polypeptide.[400] Dr. Petsko independently modeled and evaluated the L317Q substitution and also found that L317Q forms interactions with E31 and N321.[401] He did not characterize the nature of the interactions L317Q forms with E31 and N321. Dr. Park provided a more complete description of his observations of the L317Q substitution within the environment of position 317.[402]



Park



Petsko

[400] EX1004 (Park Dec.), ¶¶ 116, 119.

[401] EX2070 (Petsko Dec.), ¶ 404.

[402] EX1004 (Park Dec.), ¶¶ 112-119.

350. Dr. Park's MSA analysis showed that glutamine is the most prevalent amino acid that occurs in homologous proteins at positions corresponding to position 317. It occurs in ~30% of the homologous proteins in Dr. Park's set of 88.[403] A variety of amino acids naturally occur at positions corresponding to position 317, including amino acids that vary in polarity and size. The variety of amino acids suggests that position 317 is not well conserved and many different amino acid substitutions are likely to retain some activity.[404]

351. Dr. Park gave a glutamine substitution a score of 2, meaning it is likely to be stabilizing and thus retain activity, based on its high frequency in homologous sequences, likelihood of stabilizing the PH20 structure based on its potential to form hydrogen bonds with neighboring residues, and the lack of potentially destabilizing interactions in the solvent-exposed environment of position 317.[405] I agree with Dr. Park's analysis, as does Dr. Hecht.[406] Neither Dr. Petsko nor Dr. Simpson identified any errors with Dr. Park's assessment of the L317Q substitution. As I noted above (¶ 349), Dr. Petsko, the only other expert to

---

[403] EX1004 (Park Dec.), ¶ 106, Appendix C.

[404] EX1004 (Park Dec.), ¶ 106.

[405] EX1004 (Park Dec.), ¶¶ 106, 112-119.

[406] EX1003 (Hecht Dec.), ¶¶ 230-234.

conduct a structural analysis of PH20, confirmed Dr. Park's finding of L317Q forming stabilizing interactions with E31 and N321.[407] Dr. Simpson, the only expert who criticized Dr. Park's work, did no independent structural analysis in this case—relying instead on Dr. Petsko's work. That Dr. Petsko's analysis confirmed Dr. Park's findings highlights the validity and reproducibility of Dr. Park's methodology.

352. Dr. Park explained in his declaration and again at his depositions that his assessments were performed objectively, thoroughly, consistently and without bias.[408] His testimony (below) of the methodology he followed of working through the PH20 sequence, position-by-position, and assessing sets of substitutions at each position until he reached the end is consistent with the records of his work that I reviewed.[409] Those show individual substitutions throughout the sequence being rated with comments.[410]

---

[407] EX2070 (Petsko Dec.), ¶ 404.

[408] EX1004 (Park Dec.), ¶¶ 102-103; EX2077 (Park Dep.), 172:19-173:9 ("It's only through repetition that I developed confidence in my scoring system and can vouch that I am being objective when evaluating the substitution."); EX2078 (Park Dep. II), 288:7-16.

[409] EX2078 (Park Dep. II), 170:22-171:17.

[410] EX2176 (Park Analysis Spreadsheet).

A Absolutely. For each substitution that was suggested by the sequence alignment, starting from those substitutions involving amino acids that appear with decent frequency, which I started with 10 percent, but ultimately went down to 5 percent.

So in the end, any substitutions to an amino acid that occur with a frequency 5 or greater percentage, I considered the wild-type residue in the context of the modeled PH20 structure, as well as the individual substitutions listed in the multiple sequence alignment. And for each substitution, I scored the substitution on a scale of 1 to 3, using the metric that was demonstrated in Appendix C earlier.

Q And did you record -- did you record the results of that analysis for every substitution that you analyzed?

A Yes. I created a spreadsheet and provided that to counsel.

Q You provided a spreadsheet with your analysis for every single residue that you considered to counsel?

A Yes. I did.

Q So you said that when you went through the analysis, you would revisit the scores for each position multiple times to make sure that the way you assigned scores was consistent.

Can you explain what you mean by that?

A Yes. I would look at a substitution, give it a score, move on to the next position, consider substitutions at the next position, finish them, move on to the next position, consider substitutions.

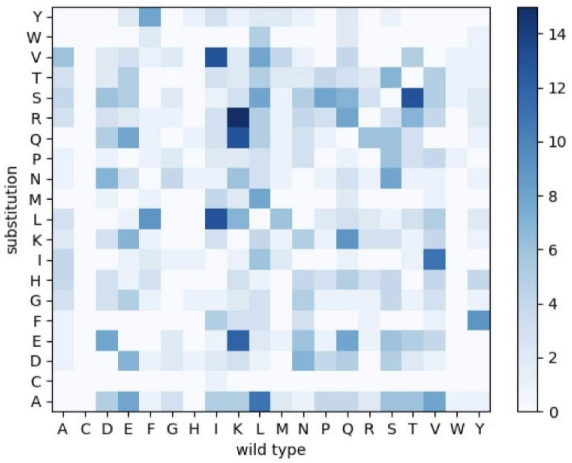
I finished this exercise all the way to the end of the structure, and just to be consistent, I went back to the beginning, and repeated that exercise to either confirm what I had found the first time around, or I might have decided, perhaps I was a bit too harsh here, and I would update the score a little bit. This I did twice for the whole set.

Q So you looked at every single substitution twice?

A Every single substitution which amounted to roughly one quarter of the complete set of possible substitutions, which as I said was about 3,000 substitutions.

353. Dr. Park also included a heatmap in his spreadsheet (below), which shows the number of wild-type vs. substituted amino acid pairs he assessed across the PH20 sequence.[411]

[411] EX2176 (Park Analysis Spreadsheet); EX2078 (Park Dep. II), 297:11-298:8, 298:18-299:2.



354. Dr. Park testified that he generated this heatmap to ensure the robustness of his model (*i.e.*, whether the model assessed many different kinds of substitutions or was a limited set of the possible universe of substitutions).[412] He also said that he used the heatmap to adjust the number of substitutions he considered because he had concluded he had not assessed a sufficient number of pairs of wild-type vs. substituted amino acid combinations.[413]

---

[412] EX2078 (Park Dep. II), 300:2-20.

[413] EX2078 (Park Dep. II), 300:2-20 ("If the heat map shows a lot of empty space, that would indicate that my coverage is limited, so I wanted to make sure that the heat map is well-populated, and that was also the reason that I went from 10 percent to 5 percent in terms of putting a limit on the substitutions."); EX2176 (Park Analysis Spreadsheet).

Q Why did you prepare this heat map?
A One of the things that I was aiming for as I was going through the substitution analysis was to develop a robust system, and I needed a way of knowing how robust my system was, and that I believed involved looking at many different kinds of substitutions; and therefore, I wanted to keep track of what sort of substitutions I was looking at.

If the heat map shows a lot of empty space, that would indicate that my coverage is limited, so I wanted to make sure that the heat map is well-populated, and that was also the reason that I went from 10 percent to 5 percent in terms of putting a limit on the substitutions.
I think at some point, I generated a heat map and didn't – I wasn't convinced that I had enough statistics in there, and felt that I might want to do more.

355. For example, Dr. Park testified that at some point in his analysis he concluded there was too much empty space in the heatmap, indicating a lack of coverage of certain substitution pairs. Based on his testimony, that appears to be a reason he adjusted his threshold down to evaluate more MSA-identified substitutions (*i.e.*, from 10% to 5%).[414] The heatmap reflects the thorough and objective nature of the substitutions Dr. Park assessed.[415]

A Well, at some point, I completed the analysis down to 10 percent, and I looked at the distribution of the substitutions, namely the heat map, which I generated at times, and I wasn't entirely satisfied with what I was seeing, thought that I should perhaps include additional substitutions, and went on to analyze more substitutions, this time down to 5 percent.
And those residues ended up taking longer than expected, actually. I think those residues that fell between 5 to 10 percent were, in fact, more challenging than the ones that had a frequency above 10 percent, which were in ways, more straightforward.

So I was going over the substitutions within the range of 5 to 10 percent. It was taking a while, and it was time for me to consult with the lawyers. And by that time, I had not completed all the evaluations. So there were some positions -- positions and substitutions that weren't included in that study.
Q Okay. So you didn't -- you didn't have time to actually go through and look at all of the substitutions that appear between 5 to 10 percent of the time; is that right?
A That's what ended up happening.

---

414 EX2078 (Park Dep. II), 300:2-20, 322:20-324:3.

415 EX2176 (Park Analysis Spreadsheet).

356. Dr. Park's use of objective metrics to obtain feedback on whether his assessments had adequately probed sufficient wild-type vs. substituted amino acid pairs reflects a good control on his methodology. The actual numbers of each pair, of course, reflect what occurs naturally in the homologous protein set.

357. My review of Dr. Park's description of the process he followed to form his opinions about single amino acid substitutions in non-essential regions of the PH20 protein demonstrates to me that it was objective, unbiased, scientifically rigorous and consistent with the practices followed by POSAs to perform similar analyses in the 2011-2012 timeframe. Moreover, it was described in sufficient detail to be replicated and any error identified. In Appendix B-12, prepared with the assistance of counsel, is a compilation of paragraphs showing where experts in this case agree that tools and techniques Dr. Park used to conduct his analysis were the ones a POSA would have used in the 2011-2012 timeframe, and that he used those tools and techniques appropriately.

358. Dr. Simpson's opinion that Dr. Park's analysis focused on position 317 only after he was asked by counsel is contradicted by the declaration and deposition. Dr. Park testified that position 317 was identified as part of his original, position-agnostic analysis that evaluated non-essential residues and predicted which single amino acid substitutions would be tolerated using pre-2011 tools and

methods.[416] My review of the records of his work is consistent with his explanation. Only after that analysis was completed did counsel ask Dr. Park to elaborate on particular residues, including position 317, for purposes of explanation and presentation. In my view, his comments on L317 reflect a summary of the results already generated by Dr. Park's independent analysis, not a counsel-directed selection of that position.[417] Similar summaries could have been prepared by a POSA for any residue from the results in Dr. Park's spreadsheet.

15 Q Are there residues that you -- for which
16 you considered substitutions that appear less
17 than 5 percent of the time?

5 Perhaps a better example would be later on
6 at position 320 for which I wrote a detailed
7 description as part of the declaration. There
8 were other positions, other substitutions, that
9 had less than 5 percent, which were evaluated.
10 Q So why did you evaluate those
11 substitutions?

12 A By that time, I had already completed all
13 my analysis of all of positions, and provided a
14 spreadsheet to counsel, who then asked me to
15 elaborate on certain positions, including 320,
16 317, and so on. And since I'm being asked to
17 provide detailed description of the substitutions
18 and evaluation at those positions, I went back and
19 finished the analysis with all the substitutions
20 that were available at those positions.
21 Q So am I understanding correctly that you
22 completed your analysis and provided the

1 A The substitutions that fell between 5
2 to 10 percent was included in the analysis at a
3 later time, and it wasn't meaning after the 10
4 percent, cutoff analysis. And then I went back
5 and decided to lower the threshold hold to 5
6 percent. It's the second stage of the analysis,
7 and that study wasn't done with 100 percent
8 coverage. There were some positions that were
9 left out.
10 I want to say that that omission was
11 small. I tried to cover as many as possible, but
12 given the time that was allowed to me, I think
13 some positions ended up being left out. But it's
14 a small percentage, and I would argue that most of
15 them were covered. Most of the substitutions
16 between 5 to 10 were included in the analysis by
17 the time I provided the spreadsheet to counsel.

---

[416] EX2078, (Park Dep. II), 319:15-320:20.

[417] EX1004 (Park Dec.), § V.

359. Dr. Park testified that he has never seen the contents of the ’758 patent or other patents.[418] As reflected in the table below comparing Table 9 of the common disclosure to Dr. Park’s analysis, multiple substitutions that Dr. Park independently rated as “3” (stabilizing) were experimentally confirmed to retain activity, with several exhibiting activity comparable to or greater than wild-type PH20. These results show that Dr. Park’s methodology and analysis was sound and accurate.

[418] EX2077 (Park Dep.), 29:14-20, 33:15-19, 34:13-14, 35:21-22.

Position	Wildtype	Substitution	Dr. Park Rating	Table 9 Activity
27	G	E	3	1.18
31	E	V	3	4.57
43	S	T	3	0.43
70	I	T	3	5.43
73	V	T	3	1.34
97	T	L	3	1.22
97	T	E	3	1.19
98	F	H	3	1.06
98	F	A	3	0.6
105	L	S	3	0.61
139	Q	K	3	0.73
142	V	P	3	0.88
169	I	L	3	2.08
204	F	P	3	0.63
237	V	N	3	0.67
251	I	L	3	0.58
273	F	Y	3	0.9
273	F	H	3	1.41
313	M	K	3	2.85
313	M	R	3	2.3
317	L	R	3	2.41
320	D	K	3	6.42
320	D	R	3	3.19
324	E	R	3	2.28
357	W	S	3	0.41
387	T	S	3	0.93
389	D	K	3	1.91
393	Q	R	3	0.74

360. Consistent with Dr. Hecht's testimony, and based on my review, the '429 Patent (below) informs POSAs what positions to modify in $PH20_{1-447}$ (those occurring in non-essential regions) as well as the kinds of substitutions to make (single-substitutions that retain the biological activity of the protein).[419] Dr. Park's analysis is consistent with the approach a POSA would have taken in 2011-2012,

[419] EX1003 (Hecht Dec.), ¶¶ 206-207; EX1005 ('429 Patent), 16:14-22.

and reveals that many single-amino acid substitutions would not cause local interactions at the site of the substitution sufficient to destabilize that region of the protein structure, and that some would be stabilizing. For those, a POSA would be justified in expecting that the protein with that type of mutation (one rated "2" or "3") would tolerate the mutation and retain at least some hyaluronidase activity (the "biological activity" of the protein referenced by the '429 Patent). I therefore agree with Dr. Hecht that before December of 2011, a POSA who was trying to determine what the single amino acid substitutions in non-essential positions being referenced in the '429 Patent were (below), that person would have followed the methodology Dr. Park used and would have concluded that L317Q would have been one of many readily identified substitutions in $PH20_{1-447}$ that would have been tolerated.

> Suitable conservative substitutions of amino acids are known to those of skill in this art and can be made generally without altering the biological activity, for example enzymatic activity, of the resulting molecule. Those of skill in this art recognize that, in general, single amino acid substitutions in non-essential regions of a polypeptide do not substantially alter biological activity (see, e.g., Watson et al. Molecular Biology of the Gene, 4th Edition, 1987, The Benjamin/Cummings Pub. co., p.224). Also included within the definition, is

I, James Naismith, do hereby declare and state, that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, under Section 1001 of Title 18 of the United States Code.



Executed on: **28 April 2026**

Appendix A - Exhibit List

The list of exhibits that I relied upon in forming my opinions is below.

No.	Exhibit Description
1001	U.S. Patent No. 12,054,758
1003	Declaration of Dr. Michael Hecht
1004	Declaration of Dr. Sheldon Park
1005	U.S. Patent No. 7,767,429
1006	Chao et al., "Structure of Human Hyaluronidase-1, a Hyaluronan Hydrolyzing Enzyme Involved in Tumor Growth and Angiogenesis," Biochemistry, 46:6911-6920 (2007)
1008	Stern et al., "The Hyaluronidases: Their Genomics, Structures, and Mechanisms of Action," Chem. Rev. 106:818-839 (2006)
1010	Zhang et al., "Hyaluronidase Activity of Human Hyal1 Requires Active Site Acidic and Tyrosine Residues," J. Biol. Chem., 284(14):9433-9442 (2009)
1011	Arming et al., "In vitro mutagenesis of PH-20 hyaluronidase from human sperm," Eur. J. Biochem., 247:810-814 (1997)
1012	Bordoli et al., "Protein structure homology modeling using SWISS-MODEL workspace," Nature Protocols, 4(1):1-13 (2008)
1014	Branden & Tooze, "Introduction to Protein Structure," Second Ed., Chapters 1-6, 11-12, 17-18 (1999)
1016	Steipe, "Consensus-Based Engineering of Protein Stability: From Intrabodies to Thermostable Enzymes," Methods in Enzymology, 388:176-186 (2004)
1017	Green, "Computer Graphics, Homology Modeling, and Bioinformatics," Protein Eng'g & Design, Ch. 10, 223-237 (2010)
1026	U.S. Patent Application No. 13/694,731
1027	2024 Chemistry Nobel Prize Background
1029	Gmachl et al., "The human sperm protein PH-20 has hyaluronidase activity," FEBS Letters, 3:545-548 (1993)
1031	Yue et al., "Loss of Protein Structure Stability as a Major Causative Factor in Monogenic Disease," J. Mol. Biol., 353:459-473 (2005)

1032	Wang & Moult, "SNPs, Protein Structure, and Disease," Hum. Mutation, 17:263-270 (2001)
1033	Marković-Housley et al., "Crystal Structure of Hyaluronidase, a Major Allergen of Bee Venom," Structure, 8:1025-1035 (2000)
1039	Alberts, "Molecular Biology of the Cell," Fifth Edition, Chapter 3 (2007)
1045	Redline Comparison of the '731 App. and '758 Specifications
1051	U.S. Patent Application No. 61/631,313
1052	U.S. Patent Application No. 61/796,208
1126	Im, S. et al., "Cryo-EM Structure of Human Hyaluronidase PH20," Proteins: Structure, Function, and Bioinformatics 93:1067-1073 (2025)
1130	Melanie Ann Simpson, Ph.D. Deposition Transcript
1131	Gregory Petsko, Ph.D. Deposition Transcript
1135	Dr. Naismith Analysis Spreadsheet
1136	Dobson, "Principles of protein folding, misfolding and aggregation" Semin. Cell. Dev. Biol. 15(1):3-16 (2004)
1137	Dobson, "Protein folding and misfolding" Nature. 18;426(6968):884-90 (2003)
1138	Dobson, "Experimental investigation of protein folding and misfolding" Methods 34(1):4-14 (2004)
1139	Chiti & Dobson, "Protein Misfolding, Functional Amyloid, and Human Disease" Annu. Rev. Biochem. 75:333–66 (2006)
1140	Dobson, "Protein misfolding, evolution and Disease" Trends Biochem. Sci. (9):329-32 (1999)
1141	Goldsmith & Tawfik, "Directed enzyme evolution: beyond the low-hanging fruit" Curr. Opin. Struct. Biol. 22(4):406-12 (2012)
1142	Farber & Petsko, "The evolution of alpha/beta barrel enzymes" Trends Biochem. Sci. 15(6):228-34 (1990)
1143	Nagano et al., "One Fold with Many Functions: The Evolutionary Relationships between TIM Barrel Families Based on their Sequences, Structures and Functions" J. Mol. Biol. 30;321(5):741-65 (2002)

1144	Ochoa-Levya et al., "Protein Design through Systematic Catalytic Loop Exchange in the (beta/alpha)8 Fold" J. Mol. Biol. 10 387(4):949-64 (2009)
1145	AutoDock Vina Tutorial 2011
1146	Pal et al., "Comprehensive and Quantitative Mapping of Energy Landscapes for Protein-Protein Interactions by Rapid Combinatorial Scanning" J. Biol. Chem. 4;281(31):22378-22385 (2006)
1147	van Rossum et al., "Reporter-based screening and selection of enzymes" FEBS J. 280:2979-96 (2013)
1148	Reed et al., "Use of Coefficient of Variation in Assessing Variability of Quantitative Assays" Clin. Diagn. Lab Immunol. 9(6):1235-9 (2002)
1149	Ten Have et al., "Cloning and characterization of the cDNA encoding the PH20 protein in the European red fox Vulpes vulpes" Reprod. Fertil. Dev. 10(2):165-72 (1998)
1150	Chowpongpang et al., "Cloning and characterization of the bovine testicular PH-20 hyaluronidase core domain" Biotechnol. Lett. 26(15):1247-52 (2004)
1151	Fowler et al., "High-resolution mapping of protein sequence-function relationships" Nature Methods 7, 741–746 (2010)
1152	Ringe & Petsko, "Q&A: What are pharmacological chaperones and why are they interesting?" J. Biol. 8:80 (2009)
1153	Petsko, "Herding CATS" Sci. Transl. Med. 24;3(97):97 (2011)
1154	Second Table Associating Citations from the '758 Patent (EX1001) to Corresponding Citations in the '731 Application (EX1026)
1155	Calloni et al., "Investigating the Effects of Mutations on Protein Aggregation in the Cell" J. Biol. Chem. 10607-13 (2005)
1156	Chiti et al., "Studies of aggregation of mutant proteins in vitro provided insights into genetics of amyloid diseases" PNAS 99:16419-26 (2002)
1157	Chiti et al., "Rationalization of the effects of mutations on peptide and protein aggregation rates" Nature, 424:805-808 (2003)

1158	Wang et al., "Immunogenicity of protein aggregates-Concerns and Realities" Int. J. Pharmaceutics 431:1-11, (2012)
1159	Brachmann et al., "Genetic selection of intragenic suppressor mutations that reverse the effect of common p53 cancer mutations" EMBO Journal 17:1847–59 (1998)
1160	Poelwijk, Frank J., et al., "Empirical fitness landscapes reveal accessible evolutionary paths" Nature 445.7126 (2007): 383-386
2068	Declaration of Melanie A. Simpson, Ph.D. in Support of Patent Owner's Response
2070	Declaration of Gregory A. Petsko, Ph.D. in Support of Patent Owner's Response
2076	Transcript of the Deposition of Michael Hecht, Ph.D., August 26, 2025
2077	Transcript of the Deposition of Dr. Sheldon Park, August 7, 2025
2078	Transcript of the Deposition of Dr. Sheldon Park, August 21, 2025
2164	Petsko, G., et al., Protein Structure and Function, New Science Press (2004)
2165	U.S. Patent Application Publication No. 2010/0143457 A1 to Wei et al., published June 10, 2010
2166	Petsko Patent Analysis Data
2176	Park Analysis Spreadsheet

Appendix B

Tables

Appendix B-1

Counts and Classifications of $PH20_{1-447}$ Mutants

Source(s): EX1026 ('731 App.), at Tables 3, 5, 9, and 10; EX2166 (Petsko Patent Analysis Data).

Position	WT	Assigned Color	% Active of Tested	Active Subs (Col R)	Inactive Subs (Col S)	Tested Subs	% Active of Tested (w/o errors, >40%)	Total Tested (w/o errors)	Total (difference)	Class Changed?	% Diff Active	Errors	Active subs (20%<x<40%)	Active subs (>40%)	Active Subs (P plus Q)	Inactive subs (<20%)
1	L	4	100.0%	15	0	15	100.0%	15	0		0.0%			ACEFGHKNPQRSTVW	ACEFGHKNPQRSTVW	
2	N	3	71.4%	10	4	14	69.2%	13	1	N	2.2%		I	ACGLPQSTV	IACGLPQSTV	HKWY
3	F		40.0%	4	6	10	40.0%	10	0		0.0%			EHLY	EHLY	AGKPTV
4	R		38.5%	5	8	13	38.5%	13	0		0.0%			AISTV	AISTV	DEFGLPWY
5	A	2	7.1%	1	13	14	7.1%	14	0		0.0%			H	H	DGILMNPQRTVWY
6	P		58.3%	7	5	12	58.3%	12	0		0.0%			AHKLNQR	AHKLNQR	EFTVY
7	P	2	6.7%	1	14	15	6.7%	15	0		0.0%			M	M	CDFGHIKLQRSTWY
8	V		33.3%	4	8	12	27.3%	11	1	Y	6.1%		P	ILM	PILM	DEGHNRSW
9	I		50.0%	6	6	12	50.0%	12	0		0.0%			KLQRSV	KLQRSV	CDEGNP
10	P	3	64.3%	9	5	14	64.3%	14	0		0.0%			DEGHNQRSW	DEGHNQRSW	FILMY
11	N		35.7%	5	9	14	30.8%	13	1	N	4.9%		S	DGHK	SDGHK	ACFILPTWY
12	V	3	75.0%	9	3	12	75.0%	12	0		0.0%			AEIKLNRST	AEIKLNRST	GHW
13	P		40.0%	4	6	10	40.0%	10	0		0.0%			HSTY	HSTY	EGILMV
14	F		30.8%	4	9	13	30.8%	13	0		0.0%			DIMV	DIMV	AEGHKNPQW
15	L	2	23.1%	3	10	13	23.1%	13	0		0.0%			AMV	AMV	EFGKNPQRSY
16	W	1	0.0%	0	14	14	0.0%	14	0		0.0%					ACDEFGHKMPRSTY
17	A	1	0.0%	0	15	15	0.0%	15	0		0.0%					DEGHILNPQRSTVWY
18	W	1	0.0%	0	14	14	0.0%	14	0		0.0%					CDFGHILMPQSTVY
19	N	1	0.0%	0	15	15	0.0%	15	0		0.0%					ACFGHILMPQRSVWY
20	A	2	7.7%	1	12	13	7.7%	13	0		0.0%			S	S	DEFHKLNPRTVY
21	P	1	0.0%	0	14	14	0.0%	14	0		0.0%					ACDEGHILMRSTVW
22	S		44.4%	4	5	9	44.4%	9	0		0.0%			HMTY	HMTY	CEGKP
23	E	2	9.1%	1	10	11	9.1%	11	0		0.0%			D	D	AFLMNPRSTV
24	F	4	100.0%	13	0	13	100.0%	13	0		0.0%			AEGHIKLMNRTVY	AEGHIKLMNRTVY	
25	C	1	0.0%	0	15	15	0.0%	15	0		0.0%					DEFGHIKLNPRSTVY
26	L	4	100.0%	15	0	15	100.0%	15	0		0.0%			AEGHIKMPQRSTVWY	AEGHIKMPQRSTVWY	
27	G	3	93.3%	14	1	15	93.3%	15	0		0.0%			ADEFHIKLPQRSTW	ADEFHIKLPQRSTW	C
28	K	4	100.0%	14	0	14	100.0%	14	0		0.0%			ADEFILMNPRSTVW	ADEFILMNPRSTVW	
29	F	4	100.0%	14	0	14	100.0%	14	0		0.0%			AEGHIKLMPRSTVW	AEGHIKLMPRSTVW	
30	D	4	100.0%	14	0	14	100.0%	14	0		0.0%			AFGHKLMPQRSTVW	AFGHKLMPQRSTVW	
31	E	4	100.0%	14	0	14	100.0%	14	0		0.0%			ACGHIKLPRSTVWY	ACGHIKLPRSTVWY	
32	P	4	100.0%	16	0	16	100.0%	16	0		0.0%			ACFGHKLMNQRSTVWY	ACFGHKLMNQRSTVWY	
33	L		57.1%	8	6	14	57.1%	14	0		0.0%			GMPQRSTW	GMPQRSTW	CDHNVY
34	D		53.8%	7	6	13	50.0%	12	1	N	3.8%		A	EHKQRW	AEHKQRW	ILNSTV
35	M		53.8%	7	6	13	45.5%	11	2	N	8.4%		QV	FHLTY	QVFHLTY	ADGPRS
36	S	3	64.3%	9	5	14	58.3%	12	2	Y	6.0%		DN	AGHKLRT	DNAGHKLRT	CFVWY
37	L	3	61.5%	8	5	13	58.3%	12	1	Y	3.2%		W	FIKMPRV	WFIKMPRV	CEGNS
38	F	2	10.0%	1	9	10	10.0%	10	0		0.0%			Y	Y	EGKLNQRTW
39	S	3	63.6%	7	4	11	63.6%	11	0		0.0%			ALNQRTY	ALNQRTY	CDFW
40	F	2	16.7%	2	10	12	16.7%	12	0		0.0%			LW	LW	ADEGKNRSTV
41	I	3	90.9%	10	1	11	90.9%	11	0		0.0%			ACDEGHNTVW	ACDEGHNTVW	Q
42	G	2	7.1%	1	13	14	7.1%	14	0		0.0%			A	A	DEHIKLMPQRSTV
43	S	2	16.7%	2	10	12	9.1%	11	1	N	7.6%		N	T	NT	AEFGIKLQRV
44	P	2	6.7%	1	14	15	6.7%	15	0		0.0%			E	E	ACFGHILNQRSTWY
45	R	2	25.0%	2	6	8	25.0%	8	0		0.0%			IK	IK	ADFGPW
46	I	3	86.7%	13	2	15	85.7%	14	1	N	1.0%		C	AEFHLMNRSTVY	CAEFHLMNRSTVY	PW
47	N	3	92.9%	13	1	14	92.9%	14	0		0.0%			ADFGHKMQRSTWY	ADFGHKMQRSTWY	V
48	A	3	92.3%	12	1	13	92.3%	13	0		0.0%			FGHIKMNQRSVY	FGHIKMNQRSVY	P
49	T		50.0%	5	5	10	50.0%	10	0		0.0%			IKRSV	IKRSV	CDGHP
50	G	4	100.0%	13	0	13	100.0%	11	2	N	0.0%		V	ACDEHLMQRSVY	VACDEHLMQRSVY	
51	Q		33.3%	4	8	12	20.0%	10	2	Y	13.3%		AR	NS	ARNS	CFIMPTWY
52	G		54.5%	6	5	11	54.5%	11	0		0.0%			NPQRST	NPQRST	CEFWY
53	V	1	0.0%	0	15	15	0.0%	15	0		0.0%					ACDEGHLNPQRSTWY
54	T		50.0%	6	6	12	50.0%	12	0		0.0%			AFNQSV	AFNQSV	DEGPRY
55	I	1	0.0%	0	11	11	0.0%	11	0		0.0%					ADGHNPQRTVY
56	F	1	0.0%	0	14	14	0.0%	14	0		0.0%					ACEGHIKLPRSTVW
57	Y	1	0.0%	0	11	11	0.0%	12	-1	N	0.0%					ADFGIMPQRVW
58	V	3	92.9%	13	1	14	92.9%	14	0		0.0%			CGHIKLNPQRSWY	CGHIKLNPQRSWY	A
59	D	2	15.4%	2	11	13	15.4%	13	0		0.0%			QN	QN	AEILMPRTVWY
60	R	2	6.7%	1	14	15	6.7%	15	0		0.0%			K	K	ADFGHILNPQSTVY
61	L	2	25.0%	4	12	16	21.4%	14	2	N	3.6%	F	F	IMV	FIMV	AEFGHNPQRTWY
62	G	1	0.0%	0	15	15	0.0%	15	0		0.0%					ACDFIKLMPQRSTVY

Position	WT	Assigned Color	% Active of Tested	Active Subs (Col R)	Inactive Subs (Col S)	Tested Subs	% Active of Tested (w/o errors, >40%)	Total Tested (w/o errors)	Total (difference)	Class Changed?	% Diff Active	Errors	Active subs (20%<x<40%)	Active subs (>40%)	Active Subs (P plus Q)	Inactive subs (<20%)
63	Y	3	80.0%	12	3	15	80.0%	15	0		0.0%			AHIKLMNRSTVW	AHIKLMNRSTVW	CGP
64	Y	1	0.0%	0	17	17	0.0%	17	0		0.0%					ACDEFGHIKLPQRSTVW
65	P	2	7.1%	1	13	14	0.0%	13	1	N	7.1%	R		R	R	ACDGHIKNSTVWY
66	Y	2	13.3%	2	13	15	13.3%	15	0		0.0%			HR	HR	ACDEGIKLNPSTV
67	I		41.7%	5	7	12	40.0%	10	2	N	1.7%	R	R	FLVY	RFLVY	DEGPRTW
68	D	3	66.7%	10	5	15	58.3%	12	3	Y	8.3%	GLP	G	EHKLPQRST	GEHKLPQRST	ACIVY
69	S	3	92.9%	13	1	14	92.3%	13	1	N	0.5%	T		ACEFGILMPRTWY	ACEFGILMPRTWY	N
70	I	3	93.3%	14	1	15	93.3%	15	0		0.0%			ACFGHKLNPRSTVY	ACFGHKLNPRSTVY	Q
71	T	3	90.9%	10	1	11	90.9%	11	0		0.0%			ADGHLMNQRS	ADGHLMNQRS	P
72	G	3	61.1%	11	7	18	60.0%	15	3	N	1.1%	H	Y	ADEHKLMQRS	YADEHKLMQRS	CFHIPVW
73	V	3	92.9%	13	1	14	92.9%	14	0		0.0%			ACDGHKLMQRSTW	ACDGHKLMQRSTW	P
74	T	4	100.0%	15	0	15	100.0%	15	0		0.0%			ACEFGHKLMNPRSVW	ACEFGHKLMNPRSVW	
75	V	3	80.0%	12	3	15	80.0%	15	0		0.0%			ACFHLMNQRSTY	ACFHLMNQRSTY	DGP
76	N	1	0.0%	0	14	14	0.0%	14	0		0.0%					ACFGIKLPQRSTVW
77	G	2	20.0%	2	8	10	0.0%	8	2	Y	20.0%		HK		HK	DELPQRTV
78	G	1	0.0%	0	7	7	0.0%	7	0		0.0%					ADIMPTY
79	I	2	21.4%	3	11	14	21.4%	14	0		0.0%			LTV	LTV	ADFGHKNPSWY
80	P	1	0.0%	0	15	15	0.0%	15	0		0.0%					ADEFGIKLMNRSTVY
81	Q	2	7.7%	1	12	13	0.0%	11	2	Y	7.7%	P		P	P	ACEGHLNPSVWY
82	K	3	93.3%	14	1	15	92.9%	14	1	N	0.5%	Y	Y	AEGHILMNQRSTV	YAEGHILMNQRSTV	K
83	I	3	91.7%	11	1	12	90.9%	11	1	N	0.8%		K	FGHLNQRSTV	KFGHLNQRSTV	E
84	S	4	100.0%	15	0	15	100.0%	14	1	N	0.0%		Y	DEFGHILMNPQRTW	YDEFGHILMNPQRTW	
85	L	2	8.3%	1	11	12	8.3%	12	0		0.0%			V	V	ACDEFGHNQST
86	Q	3	94.4%	17	1	18	94.1%	17	1	N	0.3%	P		ADEFGHIKLMNPRSTVW	ADEFGHIKLMNPRSTVW	C
87	D	4	100.0%	15	0	15	100.0%	14	1	N	0.0%	P	P	ACEGHILMQRSTVY	PACEGHILMQRSTVY	
88	H	1	0.0%	0	15	15	0.0%	15	0		0.0%					ACEFGIKLMPRSTVY
89	L		40.0%	6	9	15	27.3%	11	4	Y	12.7%	W	RPW	CKM	RPWCKM	ADEGQSTWY
90	D	3	92.9%	13	1	14	92.3%	13	1	N	0.5%	G		AEGHIKLNQRSTW	AEGHIKLNQRSTW	C
91	K	2	27.3%	3	8	11	25.0%	12	-1	N	2.3%	N		AQR	AQR	DEFGHILT
92	A		42.9%	6	8	14	38.5%	13	1	N	4.4%	H	H	CLMTV	HCLMTV	EFKPQRWY
93	K	4	100.0%	15	0	15	100.0%	13	2	N	0.0%		PV	DEFGHILMNQRST	PVDEFGHILMNQRST	
94	K	3	86.7%	13	2	15	85.7%	14	1	N	1.0%		C	ADEFHLMNQRST	CADEFHLMNQRST	GP
95	D	1	0.0%	0	15	15	0.0%	15	0		0.0%					ACEFGHKLMPQSVWY
96	I	2	30.0%	3	7	10	25.0%	12	-2	N	5.0%	CG		DLV	DLV	SHPRSTW
97	T	4	100.0%	15	0	15	100.0%	15	0		0.0%			ACDEFGILNPQRSWY	ACDEFGILNPQRSWY	
98	F	3	92.9%	13	1	14	92.9%	14	0		0.0%			ACDEHILMQRSVW	ACDEHILMQRSVW	P
99	Y	2	27.3%	3	8	11	20.0%	10	1	N	7.3%		A	RS	ARS	CEGINPVW
100	M	1	0.0%	0	11	11	0.0%	11	0		0.0%					CEFGNPRSTWY
101	P	1	0.0%	0	13	13	0.0%	13	0		0.0%					ACFHIKLMNQRST
102	V	3	93.3%	14	1	15	93.3%	15	0		0.0%			ACEGHKLMNQRSTW	ACEGHKLMNQRSTW	P
103	D	2	7.1%	1	13	14	0.0%	13	1	Y	7.1%		N		N	AEFGHILQRTVWY
104	N	3	75.0%	9	3	12	72.7%	11	1	N	2.3%		I	ACGKMRST	IACGKMRST	FPW
105	L	3	85.7%	12	2	14	81.8%	11	3	N	3.9%	CG	WCH	AGIPQRSTV	WCHAGIPQRSTV	MN
106	G	2	7.7%	1	12	13	7.7%	13	0		0.0%			V	V	ACDFHLMNPSWY
107	M	2	25.0%	3	9	12	25.0%	12	0		0.0%			FIL	FIL	ACHKPQSVW
108	A	2	8.3%	1	11	12	8.3%	12	0		0.0%			G	G	DEFKLMPQTVY
109	V	1	0.0%	0	8	8	0.0%	8	0		0.0%					CDELMRTW
110	I	2	14.3%	1	6	7	14.3%	7	0		0.0%			V	V	FKLMPW
111	D	1	0.0%	0	3	3	0.0%	3	0		0.0%					HIQ
112	W	1	0.0%	0	8	8	0.0%	8	0		0.0%					CEGHLNPS
113	E	1	0.0%	0	2	2	0.0%	2	0		0.0%					RV
114	E		50.0%	5	5	10	50.0%	10	0		0.0%			AGHMS	AGHMS	ILPTV
115	W	1	0.0%	0	14	14	0.0%	14	0		0.0%					ACDFGHIKLMRSVY
116	R	1	0.0%	0	14	14	0.0%	14	0		0.0%					ACDEGHILNPQSVW
117	P	2	10.0%	1	9	10	0.0%	9	1	Y	10.0%	D		D	D	GIKNQRSVW
118	T		46.7%	7	8	15	46.7%	15	0		0.0%			HKLMNQV	HKLMNQV	CDEGPRWY
119	W		40.0%	4	6	10	33.3%	9	1	N	6.7%	P	P	FQY	PFQY	AKILNR
120	A	4	100.0%	14	0	14	100.0%	14	0		0.0%			DFGHILNPRSTVWY	DFGHILNPRSTVWY	
121	R	1	0.0%	0	12	12	0.0%	12	0		0.0%					ACEFGHKLMPWY
122	N	2	8.3%	1	11	12	8.3%	12	0		0.0%			M	M	ACEFIKQRSTV
123	W	1	0.0%	0	14	14	0.0%	14	0		0.0%	C				ACDEHLMPQRSTVY
124	K		37.5%	3	5	8	16.7%	6	2	Y	20.8%		LH	R	LHR	CDEFN
125	P		40.0%	4	6	10	33.3%	9	1	N	6.7%		A	HRS	AHRS	CDGLNW
126	K	1	0.0%	0	7	7	0.0%	7	0		0.0%					FHILNPY
127	D	3	92.9%	13	1	14	92.9%	14	0		0.0%			AEGHLMNQRSTVW	AEGHLMNQRSTVW	K
128	V	3	83.3%	10	2	12	83.3%	12	0		0.0%			ACGIKLQRSW	ACGIKLQRSW	EP
129	Y	1	0.0%	0	13	13	0.0%	13	0		0.0%					ACDEGHLPQSTVW
130	K	2	16.7%	2	10	12	16.7%	12	0		0.0%			IR	IR	CDGHLNSTWY
131	N	3	93.3%	14	1	15	93.3%	15	0		0.0%			CEFGHILMQRSTVY	CEFGHILMQRSTVY	P

Position	WT	Assigned Color	% Active of Tested	Active Subs (Col R)	Inactive Subs (Col S)	Tested Subs	% Active of Tested (w/o errors, >40%)	Total Tested (w/o errors)	Total (difference)	Class Changed?	% Diff Active	Errors	Active subs (20%<x<40%)	Active subs (>40%)	Active Subs (P plus Q)	Inactive subs (<20%)
132	R	3	93.3%	14	1	15	93.3%	15	0		0.0%			ACEFHIKLNQSTVY	ACEFHIKLNQSTVY	P
133	S	2	7.1%	1	13	14	7.1%	14	0		0.0%	C		I	I	DEFGHLMNPRTVW
134	I	2	20.0%	3	12	15	20.0%	15	0		0.0%			LTV	LTV	ACDFGHKPQRSW
135	E	3	93.3%	14	1	15	93.3%	15	0		0.0%			ACDFGHKLNQRSWY	ACDFGHKLNQRSWY	P
136	L	3	92.9%	13	1	14	92.9%	14	0		0.0%			ACDFHIMNQRSTW	ACDFHIMNQRSTW	P
137	V		42.9%	6	8	14	27.3%	11	3	Y	15.6%	V	CSL	AIT	CSLAIT	FGHPRNWY
138	Q	3	91.7%	11	1	12	91.7%	12	0		0.0%			ACHILMNRSWY	ACHILMNRSWY	V
139	Q	3	93.3%	14	1	15	93.3%	15	0		0.0%	K		ACDEFGHKLMRSTV	ACDEFGHKLMRSTV	P
140	Q	4	100.0%	14	0	14	100.0%	14	0		0.0%	E		ACDFGHIKLMRVWY	ACDFGHIKLMRVWY	
141	N	4	100.0%	15	0	15	100.0%	15	0		0.0%			ADEFGHLMQRSTVWY	ADEFGHLMQRSTVWY	
142	V	4	100.0%	15	0	15	100.0%	15	0		0.0%			CDEGHIKLMNPQRST	CDEGHIKLMNPQRST	
143	Q	3	61.5%	8	5	13	58.3%	12	1	Y	3.2%	S	C	EGIKLNV	CEGIKLNV	HPRST
144	L	2	23.1%	3	10	13	16.7%	12	1	N	6.4%		R	TW	RTW	AEFIKPQSVY
145	S	3	84.6%	11	2	13	84.6%	13	0		0.0%	W		ACDEGHLMNPR	ACDEGHLMNPR	TW
146	L	4	100.0%	15	0	15	100.0%	15	0		0.0%			ACEGHIKNPQRSTVY	ACEGHIKNPQRSTVY	
147	T	4	100.0%	15	0	15	100.0%	15	0		0.0%			ACDFGILMPQRSVWY	ACDFGILMPQRSVWY	
148	E	4	100.0%	14	0	14	100.0%	14	0		0.0%			CFGHIKLQRSTVWY	CFGHIKLQRSTVWY	
149	A	3	83.3%	10	2	12	83.3%	12	0		0.0%	W		CGKLMQRSTY	CGKLMQRSTY	EP
150	T	3	93.3%	14	1	15	93.3%	15	0		0.0%	V		ACDEFGILNPRSWY	ACDEFGILNPRSWY	V
151	E	4	100.0%	15	0	15	100.0%	15	0		0.0%			ACGHKLMNQRSTVWY	ACGHKLMNQRSTVWY	
152	K	3	90.9%	10	1	11	90.0%	10	1	N	0.9%		W	ACFIMRTVY	WACFIMRTVY	L
153	A	2	30.0%	3	7	10	22.2%	9	1	N	7.8%	W	S	IL	SIL	EFMPRTV
154	K		36.4%	4	7	11	30.0%	10	1	Y	6.4%		I	RTV	IRTV	DEGPSWY
155	Q	3	87.5%	14	2	16	87.5%	16	0		0.0%			ACDFGHKLMRSTVW	ACDFGHKLMRSTVW	PY
156	E	3	92.9%	13	1	14	92.3%	13	1	N	0.5%	V	C	ADGILMQRSTVW	CADGILMQRSTVW	P
157	F	2	5.9%	1	16	17	5.9%	17	0		0.0%			W	W	ACDEGHIKLMPQRSTV
158	E		58.3%	7	5	12	54.5%	11	1	N	3.8%		G	AFHLQS	GAFHLQS	DKPRY
159	K	3	85.7%	12	2	14	84.6%	13	1	N	1.1%		G	ADEHLMNQRSV	GADEHLMNQRSV	WY
160	A	4	100.0%	15	0	15	100.0%	14	1	N	0.0%		W	CFGHIKLMNQRSVY	WCFGHIKLMNQRSVY	
161	G	3	87.5%	7	1	8	87.5%	8	0		0.0%			ACDERSV	ACDERSV	W
162	K	4	100.0%	14	0	14	100.0%	14	0		0.0%			ADEGHLMPQRSVWY	ADEGHLMPQRSVWY	
163	D	3	85.7%	12	2	14	83.3%	12	2	N	2.4%		WF	AEGKLQRSTV	WFAEGKLQRSTV	CP
164	F	2	28.6%	4	10	14	28.6%	14	0		0.0%	LV		LMVW	LMVW	ACDEGHNPQR
165	L	3	76.9%	10	3	13	75.0%	12	1	N	1.9%		C	ADFNRSVWY	CADFNRSVWY	HPT
166	V	3	92.9%	13	1	14	92.9%	14	0		0.0%	A		ACEFGHLNQRTWY	ACEFGHLNQRTWY	D
167	E	3	92.9%	13	1	14	92.3%	13	1	N	0.5%		F	ADGHKMNPRSTY	FADGHKMNPRSTY	V
168	T	2	6.7%	1	14	15	6.7%	15	0		0.0%			H	H	ACDEFGKLPRSVWY
169	I	2	20.0%	3	12	15	20.0%	15	0		0.0%			LRV	LRV	ADFGHKNPQSTY
170	K		38.5%	5	8	13	38.5%	13	0		0.0%			AQNRV	AQNRV	CDEGMPWY
171	L	2	18.2%	2	9	11	18.2%	11	0		0.0%			IV	IV	CDHMNRSWY
172	G	2	16.7%	2	10	12	16.7%	12	0		0.0%	S		AC	AC	DEILPQTWY
173	K	2	20.0%	3	12	15	14.3%	14	1	N	5.7%	C	Q	NR	QNR	DEGHILMPSVWY
174	L	3	92.9%	13	1	14	92.3%	13	1	N	0.5%	R	H	AGKMNQRSTVWY	HAGKMNQRSTVWY	P
175	L		41.7%	5	7	12	41.7%	12	0		0.0%	D		EHTVY	EHTVY	CDGKPRS
176	R	2	13.3%	2	13	15	13.3%	15	0		0.0%	GI		KL	KL	ACEFGHIPQSTVW
177	P	2	7.1%	1	13	14	0.0%	13	1	Y	7.1%		V		V	ACDFGHLMQRSTW
178	N		40.0%	4	6	10	40.0%	10	0		0.0%			GKMR	GKMR	EILVWY
179	H	3	93.3%	14	1	15	93.3%	15	0		0.0%	F		ACEGIKLMNPRSTV	ACEGIKLMNPRSTV	W
180	L		45.5%	5	6	11	40.0%	10	1	N	5.5%	G	I	FGKM	IFGKM	ACEPRS
181	W	2	23.1%	3	10	13	15.4%	13	0	N	7.7%		K	MQ	KMQ	ACDEFHILRS
182	G	2	7.1%	1	13	14	7.1%	14	0		0.0%	V		L	L	ACDEHNPQRSTVY
183	Y	2	15.4%	2	11	13	8.3%	12	1	N	7.1%		E	L	EL	CDGIKNPQRSV
184	Y	2	6.7%	1	14	15	0.0%	14	1	Y	6.7%		W		W	ACDEFGHKLMPRSV
185	L	1	0.0%	0	14	14	0.0%	14	0		0.0%	CE				ADEFGIKPRSTVWY
186	F	2	6.7%	1	14	15	6.7%	15	0		0.0%			Y	Y	ADGHIKLNPQRSVW
187	P	1	0.0%	0	15	15	0.0%	15	0		0.0%					AFGHILMNQRSTVWY
188	D	1	0.0%	0	15	15	0.0%	15	0		0.0%	C				ACFGHLMNPQRSTVW
189	C	1	0.0%	0	15	15	0.0%	15	0		0.0%					AEGHKLMNPRSTVWY
190	Y	1	0.0%	0	14	14	0.0%	14	0		0.0%					CEFGHKLNQRSTVW
191	N	1	0.0%	0	15	15	0.0%	15	0		0.0%					AEFGKLMPQRSTVWY
192	H	2	13.3%	2	13	15	13.3%	15	0		0.0%			ST	ST	CFGKLMNPQRVWY
193	H		46.2%	6	7	13	36.4%	11	2	N	9.8%		RF	GQSY	RFGQSY	ADKLMPV
194	Y	1	0.0%	0	8	8	0.0%	8	0		0.0%					ACILPSTV
195	K	4	100.0%	12	0	12	100.0%	10	2	N	0.0%		V	AGHILNQRSTW	VAGHILNQRSTW	
196	K	4	100.0%	9	0	9	100.0%	7	2	N	0.0%		NY	EGLRSTW	NYEGLRSTW	
197	P	3	93.3%	14	1	15	92.9%	14	1		0.5%		W	ADEFGHKLMQRST	WADEFGHKLMQRST	C
198	G	3	92.3%	12	1	13	91.7%	12	1	N	0.6%		W	ADEHLNQRSTY	WADEHLNQRSTY	V
199	Y	1	0.0%	0	10	10	0.0%	10	0		0.0%					EGHIKLPRSW
200	N	2	13.3%	2	13	15	7.1%	14	1		6.2%		T	D	TD	AFGHKLMPQRSWY

Position	WT	Assigned Color	% Active of Tested	Active Subs (Col R)	Inactive Subs (Col S)	Tested Subs	% Active of Tested (w/o errors, >40%)	Total Tested (w/o errors)	Total (difference)	Class Changed?	% Diff Active	Errors	Active subs (20%<x<40%)	Active subs (>40%)	Active Subs (P plus Q)	Inactive subs (<20%)
201	G	1	0.0%	0	11	11	0.0%	11	0		0.0%					AFLMNPRSTVW
202	S	2	7.1%	1	13	14	7.1%	14	0		0.0%			M	M	AEFGHKNPQRVWY
203	C	1	0.0%	0	13	13	0.0%	13	0		0.0%	L				ADEGHLMNQRSTV
204	F	2	15.4%	2	11	13	8.3%	12	1	N	7.1%		W	P	WP	ACEGHIKQRST
205	N	4	100.0%	7	0	7	100.0%	14	-7	N	0.0%		L	RSTVWY	LRSTVWY	
206	V	3	60.0%	9	6	15	57.1%	14	1		2.9%		Q	HIKLMRST	QHIKLMRST	CDFGPY
207	E	1	0.0%	0	11	11	0.0%	11	0		0.0%					AFGMPQRSTVW
208	I	3	71.4%	10	4	14	71.4%	14	0		0.0%	H		ACKLMQRSTV	ACKLMQRSTV	DGPW
209	K	3	81.8%	9	2	11	80.0%	10	1	N	1.8%	HK	L	AEFGNRST	LAEFGNRST	CP
210	R	1	0.0%	0	14	14	0.0%	14	0		0.0%	K				ACDEGKMNPSTVWY
211	N	2	14.3%	2	12	14	7.7%	13	1	N	6.6%			LW	LW	CFGHIKMPRSTV
212	D	2	23.1%	3	10	13	23.1%	13	0		0.0%	Y		NST	NST	AGHIKLMPVW
213	D	3	86.7%	13	2	15	86.7%	15	0		0.0%			AEGHKLMNQRVWY	AEGHKLMNQRVWY	PS
214	L	2	7.1%	1	13	14	7.1%	14	0		0.0%			Q	Q	ACDEGHKNPRSTY
215	S	3	86.7%	13	2	15	86.7%	15	0		0.0%			ADEGHKLMQRTVW	ADEGHKLMQRTVW	CP
216	W	2	6.7%	1	14	15	6.7%	15	0		0.0%			Y	Y	DEGHIKLMNPQRTV
217	L	2	9.1%	1	10	11	9.1%	11	0		0.0%			M	M	ACGHPQSTVW
218	W		33.3%	3	6	9	14.3%	7	2	Y	19.0%		MV	F	MVF	AIKLPS
219	N	3	92.9%	13	1	14	92.9%	14	0		0.0%			ACDEHIKLMRSTW	ACDEHIKLMRSTW	P
220	E	3	60.0%	9	6	15	53.8%	13	2	Y	6.2%		DM	AHILSTV	DMAHILSTV	GKNPRW
221	S		53.8%	7	6	13	50.0%	12	1	N	3.8%		I	ACMQTV	IACMQTV	DEHKPR
222	T	3	83.3%	10	2	12	83.3%	12	0		0.0%			DFGIKLNRSV	DFGIKLNRSV	PY
223	A	1	0.0%	0	15	15	0.0%	15	0		0.0%					CDEGHKLPQRSTVWY
224	L	2	7.1%	1	13	14	7.1%	14	0		0.0%			I	I	ADEFGMPQRSTWY
225	Y	1	0.0%	0	12	12	0.0%	12	0		0.0%					ADEGHKPQRTVW
226	P	2	6.7%	1	14	15	0.0%	14	1	Y	6.7%			W	W	ACDEFGLNQRSTVY
227	S	1	0.0%	0	15	15	0.0%	15	0		0.0%					AFGHIKLMPQRTVWY
228	I	1	0.0%	0	13	13	0.0%	13	0		0.0%					AEFGHLMNPRSTW
229	Y	1	0.0%	0	10	10	0.0%	10	0		0.0%					EFGKLPQTVW
230	L	2	6.7%	1	14	15	6.7%	15	0		0.0%			I	I	AEGHKMNPRSTVWY
231	N	2	6.7%	1	14	15	6.7%	15	0		0.0%			T	T	ACDFGHIKLPQRSV
232	T	2	16.7%	2	10	12	16.7%	12	0		0.0%			FS	FS	CGHKLNPQVY
233	Q		58.3%	7	5	12	58.3%	12	0		0.0%			AFGKLRY	AFGKLRY	DIPST
234	Q	2	15.4%	2	11	13	15.4%	13	0		0.0%			LM	LM	ADEGHNPSTVW
235	S		50.0%	6	6	12	50.0%	12	0		0.0%			AEGHKT	AEGHKT	FLMRWY
236	P		46.2%	6	7	13	46.2%	13	0		0.0%			AGHKRS	AGHKRS	CILNQTY
237	V	4	100.0%	12	0	12	100.0%	11	1	N	0.0%		C	AEFHLNQRSTW	CAEFHLNQRSTW	
238	A		53.3%	8	7	15	53.3%	15	0		0.0%			DEHKQRST	DEHKQRST	FGLPVWY
239	A	2	7.1%	1	13	14	0.0%	13	1	Y	7.1%		N		N	CFGHILPRSTVWY
240	T		57.1%	8	6	14	57.1%	14	0		0.0%			KAMPQRSV	KAMPQRSV	EFGNWY
241	L	1	0.0%	0	12	12	0.0%	12	0		0.0%					ACDEGIPRSTVW
242	Y	2	7.1%	1	13	14	7.1%	14	0		0.0%			F	F	ACDGILMPRSTVW
243	V	1	0.0%	0	13	13	0.0%	13	0		0.0%					CDFGHLMPQRSWY
244	R	1	0.0%	0	6	6	0.0%	6	0		0.0%					ADGIVY
245	N	2	9.1%	1	10	11	9.1%	11	0		0.0%			H	H	ACFLPQRSTV
246	R	1	0.0%	0	15	15	0.0%	15	0		0.0%					ACDEGHIKLMPSTVW
247	V	2	20.0%	3	12	15	20.0%	15	0		0.0%			ILM	ILM	ACFHNPQRSTWY
248	R		33.3%	4	8	12	33.3%	12	0		0.0%			AHWY	AHWY	CDEGIMPT
249	E	2	10.0%	1	9	10	10.0%	10	0	Y	0.0%			V	V	AGHIKMQSY
250	A	1	0.0%	0	15	15	0.0%	15	0		0.0%					CFGHKLMNPQRSTVW
251	I	2	25.0%	3	9	12	18.2%	11	1	N	6.8%		Y	LM	YLM	DFGHKPSTW
252	R	1	0.0%	0	14	14	0.0%	14	0		0.0%					ADEFGHIKLNPSTY
253	V	2	7.7%	1	12	13	7.7%	13	0		0.0%			I	I	ADEGHLMNQRSW
254	S	1	0.0%	0	14	14	0.0%	14	0		0.0%					CDEGIKLPQRTVWY
255	K		50.0%	6	6	12	45.5%	11	1	N	4.5%		G	ANQRS	GANQRS	CDLPVW
256	I		44.4%	4	5	9	44.4%	9	0		0.0%			AHLV	AHLV	CDEGP
257	P	3	92.3%	12	1	13	91.7%	12	1	N	0.6%		C	AGIKLMNQRTV	CAGIKLMNQRTV	D
258	D		55.6%	5	4	9	50.0%	8	1	N	5.6%		G	HNRS	GHNRS	LPVW
259	A	4	100.0%	14	0	14	100.0%	14	0		0.0%			EGIKLNPQRSTVWY	EGIKLNPQRSTVWY	
260	K	3	84.6%	11	2	13	83.3%	12	1	N	1.3%		G	ADEHLMQRSY	GADEHLMQRSY	CP
261	S	3	90.9%	10	1	11	90.9%	11	0		0.0%			AFKMNQRTVW	AFKMNQRTVW	P
262	P	1	0.0%	0	14	14	0.0%	15	-1	N	0.0%					ADEGHIKQRSTVWY
263	L		58.3%	7	5	12	54.5%	11	1		3.8%		H	AKMRTV	HAKMRTV	EFPQW
264	P	2	15.4%	2	11	13	15.4%	13	0		0.0%			AH	AH	DEFGLMRTVWY
265	V	2	7.7%	1	12	13	7.7%	13	0		0.0%			I	I	ADFGHKLMNQRS
266	F	2	7.7%	1	12	13	7.7%	13	0		0.0%			Y	Y	ACGHMPQRSTVW
267	A	2	18.2%	2	9	11	18.2%	11	0		0.0%			MT	MT	DGHIKNRSW
268	Y	1	0.0%	0	14	14	0.0%	14	0		0.0%					ACFGHKLNPQSTVW
269	T		33.3%	4	8	12	33.3%	12	0		0.0%			ACDS	ACDS	EKLMNPQR

Position	WT	Assigned Color	% Active of Tested	Active Subs (Col R)	Inactive Subs (Col S)	Tested Subs	% Active of Tested (w/o errors, >40%)	Total Tested (w/o errors)	Total (difference)	Class Changed?	% Diff Active	Errors	Active subs (20%<x<40%)	Active subs (>40%)	Active Subs (P plus Q)	Inactive subs (<20%)
270	R		30.8%	4	9	13	30.8%	13	0		0.0%			MNST	MNST	ACEFGHIPY
271	I		46.2%	6	7	13	46.2%	13	0		0.0%			FGLMSV	FGLMSV	ADEHKTW
272	V		54.5%	6	5	11	40.0%	10	1	N	14.5%		EM	DRST	EMDRST	HLNPW
273	F	2	21.4%	3	11	14	21.4%	14	0		0.0%			HTY	HTY	ACDGILPQSVW
274	T	2	27.3%	3	8	11	27.3%	11	0		0.0%			AFS	AFS	CEGHNQWY
275	D	2	18.2%	2	9	11	0.0%	9	2	Y	18.2%		L	V	LV	AFGIKMQTW
276	Q	3	78.6%	11	3	14	76.9%	13	1	N	1.6%		G	CDEHILMRSY	GCDEHILMRSY	FPW
277	V	4	100.0%	14	0	14	100.0%	14	0		0.0%			ACDEGHKMNQRSTY	ACDEGHKMNQRSTY	
278	L	3	86.7%	13	2	15	86.7%	15	0		0.0%			AEFGHIKNRSTVY	AEFGHIKNRSTVY	MP
279	K		45.5%	5	6	11	40.0%	10	1	N	5.5%		A	HQRT	AHQRT	CFGLWY
280	F	2	18.2%	2	9	11	16.7%	12	-1	N	1.5%			GQ	GQ	DIMNRSTVW
281	L	1	0.0%	0	13	13	0.0%	13	0		0.0%					ADGHIKNPQRSVW
282	S		44.4%	4	5	9	44.4%	9	0		0.0%			DGMQ	DGMQ	FLVWY
283	Q		50.0%	5	5	10	50.0%	10	0		0.0%			EPRST	EPRST	ACDFW
284	D	3	78.6%	11	3	14	75.0%	12	2	N	3.6%		TY	AEGHLMNQS	TYAEGHLMNQS	CIP
285	E	3	61.5%	8	5	13	58.3%	12	1	Y	3.2%		A	FGHMNQY	AFGHMNQY	KPRTV
286	L	2	23.1%	3	10	13	16.7%	12	1	N	6.4%		W	RS	WRS	ACDFHKMPTY
287	V	2	21.4%	3	11	14	15.4%	13	1	N	6.0%		N	IT	NIT	ACDEGKLPQRS
288	Y	2	16.7%	2	10	12	16.7%	12	0		0.0%			LW	LW	DEFGHIKPRT
289	T	2	16.7%	2	10	12	9.1%	11	1	N	7.6%			KS	KS	ACEGHLPQRY
290	F		40.0%	2	3	5	40.0%	5	0		0.0%			IM	IM	DQY
291	G		35.7%	5	9	14	30.8%	13	1	N	4.9%		C	QRSV	CQRSV	ADEFMNTWY
292	E	3	78.6%	11	3	14	78.6%	14	0		0.0%			ACFGHKNPRVW	ACFGHKNPRVW	ILT
293	T	3	86.7%	13	2	15	86.7%	15	0		0.0%			ACDFGKLMPQSVY	ACDFGKLMPQSVY	EN
294	V	2	7.1%	1	13	14	7.1%	14	0		0.0%			M	M	AEGHKLNPQRSTW
295	A	1	0.0%	0	10	10	0.0%	10	0		0.0%					CGHILNPTVY
296	L	1	0.0%	0	13	13	0.0%	13	0		0.0%					CFGIKMQRSTVWY
297	G	2	16.7%	2	10	12	9.1%	11	1	N	7.6%		R	A	RA	CEHLNPQSTY
298	A	2	15.4%	2	11	13	15.4%	13	0		0.0%			GI	GI	CELMNPQSTWY
299	S	1	0.0%	0	11	11	0.0%	11	0		0.0%					ACDFGHLMPQT
300	G	2	6.7%	1	14	15	6.7%	15	0		0.0%			R	R	ACDEFLMNPQSTVW
301	I	2	14.3%	2	12	14	14.3%	14	0		0.0%			AV	AV	EGHKMNPQRSWY
302	V	2	13.3%	2	13	15	13.3%	15	0		0.0%			IW	IW	CDEFGHLMPRSTY
303	I	2	15.4%	2	11	13	8.3%	12	1	N	7.1%		D	V	DV	ACEFGKLMRWY
304	W	2	15.4%	2	11	13	0.0%	11	2	Y	15.4%			GI	GI	ACDMNPQSTVY
305	G	2	27.3%	3	8	11	20.0%	10	1	N	7.3%		N	DE	NDE	LPQRSTVY
306	T	2	27.3%	3	8	11	27.3%	11	0		0.0%			DES	DES	ACHILVWY
307	L	3	75.0%	9	3	12	72.7%	11	1	N	2.3%		G	KNQSTVWY	GKNQSTVWY	CIP
308	S		53.3%	8	7	15	53.3%	15	0		0.0%			DGHKNPRT	DGHKNPRT	CFLMVWY
309	I	4	100.0%	14	0	14	100.0%	14	0		0.0%			DEGHKLMNQRSTVW	DEGHKLMNQRSTVW	
310	M	3	66.7%	8	4	12	60.0%	10	2	N	6.7%		FY	AGQRSV	FYAGQRSV	CEKL
311	R		42.9%	6	8	14	42.9%	14	0	N	0.0%			GHKQST	GHKQST	CEFILPVW
312	S		50.0%	5	5	10	37.5%	8	2	Y	12.5%		KL	GNT	KLGNT	CEMVW
313	M	3	92.3%	12	1	13	92.3%	13	0		0.0%			AEGHKLPRSTVY	AEGHKLPRSTVY	C
314	K	3	76.9%	10	3	13	76.9%	13	0		0.0%			ADHINQRSTY	ADHINQRSTY	CLW
315	S	3	76.9%	10	3	13	76.9%	13	0		0.0%			AEGHKLMRTY	AEGHKLMRTY	CIV
316	C	2	7.1%	1	13	14	7.1%	14	0		0.0%			D	D	EGIKLMPRSTVWY
317	L	3	85.7%	12	2	14	85.7%	14	0		0.0%			ADHIKMNQRSTW	ADHIKMNQRSTW	GP
318	L	3	80.0%	12	3	15	80.0%	15	0		0.0%			DFGHIKMNQRST	DFGHIKMNQRST	CPW
319	L	1	0.0%	0	15	15	0.0%	15	0		0.0%					CEFGHIKMPQRSVWY
320	D	3	86.7%	13	2	15	84.6%	13	2	N	2.1%		LV	EGHIKMNRSWY	LVEGHIKMNRSWY	CP
321	N	3	72.7%	8	3	11	72.7%	11	0		0.0%			ADHKRSTY	ADHKRSTY	EMP
322	Y	1	0.0%	0	13	13	0.0%	13	0		0.0%					CDEGILNPRSTVW
323	M	2	21.4%	3	11	14	21.4%	14	0		0.0%			FIL	FIL	ACEGHKNRSTV
324	E		53.8%	7	6	13	53.8%	13	0		0.0%			ADHMNRS	ADHMNRS	CFPVWY
325	T	3	85.7%	12	2	14	85.7%	14	0		0.0%			ADEGHKMNQSVW	ADEGHKMNQSVW	CR
326	I		54.5%	6	5	11	50.0%	10	1	N	4.5%		C	KLSVY	CKLSVY	EGHNW
327	L	2	7.1%	1	13	14	7.1%	14	0		0.0%			M	M	AEFGHNQRSTVWY
328	N	4	100.0%	14	0	14	100.0%	13	1	N	0.0%		W	ACGHIKLQRSTVY	WACGHIKLQRSTVY	
329	P	1	0.0%	0	15	15	0.0%	15	0		0.0%					CFGHIKLNQRSTVWY
330	Y	1	0.0%	0	14	14	0.0%	14	0		0.0%					ACDEGILMNPRSVW
331	I	2	20.0%	3	12	15	8.3%	12	3	N	11.7%		CE	V	CEV	ADEFHKQRSTWY
332	I	1	0.0%	0	15	15	0.0%	15	0		0.0%					ACDEFGHKLNPRSTY
333	N	1	0.0%	0	10	10	0.0%	10	0		0.0%					GHIKPRSTWY
334	V	2	18.2%	2	9	11	10.0%	10	1	N	8.2%		T	P	TP	ACDEGMNRS
335	T	2	9.1%	1	10	11	9.1%	11	0		0.0%			S	S	FGHIKLPVWY
336	L	1	0.0%	0	13	13	0.0%	13	0		0.0%					AEFGKNPRSTVWY
337	A	1	0.0%	0	10	10	0.0%	10	0		0.0%					CFGIKLMRTW
338	A	2	7.1%	1	13	14	7.1%	14	0		0.0%			Q	Q	CDEFGHIKLPRTV

Position	WT	Assigned Color	% Active of Tested	Active Subs (Col R)	Inactive Subs (Col S)	Tested Subs	% Active of Tested (w/o errors, >40%)	Total Tested (w/o errors)	Total (difference)	Class Changed?	% Diff Active	Errors	Active subs (20%<x<40%)	Active subs (>40%)	Active Subs (P plus Q)	Inactive subs (<20%)
339	K	2	7.1%	1	13	14	7.1%	14	0		0.0%			M	M	DEFGHLNPSTVWY
340	M	1	0.0%	0	14	14	0.0%	14	0		0.0%					ACDEFGHKPRSTVW
341	C	1	0.0%	0	14	14	0.0%	14	0		0.0%					AEGHKLMNQRSTVY
342	S	2	7.7%	1	12	13	7.7%	13	0		0.0%			A	A	DEFHKLMPQRTY
343	Q	2	25.0%	2	6	8	25.0%	8	0		0.0%			TV	TV	CDFIPW
344	V	1	0.0%	0	13	13	0.0%	13	0		0.0%					FGHLMNPQRSTWY
345	L	1	0.0%	0	11	11	0.0%	11	0		0.0%					ACEHKNQRTVY
346	C	1	0.0%	0	14	14	0.0%	14	0		0.0%					ADFGIKLMPRSTVW
347	Q		50.0%	7	7	14	46.2%	13	1	N	3.8%		L	AEGMRS	LAEGMRS	CFIPTVW
348	E	2	21.4%	3	11	14	21.4%	14	0		0.0%			DGS	DGS	CHILPQRTVWY
349	Q		50.0%	7	7	14	50.0%	14	0		0.0%			AEKMNRT	AEKMNRT	DFGPVWY
350	G	1	0.0%	0	15	15	0.0%	15	0		0.0%					ADEFHKLMNPRSTVY
351	V		38.5%	5	8	13	20.0%	10	3	Y	18.5%		CIQ	AS	CIQAS	DEFHNRWY
352	C	1	0.0%	0	15	15	0.0%	15	0		0.0%					ADEFGKMPQRSTVWY
353	I	2	15.4%	2	11	13	15.4%	13	0		0.0%			TV	TV	CFGHKLMQRSW
354	R	1	0.0%	0	15	15	0.0%	15	0		0.0%					CDEGHIKLMPQSVWY
355	K	1	0.0%	0	15	15	0.0%	15	0		0.0%					DFGHLMNPQRSTVWY
356	N		30.8%	4	9	13	30.8%	13	0		0.0%			ADHS	ADHS	CGKLPRTVW
357	W		38.5%	5	8	13	33.3%	12	1	N	5.1%		K	ACST	KACST	DEFGLMQR
358	N		33.3%	4	8	12	27.3%	11	1	Y	6.1%		L	CGT	LCGT	EHIKPQRW
359	S		53.8%	7	6	13	53.8%	13	0		0.0%			DEHKMTV	DEHKMTV	AFGLPW
360	S	2	7.1%	1	13	14	7.1%	14	0		0.0%			T	T	ACEFGIKLMPQRV
361	D	2	7.7%	1	12	13	0.0%	12	1	Y	7.7%		H		H	ACEGMNPQRSVW
362	Y	1	0.0%	0	15	15	0.0%	15	0		0.0%					ACEGHKLMNPRSTVW
363	L	1	0.0%	0	15	15	0.0%	15	0		0.0%					ACDEFGHIPQRSTVW
364	H	1	0.0%	0	15	15	0.0%	15	0		0.0%					ACDEFGKLMPRSTVY
365	L	1	0.0%	0	14	14	0.0%	14	0		0.0%					ACDEGMNPQRSTWY
366	N	1	0.0%	0	12	12	0.0%	12	0		0.0%					ACEFGKMPQRTW
367	P		46.2%	6	7	13	46.2%	13	0		0.0%			ACGKRS	ACGKRS	EFILMQV
368	D	3	78.6%	11	3	14	78.6%	14	0		0.0%			AEGHKLMRSTV	AEGHKLMRSTV	CPW
369	N	2	23.1%	3	10	13	23.1%	13	0		0.0%			HRS	HRS	CEFIKLPQVW
370	F	1	0.0%	0	14	14	0.0%	14	0		0.0%					ADEGHKLNPQRSVY
371	A	3	91.7%	11	1	12	83.3%	12	0	N	8.3%		G	EFHIKLMRSV	GEFHIKLMRSV	P
372	I	1	0.0%	0	15	15	0.0%	15	0		0.0%					ADEFGHKLNPRSTVW
373	Q	3	75.0%	9	3	12	75.0%	12	0		0.0%			AEFKLMRSV	AEFKLMRSV	CPW
374	L	3	85.7%	12	2	14	84.6%	13	1	N	1.1%		W	AHIMNPRSTVY	WAHIMNPRSTVY	DE
375	E	3	66.7%	10	5	15	64.3%	14	1	N	2.4%		I	AGKLMNRST	IAGKLMNRST	CFPVY
376	K	3	78.6%	11	3	14	76.9%	13	1	N	1.6%		L	ADEMQRSTVY	LADEMQRSTVY	IPW
377	G	3	66.7%	8	4	12	66.7%	12	0		0.0%			DEHKPRST	DEHKPRST	CILV
378	G	2	23.1%	3	10	13	23.1%	13	0		0.0%			KNR	KNR	DEFILMQTWY
379	K		38.5%	5	8	13	38.5%	13	0		0.0%			GHRST	GHRST	ACEFILMW
380	F		50.0%	7	7	14	41.7%	12	2	N	8.3%		VT	ILPWY	VTILPWY	CDEGQRS
381	T	3	61.5%	8	5	13	58.3%	12	1	Y	3.2%		E	HKNQRSV	EHKNQRSV	GLPWY
382	V	1	0.0%	0	14	14	0.0%	14	0		0.0%					EGHKLMNPQRSTWY
383	R	3	84.6%	11	2	13	84.6%	13	0		0.0%			AEHIKLMNSTV	AEHIKLMNSTV	GP
384	G	1	0.0%	0	6	6	0.0%	6	0		0.0%					CFMQST
385	K	3	60.0%	9	6	15	60.0%	15	0		0.0%			AGHNQRSTV	AGHNQRSTV	CLMPWY
386	P	1	0.0%	0	15	15	0.0%	15	0		0.0%					ACFGHILMNQRSTVY
387	T	2	7.7%	1	12	13	7.7%	13	0		0.0%			S	S	CEFGHILMNVWY
388	L	3	69.2%	9	4	13	64.3%	14	-1	N	4.9%			FHIMRTVWY	FHIMRTVWY	CGPQ
389	E	3	85.7%	12	2	14	85.7%	14	0		0.0%			AGHKLMPQRSTY	AGHKLMPQRSTY	FV
390	D	1	0.0%	0	15	15	0.0%	15	0		0.0%					ACEFGHLNPRSTVWY
391	L	2	6.7%	1	14	15	6.7%	15	0		0.0%			C	C	ADGHKNPQRSTVWY
392	E	3	86.7%	13	2	15	85.7%	14	1	N	1.0%		W	AFGKLMQRSTVY	WAFGKLMQRSTVY	CP
393	Q	3	84.6%	11	2	13	84.6%	13	0		0.0%			ADFHKLMNRST	ADFHKLMNRST	CP
394	F	2	13.3%	2	13	15	13.3%	15	0		0.0%			LW	LW	ADEGIKNPQRSTV
395	S	3	63.6%	7	4	11	60.0%	10	1	N	3.6%		T	AGHKRW	TAGHKRW	CLMP
396	E		57.1%	8	6	14	53.8%	13	1	N	3.3%		L	ADHQRST	LADHQRST	CFGIPY
397	K	2	7.7%	1	12	13	0.0%	12	1	Y	7.7%			R	R	ACEFGILMPQTV
398	F	2	6.7%	1	14	15	0.0%	14	1	Y	6.7%		L		L	ACEGHINPRSTVWY
399	Y	3	85.7%	12	2	14	85.7%	14	0		0.0%			ACEKMNQRSTVW	ACEKMNQRSTVW	DP
400	C	1	0.0%	0	15	15	0.0%	15	0		0.0%					ADEFGILMPQRSTVY
401	S		41.7%	5	7	12	30.0%	10	2	Y	11.7%		GQ	AEN	GQAEN	CFHKRWY
402	C	1	0.0%	0	14	14	0.0%	14	0		0.0%					ADEFLMPQRSTVWY
403	Y	2	7.1%	1	13	14	7.1%	14	0		0.0%			F	F	ACEGHKLMNPQRT
404	S	2	20.0%	3	12	15	14.3%	14	1	N	5.7%		T	AP	TAP	CDFGHLMNRVWY
405	T	3	78.6%	11	3	14	76.9%	13	1	N	1.6%		F	AGKMPQRSWY	FAGKMPQRSWY	CIV
406	L	3	85.7%	12	2	14	85.7%	14	0		0.0%			ACEFGINQSTVY	ACEFGINQSTVY	PR
407	S	4	100.0%	14	0	14	100.0%	14	0		0.0%			ADEFGHLMNPQRVW	ADEFGHLMNPQRVW	

Position	WT	Assigned Color	% Active of Tested	Active Subs (Col R)	Inactive Subs (Col S)	Tested Subs	% Active of Tested (w/o errors, >40%)	Total Tested (w/o errors)	Total (difference)	Class Changed?	% Diff Active	Errors	Active subs (20%<x<40%)	Active subs (>40%)	Active Subs (P plus Q)	Inactive subs (<20%)
408	C	1	0.0%	0	14	14	0.0%	14	0		0.0%					AEFGIKLPRSTVWY
409	K	4	100.0%	12	0	12	100.0%	12	0		0.0%			ADEGHIPQRSTV	ADEGHIPQRSTV	
410	E	3	91.7%	11	1	12	91.7%	12	0		0.0%			DKMNPQRSTVY	DKMNPQRSTVY	W
411	K	3	66.7%	8	4	12	63.6%	11	1	N	3.0%		H	ANPRSTV	HANPRSTV	DEFG
412	A	3	92.9%	13	1	14	91.7%	12	2	N	1.2%		HQ	DGILNPRSVWY	HQDGILNPRSVWY	E
413	D	3	75.0%	9	3	12	57.1%	7	5	Y	17.9%		AHQS	EKNRT	AHQSEKNRT	ILP
414	V		33.3%	4	8	12	20.0%	10	2	Y	13.3%		KL	IM	KLIM	ADEGHRST
415	K		55.6%	5	4	9	42.9%	7	2		12.7%		YV	GSW	YVGSW	CDEP
416	D	3	85.7%	12	2	14	85.7%	14	0		0.0%			FGHIKLNQRTVY	FGHIKLNQRTVY	CS
417	T	2	8.3%	1	11	12	8.3%	12	0		0.0%			I	I	ADEFGHKMPQR
418	D	4	100.0%	14	0	14	100.0%	14	0		0.0%			AEFGILMNPQRSVY	AEFGILMNPQRSVY	
419	A	3	85.7%	12	2	14	85.7%	14	0		0.0%			EFGHIKLNRSWY	EFGHIKLNRSWY	DP
420	V	2	13.3%	2	13	15	13.3%	15	0		0.0%			IP	IP	ADFGHKLNRSTWY
421	D	4	100.0%	14	0	14	100.0%	13	1	N	0.0%		I	AEGHKLMNQRSTY	IAEGHKLMNQRSTY	
422	V	2	15.4%	2	11	13	15.4%	13	0		0.0%			IT	IT	CDGHLMNQRSY
423	C	1	0.0%	0	15	15	0.0%	15	0		0.0%					ADEFGHLMPQRSTVW
424	I	1	0.0%	0	11	11	0.0%	11	0		0.0%					ACEGHNQRSWY
425	A	3	66.7%	8	4	12	63.6%	11	1	N	3.0%		Y	GIKMNRS	YGIKMNRS	ELPW
426	D	3	66.7%	8	4	12	60.0%	10	2	N	6.7%		KS	EGNPQY	KSEGNPQY	CFMR
427	G		42.9%	6	8	14	27.3%	11	3	Y	15.6%		HTQ	IKS	HTQIKS	ACFLPVWY
428	V	2	28.6%	4	10	14	28.6%	14	0		0.0%			LMPT	LMPT	ACDEGHNRSY
429	C	1	0.0%	0	11	11	0.0%	11	0		0.0%					ADKLNPSTVWY
430	I	1	0.0%	0	9	9	0.0%	9	0		0.0%					ADELMNSTV
431	D	3	93.3%	14	1	15	93.3%	15	0		0.0%			AEGHIKLNQRSVWY	AEGHIKLNQRSVWY	P
432	A		42.9%	6	8	14	38.5%	13	1	N	4.4%		H	EGNSV	HEGNSV	CFIKLMPY
433	F	4	100.0%	15	0	15	100.0%	15	0		0.0%			ACDEGHIKLPRSTVW	ACDEGHIKLPRSTVW	
434	L		45.5%	5	6	11	45.5%	11	0		0.0%			FGIMV	FGIMV	HKPQRW
435	K	4	100.0%	10	0	10	100.0%	10	0		0.0%			ACEGHRSTVY	ACEGHRSTVY	
436	P	4	100.0%	15	0	15	100.0%	14	1	N	0.0%		C	DEGHIKLMQRSTWY	CDEGHIKLMQRSTWY	
437	P	3	92.3%	12	1	13	92.3%	13	0		0.0%			ADGHIKLMQRSY	ADGHIKLMQRSY	T
438	M	3	93.3%	14	1	15	93.3%	15	0		0.0%			ACDEGLNPQRSTVW	ACDEGLNPQRSTVW	Y
439	E	3	86.7%	13	2	15	86.7%	15	0		0.0%			ACFGHKLPQSTVW	ACFGHKLPQSTVW	NR
440	T	3	93.3%	14	1	15	93.3%	15	0		0.0%			ADEFGHILMPRSVY	ADEFGHILMPRSVY	Q
441	E	3	92.9%	13	1	14	92.9%	14	0		0.0%			ADFGHKLNQSTVY	ADFGHKLNQSTVY	R
442	E	3	80.0%	12	3	15	78.6%	14	1	N	1.4%		W	CGHKLPQRTVY	WCGHKLPQRTVY	MNS
443	P	3	92.9%	13	1	14	92.9%	14	0		0.0%			AEFGHLMNQRSTW	AEFGHLMNQRSTW	D
444	Q	4	100.0%	13	0	13	100.0%	12	1	N	0.0%		M	DEFGHIKNRVWY	MDEFGHIKNRVWY	
445	I	4	100.0%	14	0	14	100.0%	14	0		0.0%			AGHLMNPQRSTVWY	AGHLMNPQRSTVWY	
446	F	4	100.0%	15	0	15	100.0%	15	0		0.0%			ACDEGHIKLMQRTVW	ACDEGHIKLMQRTVW	
447	Y	4	100.0%	14	0	14	100.0%	14	0		0.0%			DEFGILMNPQRTVW	DEFGILMNPQRTVW	

Pos	Wild Type (WT)	SASA (Å)	fSASA	Active Mutants	Mutants 20-40% Activity	Active Mutants (>40%, No errors)	Mutants 20-40% Activity (no errors)	Inactive Mutants	Inactive Mutants (no errors)	Unclassified	Not Made	Errors2	Mutants w/ 20-40% Activity + Errors
1	L			ACEFGHKNPQRSTVW		ACEFGHKNPQRSTVW				D	IMY		
2	N			IACGLPQSTV	I	ACGLPQSTV	I	HKWY	HKWY	F	DEMR		I
3	F			EHLY		EHLY		AGKPTV	AGKPTV	IMNRS	CDQW		
4	R	139.073	0.966	AISTV		AISTV		DEFGLPWY	DEFGLPWY	MN	CHKQ		
5	A	48.781	0.439	H		H		DGILMNPQRTVWY	DGILMNPQRTVWY	S	CEFK		
6	P	65.732	0.487	AHKLNQR		AHKLNQR		EFTVY	EFTVY	DGSW	CIM		
7	P	38.050	0.282	M		M		CDFGHIKLQRSTWY	CDFGHIKLQRSTWY	AV	EN		
8	V	33.672	0.226	PILM	P	ILM	P	DEGHNRSW	DEGHNRSW	AQT	CFKY		P
9	I	52.912	0.306	KLQRSV		KLQRSV		CDEGNP	CDEGNP	AHT	FMWY		
10	P	103.226	0.765	DEGHNQRSW		DEGHNQRSW		FILMY	FILMY	T	ACKV		
11	N	90.515	0.545	SDGHK	S	DGHK	S	ACFILPTWY	ACFILPTWY	E	MQRV		S
12	V	16.843	0.113	AEIKLNRST		AEIKLNRST		GHW	GHW	DMP	CFQY		
13	P	8.983	0.067	HSTY		HSTY		EGILMV	EGILMV	AFQRW	CDKN		
14	F	0.000	0.000	DIMV		DIMV		AEGHKNPQW	AEGHKNPQW	RT	CLSY		
15	L	8.367	0.047	AMV		AMV		EFGKNPQRSY	EFGKNPQRSY	TW	CDHI		
16	W	2.350	0.009					ACDEFGHKMPRSTY	ACDEFGHKMPRSTY	L	INQV		
17	A	7.617	0.069					DEGHILNPQRSTVWY	DEGHILNPQRSTVWY		CFKM		
18	W	1.440	0.006					CDFGHILMPQSTVY	CDFGHILMPQSTVY	R	AEKN		
19	N	10.312	0.062					ACFGHILMPQRSVWY	ACFGHILMPQRSVWY		DEKT		
20	A	0.706	0.006	S		S		DEFHKLNPRTVY	DEFHKLNPRTVY	GQ	CIMW		
21	P	3.031	0.022					ACDEGHILMRSTVW	ACDEGHILMRSTVW	K	FNQY		
22	S	6.223	0.050	HMTY		HMTY		CEGKP	CEGKP	ADLNRV	FIQW		
23	E	83.710	0.448	D		D		AFLMNPRSTV	AFLMNPRSTV	GHQW	CIKY		
24	F	63.014	0.303	AEGHIKLMNRTVY		AEGHIKLMNRTVY				CP	DQSW		
25	C	1.435	0.009					DEFGHIKLNPRSTVY	DEFGHIKLNPRSTVY		AMQW		
26	L	96.728	0.540	AEGHIKMPQRSTVWY		AEGHIKMPQRSTVWY					CDFN		
27	G	76.159	0.886	ADEFHIKLPQRSTW		ADEFHIKLPQRSTW		C	C		MNYV		
28	K	124.901	0.589	ADEFILMNPRSTVW		ADEFILMNPRSTVW				G	CHQY		
29	F	48.655	0.234	AEGHIKLMPRSTVW		AEGHIKLMPRSTVW				C	DNQY		
30	D	130.140	0.813	AFGHKLMPQRSTVW		AFGHKLMPQRSTVW				E	CINY		
31	E	25.495	0.136	ACGHIKLPRSTVWY		ACGHIKLPRSTVWY				N	DFMQ		
32	P	84.022	0.622	ACFGHKLMNQRSTVWY		ACFGHKLMNQRSTVWY					DEI		
33	L	16.981	0.095	GMPQRSTW		GMPQRSTW		CDHNVY	CDHNVY	I	AEFK		
34	D	44.945	0.281	AEHKQRW	A	EHKQRW	A	ILNSTV	ILNSTV	GP	CFMY		A
35	M	68.964	0.343	QVFHLTY	QV	FHLTY	QV	ADGPRS	ADGPRS	IN	CEKW		QV
36	S	65.053	0.520	DNAGHKLRT	DN	AGHKLRT	DN	CFVWY	CFVWY	P	EIMQ		DN
37	L	36.039	0.201	WFIKMPRV	W	FIKMPRV	W	CEGNS	CEGNS	AT	DHQY		W
38	F	14.561	0.070	Y		Y		EGKLNQRTW	EGKLNQRTW	ACMPS	DHIV		
39	S	51.573	0.413	ALNQRTY		ALNQRTY		CDFW	CDFW	GMPV	EHIK		
40	F	43.210	0.208	LW		LW		ADEGKNRSTV	ADEGKNRSTV	IQY	CHMP		
41	I	60.110	0.347	ACDEGHNTVW		ACDEGHNTVW		Q	Q	FPRS	KLMY		
42	G	10.926	0.127	A		A		DEHIKLMPQRSTV	DEHIKLMPQRSTV	C	FNWY		
43	S	13.844	0.111	NT	N	T	N	AEFGIKLQRV	AEFGIKLQRV	DHP	CMWY		N
44	P	20.476	0.152	E		E		ACFGHILNQRSTWY	ACFGHILNQRSTWY		DKMV		
45	R	113.880	0.456	IK		IK		ADFGPW	ADFGPW	HMNQSTVY	CEL		
46	I	60.998	0.353	CAEFHLMNRSTVY	C	AEFHLMNRSTVY	C	PW	PW		DGKQ		C
47	N	90.218	0.543	ADFGHKMQRSTWY		ADFGHKMQRSTWY		V	V	ILP	CE		
48	A	83.814	0.755	FGHIKMNQRSVY		FGHIKMNQRSVY		P	P	CELW	DT		
49	T	92.499	0.642	IKRSV		IKRSV		CDGHP	CDGHP	AFLNW	EMQY		
50	G	33.490	0.389	VACDEHLMQRSY	V	ACDEHLMQRSY	V			FPW	IKNT		V
51	Q	109.269	0.563	ARNS	AR	NS	AR	CFIMPTWY	CFIMPTWY	DHK	EGLV		AR
52	G	8.324	0.097	NPQRST		NPQRST		CEFWY	CEFWY	AHKL	DIMV		
53	V	36.816	0.247					ACDEGHLNPQRSTWY	ACDEGHLNPQRSTWY		FIKM		
54	T	12.982	0.090	AFNQSV		AFNQSV		DEGPRY	DEGPRY	HIM	CKLW		
55	I	10.222	0.059					ADGHNPQRTVY	ADGHNPQRTVY	CFLS	EKMW		
56	F	0.000	0.000					ACEGHIKLPRSTVW	ACEGHIKLPRSTVW	N	DMQY		
57	Y	64.770	0.285					ADFGILMPQRVW	ADFGILMPQRVW	EST	CHKN		
58	V	34.213	0.230	CGHIKLNPQRSWY		CGHIKLNPQRSWY		A	A	D	EFMT		
59	D	112.167	0.701	QN		QN		AEILMPRTVWY	AEILMPRTVWY	GH	CFKS		
60	R	93.608	0.374	K		K		ADFGHILNPQSTVY	ADFGHILNPQSTVY		CEMW		
61	L	0.000	0.000	FIMV	F	IMV		AEGHNPQRTWYF	AEGHNPQRTWY		CDKS	F	F
62	G	9.936	0.116					ACDFIKLMPQRSTVY	ACDFIKLMPQRSTVY		EHNW		

Pos	Wild Type (WT)	SASA (Å)	fSASA	Active Mutants	Mutants 20-40% Activity	Active Mutants (>40%, No errors)	Mutants 20-40% Activity (no errors)	Inactive Mutants	Inactive Mutants (no errors)	Unclassified	Not Made	Errors2	Mutants w/ 20-40% Activity + Errors
63	Y	92.692	0.408	AHIKLMNRSTVW		AHIKLMNRSTVW		CGP	CGP		DEFQ		
64	Y	4.748	0.021					ACDEFGHIKLPQRSTVW	ACDEFGHIKLPQRSTVW		MN		
65	P	0.720	0.005	R				ACDGHIKNSTVWYR	ACDGHIKNSTVWY	F	ELMQ	R	R
66	Y	51.764	0.228	HR		HR		ACDEGIKLNPSTV	ACDEGIKLNPSTV		FMQW		
67	I	8.003	0.046	RFLVY	R	FLVY		DEGPTWR	DEGPTW	CHNQ	AKMS	R	R
68	D	54.181	0.339	GEHKLPQRST	G	EHKQRST		ACIVYGLP	ACIVY		FMNW	GLP	GLP
69	S	111.432	0.891	ACEFGILMPRTWY		ACEFGILMPRWY		NT	N	V	DHKQ	T	T
70	I	144.199	0.834	ACFGHKLNPRSTVY		ACFGHKLNPRSTVY		Q	Q		DEMW		
71	T	88.846	0.617	ADGHLMNQRS		ADGHLMNQRS		P	P	CEVY	FIKW		
72	G	20.860	0.243	YADEHKLMQRS	Y	ADEKLMQRS	Y	CFIPVWH	CFIPVW	T	N	H	YH
73	V	79.688	0.535	ACDGHKLMQRSTW		ACDGHKLMQRSTW		P	P	F	EINY		
74	T	78.760	0.547	ACEFGHKLMNPRSVW		ACEFGHKLMNPRSVW					DIQY		
75	V	45.747	0.307	ACFHLMNQRSTY		ACFHLMNQRSTY		DGP	DGP	W	EIK		
76	N	69.106	0.416					ACFGIKLPQRSTVW	ACFGIKLPQRSTVW	D	EHMY		
77	G	28.072	0.326	HK	HK		HK	DELPQRTV	DELPQRTV	FMNSY	ACIW		HK
78	G	0.000	0.000					ADIMPTY	ADIMPTY	CHKLQRSV	EFNW		
79	I	5.990	0.035	LTV		LTV		ADFGHKNPSWY	ADFGHKNPSWY	R	CEMQ		
80	P	4.421	0.033					ADEFGIKLMNRSTVY	ADEFGIKLMNRSTVY		CHQW		
81	Q	46.361	0.239	P				ACEGHLNSVWYP	ACEGHLNSVWY	FMR	DIKT	P	P
82	K	117.628	0.555	YAEGHILMNQRSTV	Y	AEGHILMNQRSTV		WY	W	P	CDF	Y	Y
83	I	34.146	0.197	KFGHLNQRSTV	K	FGHLNQRSTV		EK	E	PY	ACDMW	K	K
84	S	47.481	0.380	YDEFGHILMNPQRTW	Y	DEFGHILMNPQRTW		Y			ACKV	Y	Y
85	L	46.614	0.260	V		V		ACDEFGHNQST	ACDEFGHNQST	KPR	IMWY		
86	Q	112.172	0.578	ADEFGHIKLMNPRSTVW		ADEFGHIKLMNRSTVW		CP	C		Y	P	P
87	D	78.456	0.490	PACEGHILMQRSTVY	P	ACEGHILMQRSTVY		P	P		FKNW	P	P
88	H	0.000	0.000					ACEFGIKLMPRSTVY	ACEFGIKLMPRSTVY		DNQW		
89	L	35.091	0.196	RPWCKM	RPW	CKM	RP	ADEGQSTYW	ADEGQSTY	N	FHIV	W	RPW
90	D	79.306	0.496	AEGHIKLNQRSTW		AEHIKLNQRSTW		CG	C	P	FMVY	G	G
91	K	67.628	0.319	AQR		AQR		DEFGHILNT	DEFGHILNT	SY	CMPVW		
92	A	1.383	0.012	HCLMTV	H	CLMTV		EFKPQRWYH	EFKPQRWY	G	DINS	H	H
93	K	114.335	0.539	PVDEFGHILMNQRST	PV	DEFGHILMNQRST	PV				ACWY		PV
94	K	141.266	0.666	CADEFHLMNQRST	C	ADEFHLMNQRST	C	GP	GP		IVWY		C
95	D	16.707	0.104					ACEFGHKLMPQSVWY	ACEFGHKLMPQSVWY		INRT		
96	I	0.795	0.005	DLV		DLV		ACGHPRSTW	ACGHPRSTW	EFN	KMQY		
97	T	60.642	0.421	ACDEFGILNPQRSWY		ACDEFGILNPQRSWY					HKMV		
98	F	170.109	0.818	ACDEHILMQRSVW		ACDEHILMQRSVW		P	P	G	KNTY		
99	Y	46.946	0.207	ARS	A	RS	A	CGINPVWE	CGINPVWE	FLQT	HKMD		A
100	M	5.098	0.025					CEFGNPRSTWY	CEFGNPRSTWY	KLQV	ADHI		
101	P	90.756	0.672					ACFHIKLMNQRST	ACFHIKLMNQRST	GY	DEVW		
102	V	57.386	0.385	ACEGHKLMNQRSTW		ACEGHKLMNQRSTW		P	P		DFIY		
103	D	104.174	0.651	N	N		N	AEFGHILQRTVWY	AEFGHILQRTVWY	S	CKMP		N
104	N	4.550	0.027	IACGKMRST	I	ACGKMRST	I	FPW	FPW	HLV	DEQY		I
105	L	99.070	0.553	WCHAGIPQRSTV	WCH	AGIPQRSTV	WH	MNC	MN	DE	FKY	C	WCH
106	G	3.543	0.041	V		V		ACDFHLMNPSWY	ACDFHLMNPSWY	EI	KQRT		
107	M	6.944	0.035	FIL		FIL		ACHKPQSVW	ACHKPQSVW	DGR	ENTY		
108	A	0.000	0.000	G		G		DEFKLMPQTVY	DEFKLMPQTVY	HNRS	CIW		
109	V	1.591	0.011					CDELMRTW	CDELMRTW	AFGHPQY	IKNS		
110	I	0.000	0.000	V		V		FKLMPW	FKLMPW	ACDGHNRS	EQTY		
111	D	30.647	0.192					HIQ	HIQ	CEGKLMPRSTVWY	AFN		
112	W	8.741	0.035					CEGHLNPS	CEGHLNPS	DFIQRVY	AKMT		
113	E	81.629	0.437					RV	RV	ACDFGHLPQSTWY	IKMN		
114	E	54.483	0.291	AGHMS		AGHMS		ILPTV	ILPTV	CDRWY	FKNQ		
115	W	5.584	0.022					ACDFGHIKLMRSVY	ACDFGHIKLMRSVY	P	ENQT		
116	R	29.305	0.117					ACDEGHILNPQSVW	ACDEGHILNPQSVW	T	FKMY		
117	P	6.733	0.050	D				GIKNQRSVWD	GIKNQRSVW	EFHTY	ACLM	D	D
118	T	7.966	0.055	HKLMNQV		HKLMNQV		CDEGPRWY	CDEGPRWY		AFIS		
119	W	22.218	0.089	PFQY	P	FQY		AIKLNRP	AIKLNR	DEGSV	CHMT	P	P
120	A	53.068	0.478	DFGHILNPRSTVWY		DFGHILNPRSTVWY				C	EKMQ		
121	R	2.963	0.012					ACEFGHKLMPWY	ACEFGHKLMPWY	DSTV	INQ		
122	N	2.114	0.013	M		M		ACEFIKQRSTV	ACEFIKQRSTV	LPW	DGHY		
123	W	91.137	0.366					ACDEHLMPQRSTVY	ACDEHLMPQRSTVY	G	FIKN		
124	K	177.903	0.839	LHR	LH	R	LH	CDEFN	CDEFN	AGIPSTVW	MQY		LH
125	P	102.480	0.759	AHRS	A	HRS	A	CDGLNW	CDGLNW	QTVY	EFIKM		A
126	K	37.646	0.178					FHILNPY	FHILNPY	ADEGMQRSTVW	C		

Pos	Wild Type (WT)	SASA (Å)	fSASA	Active Mutants	Mutants 20-40% Activity	Active Mutants (>40%, No errors)	Mutants 20-40% Activity (no errors)	Inactive Mutants	Inactive Mutants (no errors)	Unclassified	Not Made	Errors2	Mutants w/ 20-40% Activity + Errors
127	D	48.435	0.303	AEGHLMNQRSTVW		AEGHLMNQRSTVW		K	K	F	CIPY		
128	V	58.060	0.390	ACGIKLQRSW		ACGIKLQRSW		EP	EP	FHY	DMNT		
129	Y	3.493	0.015					ACDEGHLPQSTVW	ACDEGHLPQSTVW	MR	FIKN		
130	K	50.824	0.240	IR		IR		CDGHLNSTWY	CDGHLNSTWY	EQV	AFMP		
131	N	86.923	0.524	CEFGHILMQRSTVY		CEFGHILMQRSTVY		P	P		ADKW		
132	R	93.616	0.374	ACEFHIKLNQSTVY		ACEFHIKLNQSTVY		P	P		DGMW		
133	S	0.000	0.000	I		I		DEFGHLMNPRTVW	DEFGHLMNPRTVW	A	CKQY		
134	I	38.703	0.224	LTV		LTV		ACDFGHKPQRSW	ACDFGHKPQRSW		EMNY		
135	E	83.130	0.445	ACDFGHKLNQRSWY		ACDFGHKLNQRSWY		P	P		IMTV		
136	L	41.446	0.232	ACDFHIMNQRSTW		ACDFHIMNQRSTW		P	P	G	EKVY		
137	V	13.182	0.088	CSLAIT	CSL	AIT	CSL	FGHNPRWY	FGHNPRWY	EQ	DKM		CSL
138	Q	57.823	0.298	ACHILMNRSWY		ACHILMNRSWY		V	V	EFG	DKPT		
139	Q	119.753	0.617	ACDEFGHKLMRSTV		ACDEFGHKLMRSTV		P	P		INWY		
140	Q	160.155	0.826	ACDFGHIKLMRVWY		ACDFGHIKLMRVWY				S	ENPT		
141	N	65.932	0.397	ADEFGHLMQRSTVWY		ADEFGHLMQRSTVWY				P	CIK		
142	V	99.585	0.668	CDEGHIKLMNPQRST		CDEGHIKLMNPQRST					AFWY		
143	Q	165.074	0.851	CEGIKLNV	C	EGIKLNV		HPRSTC	HPRST	FMY	ADW	C	C
144	L	32.641	0.182	RTW	R	TW	R	AEFIKPQSVY	AEFIKPQSVY	GN	CDHM		R
145	S	68.577	0.549	ACDEGHLMNPR		ACDEGHLMNPR		TW	TW	FV	IKQY		
146	L	122.629	0.685	ACEGHIKNPQRSTVY		ACEGHIKNPQRSTVY					DFMW		
147	T	98.809	0.686	ACDFGILMPQRSVWY		ACDFGILMPQRSVWY					EHKN		
148	E	92.061	0.492	CFGHIKLQRSTVWY		CFGHIKLQRSTVWY				P	ADMN		
149	A	0.000	0.000	CGKLMQRSTV		CGKLMQRSTV		EP	EP	FWY	DHIN		
150	T	47.998	0.333	ACDEFGILNPRSWY		ACDEFGILNPRSWY		V	V		HKMQ		
151	E	86.410	0.462	ACGHKLMNQRSTVWY		ACGHKLMNQRSTVWY					DFIP		
152	K	102.753	0.485	WACFIMRTVY	W	ACFIMRTVY	W	L	L	GNPS	DEHQ		W
153	A	0.000	0.000	SIL	S	IL	S	EFMPRTV	EFMPRTV	CGHKQW	DNY		S
154	K	65.497	0.309	IRTV	I	RTV	I	DEGPSWY	DEGPSWY	ACHLQ	FMN		I
155	Q	111.552	0.575	ACDFGHKLMRSTVW		ACDFGHKLMRSTVW		PY	PY		EIN		
156	E	73.070	0.391	CADGILMQRSTVW	C	ADGILMQRSTVW	C	P	P	K	FHNY		C
157	F	3.035	0.015	W		W		ACDEGHIKLMPQRSTV	ACDEGHIKLMPQRSTV		NY		
158	E	43.559	0.233	GAFHLQS	G	AFHLQS	G	DKPRY	DKPRY	CNV	IMTW		G
159	K	139.479	0.658	GADEHLMNQRSV	G	ADEHLMNQRSV	G	WY	WY	F	CIPT		G
160	A	10.651	0.096	WCFGHIKLMNQRSVY	W	CFGHIKLMNQRSVY	W				DEPT		W
161	G	0.000	0.000	ACDERSV		ACDERSV		W	W	HIKLMQT	FNPY		
162	K	79.805	0.376	ADEGHLMPQRSVWY		ADEGHLMPQRSVWY				F	CINT		
163	D	55.396	0.346	WFAEGKLQRSTV	WF	AEGKLQRSTV	WF	CP	CP	H	IMNY		WF
164	F	3.866	0.019	LMVW		LMVW		ACDEGHNPQR	ACDEGHNPQR	S	IKTY		
165	L	5.098	0.028	CADFNRSVWY	C	ADFNRSVWY		HPTC	HPT	GQ	EIKM	C	C
166	V	25.229	0.169	ACEFGHLNQRTWY		ACEFGHLNQRTWY		D	D	P	IKMS		
167	E	61.558	0.329	FADGHKMNPRSTY	F	ADGHKMNPRSTY	F	V	V	L	CIQW		F
168	T	0.000	0.000	H		H		ACDEFGKLPRSVWY	ACDEFGKLPRSVWY		IMNQ		
169	I	0.720	0.004	LRV		LRV		ADFGHKNPQSTY	ADFGHKNPQSTY		CEMW		
170	K	89.258	0.421	AQNRV		AQNRV		CDEGMPWY	CDEGMPWY	IL	FHST		
171	L	29.926	0.167	IV		IV		CDHMNRSWY	CDHMNRSWY	AFKPQ	EGT		
172	G	0.000	0.000	AC		AC		DEILPQTVWY	DEILPQTVWY	MRS	FHKN		
173	K	84.793	0.400	QNR	Q	NR	Q	DEGHILMPSVWY	DEGHILMPSVWY		ACFT		Q
174	L	121.372	0.678	HAGKMNQRSTVWY	H	AGKMNQRSTVWY	H	P	P	C	DEFI		H
175	L	46.659	0.261	EHTVY		EHTVY		CDGKPRS	CDGKPRS	FNW	AIMQ		
176	R	39.598	0.158	KL		KL		ACEFGHIPQSTVW	ACEFGHIPQSTVW		DMNY		
177	P	74.735	0.554	V	V			ACDFGHLMQRSTWV	ACDFGHLMQRSTW	Y	EIKN	V	V
178	N	79.973	0.482	GKMR		GKMR		EILVWY	EILVWY	ADPST	CFHQ		
179	H	2.085	0.011	ACEGIKLMNPRSTV		ACEGIKLMNPRSTV		W	W		DFQY		
180	L	13.986	0.078	IFGKM	I	FGKM	I	ACEPRS	ACEPRS	HNTW	DQVY		I
181	W	0.000	0.000	KMQ	K	MQ		ACDEFHILRSVK	ACDEFHILRSV	N	GPTY	K	K
182	G	0.000	0.000	L		L		ACDEHNPQRSTVY	ACDEHNPQRSTVY	M	FIKW		
183	Y	5.539	0.024	EL	E	L		CDGIKNPQRSVE	CDGIKNPQRSV	AW	FHMT	E	E
184	Y	17.785	0.078	W	W		W	ACDEFGHKLMPRSV	ACDEFGHKLMPRSV		INQT		W
185	L	27.644	0.154					ADEFGIKPRSTVWY	ADEFGIKPRSTVWY	N	CHMQ		
186	F	12.206	0.059	Y		Y		ADGHIKLNPQRSVW	ADGHIKLNPQRSVW		CEMT		
187	P	4.158	0.031					AFGHILMNQRSTVWY	AFGHILMNQRSTVWY		CDEK		
188	D	9.692	0.061					ACFGHLMNPQRSTVW	ACFGHLMNPQRSTVW		EIKY		
189	C	2.092	0.013					AEGHKLMNPRSTVWY	AEGHKLMNPRSTVWY		DFIQ		
190	Y	24.249	0.107					CEFGHKLNQRSTVW	CEFGHKLNQRSTVW	P	ADIM		

Pos	Wild Type (WT)	SASA (Å)	fSASA	Active Mutants	Mutants 20-40% Activity	Active Mutants (>40%, No errors)	Mutants 20-40% Activity (no errors)	Inactive Mutants	Inactive Mutants (no errors)	Unclassified	Not Made	Errors2	Mutants w/ 20-40% Activity + Errors
191	N	3.670	0.022					AEFGKLMPQRSTVWY	AEFGKLMPQRSTVWY		CDHI		
192	H	98.584	0.516	ST		ST		CFGKLMNPQRVWY	CFGKLMNPQRVWY		ADEI		
193	H	101.781	0.533	RFGQSY	RF	GQSY	RF	ADKLMPV	ADKLMPV	CT	EINW		RF
194	Y	68.699	0.303					ACILPSTV	ACILPSTV	EFGNQRW	DHKM		
195	K	163.551	0.771	VAGHILNQRSTW	V	AGHILNQRTW	V	S		EFY	CDMP	S	VS
196	K	101.758	0.480	NYEGLRSTW	NY	EGLRSTW	NY			ACDIPV	FHMQ		NY
197	P	144.905	1.073	WADEFGHKLMQRST	W	ADEFGHKLMQRST	W	C	C		INVY		W
198	G	50.427	0.586	WADEHLNQRSTY	W	ADEHLNQRSTY		VW	V	CP	FIKM	W	W
199	Y	38.340	0.169					EGHIKLPRSW	EGHIKLPRSW	ACNQT	DFMV		
200	N	101.580	0.612	TD	T	D	T	AFGHKLMPQRSWY	AFGHKLMPQRSWY	V	CEI		T
201	G	0.720	0.008					AFLMNPRSTVW	AFLMNPRSTVW	EHKQ	CDIY		
202	S	65.334	0.523	M		M		AEFGHKNPQRVWY	AEFGHKNPQRVWY	T	CDIL		
203	C	13.271	0.085					ADEGHLMNQRSTV	ADEGHLMNQRSTV	PW	FIKY		
204	F	98.579	0.474	WP	W	P	W	ACEGHIKQRST	ACEGHIKQRST	LMV	DNY		W
205	N	98.259	0.592	ADEFGKMPRSTVWLY	L	ADEFGKMPRSTVWY	L				HQCI		L
206	V	57.311	0.385	QHIKLMRST	Q	HIKLMRST	Q	CDFGPY	CDFGPY		AENW		Q
207	E	24.338	0.130					AFGMPQRSTVW	AFGMPQRSTVW	HIKL	CDNY		
208	I	64.061	0.370	ACKLMQRSTV		ACKLMQRSTV		DGPW	DGPW	E	FHNY		
209	K	107.619	0.508	LAEFGNRST	L	AEFGNRST	L	CP	CP	DVWY	HIMQ		L
210	R	71.659	0.287					ACDEGKMNPSTVWY	ACDEGKMNPSTVWY	L	FHIQ		
211	N	1.184	0.007	LW		L		CFGHIKMPRSTVW	CFGHIKMPRSTV	A	DEQY	W	W
212	D	52.008	0.325	NST		NST		AGHIKLMPVW	AGHIKLMPVW	EQ	CFRY		
213	D	90.794	0.567	AEGHKLMNQRVWY		AEGHKLMNQRVWY		PS	PS		CFIT		
214	L	5.023	0.028	Q		Q		ACDEGHKNPRSTY	ACDEGHKNPRSTY	V	FIMW		
215	S	24.295	0.194	ADEGHKLMQRTVW		ADEGHKLMQRTVW		CP	CP		FINY		
216	W	38.551	0.155	Y		Y		DEGHIKLMNPQRTV	DEGHIKLMNPQRTV		ACFS		
217	L	2.311	0.013	M		M		ACGHPQSTVW	ACGHPQSTVW	EIRY	DFKN		
218	W	3.197	0.013	MVF	MV	F	M	AIKLPSV	AIKLPS	DGHQRT	CENY	V	MV
219	N	71.240	0.429	ACDEHIKLMRSTW		ACDEHIKLMRSTW		P	P	G	FQVY		
220	E	41.651	0.223	DMAHILSTV	DM	AHILSTV	DM	GKNPRW	GKNPRW		CFQY		DM
221	S	8.290	0.066	IACMQTV	I	ACMQTV	I	DEHKPR	DEHKPR	GL	FNWY		I
222	T	25.292	0.176	DFGIKLNRSV		DFGIKLNRSV		PY	PY	AEW	CHMQ		
223	A	0.639	0.006					CDEGHKLPQRSTVWY	CDEGHKLPQRSTVWY		FIMN		
224	L	2.235	0.012	I		I		ADEFGMPQRSTWY	ADEFGMPQRSTWY	V	CHKN		
225	Y	0.000	0.000					ADEGHKPQRTVW	ADEGHKPQRTVW	LS	CFIMN		
226	P	0.760	0.006	W				ACDEFGLNQRSTVYW	ACDEFGLNQRSTVY		HIKM	W	W
227	S	6.187	0.049					AFGHIKLMPQRTVWY	AFGHIKLMPQRTVWY		CDEN		
228	I	2.505	0.014					AEFGHLMNPRSTW	AEFGHLMNPRSTW	KQ	CDVY		
229	Y	35.171	0.155					EFGKLPQTVW	EFGKLPQTVW	HINRS	ACDM		
230	L	3.400	0.019	I		I		AEGHKMNPRSTVWY	AEGHKMNPRSTVWY		CDFQ		
231	N	69.570	0.419	T		T		ACDFGHIKLPQRSV	ACDFGHIKLPQRSV		EMWY		
232	T	56.716	0.394	FS		FS		CGHKLNPQVY	CGHKLNPQVY	AR	DEIMW		
233	Q	94.120	0.485	AFGKLRY		AFGKLRY		DIPST	DIPST	CVW	EHMN		
234	Q	10.225	0.053	LM		LM		ADEGHNPSTVW	ADEGHNPSTVW	CR	FIKY		
235	S	102.880	0.823	AEGHKT		AEGHKT		FLMRWY	FLMRWY	PQV	CDIN		
236	P	58.304	0.432	AGHKRS		AGHKRS		CILNQTY	CILNQTY	EW	DFMV		
237	V	135.527	0.910	CAEFHLNQRSTW	C	AEFHLNQRSTW	C			GPY	DIKM		C
238	A	12.639	0.114	DEHKQRST		DEHKQRST		FGLPVWY	FGLPVWY		CIMN		
239	A	2.243	0.020	N	N		N	CFGHILPRSTVWY	CFGHILPRSTVWY	K	DEMQ		N
240	T	24.367	0.169	KAMPQRSV		KAMPQRSV		EFGNWY	EFGNWY	L	CDHI		
241	L	57.938	0.324					ACDEGIPRSTVW	ACDEGIPRSTVW	FKQ	HMNY		
242	Y	10.153	0.045	F		F		ACDGILMPRSTVW	ACDGILMPRSTVW	K	EHNQ		
243	V	0.000	0.000					CDFGHLMPQRSWY	CDFGHLMPQRSWY	AT	EIKN		
244	R	68.249	0.273					ADGIVY	ADGIVY	HKMNPQSTW	CEFL		
245	N	3.179	0.019	H		H		ACFLPQRSTV	ACFLPQRSTV	GIKW	DEMY		
246	R	7.459	0.030					ACDEGHIKLMPSTVW	ACDEGHIKLMPSTVW		FNQY		
247	V	0.000	0.000	ILM		ILM		ACFHNPQRSTWY	ACFHNPQRSTWY		DEGK		
248	R	89.071	0.356	AHWY		AHWY		CDEGIMPT	CDEGIMPT	LSV	FKNQ		
249	E	2.065	0.011	V		V		AGHIKMQSY	AGHIKMQSY	LPRTW	CDFN		
250	A	0.000	0.000					CFGHKLMNPQRSTVW	CFGHKLMNPQRSTVW		DEIY		
251	I	38.094	0.220	YLM	Y	LM	Y	DFGHKPSTW	DFGHKPSTW	CQV	AENR		Y
252	R	58.364	0.233					ADEFGHIKLNPSTY	ADEFGHIKLNPSTY	V	CMQW		
253	V	0.000	0.000	I		I		ADEGHLMNQRSW	ADEGHLMNQRSW	PT	CFKY		
254	S	2.200	0.018					CDEGIKLPQRTVWY	CDEGIKLPQRTVWY	N	AFHM		

Pos	Wild Type (WT)	SASA (Å)	fSASA	Active Mutants	Mutants 20-40% Activity	Active Mutants (>40%, No errors)	Mutants 20-40% Activity (no errors)	Inactive Mutants	Inactive Mutants (no errors)	Unclassified	Not Made	Errors2	Mutants w/ 20-40% Activity + Errors
255	K	117.817	0.556	GANQRS	G	ANQRS	G	CDLPVW	CDLPVW	HTY	EFIM		G
256	I	54.806	0.317	AHLV		AHLV		CDEGP	CDEGP	MNQRTW	FKSY		
257	P	84.209	0.624	CAGIKLMNQRTV	C	AGIKLMNQRTV	C	D	D	SW	EFHY		C
258	D	92.576	0.579	GHNRS	G	HNRS	G	LPVW	LPVW	AEIQTY	CFKM		G
259	A	21.353	0.192	EGIKLNPQRSTVWY		EGIKLNPQRSTVWY				M	CDFH		
260	K	175.390	0.827	GADEHLMQRSY	G	ADEHLMQRSY	G	CP	CP	VW	FINT		G
261	S	68.849	0.551	AFKMNQRTVW		AFKMNQRTVW		P	P	EGIL	CDHY		
262	P	57.941	0.429					ADEFGHIKQRSTVWY	ADEFGHIKQRSTVWY		CLMN		
263	L	25.418	0.142	HAKMRTV	H	AKMRTV	H	EFPQW	EFPQW	GNS	CDIY		H
264	P	16.735	0.124	AH		AH		DEFGLMRTVWY	DEFGLMRTVWY	NS	CIKQ		
265	V	6.141	0.041	I		I		ADFGHKLMNQRS	ADFGHKLMNQRS	CEPWY	T		
266	F	2.350	0.011	Y		Y		ACGHMPQRSTVW	ACGHMPQRSTVW	DL	EIKN		
267	A	0.760	0.007	MT		MT		DGHIKNRSW	DGHIKNRSW	ELPV	CFQY		
268	Y	12.220	0.054					ACFGHKLNPQSTVW	ACFGHKLNPQSTVW	R	DEIM		
269	T	2.823	0.020	ACDS		ACDS		EKLMNPQR	EKLMNPQR	GVY	FHIW		
270	R	13.176	0.053	MNST		MNST		ACEFGHIPY	ACEFGHIPY	DQV	KLW		
271	I	7.853	0.045	FGLMSV		FGLMSV		ADEHKTW	ADEHKTW	PR	CNQY		
272	V	2.114	0.014	EMDRST	EM	DRST	EM	AHLNPW	AHLNPW	CGK	FIQY		EM
273	F	26.679	0.128	HTY		HTY		ACDGILPQSVW	ACDGILPQSVW	R	EKMN		
274	T	23.215	0.161	AFS		AFS		CEGHNQWY	CEGHNQWY	LPRV	DIKM		
275	D	111.462	0.697	LV	L			AFGIKMQTWLV	AFGIKMQTW	CERS	HNPY	LV	LV
276	Q	46.527	0.240	GCDEHILMRSY	G	CDEHILMRSY	G	FPW	FPW	V	AKNT		G
277	V	53.292	0.358	ACDEGHKMNQRSTY		ACDEGHKMNQRSTY				L	FIPW		
278	L	110.288	0.616	AEFGHIKNRSTVY		AEFGHIKNRSTVY		MP	MP		CDQW		
279	K	105.162	0.496	AHQRT	A	HQRT		CFGLWYA	CFGLWY	DPSV	EIMN	A	A
280	F	57.997	0.279	GQ		GQ		DILMNRSTVW	DILMNRSTVW	EHP	ACKY		
281	L	5.232	0.029					ADGHIKNPQRSVW	ADGHIKNPQRSVW	FY	CEMT		
282	S	53.343	0.427	DGMQ		DGMQ		FLVWY	FLVWY	ACEPRT	HIKN		
283	Q	106.552	0.549	EPRST		EPRST		ACDFW	ACDFW	GHLNY	IKMV		
284	D	79.206	0.495	TYAEGHLMNQS	TY	AEGHLMNQS	TY	CIP	CIP	V	FKRW		TY
285	E	11.785	0.063	AFGHMNQY	A	FGHMNQY	A	KPRTV	KPRTV	SW	CDIL		A
286	L	0.000	0.000	WRS	W	RS	W	ACDFHKMPTY	ACDFHKMPTY	EG	INQV		W
287	V	41.730	0.280	NIT	N	IT		ACDEGKLPQRSN	ACDEGKLPQRS	F	HMWY	N	N
288	Y	63.463	0.280	LW		LW		DEFGHIKPRT	DEFGHIKPRT	QSV	ACMN		
289	T	3.830	0.027	KS		K		ACEGHLPQRYS	ACEGHLPQRY	MNV	DFIW	S	S
290	F	0.000	0.000	IM		IM		DQY	DQY	ACGHKLRSTV	ENPW		
291	G	3.146	0.037	CQRSV	C	QRSV		ADEFMNTWYC	ADEFMNTWY	HLP	IK	C	C
292	E	1.380	0.007	ACFGHKNPRVW		ACFGHKNPRVW		ILT	ILT	Q	DMSY		
293	T	0.000	0.000	ACDFGKLMPQSVY		ACDFGKLMPQSVY		EN	EN		HIRW		
294	V	1.829	0.012	M		M		AEGHKLNPQRSTW	AEGHKLNPQRSTW	C	DFIY		
295	A	0.000	0.000					CGHILNPTVY	CGHILNPTVY	DEFQS	KMRW		
296	L	28.766	0.161					CFGIKMQRSTVWY	CFGIKMQRSTVWY	AP	DEHN		
297	G	23.523	0.274	RA	R	A	R	CEHLNPQRSTY	CEHLNPQSTY	IVW	DFKM		R
298	A	2.079	0.019	GI		GI		CELMNPQSTWY	CELMNPQSTWY	RV	DFHK		
299	S	26.443	0.212					ACDFGHLMPQT	ACDFGHLMPQT	EIRY	KNVW		
300	G	0.000	0.000	R		R		ACDEFLMNPQSTVW	ACDEFLMNPQSTVW		HIKY		
301	I	2.018	0.012	AV		AV		EGHKMNPQRSWY	EGHKMNPQRSWY	L	CDFT		
302	V	2.862	0.019	IW		IW		CDEFGHLMPRSTY	CDEFGHLMPRSTY		AKNQ		
303	I	0.000	0.000	DV	D	V		ACEFGKLMRWYD	ACEFGKLMRWY	PS	HNQT	D	D
304	W	69.383	0.279	GI				ACDMNPQSTVYGI	ACDMNPQSTVY	LR	EFHK	GI	GI
305	G	5.680	0.066	NDE	N	DE	N	LPQRSTVY	LPQRSTVY	CFHK	AIMW		N
306	T	19.027	0.132	DES		DES		ACHILVWY	ACHILVWY	FGPR	KMNQ		
307	L	112.244	0.627	GKNQSTVWY	G	KNQSTVWY	G	CIP	CIP	EFR	ADHM		G
308	S	64.223	0.514	DGHKNPRT		DGHKNPRT		CFLMVWY	CFLMVWY		AEIQ		
309	I	7.201	0.042	DEGHKLMNQRSTVW		DEGHKLMNQRSTVW				Y	ACFP		
310	M	1.383	0.007	FYAGQRSV	FY	AGQRSV	Y	CEKLF	CEKL	NPW	DHIT	F	FY
311	R	177.550	0.710	GHKQST		GHKQST		CEFILPVW	CEFILPVW	A	DMNY		
312	S	42.440	0.340	KLGNT	KL	GNT	KL	CEMVW	CEMVW	AFHPQR	DIY		KL
313	M	107.328	0.534	AEGHKLPRSTVY		AEGHKLPRSTVY		C	C	DF	INQW		
314	K	136.842	0.645	ADHINQRSTY		ADHINQRSTY		CLW	CLW	PV	EFGM		
315	S	11.506	0.092	AEGHKLMRTY		AEGHKLMRTY		CIV	CIV	PW	DFNQ		
316	C	2.591	0.017	D		D		EGIKLMPRSTVWY	EGIKLMPRSTVWY	A	FHNQ		
317	L	75.812	0.424	ADHIKMNQRSTW		ADHIKMNQRSTW		GP	GP	C	EFVY		
318	L	72.002	0.402	DFGHIKMNQRST		DFGHIKMNQRST		CPW	CPW		AEVY		

Pos	Wild Type (WT)	SASA (Å)	fSASA	Active Mutants	Mutants 20-40% Activity	Active Mutants (>40%, No errors)	Mutants 20-40% Activity (no errors)	Inactive Mutants	Inactive Mutants (no errors)	Unclassified	Not Made	Errors2	Mutants w/ 20-40% Activity + Errors
319	L	2.100	0.012					CEFGHIKMPQRSVWY	CEFGHIKMPQRSVWY		ADNT		
320	D	23.798	0.149	LVEGHIKMNRSWY	LV	EGHIKMNRSWY	L	CPV	CP	F	AQT	V	LV
321	N	76.586	0.461	ADHKRSTY		ADHKRSTY		EMP	EMP	GILV	CFQW		
322	Y	10.085	0.044					CDEGILNPRSTVW	CDEGILNPRSTVW	FH	AKMQ		
323	M	3.715	0.018	FIL		FIL		ACEGHKNRSTV	ACEGHKNRSTV	P	DQWY		
324	E	89.034	0.476	ADHMNRS		ADHMNRS		CFPVWY	CFPVWY	GL	IKQT		
325	T	90.864	0.631	ADEGHKMNQSVW		ADEGHKMNQSVW		CR	CR	I	FLPY		
326	I	43.424	0.251	CKLSVY	C	KLSVY	C	EGHNW	EGHNW	ADPR	FMQT		C
327	L	0.000	0.000	M		M		AEFGHNQRSTVWY	AEFGHNQRSTVWY	D	CIKP		
328	N	0.652	0.004	WACGHIKLQRSTVY	W	ACGHIKLQRSTVY	W			D	EFMP		W
329	P	32.519	0.241					CFGHIKLNQRSTVWY	CFGHIKLNQRSTVWY		ADEM		
330	Y	32.104	0.141					ACDEGILMNPRSVW	ACDEGILMNPRSVW	F	HKQT		
331	I	0.720	0.004	CEV	CE	V		ADFHKQRSTWYCE	ADFHKQRSTWY	L	GMNP	CE	CE
332	I	31.326	0.181					ACDEFGHKLNPRSTY	ACDEFGHKLNPRSTY		MQVW		
333	N	35.677	0.215					GHIKPRSTWY	GHIKPRSTWY	AELMV	CDFQ		
334	V	0.720	0.005	TP	T	P	T	ACDEGMNRS	ACDEGMNRS	HLQY	FIKW		T
335	T	33.546	0.233	S		S		FGHIKLPVWY	FGHIKLPVWY	ACNQ	DEMR		
336	L	29.148	0.163					AEFGKNPRSTVWY	AEFGKNPRSTVWY	HM	CDIQ		
337	A	0.000	0.000					CFGIKLMRTW	CFGIKLMRTW	HNPSV	DEQY		
338	A	3.664	0.033	Q		Q		CDEFGHIKLPRTV	CDEFGHIKLPRTV	S	MNWY		
339	K	57.662	0.272	M		M		DEFGHLNPSTVWY	DEFGHLNPSTVWY	R	ACIQ		
340	M	33.489	0.167					ACDEFGHKPRSTVW	ACDEFGHKPRSTVW	L	INQY		
341	C	0.000	0.000					AEGHKLMNQRSTVY	AEGHKLMNQRSTVY	W	DFIP		
342	S	5.448	0.044	A		A		DEFHKLMPQRTY	DEFHKLMPQRTY	GI	CNVW		
343	Q	58.968	0.304	TV		TV		CDFIPW	CDFIPW	EGLMRSY	AHKN		
344	V	49.922	0.335					FGHLMNPQRSTWY	FGHLMNPQRSTWY	EI	ACDK		
345	L	31.197	0.174					ACEHKNQRTVY	ACEHKNQRTVY	DGPW	FIMS		
346	C	0.000	0.000					ADFGIKLMPRSTVW	ADFGIKLMPRSTVW	Q	EHNY		
347	Q	86.748	0.447	LAEGMRS	L	AEGMRS	L	CFIPTVW	CFIPTVW	Y	DHKN		L
348	E	101.491	0.543	DGS		DGS		CHILPQRTVWY	CHILPQRTVWY	M	AFKN		
349	Q	70.960	0.366	AEKMNRT		AEKMNRT		DFGPVWY	DFGPVWY	HLS	CI		
350	G	0.000	0.000					ADEFHKLMNPRSTVY	ADEFHKLMNPRSTVY		CIQW		
351	V	18.705	0.126	CIQAS	CIQ	AS	IQ	DEFHNRWYC	DEFHNRWY	GL	KMPT	C	CIQ
352	C	0.652	0.004					ADEFGKMPQRSTVWY	ADEFGKMPQRSTVWY		HILN		
353	I	42.834	0.248	TV		TV		CFGHKLMQRSW	CFGHKLMQRSW	AE	DNPY		
354	R	20.892	0.084					CDEGHIKLMPQSVWY	CDEGHIKLMPQSVWY		AFNT		
355	K	54.632	0.258					DFGHLMNPQRSTVWY	DFGHLMNPQRSTVWY		ACEI		
356	N	147.976	0.891	ADHS		ADHS		CGKLPRTVW	CGKLPRTVW	FQ	EIMY		
357	W	219.026	0.880	KACST	K	ACST	K	DEFGLMQR	DEFGLMQR	PV	HINY		K
358	N	54.014	0.325	LCGT	L	CGT	L	EHIKPQRW	EHIKPQRW	DV	AFMSY		L
359	S	72.418	0.579	DEHKMTV		DEHKMTV		AFGLPW	AFGLPW	CR	INQY		
360	S	71.417	0.571	T		T		ACEFGIKLMPQRV	ACEFGIKLMPQRV	N	DHWY		
361	D	29.415	0.184	H	H		H	ACEGMNPQRSVW	ACEGMNPQRSVW	LY	FIKT		H
362	Y	39.880	0.176					ACEGHKLMNPRSTVW	ACEGHKLMNPRSTVW		DFIQ		
363	L	6.819	0.038					ACDEFGHIPQRSTVW	ACDEFGHIPQRSTVW		KMNY		
364	H	91.537	0.479					ACDEFGKLMPRSTVY	ACDEFGKLMPRSTVY		INQW		
365	L	11.873	0.066					ACDEGMNPQRSTWY	ACDEGMNPQRSTWY	IV	FHK		
366	N	46.049	0.277					ACEFGKMPQRTW	ACEFGKMPQRTW	LSV	DHIY		
367	P	105.954	0.785	ACGKRS		ACGKRS		EFILMQV	EFILMQV	HW	DNTY		
368	D	140.876	0.880	AEGHKLMRSTV		AEGHKLMRSTV		CPW	CPW	Y	FINQ		
369	N	25.624	0.154	HRS		HRS		CEFIKLPQVW	CEFIKLPQVW	AT	DGMY		
370	F	20.643	0.099					ADEGHKLNPQRSVY	ADEGHKLNPQRSVY	I	CMTW		
371	A	35.895	0.323	GEFHIKLMRSV	G	EFHIKLMRSV	G	PW	PW	CT	DNQY		G
372	I	44.651	0.258					ADEFGHKLNPRSTVW	ADEFGHKLNPRSTVW		CMQY		
373	Q	103.495	0.533	AEFKLMRSV		AEFKLMRSV		CPW	CPW	GHNT	DIY		
374	L	73.823	0.412	WAHIMNPRSTVY	W	AHIMNPRSTVY	W	DE	DE	G	CFKQ		W
375	E	106.861	0.571	IAGKLMNRST	I	AGKLMNRST	I	CFPVY	CFPVY		DHQW		I
376	K	205.579	0.970	LADEMQRSTVY	L	ADEMQRSTVY	L	IPW	IPW	G	CFHN		L
377	G	78.889	0.917	DEHKPRST		DEHKPRST		CILV	CILV	FMY	ANQW		
378	G	28.816	0.335	KNR		KNR		DEFILMQTWY	DEFILMQTWY	SV	ACHP		
379	K	131.483	0.620	GHRST		GHRST		ACEFILMW	ACEFILMW	NV	DPQY		
380	F	19.524	0.094	VTILPWY	VT	ILPWY	VT	CDEGRSQ	CDEGRSQ	A	HKMN		VT
381	T	23.941	0.166	EHKNQRSV	E	HKNQRSV	E	GLPWY	GLPWY	AF	CDIM		E
382	V	41.918	0.281					EGHKLMNPQRSTWY	EGHKLMNPQRSTWY	I	ACDF		

Pos	Wild Type (WT)	SASA (Å)	fSASA	Active Mutants	Mutants 20-40% Activity	Active Mutants (>40%, No errors)	Mutants 20-40% Activity (no errors)	Inactive Mutants	Inactive Mutants (no errors)	Unclassified	Not Made	Errors2	Mutants w/ 20-40% Activity + Errors
383	R	140.040	0.560	AEHIKLMNSTV		AEHIKLMNSTV		GP	GP	FW	CDQY		
384	G	61.768	0.718					CFMQST	CFMQST	ADEHIKLPR	NVWY		
385	K	135.441	0.639	AGHNQRSTV		AGHNQRSTV		CLMPWY	CLMPWY		DEFI		
386	P	50.489	0.374					ACFGHILMNQRSTVY	ACFGHILMNQRSTVY		DEKW		
387	T	51.574	0.358	S		S		CEFGHILMNVWY	CEFGHILMNVWY	KQ	ADPR		
388	L	132.293	0.739	FHIMRTVWY		FHIMRTVWY		CGPQS	CGPQS	A	DEKN		
389	D	134.042	0.717	AGHKLMPQRSTY		AGHKLMPQRSTY		FV	FV	I	CENW		
390	D	19.943	0.125					ACEFGHLNPRSTVWY	ACEFGHLNPRSTVWY		IKMQ		
391	L	55.602	0.311	C		C		ADGHKNPQRSTVWY	ADGHKNPQRSTVWY		EFIM		
392	E	122.480	0.655	WAFGKLMQRSTVY	W	AFGKLMQRSTVY	W	CP	CP		DHIN		W
393	Q	93.910	0.484	ADFHKLMNRST		ADFHKLMNRST		CP	CP	GI	EVWY		
394	F	16.147	0.078	LW		LW		ADEGIKNPQRSTV	ADEGIKNPQRSTV		CHMY		
395	S	52.380	0.419	TAGHKRW	T	AGHKRW	T	CLMP	CLMP	DEVY	FINQ		T
396	E	123.156	0.659	LADHQRST	L	ADHQRST	L	CFGIPY	CFGIPY	V	KMNW		L
397	K	28.635	0.135					ACEFGILMPQTV	ACEFGILMPQTV	NRS	DHWY		
398	F	0.000	0.000	L	L			ACEGHINPRSTVWYL	ACEGHINPRSTVWY		DKMQ	L	L
399	Y	119.359	0.526	ACEKMNQRSTVW		ACEKMNQRSTVW		DP	DP	G	FHIL		
400	C	20.202	0.129					ADEFGILMPQRSTVY	ADEFGILMPQRSTVY		HKNW		
401	S	55.944	0.448	GQAEN	GQ	AEN	GQ	CFHKRWY	CFHKRWY	DLT	IMPV		GQ
402	C	40.505	0.258					ADEFLMPQRSTVWY	ADEFLMPQRSTVWY	G	HIKN		
403	Y	13.826	0.061	F		F		ACEGHKLMNPQRT	ACEGHKLMNPQRT	S	DIVW		
404	S			TAP	T	AP	T	CDFGHLMNRVWY	CDFGHLMNRVWY		EIKQ		T
405	T			FAGKMPQRSWY	F	AGKMPQRSWY	F	CIV	CIV	L	DEHN		F
406	L			ACEFGINQSTVY		ACEFGINQSTVY		PR	PR	D	HKMW		
407	S			ADEFGHLMNPQRVW		ADEFGHLMNPQRVW				T	CIKY		
408	C							AEFGIKLPRSTVWY	AEFGIKLPRSTVWY	N	DHMQ		
409	K			ADEGHIPQRSTV		ADEGHIPQRSTV				CLW	FMNY		
410	E			DKMNPQRSTVY		DKMNPQRSTVY		W	W	GIL	ACFH		
411	K			HANPRSTV	H	ANPRSTV	H	DEFG	DEFG	ILW	CMQY		H
412	A			HQDGILNPRSVWY	HQ	DGILNPRSVWY	Q	EH	E		CFKMT	H	HQ
413	D			AHSEKNRTQ	AHSQ	ENRT	ASQ	ILPHK	ILP	FGW	CMVY	HK	AHSKQ
414	V			KLIM	KL	IM	L	ADEGHRSTK	ADEGHRST	FQY	CNPW	K	KL
415	K			YVGSW	YV	GSW	YV	CDEP	CDEP	LMQRT	AFHIN		YV
416	D			FGHIKLNQRTVY		FGHIKLNQRTVY		CS	CS	W	AEMP		
417	T			I		I		ADEFGHKMPQR	ADEFGHKMPQR	LSW	CNVY		
418	D			AEFGILMNPQRSVY		AEFGILMNPQRSVY				C	HKTW		
419	A			EFGHIKLNRSWY		EFGHIKLNRSWY		DP	DP	T	CMQV		
420	V			IP		IP		ADFGHKLNRSTWY	ADFGHKLNRSTWY		CEMQ		
421	D			AEGHKLMNQRSTYI	I	AEGHKLMNQRSTY	I			W	CFPV		I
422	V			IT		IT		CDGHLMNQRSY	CDGHLMNQRSY	AEPW	FK		
423	C							ADEFGHLMPQRSTVW	ADEFGHLMPQRSTVW		IKNY		
424	I							ACEGHNQRSWY	ACEGHNQRSWY	KLTV	DFMP		
425	A			YGIKMNRS	Y	GIKMNRS		ELPWY	ELPW	CDV	FHQT	Y	Y
426	D			KSEGNPQY	KS	EGNPQY	KS	CFMR	CFMR	AIL	HTVW		KS
427	G			HTQIKS	HTQ	IKS	HTQ	ACFLPVWY	ACFLPVWY	R	DEMN		HTQ
428	V			LMPT		LMPT		ACDEGHNRSY	ACDEGHNRSY	F	IKQW		
429	C							ADKLNPSTVWY	ADKLNPSTVWY	GIMR	EFHQ		
430	I							ADELMNSTV	ADELMNSTV	GHKPRW	CFQY		
431	D			AEGHIKLNQRSVWY		AEGHIKLNQRSVWY		P	P		CFMT		
432	A			HEGNSV	H	EGNSV	H	CFIKLMPY	CFIKLMPY	R	DQTW		H
433	F			ACDEGHIKLPRSTVW		ACDEGHIKLPRSTVW					MNQY		
434	L			FGIMV		FGIMV		HKPQRW	HKPQRW	NST	ACDEY		
435	K			ACEGHRSTVY		ACEGHRSTVY				FILPW	DMNQ		
436	P			CDEGHIKLMQRSTWY	C	DEGHIKLMQRSTWY	C				AFNV		C
437	P			ADGHIKLMQRSY		ADGHIKLMQRSY		T	T	FW	CENV		
438	M			ACDEGLNPQRSTVW		ACDEGLNPQRSTVW		Y	Y		FHIK		
439	E			ACFGHKLPQSTVW		ACFGHKLPQSTVW		NR	NR		DIMY		
440	T			ADEFGHILMPRSVY		ADEFGHILMPRSVY		Q	Q		CKNW		
441	E			ADFGHKLNQSTVY		ADFGHKLNQSTVY		R	R	C	IMPW		
442	E			WCGHKLPQRTVY	W	CGHKLPQRTVY	W	MNS	MNS		ADFI		W
443	P			AEFGHLMNQRSTW		AEFGHLMNQRSTW		D	D	I	CKVY		
444	Q			MDEFGHIKNRVWY	M	DEFGHIKNRVWY	M			CL	APST		M
445	I			AGHLMNPQRSTVWY		AGHLMNPQRSTVWY				CDFK	E		
446	F			ACDEGHIKLMQRTVW		ACDEGHIKLMQRTVW					NPSY		

Pos	Wild Type (WT)	SASA (Å)	fSASA	Active Mutants	Mutants 20-40% Activity	Active Mutants (>40%, No errors)	Mutants 20-40% Activity (no errors)	Inactive Mutants	Inactive Mutants (no errors)	Unclassified	Not Made	Errors2	Mutants w/ 20-40% Activity + Errors
447	Y			DEFGILMNPQRTVW		DEFGILMNPQRTVW				K	ACHS		

Position #	Count Active	Count Active (>40%, no errors)	Count Inactive	Count Inactive (no errors)	Count Unclassified	Count Not Made	Count Mutants 20-40% Activity	Count Mutants 20-40% Activity (no errors)	Count Errors	Count Mutants 20-40% Activity + Errors	Tested (w/o errors)	Total
1	15	15	0	0	1	3	0	0	0	0	15	19
2	10	9	4	4	1	4	1	1	0	1	13	19
3	4	4	6	6	5	4	0	0	0	0	10	19
4	5	5	8	8	2	4	0	0	0	0	13	19
5	1	1	13	13	1	4	0	0	0	0	14	19
6	7	7	5	5	4	3	0	0	0	0	12	19
7	1	1	14	14	2	2	0	0	0	0	15	19
8	4	3	8	8	3	4	1	1	0	1	11	19
9	6	6	6	6	3	4	0	0	0	0	12	19
10	9	9	5	5	1	4	0	0	0	0	14	19
11	5	4	9	9	1	4	1	1	0	1	13	19
12	9	9	3	3	3	4	0	0	0	0	12	19
13	4	4	6	6	5	4	0	0	0	0	10	19
14	4	4	9	9	2	4	0	0	0	0	13	19
15	3	3	10	10	2	4	0	0	0	0	13	19
16	0	0	14	14	1	4	0	0	0	0	14	19
17	0	0	15	15	0	4	0	0	0	0	15	19
18	0	0	14	14	1	4	0	0	0	0	14	19
19	0	0	15	15	0	4	0	0	0	0	15	19
20	1	1	12	12	2	4	0	0	0	0	13	19
21	0	0	14	14	1	4	0	0	0	0	14	19
22	4	4	5	5	6	4	0	0	0	0	9	19
23	1	1	10	10	4	4	0	0	0	0	11	19
24	13	13	0	0	2	4	0	0	0	0	13	19
25	0	0	15	15	0	4	0	0	0	0	15	19
26	15	15	0	0	0	4	0	0	0	0	15	19
27	14	14	1	1	0	4	0	0	0	0	15	19
28	14	14	0	0	1	4	0	0	0	0	14	19
29	14	14	0	0	1	4	0	0	0	0	14	19
30	14	14	0	0	1	4	0	0	0	0	14	19
31	14	14	0	0	1	4	0	0	0	0	14	19
32	16	16	0	0	0	3	0	0	0	0	16	19
33	8	8	6	6	1	4	0	0	0	0	14	19
34	7	6	6	6	2	4	1	1	0	1	12	19
35	7	5	6	6	2	4	2	2	0	2	11	19
36	9	7	5	5	1	4	2	2	0	2	12	19
37	8	7	5	5	2	4	1	1	0	1	12	19
38	1	1	9	9	5	4	0	0	0	0	10	19
39	7	7	4	4	4	4	0	0	0	0	11	19
40	2	2	10	10	3	4	0	0	0	0	12	19
41	10	10	1	1	4	4	0	0	0	0	11	19
42	1	1	13	13	1	4	0	0	0	0	14	19
43	2	1	10	10	3	4	1	1	0	1	11	19
44	1	1	14	14	0	4	0	0	0	0	15	19
45	2	2	6	6	8	3	0	0	0	0	8	19
46	13	12	2	2	0	4	1	1	0	1	14	19
47	13	13	1	1	3	2	0	0	0	0	14	19
48	12	12	1	1	4	2	0	0	0	0	13	19
49	5	5	5	5	5	4	0	0	0	0	10	19
50	12	11	0	0	3	4	1	1	0	1	11	19
51	4	2	8	8	3	4	2	2	0	2	10	19
52	6	6	5	5	4	4	0	0	0	0	11	19
53	0	0	15	15	0	4	0	0	0	0	15	19
54	6	6	6	6	3	4	0	0	0	0	12	19
55	0	0	11	11	4	4	0	0	0	0	11	19
56	0	0	14	14	1	4	0	0	0	0	14	19
57	0	0	12	12	3	4	0	0	0	0	12	19
58	13	13	1	1	1	4	0	0	0	0	14	19
59	2	2	11	11	2	4	0	0	0	0	13	19
60	1	1	14	14	0	4	0	0	0	0	15	19
61	4	3	12	11	0	4	1	0	1	1	14	19
62	0	0	15	15	0	4	0	0	0	0	15	19

Position #	Count Active	Count Active (>40%, no errors)	Count Inactive	Count Inactive (no errors)	Count Unclassified	Count Not Made	Count Mutants 20-40% Activity	Count Mutants 20-40% Activity (no errors)	Count Errors	Count Mutants 20-40% Activity + Errors	Tested (w/o errors)	Total
63	12	12	3	3	0	4	0	0	0	0	15	19
64	0	0	17	17	0	2	0	0	0	0	17	19
65	1	0	14	13	1	4	0	0	1	1	13	19
66	2	2	13	13	0	4	0	0	0	0	15	19
67	5	4	7	6	4	4	1	0	1	1	10	19
68	10	7	8	5	0	4	1	0	3	3	12	19
69	13	12	2	1	1	4	0	0	1	1	13	19
70	14	14	1	1	0	4	0	0	0	0	15	19
71	10	10	1	1	4	4	0	0	0	0	11	19
72	11	9	7	6	1	1	1	1	1	2	15	19
73	13	13	1	1	1	4	0	0	0	0	14	19
74	15	15	0	0	0	4	0	0	0	0	15	19
75	12	12	3	3	1	3	0	0	0	0	15	19
76	0	0	14	14	1	4	0	0	0	0	14	19
77	2	0	8	8	5	4	2	2	0	2	8	19
78	0	0	7	7	8	4	0	0	0	0	7	19
79	3	3	11	11	1	4	0	0	0	0	14	19
80	0	0	15	15	0	4	0	0	0	0	15	19
81	1	0	12	11	3	4	0	0	1	1	11	19
82	14	13	2	1	1	3	1	0	1	1	14	19
83	11	10	2	1	2	5	1	0	1	1	11	19
84	15	14	1	0	0	4	1	0	1	1	14	19
85	1	1	11	11	3	4	0	0	0	0	12	19
86	17	16	2	1	0	1	0	0	1	1	17	19
87	15	14	1	0	0	4	1	0	1	1	14	19
88	0	0	15	15	0	4	0	0	0	0	15	19
89	6	3	9	8	1	4	3	2	1	3	11	19
90	13	12	2	1	1	4	0	0	1	1	13	19
91	3	3	9	9	2	5	0	0	0	0	12	19
92	6	5	9	8	1	4	1	0	1	1	13	19
93	15	13	0	0	0	4	2	2	0	2	13	19
94	13	12	2	2	0	4	1	1	0	1	14	19
95	0	0	15	15	0	4	0	0	0	0	15	19
96	3	3	9	9	3	4	0	0	0	0	12	19
97	15	15	0	0	0	4	0	0	0	0	15	19
98	13	13	1	1	1	4	0	0	0	0	14	19
99	3	2	8	8	4	4	1	1	0	1	10	19
100	0	0	11	11	4	4	0	0	0	0	11	19
101	0	0	13	13	2	4	0	0	0	0	13	19
102	14	14	1	1	0	4	0	0	0	0	15	19
103	1	0	13	13	1	4	1	1	0	1	13	19
104	9	8	3	3	3	4	1	1	0	1	11	19
105	12	9	3	2	2	3	3	2	1	3	11	19
106	1	1	12	12	2	4	0	0	0	0	13	19
107	3	3	9	9	3	4	0	0	0	0	12	19
108	1	1	11	11	4	3	0	0	0	0	12	19
109	0	0	8	8	7	4	0	0	0	0	8	19
110	1	1	6	6	8	4	0	0	0	0	7	19
111	0	0	3	3	13	3	0	0	0	0	3	19
112	0	0	8	8	7	4	0	0	0	0	8	19
113	0	0	2	2	13	4	0	0	0	0	2	19
114	5	5	5	5	5	4	0	0	0	0	10	19
115	0	0	14	14	1	4	0	0	0	0	14	19
116	0	0	14	14	1	4	0	0	0	0	14	19
117	1	0	10	9	5	4	0	0	1	1	9	19
118	7	7	8	8	0	4	0	0	0	0	15	19
119	4	3	7	6	5	4	1	0	1	1	9	19
120	14	14	0	0	1	4	0	0	0	0	14	19
121	0	0	12	12	4	3	0	0	0	0	12	19
122	1	1	11	11	3	4	0	0	0	0	12	19
123	0	0	14	14	1	4	0	0	0	0	14	19
124	3	1	5	5	8	3	2	2	0	2	6	19
125	4	3	6	6	4	5	1	1	0	1	9	19
126	0	0	7	7	11	1	0	0	0	0	7	19
127	13	13	1	1	1	4	0	0	0	0	14	19
128	10	10	2	2	3	4	0	0	0	0	12	19
129	0	0	13	13	2	4	0	0	0	0	13	19
130	2	2	10	10	3	4	0	0	0	0	12	19
131	14	14	1	1	0	4	0	0	0	0	15	19
132	14	14	1	1	0	4	0	0	0	0	15	19
133	1	1	13	13	1	4	0	0	0	0	14	19
134	3	3	12	12	0	4	0	0	0	0	15	19
135	14	14	1	1	0	4	0	0	0	0	15	19

Position #	Count Active	Count Active (>40%, no errors)	Count Inactive	Count Inactive (no errors)	Count Unclassified	Count Not Made	Count Mutants 20-40% Activity	Count Mutants 20-40% Activity (no errors)	Count Errors	Count Mutants 20-40% Activity + Errors	Tested (w/o errors)	Total
136	13	13	1	1	1	4	0	0	0	0	14	19
137	6	3	8	8	2	3	3	3	0	3	11	19
138	11	11	1	1	3	4	0	0	0	0	12	19
139	14	14	1	1	0	4	0	0	0	0	15	19
140	14	14	0	0	1	4	0	0	0	0	14	19
141	15	15	0	0	1	3	0	0	0	0	15	19
142	15	15	0	0	0	4	0	0	0	0	15	19
143	8	7	6	5	3	3	1	0	1	1	12	19
144	3	2	10	10	2	4	1	1	0	1	12	19
145	11	11	2	2	2	4	0	0	0	0	13	19
146	15	15	0	0	0	4	0	0	0	0	15	19
147	15	15	0	0	0	4	0	0	0	0	15	19
148	14	14	0	0	1	4	0	0	0	0	14	19
149	10	10	2	2	3	4	0	0	0	0	12	19
150	14	14	1	1	0	4	0	0	0	0	15	19
151	15	15	0	0	0	4	0	0	0	0	15	19
152	10	9	1	1	4	4	1	1	0	1	10	19
153	3	2	7	7	6	3	1	1	0	1	9	19
154	4	3	7	7	5	3	1	1	0	1	10	19
155	14	14	2	2	0	3	0	0	0	0	16	19
156	13	12	1	1	1	4	1	1	0	1	13	19
157	1	1	16	16	0	2	0	0	0	0	17	19
158	7	6	5	5	3	4	1	1	0	1	11	19
159	12	11	2	2	1	4	1	1	0	1	13	19
160	15	14	0	0	0	4	1	1	0	1	14	19
161	7	7	1	1	7	4	0	0	0	0	8	19
162	14	14	0	0	1	4	0	0	0	0	14	19
163	12	10	2	2	1	4	2	2	0	2	12	19
164	4	4	10	10	1	4	0	0	0	0	14	19
165	10	9	4	3	2	4	1	0	1	1	12	19
166	13	13	1	1	1	4	0	0	0	0	14	19
167	13	12	1	1	1	4	1	1	0	1	13	19
168	1	1	14	14	0	4	0	0	0	0	15	19
169	3	3	12	12	0	4	0	0	0	0	15	19
170	5	5	8	8	2	4	0	0	0	0	13	19
171	2	2	9	9	5	3	0	0	0	0	11	19
172	2	2	10	10	3	4	0	0	0	0	12	19
173	3	2	12	12	0	4	1	1	0	1	14	19
174	13	12	1	1	1	4	1	1	0	1	13	19
175	5	5	7	7	3	4	0	0	0	0	12	19
176	2	2	13	13	0	4	0	0	0	0	15	19
177	1	0	14	13	1	4	1	0	1	1	13	19
178	4	4	6	6	5	4	0	0	0	0	10	19
179	14	14	1	1	0	4	0	0	0	0	15	19
180	5	4	6	6	4	4	1	1	0	1	10	19
181	3	2	12	11	1	4	1	0	1	1	13	19
182	1	1	13	13	1	4	0	0	0	0	14	19
183	2	1	12	11	2	4	1	0	1	1	12	19
184	1	0	14	14	0	4	1	1	0	1	14	19
185	0	0	14	14	1	4	0	0	0	0	14	19
186	1	1	14	14	0	4	0	0	0	0	15	19
187	0	0	15	15	0	4	0	0	0	0	15	19
188	0	0	15	15	0	4	0	0	0	0	15	19
189	0	0	15	15	0	4	0	0	0	0	15	19
190	0	0	14	14	1	4	0	0	0	0	14	19
191	0	0	15	15	0	4	0	0	0	0	15	19
192	2	2	13	13	0	4	0	0	0	0	15	19
193	6	4	7	7	2	4	2	2	0	2	11	19
194	0	0	8	8	7	4	0	0	0	0	8	19
195	12	10	1	0	3	4	1	1	1	2	10	19
196	9	7	0	0	6	4	2	2	0	2	7	19
197	14	13	1	1	0	4	1	1	0	1	14	19
198	12	11	2	1	2	4	1	0	1	1	12	19
199	0	0	10	10	5	4	0	0	0	0	10	19
200	2	1	13	13	1	3	1	1	0	1	14	19
201	0	0	11	11	4	4	0	0	0	0	11	19
202	1	1	13	13	1	4	0	0	0	0	14	19
203	0	0	13	13	2	4	0	0	0	0	13	19
204	2	1	11	11	3	3	1	1	0	1	12	19
205	15	14	0	0	0	4	1	1	0	1	14	19
206	9	8	6	6	0	4	1	1	0	1	14	19
207	0	0	11	11	4	4	0	0	0	0	11	19
208	10	10	4	4	1	4	0	0	0	0	14	19

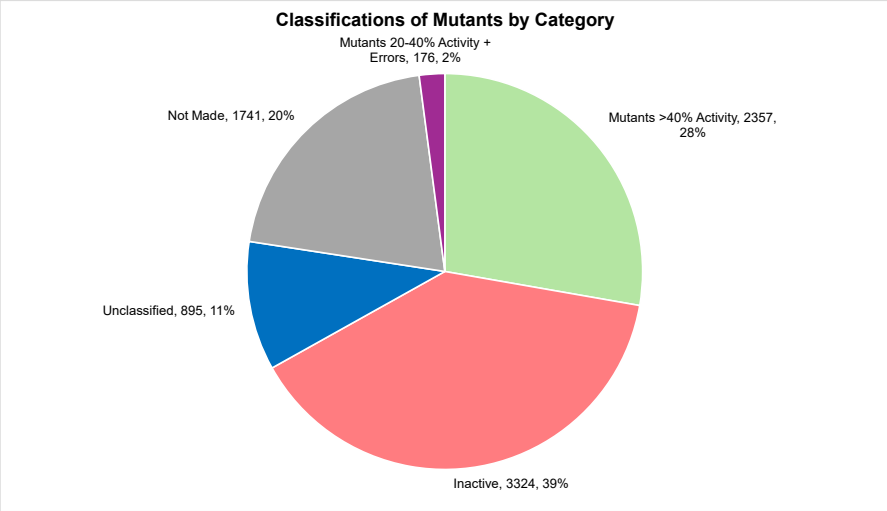
Position #	Count Active	Count Active (>40%, no errors)	Count Inactive	Count Inactive (no errors)	Count Unclassified	Count Not Made	Count Mutants 20-40% Activity	Count Mutants 20-40% Activity (no errors)	Count Errors	Count Mutants 20-40% Activity + Errors	Tested (w/o errors)	Total
209	9	8	2	2	4	4	1	1	0	1	10	19
210	0	0	14	14	1	4	0	0	0	0	14	19
211	2	1	13	12	1	4	0	0	1	1	13	19
212	3	3	10	10	2	4	0	0	0	0	13	19
213	13	13	2	2	0	4	0	0	0	0	15	19
214	1	1	13	13	1	4	0	0	0	0	14	19
215	13	13	2	2	0	4	0	0	0	0	15	19
216	1	1	14	14	0	4	0	0	0	0	15	19
217	1	1	10	10	4	4	0	0	0	0	11	19
218	3	1	7	6	6	4	2	1	1	2	7	19
219	13	13	1	1	1	4	0	0	0	0	14	19
220	9	7	6	6	0	4	2	2	0	2	13	19
221	7	6	6	6	2	4	1	1	0	1	12	19
222	10	10	2	2	3	4	0	0	0	0	12	19
223	0	0	15	15	0	4	0	0	0	0	15	19
224	1	1	13	13	1	4	0	0	0	0	14	19
225	0	0	12	12	2	5	0	0	0	0	12	19
226	1	0	15	14	0	4	0	0	1	1	14	19
227	0	0	15	15	0	4	0	0	0	0	15	19
228	0	0	13	13	2	4	0	0	0	0	13	19
229	0	0	10	10	5	4	0	0	0	0	10	19
230	1	1	14	14	0	4	0	0	0	0	15	19
231	1	1	14	14	0	4	0	0	0	0	15	19
232	2	2	10	10	2	5	0	0	0	0	12	19
233	7	7	5	5	3	4	0	0	0	0	12	19
234	2	2	11	11	2	4	0	0	0	0	13	19
235	6	6	6	6	3	4	0	0	0	0	12	19
236	6	6	7	7	2	4	0	0	0	0	13	19
237	12	11	0	0	3	4	1	1	0	1	11	19
238	8	8	7	7	0	4	0	0	0	0	15	19
239	1	0	13	13	1	4	1	1	0	1	13	19
240	8	8	6	6	1	4	0	0	0	0	14	19
241	0	0	12	12	3	4	0	0	0	0	12	19
242	1	1	13	13	1	4	0	0	0	0	14	19
243	0	0	13	13	2	4	0	0	0	0	13	19
244	0	0	6	6	9	4	0	0	0	0	6	19
245	1	1	10	10	4	4	0	0	0	0	11	19
246	0	0	15	15	0	4	0	0	0	0	15	19
247	3	3	12	12	0	4	0	0	0	0	15	19
248	4	4	8	8	3	4	0	0	0	0	12	19
249	1	1	9	9	5	4	0	0	0	0	10	19
250	0	0	15	15	0	4	0	0	0	0	15	19
251	3	2	9	9	3	4	1	1	0	1	11	19
252	0	0	14	14	1	4	0	0	0	0	14	19
253	1	1	12	12	2	4	0	0	0	0	13	19
254	0	0	14	14	1	4	0	0	0	0	14	19
255	6	5	6	6	3	4	1	1	0	1	11	19
256	4	4	5	5	6	4	0	0	0	0	9	19
257	12	11	1	1	2	4	1	1	0	1	12	19
258	5	4	4	4	6	4	1	1	0	1	8	19
259	14	14	0	0	1	4	0	0	0	0	14	19
260	11	10	2	2	2	4	1	1	0	1	12	19
261	10	10	1	1	4	4	0	0	0	0	11	19
262	0	0	15	15	0	4	0	0	0	0	15	19
263	7	6	5	5	3	4	1	1	0	1	11	19
264	2	2	11	11	2	4	0	0	0	0	13	19
265	1	1	12	12	5	1	0	0	0	0	13	19
266	1	1	12	12	2	4	0	0	0	0	13	19
267	2	2	9	9	4	4	0	0	0	0	11	19
268	0	0	14	14	1	4	0	0	0	0	14	19
269	4	4	8	8	3	4	0	0	0	0	12	19
270	4	4	9	9	3	3	0	0	0	0	13	19
271	6	6	7	7	2	4	0	0	0	0	13	19
272	6	4	6	6	3	4	2	2	0	2	10	19
273	3	3	11	11	1	4	0	0	0	0	14	19
274	3	3	8	8	4	4	0	0	0	0	11	19
275	2	0	11	9	4	4	1	0	2	2	9	19
276	11	10	3	3	1	4	1	1	0	1	13	19
277	14	14	0	0	1	4	0	0	0	0	14	19
278	13	13	2	2	0	4	0	0	0	0	15	19
279	5	4	7	6	4	4	1	0	1	1	10	19
280	2	2	10	10	3	4	0	0	0	0	12	19
281	0	0	13	13	2	4	0	0	0	0	13	19

Position #	Count Active	Count Active (>40%, no errors)	Count Inactive	Count Inactive (no errors)	Count Unclassified	Count Not Made	Count Mutants 20-40% Activity	Count Mutants 20-40% Activity (no errors)	Count Errors	Count Mutants 20-40% Activity + Errors	Tested (w/o errors)	Total
282	4	4	5	5	6	4	0	0	0	0	9	19
283	5	5	5	5	5	4	0	0	0	0	10	19
284	11	9	3	3	1	4	2	2	0	2	12	19
285	8	7	5	5	2	4	1	1	0	1	12	19
286	3	2	10	10	2	4	1	1	0	1	12	19
287	3	2	12	11	1	4	1	0	1	1	13	19
288	2	2	10	10	3	4	0	0	0	0	12	19
289	2	1	11	10	3	4	0	0	1	1	11	19
290	2	2	3	3	10	4	0	0	0	0	5	19
291	5	4	10	9	3	2	1	0	1	1	13	19
292	11	11	3	3	1	4	0	0	0	0	14	19
293	13	13	2	2	0	4	0	0	0	0	15	19
294	1	1	13	13	1	4	0	0	0	0	14	19
295	0	0	10	10	5	4	0	0	0	0	10	19
296	0	0	13	13	2	4	0	0	0	0	13	19
297	2	1	11	10	3	4	1	1	0	1	11	19
298	2	2	11	11	2	4	0	0	0	0	13	19
299	0	0	11	11	4	4	0	0	0	0	11	19
300	1	1	14	14	0	4	0	0	0	0	15	19
301	2	2	12	12	1	4	0	0	0	0	14	19
302	2	2	13	13	0	4	0	0	0	0	15	19
303	2	1	12	11	2	4	1	0	1	1	12	19
304	2	0	13	11	2	4	0	0	2	2	11	19
305	3	2	8	8	4	4	1	1	0	1	10	19
306	3	3	8	8	4	4	0	0	0	0	11	19
307	9	8	3	3	3	4	1	1	0	1	11	19
308	8	8	7	7	0	4	0	0	0	0	15	19
309	14	14	0	0	1	4	0	0	0	0	14	19
310	8	6	5	4	3	4	2	1	1	2	10	19
311	6	6	8	8	1	4	0	0	0	0	14	19
312	5	3	5	5	6	3	2	2	0	2	8	19
313	12	12	1	1	2	4	0	0	0	0	13	19
314	10	10	3	3	2	4	0	0	0	0	13	19
315	10	10	3	3	2	4	0	0	0	0	13	19
316	1	1	13	13	1	4	0	0	0	0	14	19
317	12	12	2	2	1	4	0	0	0	0	14	19
318	12	12	3	3	0	4	0	0	0	0	15	19
319	0	0	15	15	0	4	0	0	0	0	15	19
320	13	11	3	2	1	3	2	1	1	2	13	19
321	8	8	3	3	4	4	0	0	0	0	11	19
322	0	0	13	13	2	4	0	0	0	0	13	19
323	3	3	11	11	1	4	0	0	0	0	14	19
324	7	7	6	6	2	4	0	0	0	0	13	19
325	12	12	2	2	1	4	0	0	0	0	14	19
326	6	5	5	5	4	4	1	1	0	1	10	19
327	1	1	13	13	1	4	0	0	0	0	14	19
328	14	13	0	0	1	4	1	1	0	1	13	19
329	0	0	15	15	0	4	0	0	0	0	15	19
330	0	0	14	14	1	4	0	0	0	0	14	19
331	3	1	13	11	1	4	2	0	2	2	12	19
332	0	0	15	15	0	4	0	0	0	0	15	19
333	0	0	10	10	5	4	0	0	0	0	10	19
334	2	1	9	9	4	4	1	1	0	1	10	19
335	1	1	10	10	4	4	0	0	0	0	11	19
336	0	0	13	13	2	4	0	0	0	0	13	19
337	0	0	10	10	5	4	0	0	0	0	10	19
338	1	1	13	13	1	4	0	0	0	0	14	19
339	1	1	13	13	1	4	0	0	0	0	14	19
340	0	0	14	14	1	4	0	0	0	0	14	19
341	0	0	14	14	1	4	0	0	0	0	14	19
342	1	1	12	12	2	4	0	0	0	0	13	19
343	2	2	6	6	7	4	0	0	0	0	8	19
344	0	0	13	13	2	4	0	0	0	0	13	19
345	0	0	11	11	4	4	0	0	0	0	11	19
346	0	0	14	14	1	4	0	0	0	0	14	19
347	7	6	7	7	1	4	1	1	0	1	13	19
348	3	3	11	11	1	4	0	0	0	0	14	19
349	7	7	7	7	3	2	0	0	0	0	14	19
350	0	0	15	15	0	4	0	0	0	0	15	19
351	5	2	9	8	2	4	3	2	1	3	10	19
352	0	0	15	15	0	4	0	0	0	0	15	19
353	2	2	11	11	2	4	0	0	0	0	13	19
354	0	0	15	15	0	4	0	0	0	0	15	19

Position #	Count Active	Count Active (>40%, no errors)	Count Inactive	Count Inactive (no errors)	Count Unclassified	Count Not Made	Count Mutants 20-40% Activity	Count Mutants 20-40% Activity (no errors)	Count Errors	Count Mutants 20-40% Activity + Errors	Tested (w/o errors)	Total
355	0	0	15	15	0	4	0	0	0	0	15	19
356	4	4	9	9	2	4	0	0	0	0	13	19
357	5	4	8	8	2	4	1	1	0	1	12	19
358	4	3	8	8	2	5	1	1	0	1	11	19
359	7	7	6	6	2	4	0	0	0	0	13	19
360	1	1	13	13	1	4	0	0	0	0	14	19
361	1	0	12	12	2	4	1	1	0	1	12	19
362	0	0	15	15	0	4	0	0	0	0	15	19
363	0	0	15	15	0	4	0	0	0	0	15	19
364	0	0	15	15	0	4	0	0	0	0	15	19
365	0	0	14	14	2	3	0	0	0	0	14	19
366	0	0	12	12	3	4	0	0	0	0	12	19
367	6	6	7	7	2	4	0	0	0	0	13	19
368	11	11	3	3	1	4	0	0	0	0	14	19
369	3	3	10	10	2	4	0	0	0	0	13	19
370	0	0	14	14	1	4	0	0	0	0	14	19
371	11	10	2	2	2	4	1	1	0	1	12	19
372	0	0	15	15	0	4	0	0	0	0	15	19
373	9	9	3	3	4	3	0	0	0	0	12	19
374	12	11	2	2	1	4	1	1	0	1	13	19
375	10	9	5	5	0	4	1	1	0	1	14	19
376	11	10	3	3	1	4	1	1	0	1	13	19
377	8	8	4	4	3	4	0	0	0	0	12	19
378	3	3	10	10	2	4	0	0	0	0	13	19
379	5	5	8	8	2	4	0	0	0	0	13	19
380	7	5	7	7	1	4	2	2	0	2	12	19
381	8	7	5	5	2	4	1	1	0	1	12	19
382	0	0	14	14	1	4	0	0	0	0	14	19
383	11	11	2	2	2	4	0	0	0	0	13	19
384	0	0	6	6	9	4	0	0	0	0	6	19
385	9	9	6	6	0	4	0	0	0	0	15	19
386	0	0	15	15	0	4	0	0	0	0	15	19
387	1	1	12	12	2	4	0	0	0	0	13	19
388	9	9	5	5	1	4	0	0	0	0	14	19
389	12	12	2	2	1	4	0	0	0	0	14	19
390	0	0	15	15	0	4	0	0	0	0	15	19
391	1	1	14	14	0	4	0	0	0	0	15	19
392	13	12	2	2	0	4	1	1	0	1	14	19
393	11	11	2	2	2	4	0	0	0	0	13	19
394	2	2	13	13	0	4	0	0	0	0	15	19
395	7	6	4	4	4	4	1	1	0	1	10	19
396	8	7	6	6	1	4	1	1	0	1	13	19
397	0	0	12	12	3	4	0	0	0	0	12	19
398	1	0	15	14	0	4	1	0	1	1	14	19
399	12	12	2	2	1	4	0	0	0	0	14	19
400	0	0	15	15	0	4	0	0	0	0	15	19
401	5	3	7	7	3	4	2	2	0	2	10	19
402	0	0	14	14	1	4	0	0	0	0	14	19
403	1	1	13	13	1	4	0	0	0	0	14	19
404	3	2	12	12	0	4	1	1	0	1	14	19
405	11	10	3	3	1	4	1	1	0	1	13	19
406	12	12	2	2	1	4	0	0	0	0	14	19
407	14	14	0	0	1	4	0	0	0	0	14	19
408	0	0	14	14	1	4	0	0	0	0	14	19
409	12	12	0	0	3	4	0	0	0	0	12	19
410	11	11	1	1	3	4	0	0	0	0	12	19
411	8	7	4	4	3	4	1	1	0	1	11	19
412	13	11	2	1	0	5	2	1	1	2	12	19
413	9	4	5	3	3	4	4	3	2	5	7	19
414	4	2	9	8	3	4	2	1	1	2	10	19
415	5	3	4	4	5	5	2	2	0	2	7	19
416	12	12	2	2	1	4	0	0	0	0	14	19
417	1	1	11	11	3	4	0	0	0	0	12	19
418	14	14	0	0	1	4	0	0	0	0	14	19
419	12	12	2	2	1	4	0	0	0	0	14	19
420	2	2	13	13	0	4	0	0	0	0	15	19
421	14	13	0	0	1	4	1	1	0	1	13	19
422	2	2	11	11	4	2	0	0	0	0	13	19
423	0	0	15	15	0	4	0	0	0	0	15	19
424	0	0	11	11	4	4	0	0	0	0	11	19
425	8	7	5	4	3	4	1	0	1	1	11	19
426	8	6	4	4	3	4	2	2	0	2	10	19
427	6	3	8	8	1	4	3	3	0	3	11	19

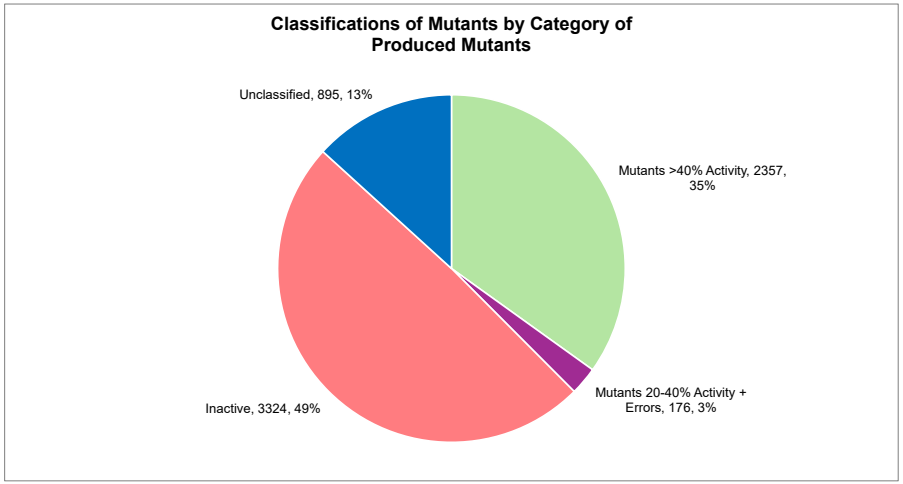
Position #	Count Active	Count Active (>40%, no errors)	Count Inactive	Count Inactive (no errors)	Count Unclassified	Count Not Made	Count Mutants 20-40% Activity	Count Mutants 20-40% Activity (no errors)	Count Errors	Count Mutants 20-40% Activity + Errors	Tested (w/o errors)	Total
428	4	4	10	10	1	4	0	0	0	0	14	19
429	0	0	11	11	4	4	0	0	0	0	11	19
430	0	0	9	9	6	4	0	0	0	0	9	19
431	14	14	1	1	0	4	0	0	0	0	15	19
432	6	5	8	8	1	4	1	1	0	1	13	19
433	15	15	0	0	0	4	0	0	0	0	15	19
434	5	5	6	6	3	5	0	0	0	0	11	19
435	10	10	0	0	5	4	0	0	0	0	10	19
436	15	14	0	0	0	4	1	1	0	1	14	19
437	12	12	1	1	2	4	0	0	0	0	13	19
438	14	14	1	1	0	4	0	0	0	0	15	19
439	13	13	2	2	0	4	0	0	0	0	15	19
440	14	14	1	1	0	4	0	0	0	0	15	19
441	13	13	1	1	1	4	0	0	0	0	14	19
442	12	11	3	3	0	4	1	1	0	1	14	19
443	13	13	1	1	1	4	0	0	0	0	14	19
444	13	12	0	0	2	4	1	1	0	1	12	19
445	14	14	0	0	4	1	0	0	0	0	14	19
446	15	15	0	0	0	4	0	0	0	0	15	19
447	14	14	0	0	1	4	0	0	0	0	14	19
	2533	2357	3375	3324	895	1741	159	126	50	176	5681	8493

Mutants >40% Activity	Inactive	Unclassified	Not Made	Mutants 20-40% Activity + Errors	Errors	Total
2357	3324	895	1741	176	50	8493



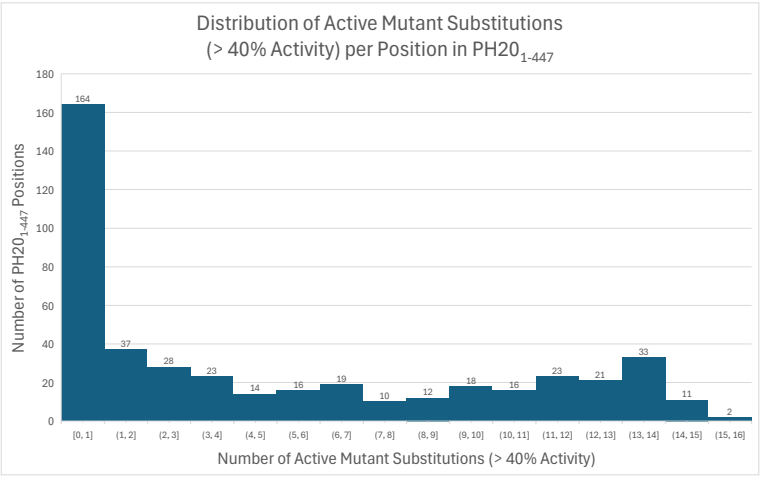
Numbers of mutants by Category (made)

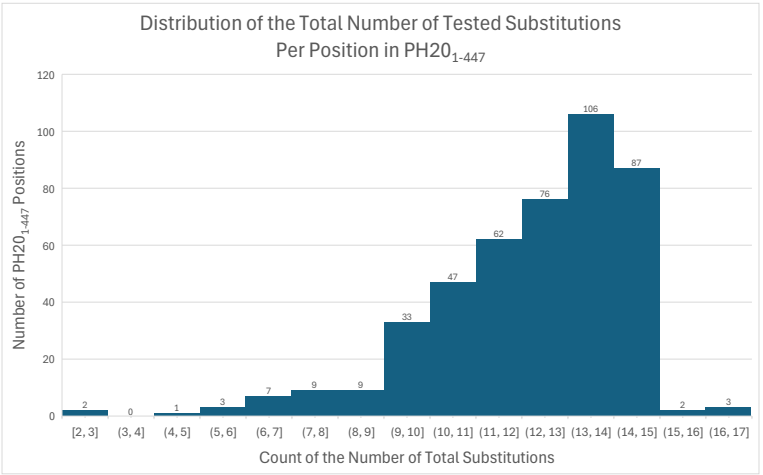
Mutants >40% Activity	Mutants 20-40% Activity + Errors	Inactive	Unclassified
2357	176	3324	895



Note:
Information in this sheet was obtained from EX1026 ('731 App.), at Tables 3, 5, 9, and 10.

Information in this sheet was also obtained from EX2166 (Petsko Patent Analysis Data).





Appendix B-2

$PH20_{1-447}$ Physiochemical and Activity Data

Source(s): EX1026 ('731 App.), at Tables 3, 5, 9, and 10; EX2166 (Petsko Patent Analysis Data); EX1004 (Park Dec.); EX2176 (Park Analysis Spreadsheet)

Position	Wildtype (WT)	G	A	P	V	L	I	M	C	T	S	N	Q	D	E	H	K	R	Y	F	W
1	L	0.62	0.95	0.92	1.00				0.89	0.99	0.74	0.87	3.27		0.55	1.90	1.39	0.72		0.41	0.88
2	N	0.44	0.61	0.54	0.65	0.46			0.40	1.05	0.78		0.84								
3	F					0.59									0.42	0.68			0.50		
4	R		0.73		1.09		0.54			0.66	0.60										
5	A															0.44					
6	P		0.78			0.76						0.40	0.89			0.58	0.80	0.56			
7	P							0.57													
8	V					0.53	1.17	0.47													
9	I				0.84	1.08					0.98		0.40				0.69	0.53			
10	P	0.55									0.55	0.55	0.89	0.62	0.66	0.43		0.73			0.59
11	N	0.45												0.54		0.69	0.58				
12	V		0.56			0.44	0.68			1.50	0.75	0.46			1.86		0.65	0.50			
13	P									0.90	0.68					0.46			0.51		
14	F				0.46		0.42	0.47						0.64							
15	L		0.65		2.20			0.45													
16	W																				
17	A																				
18	W																				
19	N																				
20	A										0.50										
21	P																				
22	S							0.49		0.48						0.57			0.45		
23	E													0.97							
24	F	0.75	0.69		1.15	0.62	0.70	0.85		1.18		0.60			3.99	2.07	0.96	1.22	0.90		
25	C																				
26	L	0.81	1.30	0.55	0.52		0.51	1.43		0.87	0.78		1.44		3.22	0.97	1.88	1.43	0.52		0.53
27	G		0.79	0.46		0.76	0.41			0.61	0.94		1.12	1.22	1.18	1.11	2.71	1.88		0.61	0.76
28	K		0.78	0.40	0.76	0.51	0.55	0.67		0.68	0.46	0.58		0.62	0.54			0.71		0.75	0.51
29	F	1.05	0.90	3.79	1.65	2.36	1.53	2.08		0.85	2.21				4.03	0.82	1.34	1.24			0.48
30	D	2.02	1.12	1.19	0.46	1.32		1.85		0.57	1.62		0.84			1.69	2.63	1.82		0.84	0.62
31	E	1.27	2.05	1.51	4.57	2.62	3.89		2.95	3.96	1.70					2.74	3.13	2.27	1.13		1.26
32	P	1.60	0.92		0.81	0.82		0.67	0.40	0.77	1.01	0.70	1.11			2.08	1.04	1.17	1.01	2.71	0.54
33	L	0.57		0.87				0.69		0.45	0.48		0.45					0.61			1.58
34	D												0.59		0.58	0.41	0.54	1.17			0.46
35	M					0.52				0.83						0.60			0.78	0.87	
36	S	0.64	0.45			0.71				0.51						0.54	0.83	1.09			
37	L			0.63	0.57		0.62	1.46									0.43	0.51		3.33	
38	F																		1.29		
39	S		1.06			0.80				1.57		2.32	1.10					0.56	0.56		
40	F					0.92															1.11
41	I	0.76	0.67		0.73				0.53	1.47		0.40		0.78	0.51	0.77					0.66
42	G		0.64																		
43	S									0.43											
44	P														0.59						
45	R						0.45										0.53				
46	I		1.04		1.01	1.08		1.00		0.55	0.64	0.66			0.43	0.82		2.29	0.76	0.73	
47	N	0.82	0.48					0.77		1.49	0.85		0.69	0.82		1.16	0.67	0.84	0.45	1.32	0.63
48	A	0.83			0.60		0.64	0.76			1.06	4.25	1.05			1.99	1.28	0.66	0.81	2.51	
49	T				0.45		0.42				0.92						0.85	1.41			
50	G		0.93			0.43		0.47	0.41		1.24		0.86	1.37	0.78	0.74		0.86	0.58		
51	Q										0.46	0.60									
52	G			0.43						0.49	1.32	0.89	3.71					0.53			
53	V																				
54	T		0.43		0.66						0.70	0.48	0.91							0.56	
55	I																				
56	F																				
57	Y																				
58	V	0.54		0.90		1.54	0.57		0.55		0.83	0.49	4.54			1.09	4.08	1.92	1.07		0.65
59	D											1.27	0.40								
60	R																0.69				
61	L				0.59		0.42	0.73													
62	G																				
63	Y		0.63		0.43	1.33	1.03	1.32		1.07	1.00	0.96				1.07	1.36	1.40			1.53
64	Y																				
65	P																	0.57			
66	Y															0.47		0.51			
67	I				1.80	0.45													0.55	1.00	
68	D									0.75	0.81		1.67		0.72	2.06	1.08	0.70			
69	S	6.06	22.06	8.14		3.44	3.12	2.67	1.97						1.48			14.06	2.71	8.75	2.18
70	I	6.22	27.00	3.03	6.34	3.05			2.57	5.43	3.63	6.19				9.09	14.64	13.95	1.26	5.69	
71	T	1.41	0.86			1.09		0.89			1.54	1.21	0.68	0.50		0.93		2.17			
72	G		0.45			0.43		3.11			0.51		2.33	0.60	0.69		1.39	0.65			
73	V	1.17	1.38			1.59		0.68	0.84	1.34	0.86		0.96	0.94		1.54	1.42	0.72			1.91
74	T	2.75	2.28	2.45	2.27	1.43		0.52	2.18		1.80	2.12			1.38	1.40	1.29	2.22		1.43	2.13
75	V		0.71			5.22		1.16	0.46	4.34	0.76	0.81	1.51			0.62		3.02	0.63	2.00	
76	N																				
77	G																				
78	G																				
79	I				1.01	1.44				0.79											

Position	Wildtype (WT)	G	A	P	V	L	I	M	C	T	S	N	Q	D	E	H	K	R	Y	F	W
80	P																				
81	Q																				
82	K	0.64	0.94		0.57	0.87	1.01	0.58		0.56	0.62	0.96	0.76		0.50	0.44		0.85			
83	I	1.05			0.99	0.93				0.95	0.79	0.82	1.07			0.40		0.45		0.57	
84	S	8.68		0.57		0.92	0.90	0.77		0.82		0.89	0.86	0.98	0.52	0.96		1.89		0.72	0.86
85	L				0.42																
86	Q	1.02	2.70		0.97	0.92	0.65	1.06		0.58	0.85	1.28		0.88	1.18	1.70	0.97	0.93		0.54	1.21
87	D	1.00	1.00		0.66	0.55	0.53	0.58	1.77	1.70	0.99		1.05		0.86	0.72		1.28	2.72		
88	H																				
89	L							0.63	1.46								0.45				
90	D		1.48			1.15	1.10			1.02	1.15	1.18	1.11		1.15	1.24	1.36	1.49			0.81
91	K		0.89										0.43					0.67			
92	A				1.09	1.29		0.86	1.97	0.70											
93	K	0.97				1.53	3.25	0.70		3.93	1.25	0.71	0.84	0.71	0.83	0.61		1.52		0.50	
94	K		0.64			0.52		0.66		1.14	0.94	0.99	1.22	0.93	0.79	0.72		3.94		0.59	
95	D																				
96	I				0.68	0.46								0.69							
97	T	4.84	1.25	0.62		1.22	0.85		0.53		1.21	1.10	1.17	1.31	1.19			0.95	0.74	0.75	0.53
98	F		0.60		0.46	0.58	0.52	0.87	0.58		0.56		0.65	0.47	0.44	1.06		0.72			0.81
99	Y										0.43							0.53			
100	M																				
101	P																				
102	V	0.67	0.83			0.71		0.77	0.69	1.26	1.41	1.02	1.03		0.90	0.88	1.03	0.94			0.76
103	D																				
104	N	0.48	0.69					0.61	0.41	0.71	1.03						0.88	1.25			
105	L	0.51	0.54	0.84	0.99		0.94			0.51	0.61		0.90					0.65			
106	G				0.43																
107	M					1.32	0.67													0.91	
108	A	0.47																			
109	V																				
110	I				0.51																
111	D																				
112	W																				
113	E																				
114	E	0.73	1.44					0.44			0.69					0.75					
115	W																				
116	R																				
117	P													0.56							
118	T				0.79	1.09		0.53				0.67	3.37			0.47	0.53				
119	W												0.72						1.08	0.53	
120	A	1.03		0.42	1.53	1.25	1.33			0.62	1.21	0.81		0.76		1.11		0.82	1.95	2.62	0.59
121	R																				
122	N							0.56													
123	W																				
124	K																	0.62			
125	P										0.54					0.43		0.63			
126	K																				
127	D	0.97	0.89		0.56	0.84		0.40		1.11	0.77	1.69	1.21		1.31	2.33		0.51			0.44
128	V	0.49	0.53			0.95	1.25		0.68		0.53		0.55				1.16	0.74			0.50
129	Y																				
130	K						0.50											1.42			
131	N	2.47			2.08	0.82	1.40	0.99	0.60	1.02	0.76		1.24		0.44	0.80		2.81	0.85	0.63	
132	R		0.68		0.73	0.76	0.56		0.58	0.61	0.79	1.28	0.69		0.70	0.66	1.05		0.78	0.60	
133	S						0.54														
134	I				1.08	1.04				0.60											
135	E	2.79	0.99			0.82			0.77		1.13	0.56	1.59	2.68		0.79	1.15	2.08	0.50	0.73	0.63
136	L		0.73				0.65	1.05	0.56	0.72	0.80	0.48	0.61	0.47		1.00		0.74		0.96	1.11
137	V		0.48				1.01			0.51											
138	Q		0.69			0.59	0.54	0.68	0.65		0.48	0.61				0.71		0.53	0.60		0.41
139	Q	0.65	0.92		0.77	0.70		0.95	0.44	1.31	0.81			0.48	0.94	0.56	0.73	0.79		0.53	
140	Q	0.73	0.96		0.61	0.51	0.75	0.80	0.50					0.59		0.84	0.93	0.85	0.41	0.66	0.59
141	N	1.15	1.12		0.50	0.61		0.48		0.45	0.72		1.16	1.09	0.67	2.03		1.40	1.55	0.81	0.83
142	V	0.98		0.88		0.75	0.81	0.76	0.61	1.19	0.93	0.98	1.04	0.71	0.87	1.11	1.40	1.53			
143	Q	0.62			0.57	0.56	0.44					0.73			0.77		1.30				
144	L									1.02											0.79
145	S	0.94	0.58	1.04		0.44		0.56	0.44			0.58		0.48	0.56	0.56		0.97			
146	L	0.62	0.52	0.93	0.84		0.82		0.42	0.74	0.71	0.57	0.84		0.50	0.78	0.84	1.47	0.80		
147	T	1.05	1.20	1.09	2.04	1.30	0.85	0.79	0.47		1.27		1.29	0.71				2.11	1.04	1.24	0.97
148	E	1.05			0.99	0.85	0.73		0.66	0.82	1.15		1.44			1.24	1.63	0.97	0.95	0.42	0.43
149	A	0.52			1.34	0.88		0.88	1.15	1.24	1.08		1.15				0.51	1.02			
150	T	2.19	1.21	0.88		0.70	0.52		0.70		0.92	0.91		1.24	1.05			0.90	1.36	0.71	1.25
151	E	1.06	1.27		1.38	1.03		1.26	1.00	1.21	1.28	0.95	2.01			1.34	2.05	1.61	1.31		1.31
152	K		0.51		0.82		0.65	0.75	0.52	1.20								1.85	0.67	0.61	
153	A					0.51	0.93														
154	K				0.46					0.83								0.86			
155	Q	1.61	0.91		0.73	0.86		0.97	0.60	0.76	0.77			1.49		1.03	1.57	1.27		0.70	0.91
156	E	0.49	0.79		0.45	0.43	0.51	0.87		0.69	0.62		0.84	1.95				0.43			0.49
157	F																				0.61
158	E		0.56			0.44					0.95		1.25			0.54				0.51	

Position	Wildtype (WT)	G	A	P	V	L	I	M	C	T	S	N	Q	D	E	H	K	R	Y	F	W
159	K		0.64		0.41	0.62		0.66			0.67	0.73	0.92	0.52	0.49	0.74		0.88			
160	A	0.75			0.73	0.67	0.43	0.77	0.61		1.35	0.56	0.65			0.47	0.91	0.89	1.07	0.79	
161	G		0.99		0.42				0.44		0.77			0.86	0.49			0.48			
162	K	0.56	0.50	0.64	0.52	0.54		1.04			0.47		0.58	0.77	0.51	0.62		0.52	0.72		1.01
163	D	1.15	1.52		0.76	1.18				1.13	1.34		1.40		1.63		1.90	1.80			
164	F				1.23	1.13		1.66													0.72
165	L		0.48		1.22						1.31	2.19		5.79				0.59	0.66	1.23	1.14
166	V	1.11	2.85			4.38			1.16	4.26		0.47	3.61		1.28	1.74		5.56	2.08	1.67	1.26
167	E	0.60	0.84	0.58				0.87		0.59	1.17	0.83		0.69		0.89	0.91	1.02	0.55		
168	T															0.46					
169	I				0.74	2.08												0.54			
170	K		0.40		0.58							0.72	0.40					2.58			
171	L				0.64		0.73														
172	G		1.20						1.03												
173	K											0.44						0.82			
174	L	0.40	1.20		0.62			0.79		1.12	0.85	1.36	0.99				2.39	1.50	1.06		0.78
175	L				0.94					1.43					0.43	0.57			0.66		
176	R					0.40											0.67				
177	P																				
178	N	0.85						0.88									0.85	1.10			
179	H	0.86	1.06	0.44	0.42	0.73	0.90	0.63	0.94	0.43	0.51	0.96			0.62		1.39	0.96			
180	L	0.62						0.64									0.44			0.59	
181	W							0.88					0.88								
182	G					0.90															
183	Y					0.70															
184	Y																				
185	L																				
186	F																		0.59		
187	P																				
188	D																				
189	C																				
190	Y																				
191	N																				
192	H									0.50	0.49										
193	H	0.68									0.42		0.82						0.58		
194	Y																				
195	K	0.45	0.51			0.45	0.50			0.58		0.74	0.71			0.45		0.85			0.49
196	K	0.41				0.65				1.18	0.68				0.43			0.58			0.55
197	P	0.75	0.81			0.56		1.03		0.41	0.70		0.69	0.58	0.52	0.62	0.99	0.58		0.48	
198	G		0.80			0.48				0.41	0.76	0.80	0.55	1.99	0.49	0.84		0.58	0.81		
199	Y																				
200	N													0.46							
201	G																				
202	S							0.40													
203	C																				
204	F			0.63																	
205	N	0.79	1.30	0.75	0.49			0.58		0.85	0.80			0.85	1.94		0.76	0.54	0.40	0.52	0.41
206	V					1.57	0.94	0.43		0.59	0.72					0.50	1.75	1.30			
207	E																				
208	I		0.62		1.07	0.84		0.88	0.48	1.01	0.62		0.77				0.91	1.14			
209	K	0.44	0.53							0.50	0.50	0.50			0.46			0.68		0.40	
210	R																				
211	N					0.41															
212	D									0.76	0.93	1.52									
213	D	0.81	0.85		0.47	0.56		1.56				1.53	1.04		0.79	0.75	0.82	0.92	0.49		0.49
214	L												0.57								
215	S	0.88	0.74		0.69	0.60		1.77		0.80			0.79	0.62	0.74	0.91	0.99	0.71			0.52
216	W																		0.48		
217	L							0.51													
218	W																			0.57	
219	N		1.29			0.72	0.60	1.02	0.43	0.82	2.48			0.75	0.95	0.97	1.45	1.10			0.48
220	E		0.75		1.35	1.45	1.34			0.91	0.62					1.40					
221	S		0.72		1.04			0.46	0.59	0.94			1.37								
222	T	0.49			0.79	0.64	0.40				0.63	0.80		0.43			0.75	0.75		0.43	
223	A																				
224	L						0.61														
225	Y																				
226	P																				
227	S																				
228	I																				
229	Y																				
230	L						0.87														
231	N									1.10											
232	T										0.76									0.73	
233	Q	0.46	0.71			0.69											1.69	1.50	0.50	0.53	
234	Q					0.40		1.65													
235	S	0.95	0.47							0.66					1.00	0.44	0.53				
236	P	1.09	1.07								0.91					0.46	0.71	3.09			
237	V		0.90			1.12				1.01	1.03	0.67	1.46		1.93	0.75		0.71		0.41	0.52

Position	Wildtype (WT)	G	A	P	V	L	I	M	C	T	S	N	Q	D	E	H	K	R	Y	F	W
238	A									0.44	2.62		1.02	0.75	0.59	0.60	0.60	0.49			
239	A																				
240	T		0.48	0.56	0.77			0.48			0.74		0.75				1.13	0.91			
241	L																				
242	Y																			1.08	
243	V																				
244	R																				
245	N															0.50					
246	R																				
247	V					0.83	2.01	0.52													
248	R		0.43													0.40			0.67		0.52
249	E																				
250	A																				
251	I					0.58		0.43													
252	R																				
253	V						0.76														
254	S																				
255	K		0.40								0.43	0.52	0.91					0.71			
256	I		0.42		0.51	0.64										0.51					
257	P	0.51	0.82		0.88	0.69	1.07	0.90		2.04		0.69	0.61				0.92	1.38			
258	D										1.44	1.44				0.84		0.45			
259	A	0.68		1.54	0.41	0.53	0.46			0.51	0.63	0.49	0.70		0.85		0.76	0.72	0.51		0.55
260	K		0.66			0.60		0.85			0.66		0.58	0.41	0.58	0.87		0.83	1.73		
261	S		0.74		0.48			0.56		0.66		1.98	0.76				2.54	1.19		0.73	0.44
262	P																				
263	L		0.76		0.75			0.89		0.49							2.73	1.63			
264	P		0.43													0.60					
265	V						0.58														
266	F																		0.58		
267	A							0.45		1.34											
268	Y																				
269	T		1.63						0.75		1.01			0.76							
270	R							0.46		0.40	0.69	0.52									
271	I	1.29			1.05	10.62		3.24			0.42									0.72	
272	V									1.61	0.96			1.36				0.74			
273	F									0.48						1.41			0.90		
274	T		0.51								0.62									1.28	
275	D																				
276	Q					0.48	0.51	1.14	0.88		1.63			1.69	1.05	1.20		1.30	1.94		
277	V	1.18	0.65					0.94	0.41	1.94	0.83	1.15	0.82	0.79	1.02	1.09	1.51	1.63	0.66		
278	L	1.33	1.13		0.44		0.93			1.66	1.67	1.74			1.03	4.50	1.75	5.87	1.51	1.26	
279	K									0.86			0.84			0.44		1.10			
280	F	0.47											0.43								
281	L																				
282	S	0.54						2.64					0.41	0.41							
283	Q			1.18						0.65	1.73				0.63			0.59			
284	D	0.60	0.58			0.50		0.56			0.99	0.40	0.95		1.21	0.51					
285	E	0.52						0.43				0.40	0.59			1.30			0.99	0.47	
286	L										0.46							0.53			
287	V						0.51			0.50											
288	Y					0.79															0.49
289	T																0.75				
290	F						0.41	1.03													
291	G				1.63						0.45		0.80					0.45			
292	E	0.41	0.66	1.05	1.28				0.71			0.99				1.26	1.27	0.42		0.90	0.83
293	T	1.00	1.90	1.64	2.15	1.00		2.29	1.67		0.89		1.83	1.46			1.35		1.49	1.94	
294	V							0.41													
295	A																				
296	L																				
297	G		0.57																		
298	A	0.43					0.41														
299	S																				
300	G																	0.42			
301	I		0.88		0.88																
302	V						0.45														0.46
303	I				0.47																
304	W																				
305	G													1.00	1.62						
306	T										1.02			0.76	0.52						
307	L				0.48					1.08	0.86	0.76	0.61				0.43		0.60		0.64
308	S	1.73		0.65						0.72		2.33		0.92		1.15	1.33	1.34			
309	I	1.44			0.60	1.72		1.47		2.09	1.16	3.11	1.64	0.72	1.99	1.30	0.98	2.27			0.88
310	M	2.73	1.50		0.70						1.61		0.59					1.61			
311	R									0.52	0.84		0.43			0.48	0.72				
312	S	0.49								0.75		1.26									
313	M	0.56	1.34	1.11	0.99	1.05				0.67	0.88				0.63	1.23	2.85	2.30	1.12		
314	K		0.82				0.54			0.61	0.61	0.57	0.62	0.53		1.10		0.95	0.45		
315	S	0.72	0.85			0.42		0.63		0.97					0.41	2.04	0.62	1.04	0.50		
316	C													0.41							

Position	Wildtype (WT)	G	A	P	V	L	I	M	C	T	S	N	Q	D	E	H	K	R	Y	F	W
317	L		1.27				1.76	1.20		0.93	1.03	0.73	1.67	0.61		1.05	5.11	2.41			0.84
318	L	0.49					0.70	1.68		0.63	0.71	0.52	0.71	0.46		0.45	1.36	1.34		0.51	
319	L																				
320	D	0.83					1.00	0.79			1.19	0.52			0.78	1.75	6.42	3.19	0.86		0.40
321	N		1.01							0.64	1.26			1.25		0.92	1.29	1.23	0.40		
322	Y																				
323	M					0.55	0.55													0.64	
324	E		0.59					0.50			0.62	1.01		1.15		0.79		2.28			
325	T	4.21	1.87		1.24			2.11			3.19	4.64	5.08	1.78	4.03	3.45	4.37				0.62
326	I				6.29	1.50					0.95						0.95		0.77		
327	L							0.52													
328	N	0.56	0.67		1.16	2.01	1.85		1.25	0.59	2.22		1.13			0.88	2.12	0.68	1.66		
329	P																				
330	Y																				
331	I				0.94																
332	I																				
333	N																				
334	V			0.46																	
335	T										0.47										
336	L																				
337	A																				
338	A												0.63								
339	K							0.61													
340	M																				
341	C																				
342	S		0.68																		
343	Q				0.51					0.49											
344	V																				
345	L																				
346	C																				
347	Q	2.68	0.78					0.61			2.38				0.78			0.55			
348	E	0.55									0.44			0.67							
349	Q		0.47					0.70		0.49		0.44			0.83		0.93	0.73			
350	G																				
351	V		1.14								0.92										
352	C																				
353	I				1.61					0.42											
354	R																				
355	K																				
356	N		0.41								0.46			0.79		0.82					
357	W		0.80						0.67	0.62	0.41										
358	N	0.41							0.66	0.58											
359	S				0.65			0.63		2.11				0.45	1.05	0.44	0.66				
360	S									0.50											
361	D																				
362	Y																				
363	L																				
364	H																				
365	L																				
366	N																				
367	P	0.47	0.55						0.83		0.52						0.57	0.46			
368	D	0.49	1.34		0.41	0.64		0.78		0.80	0.93				1.28	0.96	1.31	1.31			
369	N										0.54					1.33		0.55			
370	F																				
371	A				0.94	0.57	0.50	0.57			1.45				1.05	1.20	1.76	1.51		0.52	
372	I																				
373	Q		0.65		1.05	0.84		1.43			0.87				0.81		0.73	0.68		0.62	
374	L		0.60	0.43	0.56		0.80	1.11		0.47	0.58	0.43				1.42		0.83	0.66		
375	E	0.90	0.42			0.46		0.54		1.17	0.77	0.81					1.49	0.43			
376	K		0.95		0.58			0.46		0.53	0.80		0.69	0.78	0.88			0.67	0.42		
377	G			2.30						3.83	1.80			1.35	0.59	1.49	1.50	1.28			
378	G											0.64					1.22	1.03			
379	K	0.52								0.40	0.46					0.57		0.74			
380	F			0.47		0.67	0.56												1.50		2.15
381	T				0.89						0.87	0.51	0.84			0.48	1.06	0.87			
382	V																				
383	R		0.61		1.05	1.31	0.71	0.61		0.98	0.87	0.77			0.51	0.71	1.30				
384	G																				
385	K	0.62	1.12		0.43					0.46	1.05	0.41	0.73			0.50		0.94			
386	P																				
387	T										0.93										
388	L				0.78		0.98	0.79		0.51						0.47		0.60	1.18	0.92	0.77
389	E	0.91	1.14	0.75		0.65		0.60		0.70	1.08		0.69			1.17	1.91	0.94	0.77		
390	D																				
391	L								0.90												
392	E	1.00	0.58		1.27	0.80		1.54		0.72	0.52		1.01				0.66	0.66	0.92	0.54	
393	Q		1.26			0.91		0.80		0.41	1.15	0.72		0.45		1.05	0.80	0.74		1.23	
394	F					0.56															0.41
395	S	0.77	1.10													0.56	0.96	1.98			0.40

Position	Wildtype (WT)	G	A	P	V	L	I	M	C	T	S	N	Q	D	E	H	K	R	Y	F	W
396	E		0.52							0.89	0.61		0.73	0.64		0.47		0.61			
397	K																				
398	F																				
399	Y		1.01		1.44			2.70	0.46	2.40	1.01	0.52	1.18		1.49		1.94	1.20			1.92
400	C																				
401	S		0.82									0.42			0.46						
402	C																				
403	Y																			0.62	
404	S		0.63	0.64																	
405	T	2.32	0.56	0.64				0.48			0.94		0.75				0.74	0.60	0.44		0.73
406	L	1.00	0.70		0.87		0.61		0.98	0.83	0.47	0.76	0.93		0.73				0.74	1.42	
407	S	0.75	1.16	0.94	0.56	0.40		0.74				0.72	1.71	1.52	1.38	1.34		1.04		1.42	0.41
408	C																				
409	K	0.50	2.18	0.48	0.48		0.51			0.63	0.72		3.33	0.65	0.62	0.64		0.84			
410	E			0.73	0.65			0.42		1.54	0.81	0.67	0.85	0.47			0.70	0.61	0.62		
411	K		0.48	0.42	0.99					0.63	1.21	1.02						0.97			
412	A	0.80		0.77	0.53	0.65	0.81				0.86	0.86		0.74				0.66	0.66		0.54
413	D									0.41		0.94			0.52			0.50			
414	V						1.12	0.53													
415	K	0.40									0.42										0.42
416	D	0.67			0.59	0.75	0.63			0.85		0.73	0.83			0.57	0.76	0.46	0.40	0.41	
417	T						1.22														
418	D	0.45	0.92	2.11	1.43	1.28	0.99	1.09			0.78	0.91	1.05		1.31			1.18	0.97	0.81	
419	A	0.42				0.56	1.64				2.65	0.53			0.45	1.21	1.88	1.81	1.44	2.17	0.69
420	V			0.48			1.04														
421	D	0.62	1.28			0.73		0.94		0.80	2.12	1.89	1.54		0.81	1.98	2.42	2.21	0.66		
422	V						0.42			0.49											
423	C																				
424	I																				
425	A	1.20					0.44	0.70			0.47	0.46					1.75	0.49			
426	D	0.85		1.03								0.61	0.42		0.62				0.43		
427	G						0.54				0.42						0.52				
428	V			0.82		1.25		0.42		0.62											
429	C																				
430	I																				
431	D	0.55	2.42		1.52	0.62	1.05				1.91	1.30	2.16		1.27	3.13	1.83	2.20	0.85		0.56
432	A	0.52			0.56						0.61	0.51			0.60						
433	F	0.54	0.97	0.95	1.63	1.87	1.06		0.69	1.86	0.86			0.95	0.82	0.83	1.36	1.63			1.28
434	L	0.47			0.46		0.89	0.60												0.41	
435	K	0.64	1.08		0.44				0.53	0.73	1.03				0.78	1.05		1.01	0.50		
436	P	1.19				0.63	0.84	0.61		0.59	0.92		0.86	1.19	0.74	0.72	2.05	1.00	0.49		0.43
437	P	0.50	0.56			0.51	2.46	2.55			0.57		0.96	0.62		1.11	0.83	0.85	0.42		
438	M	0.83	0.75	0.81	0.85	0.86			0.63	3.99	0.83	1.08	0.85	0.87	0.72			0.99			0.57
439	E	1.22	1.20	1.16	1.57	0.88			0.58	1.15	1.02		1.32			0.74	1.20			1.00	0.62
440	T	0.86	1.22	0.88	1.02	0.97	1.04	1.08			1.17			1.03	1.00	3.00		1.77	1.11	0.85	
441	E	0.87	1.47		0.54	0.82				0.66	0.79	0.82	0.81	0.67		0.65	0.80		0.51	3.91	
442	E	0.51		0.91	0.65	0.40			1.38	0.61			0.74			0.76	0.73	3.94	0.60		
443	P	1.12	1.63			1.19		1.99		0.87	0.99	1.25	0.96		1.07	1.08		1.04		0.70	0.64
444	Q	0.93			1.12		0.58					1.01		0.97	1.19	0.97	1.03	0.85	0.67	0.66	0.64
445	I	0.98	0.97	1.67	1.25	1.06		1.57		1.38	1.21	1.24	1.26			1.35		1.08	0.53		0.69
446	F	1.12	1.58		0.91	0.93	1.06	1.31	0.75	0.89			0.72	1.18	1.10	1.28	0.94	0.89			1.40
447	Y	0.92		1.46	1.38	1.09	1.36	0.90		1.90		1.58	2.37	3.25	1.36			1.12		1.41	1.07

Note:
Physicochemical groups are: small hydrophobes (G, A, P, V, L, I, M), Polar Uncharged (C, T, S, N, Q), Polar Negative (D, E), Polar Positive (H, K, R), Large Hydrophobes (Y, F, W).

Surface Residues w/ ≥ 11 Active Subs. & Physicochemical Characterization of Each Class	Wildtype (WT)	Count Active (>40%, no errors)	Tested (w/o errors)
1	L	15	15
2	N	9	13
3	F	4	10
4	R	5	13
5	A	1	14
6	P	7	12
7	P	1	15
8	V	3	11
9	I	6	12
10	P	9	14
11	N	4	13
12	V	9	12
13	P	4	10
14	F	4	13
15	L	3	13
16	W	0	14
17	A	0	15
18	W	0	14
19	N	0	15
20	A	1	13
21	P	0	14
22	S	4	9
23	E	1	11
24	F	13	13
25	C	0	15
26	L	15	15
27	G	14	15
28	K	14	14
29	F	14	14
30	D	14	14
31	E	14	14
32	P	16	16
33	L	8	14
34	D	6	12
35	M	5	11
36	S	7	12
37	L	7	12
38	F	1	10
39	S	7	11
40	F	2	12
41	I	10	11
42	G	1	14
43	S	1	11
44	P	1	15
45	R	2	8
46	I	12	14
47	N	13	14
48	A	12	13
49	T	5	10
50	G	11	11
51	Q	2	10
52	G	6	11
53	V	0	15
54	T	6	12
55	I	0	11
56	F	0	14
57	Y	0	12
58	V	13	14
59	D	2	13
60	R	1	15
61	L	3	14
62	G	0	15
63	Y	12	15
64	Y	0	17
65	P	0	13
66	Y	2	15
67	I	4	10
68	D	7	12
69	S	12	13
70	I	14	15
71	T	10	11

Subset of Surface Residues	Wildtype (WT)	SASA (Å)	fSASA	Local QMEAN Score from
1	L			
2	N			
3	F			0.37
4	R	139.073	0.966	0.47
5	A	48.781	0.439	0.50
6	P	65.732	0.487	0.63
7	P	38.050	0.282	0.68
8	V	33.672	0.226	0.67
9	I	52.912	0.306	0.71
10	P	103.226	0.765	0.75
11	N	90.515	0.545	0.57
12	V	16.843	0.113	0.68
13	P	8.983	0.067	0.67
14	F	0.000	0.000	0.73
15	L	8.367	0.047	0.76
16	W	2.350	0.009	0.75
17	A	7.617	0.069	0.81
18	W	1.440	0.006	0.70
19	N	10.312	0.062	0.70
20	A	0.706	0.006	0.67
21	P	3.031	0.022	0.62
22	S	6.223	0.050	0.65
23	E	83.710	0.448	0.61
24	F	63.014	0.303	0.54
25	C	1.435	0.009	0.63
26	L	96.728	0.540	0.64
27	G	76.159	0.886	0.51
28	K	124.901	0.589	0.46
29	F	48.655	0.234	0.53
30	D	130.140	0.813	0.54
31	E	25.495	0.136	0.48
32	P	84.022	0.622	0.52
33	L	16.981	0.095	0.52
34	D	44.945	0.281	0.57
35	M	68.964	0.343	0.58
36	S	65.053	0.520	0.64
37	L	36.039	0.201	0.58
38	F	14.561	0.070	0.66
39	S	51.573	0.413	0.77
40	F	43.210	0.208	0.67
41	I	60.110	0.347	0.70
42	G	10.926	0.127	0.68
43	S	13.844	0.111	0.60
44	P	20.476	0.152	0.50
45	R	113.880	0.456	0.55
46	I	60.998	0.353	0.50
47	N	90.218	0.543	0.48
48	A	83.814	0.755	0.58
49	T	92.499	0.642	0.61
50	G	33.490	0.389	0.62
51	Q	109.269	0.563	0.58
52	G	8.324	0.097	0.69
53	V	36.816	0.247	0.68
54	T	12.982	0.090	0.68
55	I	10.222	0.059	0.80
56	F	0.000	0.000	0.82
57	Y	64.770	0.285	0.87
58	V	34.213	0.230	0.84
59	D	112.167	0.701	0.74
60	R	93.608	0.374	0.70
61	L	0.000	0.000	0.84
62	G	9.936	0.116	0.88
63	Y	92.692	0.408	0.87
64	Y	4.748	0.021	0.89
65	P	0.720	0.005	0.91
66	Y	51.764	0.228	0.83
67	I	8.003	0.046	0.85
68	D	54.181	0.339	0.76
69	S	111.432	0.891	0.79
70	I	144.199	0.834	0.71
71	T	88.846	0.617	0.67

Note:
Column AD is shaded according to the below criteria:

Orange - buried residue with fSASA score < 0.35

Blue - surface residue with a fSASA score > 0.65

Purple - intermediate, partially-buried residue measured between buried and surface exposed with fSASA between 0.35 and 0.65

In Column X, positions shaded in green are those which are surface residues that have ≥ 11 active substitutions (> 40% activity) reported in Table 9 and also have sufficient physicochemical characterization (*i.e.*, at least one residue sampled from each of the 5 physicochemical classes).

Local QMEAN scores (col. AG) were obtained from the Swiss-Model stores in the b-factor column of the pdb file values obtained from "model_01.pdb.cif" file, provided by counsel to Halozyme

Red - QMEAN score < 0.6

Blue - QMEAN score > 0.6

Surface Residues w/ ≥ 11 Active Subs. & Physicochemical Characterization of Each Class	Wildtype (WT)	Count Active (>40%, no errors)	Tested (w/o errors)
72	G	9	15
73	V	13	14
74	T	15	15
75	V	12	15
76	N	0	14
77	G	0	8
78	G	0	7
79	I	3	14
80	P	0	15
81	Q	0	11
82	K	13	14
83	I	10	11
84	S	14	14
85	L	1	12
86	Q	16	17
87	D	14	14
88	H	0	15
89	L	3	11
90	D	12	13
91	K	3	12
92	A	5	13
93	K	13	13
94	K	12	14
95	D	0	15
96	I	3	12
97	T	15	15
98	F	13	14
99	Y	2	10
100	M	0	11
101	P	0	13
102	V	14	15
103	D	0	13
104	N	8	11
105	L	9	11
106	G	1	13
107	M	3	12
108	A	1	12
109	V	0	8
110	I	1	7
111	D	0	3
112	W	0	8
113	E	0	2
114	E	5	10
115	W	0	14
116	R	0	14
117	P	0	9
118	T	7	15
119	W	3	9
120	A	14	14
121	R	0	12
122	N	1	12
123	W	0	14
124	K	1	6
125	P	3	9
126	K	0	7
127	D	13	14
128	V	10	12
129	Y	0	13
130	K	2	12
131	N	14	15
132	R	14	15
133	S	1	14
134	I	3	15
135	E	14	15
136	L	13	14
137	V	3	11
138	Q	11	12
139	Q	14	15
140	Q	14	14
141	N	15	15
142	V	15	15

Subset of Surface Residues	Wildtype (WT)	SASA (Å)	fSASA	Local QMEAN Score from
72	G	20.860	0.243	0.72
73	V	79.688	0.535	0.76
74	T	78.760	0.547	0.78
75	V	45.747	0.307	0.88
76	N	69.106	0.416	0.86
77	G	28.072	0.326	0.87
78	G	0.000	0.000	0.85
79	I	5.990	0.035	0.84
80	P	4.421	0.033	0.88
81	Q	46.361	0.239	0.80
82	K	117.628	0.555	0.70
83	I	34.146	0.197	0.75
84	S	47.481	0.380	0.79
85	L	46.614	0.260	0.78
86	Q	112.172	0.578	0.79
87	D	78.456	0.490	0.81
88	H	0.000	0.000	0.82
89	L	35.091	0.196	0.84
90	D	79.306	0.496	0.84
91	K	67.628	0.319	0.83
92	A	1.383	0.012	0.87
93	K	114.335	0.539	0.77
94	K	141.266	0.666	0.79
95	D	16.707	0.104	0.82
96	I	0.795	0.005	0.76
97	T	60.642	0.421	0.73
98	F	170.109	0.818	0.77
99	Y	46.946	0.207	0.70
100	M	5.098	0.025	0.65
101	P	90.756	0.672	0.77
102	V	57.386	0.385	0.73
103	D	104.174	0.651	0.79
104	N	4.550	0.027	0.55
105	L	99.070	0.553	0.59
106	G	3.543	0.041	0.72
107	M	6.944	0.035	0.73
108	A	0.000	0.000	0.76
109	V	1.591	0.011	0.84
110	I	0.000	0.000	0.87
111	D	30.647	0.192	0.90
112	W	8.741	0.035	0.88
113	E	81.629	0.437	0.81
114	E	54.483	0.291	0.80
115	W	5.584	0.022	0.82
116	R	29.305	0.117	0.81
117	P	6.733	0.050	0.88
118	T	7.966	0.055	0.84
119	W	22.218	0.089	0.76
120	A	53.068	0.478	0.78
121	R	2.963	0.012	0.77
122	N	2.114	0.013	0.76
123	W	91.137	0.366	0.69
124	K	177.903	0.839	0.77
125	P	102.480	0.759	0.74
126	K	37.646	0.178	0.74
127	D	48.435	0.303	0.82
128	V	58.060	0.390	0.83
129	Y	3.493	0.015	0.82
130	K	50.824	0.240	0.80
131	N	86.923	0.524	0.82
132	R	93.616	0.374	0.71
133	S	0.000	0.000	0.82
134	I	38.703	0.224	0.84
135	E	83.130	0.445	0.74
136	L	41.446	0.232	0.79
137	V	13.182	0.088	0.84
138	Q	57.823	0.298	0.81
139	Q	119.753	0.617	0.79
140	Q	160.155	0.826	0.81
141	N	65.932	0.397	0.84
142	V	99.585	0.668	0.76

Surface Residues w/ ≥ 11 Active Subs. & Physicochemical Characterization of Each Class	Wildtype (WT)	Count Active (>40%, no errors)	Tested (w/o errors)
143	Q	7	12
144	L	2	12
145	S	11	13
146	L	15	15
147	T	15	15
148	E	14	14
149	A	10	12
150	T	14	15
151	E	15	15
152	K	9	10
153	A	2	9
154	K	3	10
155	Q	14	16
156	E	12	13
157	F	1	17
158	E	6	11
159	K	11	13
160	A	14	14
161	G	7	8
162	K	14	14
163	D	10	12
164	F	4	14
165	L	9	12
166	V	13	14
167	E	12	13
168	T	1	15
169	I	3	15
170	K	5	13
171	L	2	11
172	G	2	12
173	K	2	14
174	L	12	13
175	L	5	12
176	R	2	15
177	P	0	13
178	N	4	10
179	H	14	15
180	L	4	10
181	W	2	13
182	G	1	14
183	Y	1	12
184	Y	0	14
185	L	0	14
186	F	1	15
187	P	0	15
188	D	0	15
189	C	0	15
190	Y	0	14
191	N	0	15
192	H	2	15
193	H	4	11
194	Y	0	8
195	K	10	10
196	K	7	7
197	P	13	14
198	G	11	12
199	Y	0	10
200	N	1	14
201	G	0	11
202	S	1	14
203	C	0	13
204	F	1	12
205	N	14	14
206	V	8	14
207	E	0	11
208	I	10	14
209	K	8	10
210	R	0	14
211	N	1	13
212	D	3	13
213	D	13	15

Subset of Surface Residues	Wildtype (WT)	SASA (Å)	fSASA	Local QMEAN Score from
143	Q	165.074	0.851	0.79
144	L	32.641	0.182	0.81
145	S	68.577	0.549	0.80
146	L	122.629	0.685	0.79
147	T	98.809	0.686	0.78
148	E	92.061	0.492	0.78
149	A	0.000	0.000	0.88
150	T	47.998	0.333	0.81
151	E	86.410	0.462	0.74
152	K	102.753	0.485	0.79
153	A	0.000	0.000	0.86
154	K	65.497	0.309	0.82
155	Q	111.552	0.575	0.78
156	E	73.070	0.391	0.80
157	F	3.035	0.015	0.84
158	E	43.559	0.233	0.80
159	K	139.479	0.658	0.81
160	A	10.651	0.096	0.88
161	G	0.000	0.000	0.87
162	K	79.805	0.376	0.81
163	D	55.396	0.346	0.84
164	F	3.866	0.019	0.83
165	L	5.098	0.028	0.84
166	V	25.229	0.169	0.86
167	E	61.558	0.329	0.82
168	T	0.000	0.000	0.82
169	I	0.720	0.004	0.81
170	K	89.258	0.421	0.76
171	L	29.926	0.167	0.82
172	G	0.000	0.000	0.79
173	K	84.793	0.400	0.70
174	L	121.372	0.678	0.70
175	L	46.659	0.261	0.67
176	R	39.598	0.158	0.66
177	P	74.735	0.554	0.79
178	N	79.973	0.482	0.74
179	H	2.085	0.011	0.72
180	L	13.986	0.078	0.73
181	W	0.000	0.000	0.82
182	G	0.000	0.000	0.88
183	Y	5.539	0.024	0.89
184	Y	17.785	0.078	0.87
185	L	27.644	0.154	0.87
186	F	12.206	0.059	0.84
187	P	4.158	0.031	0.87
188	D	9.692	0.061	0.79
189	C	2.092	0.013	0.75
190	Y	24.249	0.107	0.79
191	N	3.670	0.022	0.77
192	H	98.584	0.516	0.57
193	H	101.781	0.533	0.67
194	Y	68.699	0.303	0.67
195	K	163.551	0.771	0.66
196	K	101.758	0.480	0.65
197	P	144.905	1.073	0.73
198	G	50.427	0.586	0.65
199	Y	38.340	0.169	0.64
200	N	101.580	0.612	0.75
201	G	0.720	0.008	0.83
202	S	65.334	0.523	0.81
203	C	13.271	0.085	0.81
204	F	98.579	0.474	0.71
205	N	98.259	0.592	0.56
206	V	57.311	0.385	0.84
207	E	24.338	0.130	0.83
208	I	64.061	0.370	0.84
209	K	107.619	0.508	0.80
210	R	71.659	0.287	0.76
211	N	1.184	0.007	0.83
212	D	52.008	0.325	0.78
213	D	90.794	0.567	0.74

Surface Residues w/ ≥ 11 Active Subs. & Physicochemical Characterization of Each Class	Wildtype (WT)	Count Active (>40%, no errors)	Tested (w/o errors)
214	L	1	14
215	S	13	15
216	W	1	15
217	L	1	11
218	W	1	7
219	N	13	14
220	E	7	13
221	S	6	12
222	T	10	12
223	A	0	15
224	L	1	14
225	Y	0	12
226	P	0	14
227	S	0	15
228	I	0	13
229	Y	0	10
230	L	1	15
231	N	1	15
232	T	2	12
233	Q	7	12
234	Q	2	13
235	S	6	12
236	P	6	13
237	V	11	11
238	A	8	15
239	A	0	13
240	T	8	14
241	L	0	12
242	Y	1	14
243	V	0	13
244	R	0	6
245	N	1	11
246	R	0	15
247	V	3	15
248	R	4	12
249	E	1	10
250	A	0	15
251	I	2	11
252	R	0	14
253	V	1	13
254	S	0	14
255	K	5	11
256	I	4	9
257	P	11	12
258	D	4	8
259	A	14	14
260	K	10	12
261	S	10	11
262	P	0	15
263	L	6	11
264	P	2	13
265	V	1	13
266	F	1	13
267	A	2	11
268	Y	0	14
269	T	4	12
270	R	4	13
271	I	6	13
272	V	4	10
273	F	3	14
274	T	3	11
275	D	0	9
276	Q	10	13
277	V	14	14
278	L	13	15
279	K	4	10
280	F	2	12
281	L	0	13
282	S	4	9
283	Q	5	10
284	D	9	12

Subset of Surface Residues	Wildtype (WT)	SASA (Å)	fSASA	Local QMEAN Score from
214	L	5.023	0.028	0.72
215	S	24.295	0.194	0.77
216	W	38.551	0.155	0.71
217	L	2.311	0.013	0.76
218	W	3.197	0.013	0.73
219	N	71.240	0.429	0.70
220	E	41.651	0.223	0.74
221	S	8.290	0.066	0.77
222	T	25.292	0.176	0.76
223	A	0.639	0.006	0.82
224	L	2.235	0.012	0.81
225	Y	0.000	0.000	0.85
226	P	0.760	0.006	0.86
227	S	6.187	0.049	0.85
228	I	2.505	0.014	0.84
229	Y	35.171	0.155	0.78
230	L	3.400	0.019	0.74
231	N	69.570	0.419	0.58
232	T	56.716	0.394	0.61
233	Q	94.120	0.485	0.61
234	Q	10.225	0.053	0.49
235	S	102.880	0.823	0.45
236	P	58.304	0.432	0.49
237	V	135.527	0.910	0.57
238	A	12.639	0.114	0.75
239	A	2.243	0.020	0.75
240	T	24.367	0.169	0.74
241	L	57.938	0.324	0.76
242	Y	10.153	0.045	0.77
243	V	0.000	0.000	0.83
244	R	68.249	0.273	0.74
245	N	3.179	0.019	0.82
246	R	7.459	0.030	0.79
247	V	0.000	0.000	0.87
248	R	89.071	0.356	0.77
249	E	2.065	0.011	0.81
250	A	0.000	0.000	0.82
251	I	38.094	0.220	0.74
252	R	58.364	0.233	0.71
253	V	0.000	0.000	0.79
254	S	2.200	0.018	0.75
255	K	117.817	0.556	0.69
256	I	54.806	0.317	0.55
257	P	84.209	0.624	0.54
258	D	92.576	0.579	0.47
259	A	21.353	0.192	0.49
260	K	175.390	0.827	0.44
261	S	68.849	0.551	0.52
262	P	57.941	0.429	0.58
263	L	25.418	0.142	0.70
264	P	16.735	0.124	0.73
265	V	6.141	0.041	0.76
266	F	2.350	0.011	0.76
267	A	0.760	0.007	0.80
268	Y	12.220	0.054	0.82
269	T	2.823	0.020	0.84
270	R	13.176	0.053	0.74
271	I	7.853	0.045	0.80
272	V	2.114	0.014	0.71
273	F	26.679	0.128	0.65
274	T	23.215	0.161	0.59
275	D	111.462	0.697	0.54
276	Q	46.527	0.240	0.56
277	V	53.292	0.358	0.51
278	L	110.288	0.616	0.51
279	K	105.162	0.496	0.71
280	F	57.997	0.279	0.78
281	L	5.232	0.029	0.82
282	S	53.343	0.427	0.81
283	Q	106.552	0.549	0.82
284	D	79.206	0.495	0.78

Surface Residues w/ ≥ 11 Active Subs. & Physicochemical Characterization of Each Class	Wildtype (WT)	Count Active (>40%, no errors)	Tested (w/o errors)
285	E	7	12
286	L	2	12
287	V	2	13
288	Y	2	12
289	T	1	11
290	F	2	5
291	G	4	13
292	E	11	14
293	T	13	15
294	V	1	14
295	A	0	10
296	L	0	13
297	G	1	11
298	A	2	13
299	S	0	11
300	G	1	15
301	I	2	14
302	V	2	15
303	I	1	12
304	W	0	11
305	G	2	10
306	T	3	11
307	L	8	11
308	S	8	15
309	I	14	14
310	M	6	10
311	R	6	14
312	S	3	8
313	M	12	13
314	K	10	13
315	S	10	13
316	C	1	14
317	L	12	14
318	L	12	15
319	L	0	15
320	D	11	13
321	N	8	11
322	Y	0	13
323	M	3	14
324	E	7	13
325	T	12	14
326	I	5	10
327	L	1	14
328	N	13	13
329	P	0	15
330	Y	0	14
331	I	1	12
332	I	0	15
333	N	0	10
334	V	1	10
335	T	1	11
336	L	0	13
337	A	0	10
338	A	1	14
339	K	1	14
340	M	0	14
341	C	0	14
342	S	1	13
343	Q	2	8
344	V	0	13
345	L	0	11
346	C	0	14
347	Q	6	13
348	E	3	14
349	Q	7	14
350	G	0	15
351	V	2	10
352	C	0	15
353	I	2	13
354	R	0	15
355	K	0	15

Subset of Surface Residues	Wildtype (WT)	SASA (Å)	fSASA	Local QMEAN Score from
285	E	11.785	0.063	0.74
286	L	0.000	0.000	0.80
287	V	41.730	0.280	0.83
288	Y	63.463	0.280	0.82
289	T	3.830	0.027	0.81
290	F	0.000	0.000	0.84
291	G	3.146	0.037	0.88
292	E	1.380	0.007	0.80
293	T	0.000	0.000	0.82
294	V	1.829	0.012	0.81
295	A	0.000	0.000	0.83
296	L	28.766	0.161	0.79
297	G	23.523	0.274	0.72
298	A	2.079	0.019	0.75
299	S	26.443	0.212	0.75
300	G	0.000	0.000	0.81
301	I	2.018	0.012	0.83
302	V	2.862	0.019	0.84
303	I	0.000	0.000	0.80
304	W	69.383	0.279	0.79
305	G	5.680	0.066	0.71
306	T	19.027	0.132	0.64
307	L	112.244	0.627	0.66
308	S	64.223	0.514	0.70
309	I	7.201	0.042	0.66
310	M	1.383	0.007	0.70
311	R	177.550	0.710	0.67
312	S	42.440	0.340	0.73
313	M	107.328	0.534	0.66
314	K	136.842	0.645	0.68
315	S	11.506	0.092	0.73
316	C	2.591	0.017	0.63
317	L	75.812	0.424	0.70
318	L	72.002	0.402	0.74
319	L	2.100	0.012	0.72
320	D	23.798	0.149	0.64
321	N	76.586	0.461	0.74
322	Y	10.085	0.044	0.81
323	M	3.715	0.018	0.73
324	E	89.034	0.476	0.71
325	T	90.864	0.631	0.84
326	I	43.424	0.251	0.85
327	L	0.000	0.000	0.81
328	N	0.652	0.004	0.76
329	P	32.519	0.241	0.82
330	Y	32.104	0.141	0.82
331	I	0.720	0.004	0.81
332	I	31.326	0.181	0.81
333	N	35.677	0.215	0.81
334	V	0.720	0.005	0.84
335	T	33.546	0.233	0.77
336	L	29.148	0.163	0.83
337	A	0.000	0.000	0.84
338	A	3.664	0.033	0.80
339	K	57.662	0.272	0.74
340	M	33.489	0.167	0.81
341	C	0	0	0.82
342	S	5.448	0.044	0.77
343	Q	58.968	0.304	0.72
344	V	49.922	0.335	0.75
345	L	31.197	0.174	0.79
346	C	0	0	0.73
347	Q	86.748	0.447	0.72
348	E	101.491	0.543	0.70
349	Q	70.960	0.366	0.61
350	G	0.000	0.000	0.77
351	V	18.705	0.126	0.77
352	C	0.652	0.004	0.77
353	I	42.834	0.248	0.76
354	R	20.892	0.084	0.74
355	K	54.632	0.258	0.72

Surface Residues w/ ≥ 11 Active Subs. & Physicochemical Characterization of Each Class	Wildtype (WT)	Count Active (>40%, no errors)	Tested (w/o errors)
356	N	4	13
357	W	4	12
358	N	3	11
359	S	7	13
360	S	1	14
361	D	0	12
362	Y	0	15
363	L	0	15
364	H	0	15
365	L	0	14
366	N	0	12
367	P	6	13
368	D	11	14
369	N	3	13
370	F	0	14
371	A	10	12
372	I	0	15
373	Q	9	12
374	L	11	13
375	E	9	14
376	K	10	13
377	G	8	12
378	G	3	13
379	K	5	13
380	F	5	12
381	T	7	12
382	V	0	14
383	R	11	13
384	G	0	6
385	K	9	15
386	P	0	15
387	T	1	13
388	L	9	14
389	E	12	14
390	D	0	15
391	L	1	15
392	E	12	14
393	Q	11	13
394	F	2	15
395	S	6	10
396	E	7	13
397	K	0	12
398	F	0	14
399	Y	12	14
400	C	0	15
401	S	3	10
402	C	0	14
403	Y	1	14
404	S	2	14
405	T	10	13
406	L	12	14
407	S	14	14
408	C	0	14
409	K	12	12
410	E	11	12
411	K	7	11
412	A	11	12
413	D	4	7
414	V	2	10
415	K	3	7
416	D	12	14
417	T	1	12
418	D	14	14
419	A	12	14
420	V	2	15
421	D	13	13
422	V	2	13
423	C	0	15
424	I	0	11
425	A	7	11
426	D	6	10

Subset of Surface Residues	Wildtype (WT)	SASA (Å)	fSASA	Local QMEAN Score from
356	N	147.976	0.891	0.74
357	W	219.026	0.880	0.36
358	N	54.014	0.325	0.52
359	S	72.418	0.579	0.68
360	S	71.417	0.571	0.70
361	D	29.415	0.184	0.74
362	Y	39.880	0.176	0.78
363	L	6.819	0.038	0.81
364	H	91.537	0.479	0.81
365	L	11.873	0.066	0.84
366	N	46.049	0.277	0.83
367	P	105.954	0.785	0.83
368	D	140.876	0.880	0.84
369	N	25.624	0.154	0.75
370	F	20.643	0.099	0.86
371	A	35.895	0.323	0.90
372	I	44.651	0.258	0.85
373	Q	103.495	0.533	0.80
374	L	73.823	0.412	0.75
375	E	106.861	0.571	0.73
376	K	205.579	0.970	0.74
377	G	78.889	0.917	0.59
378	G	28.816	0.335	0.61
379	K	131.483	0.620	0.66
380	F	19.524	0.094	0.71
381	T	23.941	0.166	0.79
382	V	41.918	0.281	0.85
383	R	140.040	0.560	0.83
384	G	61.768	0.718	0.90
385	K	135.441	0.639	0.85
386	P	50.489	0.374	0.85
387	T	51.574	0.358	0.83
388	L	132.293	0.739	0.86
389	E	134.042	0.717	0.80
390	D	19.943	0.125	0.80
391	L	55.602	0.311	0.85
392	E	122.480	0.655	0.80
393	Q	93.910	0.484	0.77
394	F	16.147	0.078	0.80
395	S	52.380	0.419	0.84
396	E	123.156	0.659	0.80
397	K	28.635	0.135	0.72
398	F	0.000	0.000	0.79
399	Y	119.359	0.526	0.77
400	C	20.202	0.129	0.77
401	S	55.944	0.448	0.66
402	C	40.505	0.258	0.59
403	Y	13.826	0.061	0.65
404	S			0.61
405	T			0.51
406	L			0.52
407	S			0.62
408	C			0.59
409	K			0.61
410	E			0.63
411	K			0.62
412	A			0.64
413	D			0.63
414	V			0.64
415	K			0.59
416	D			
417	T			
418	D			
419	A			
420	V			
421	D			
422	V			
423	C			
424	I			
425	A			
426	D			

Surface Residues w/ ≥ 11 Active Subs. & Physicochemical Characterization of Each Class	Wildtype (WT)	Count Active (>40%, no errors)	Tested (w/o errors)
427	G	3	11
428	V	4	14
429	C	0	11
430	I	0	9
431	D	14	15
432	A	5	13
433	F	15	15
434	L	5	11
435	K	10	10
436	P	14	14
437	P	12	13
438	M	14	15
439	E	13	15
440	T	14	15
441	E	13	14
442	E	11	14
443	P	13	14
444	Q	12	12
445	I	14	14
446	F	15	15
447	Y	14	14

Subset of Surface Residues	Wildtype (WT)	SASA (Å)	fSASA	Local QMEAN Score from
427	G			
428	V			
429	C			
430	I			
431	D			
432	A			
433	F			
434	L			
435	K			
436	P			
437	P			
438	M			
439	E			
440	T			
441	E			
442	E			
443	P			
444	Q			
445	I			
446	F			
447	Y			

Position	Wildtype (WT)	Small Hydrophobes (7)	Polar Uncharged (5)	Polar Negative (2)	Polar Positive (3)	Large Hydrophobes (3)	Count Active (>40%, no errors)	Tested (w/o errors)
1	L	4	5	1	3	2	15	15
2	N	5	4	0	0	0	9	13
3	F	1	0	1	1	1	4	10
4	R	3	2	0	0	0	5	13
5	A	0	0	0	1	0	1	14
6	P	2	2	0	3	0	7	12
7	P	1	0	0	0	0	1	15
8	V	3	0	0	0	0	3	11
9	I	2	2	0	2	0	6	12
10	P	1	3	2	2	1	9	14
11	N	1	0	1	2	0	4	13
12	V	3	3	1	2	0	9	12
13	P	0	2	0	1	1	4	10
14	F	3	0	1	0	0	4	13
15	L	3	0	0	0	0	3	13
16	W	0	0	0	0	0	0	14
17	A	0	0	0	0	0	0	15
18	W	0	0	0	0	0	0	14
19	N	0	0	0	0	0	0	15
20	A	0	1	0	0	0	1	13
21	P	0	0	0	0	0	0	14
22	S	1	1	0	1	1	4	9
23	E	0	0	1	0	0	1	11
24	F	6	2	1	3	1	13	13
25	C	0	0	0	0	0	0	15
26	L	6	3	1	3	2	15	15
27	G	4	3	2	3	2	14	15
28	K	6	3	2	1	2	14	14
29	F	7	2	1	3	1	14	14
30	D	6	3	x	3	2	14	14
31	E	6	3	x	3	2	14	14
32	P	5	5	0	3	3	16	16
33	L	3	3	0	1	1	8	14
34	D	0	1	1	3	1	6	12
35	M	1	1	0	1	2	5	11
36	S	3	1	0	3	0	7	12
37	L	4	0	0	2	1	7	12
38	F	0	0	0	0	1	1	10
39	S	2	3	0	1	1	7	11
40	F	1	0	0	0	1	2	12
41	I	3	3	2	1	1	10	11
42	G	1	0	0	0	0	1	14
43	S	0	1	0	0	0	1	11
44	P	0	0	1	0	0	1	15
45	R	1	0	0	1	0	2	8
46	I	4	3	1	2	2	12	14
47	N	3	3	1	3	3	13	14
48	A	4	3	0	3	2	12	13
49	T	2	1	0	2	0	5	10
50	G	3	3	2	2	1	11	11
51	Q	0	2	0	0	0	2	10
52	G	1	4	0	1	0	6	11
53	V	0	0	0	0	0	0	15
54	T	2	3	0	0	1	6	12
55	I	0	0	0	0	0	0	11
56	F	0	0	0	0	0	0	14
57	Y	0	0	0	0	0	0	12
58	V	4	4	0	3	2	13	14
59	D	0	2	x	0	0	2	13
60	R	0	0	0	1	0	1	15
61	L	3	0	0	0	0	3	14
62	G	0	0	0	0	0	0	15
63	Y	5	3	0	3	1	12	15
64	Y	0	0	0	0	0	0	17
65	P	0	0	0	1	0	0	13
66	Y	0	0	0	2	0	2	15
67	I	2	0	0	0	2	4	10
68	D	0	3	1	3	0	7	12
69	S	6	1	1	1	3	12	13
70	I	5	4	0	3	2	14	15
71	T	4	3	1	2	0	10	11
72	G	3	2	2	2	0	9	15
73	V	4	4	1	3	1	13	14
74	T	6	3	1	3	2	15	15
75	V	3	5	0	2	2	12	15
76	N	0	0	0	0	0	0	14
77	G	0	0	0	0	0	0	8
78	G	0	0	0	0	0	0	7
79	I	2	1	0	0	0	3	14
80	P	0	0	0	0	0	0	15
81	Q	0	0	0	0	0	0	11

Note:
Column AL ("Position") and Column AM ("Wildtype (WT)") are shaded according to the sufficiecy or insufficiency of testing of alternate residues in each physicochemical group: small hydrophobes (G, A, P, V, L, I, M), Polar Uncharged (C, T, S, N, Q), Polar Negative (D, E), Polar Positive (H, K, R), Large Hydrophobes (Y, F, W). An 'x' is placed at positions in the Polar Negative class where the other polar negative residue, either D or E, was not tested, yet the WT residue was one of D or E.

Red - insufficient testing of each physicochemical group.

Yellow - 0 members of the physichochemical class were tested at the position reported in the common disclosure

Position	Wildtype (WT)	Small Hydrophobes (7)	Polar Uncharged (5)	Polar Negative (2)	Polar Positive (3)	Large Hydrophobes (3)	Count Active (>40%, no errors)	Tested (w/o errors)
82	K	6	4	1	2	0	13	14
83	I	3	4	0	2	1	10	11
84	S	5	3	2	2	2	14	14
85	L	1	0	0	0	0	1	12
86	Q	6	3	2	3	2	16	17
87	D	6	4	1	2	1	14	14
88	H	0	0	0	0	0	0	15
89	L	1	1	0	1	0	3	11
90	D	3	4	1	3	1	12	13
91	K	1	1	0	1	0	3	12
92	A	3	2	0	0	0	5	13
93	K	4	4	2	2	1	13	13
94	K	3	4	2	2	1	12	14
95	D	0	0	x	0	0	0	15
96	I	2	0	1	0	0	3	12
97	T	5	4	2	1	3	15	15
98	F	5	3	2	2	1	13	14
99	Y	0	1	0	1	0	2	10
100	M	0	0	0	0	0	0	11
101	P	0	0	0	0	0	0	13
102	V	4	5	1	3	1	14	15
103	D	0	0	x	0	0	0	13
104	N	3	3	0	2	0	8	11
105	L	5	3	0	1	0	9	11
106	G	1	0	0	0	0	1	13
107	M	2	0	0	0	1	3	12
108	A	1	0	0	0	0	1	12
109	V	0	0	0	0	0	0	8
110	I	1	0	0	0	0	1	7
111	D	0	0	x	0	0	0	3
112	W	0	0	0	0	0	0	8
113	E	0	0	x	0	0	0	2
114	E	3	1	x	1	0	5	10
115	W	0	0	0	0	0	0	14
116	R	0	0	0	0	0	0	14
117	P	0	0	1	0	0	0	9
118	T	3	2	0	2	0	7	15
119	W	0	1	0	0	2	3	9
120	A	5	3	1	2	3	14	14
121	R	0	0	0	0	0	0	12
122	N	1	0	0	0	0	1	12
123	W	0	0	0	0	0	0	14
124	K	0	0	0	1	0	1	6
125	P	0	1	0	2	0	3	9
126	K	0	0	0	0	0	0	7
127	D	5	4	1	2	1	13	14
128	V	4	3	0	2	1	10	12
129	Y	0	0	0	0	0	0	13
130	K	1	0	0	1	0	2	12
131	N	5	4	1	2	2	14	15
132	R	4	5	1	2	2	14	15
133	S	1	0	0	0	0	1	14
134	I	2	1	0	0	0	3	15
135	E	3	4	1	3	3	14	15
136	L	3	5	1	2	2	13	14
137	V	2	1	0	0	0	3	11
138	Q	4	3	0	2	2	11	12
139	Q	5	3	2	3	1	14	15
140	Q	6	1	1	3	3	14	14
141	N	5	3	2	2	3	15	15
142	V	5	5	2	3	0	15	15
143	Q	4	1	1	1	0	7	12
144	L	0	1	0	0	1	2	12
145	S	5	2	2	2	0	11	13
146	L	5	5	1	3	1	15	15
147	T	7	3	1	1	3	15	15
148	E	4	4	x	3	3	14	14
149	A	4	4	0	2	0	10	12
150	T	5	3	2	1	3	14	15
151	E	5	5	x	3	2	15	15
152	K	4	2	0	1	2	9	10
153	A	2	0	0	0	0	2	9
154	K	1	1	0	1	0	3	10
155	Q	5	3	1	3	2	14	16
156	E	6	3	1	1	1	12	13
157	F	0	0	0	0	1	1	17
158	E	2	2	x	1	1	6	11
159	K	4	3	2	2	0	11	13
160	A	5	4	0	3	2	14	14
161	G	2	2	2	1	0	7	8
162	K	6	2	2	2	2	14	14

Position	Wildtype (WT)	Small Hydrophobes (7)	Polar Uncharged (5)	Polar Negative (2)	Polar Positive (3)	Large Hydrophobes (3)	Count Active (>40%, no errors)	Tested (w/o errors)
163	D	4	3	1	2	0	10	12
164	F	3	0	0	0	1	4	14
165	L	2	2	1	1	3	9	12
166	V	3	4	1	2	3	13	14
167	E	4	3	1	3	1	12	13
168	T	0	0	0	1	0	1	15
169	I	2	0	0	1	0	3	15
170	K	2	2	0	1	0	5	13
171	L	2	0	0	0	0	2	11
172	G	1	1	0	0	0	2	12
173	K	0	1	0	1	0	2	14
174	L	4	4	0	2	2	12	13
175	L	1	1	1	1	1	5	12
176	R	1	0	0	1	0	2	15
177	P	0	0	0	0	0	0	13
178	N	2	0	0	2	0	4	10
179	H	7	4	1	2	0	14	15
180	L	2	0	0	1	1	4	10
181	W	1	1	0	0	0	2	13
182	G	1	0	0	0	0	1	14
183	Y	1	0	0	0	0	1	12
184	Y	0	0	0	0	0	0	14
185	L	0	0	0	0	0	0	14
186	F	0	0	0	0	1	1	15
187	P	0	0	0	0	0	0	15
188	D	0	0	x	0	0	0	15
189	C	0	0	0	0	0	0	15
190	Y	0	0	0	0	0	0	14
191	N	0	0	0	0	0	0	15
192	H	0	2	0	0	0	2	15
193	H	1	2	0	0	1	4	11
194	Y	0	0	0	0	0	0	8
195	K	4	3	0	2	1	10	10
196	K	2	2	1	1	1	7	7
197	P	4	3	2	3	1	13	14
198	G	2	4	2	2	1	11	12
199	Y	0	0	0	0	0	0	10
200	N	0	0	1	0	0	1	14
201	G	0	0	0	0	0	0	11
202	S	1	0	0	0	0	1	14
203	C	0	0	0	0	0	0	13
204	F	1	0	0	0	0	1	12
205	N	5	2	2	2	3	14	14
206	V	3	2	0	3	0	8	14
207	E	0	0	0	0	0	0	11
208	I	4	4	0	2	0	10	14
209	K	2	3	1	1	1	8	10
210	R	0	0	0	0	0	0	14
211	N	1	0	0	0	0	1	13
212	D	0	3	x	0	0	3	13
213	D	5	2	1	3	2	13	15
214	L	0	1	0	0	0	1	14
215	S	5	2	2	3	1	13	15
216	W	0	0	0	0	1	1	15
217	L	1	0	0	0	0	1	11
218	W	0	0	0	0	1	1	7
219	N	4	3	2	3	1	13	14
220	E	4	2	x	1	0	7	13
221	S	3	3	0	0	0	6	12
222	T	4	2	1	2	1	10	12
223	A	0	0	0	0	0	0	15
224	L	1	0	0	0	0	1	14
225	Y	0	0	0	0	0	0	12
226	P	0	0	0	0	0	0	14
227	S	0	0	0	0	0	0	15
228	I	0	0	0	0	0	0	13
229	Y	0	0	0	0	0	0	10
230	L	1	0	0	0	0	1	15
231	N	0	1	0	0	0	1	15
232	T	0	1	0	0	1	2	12
233	Q	3	0	0	2	2	7	12
234	Q	2	0	0	0	0	2	13
235	S	2	1	1	2	0	6	12
236	P	2	1	0	3	0	6	13
237	V	2	4	1	2	2	11	11
238	A	0	3	2	3	0	8	15
239	A	0	0	0	0	0	0	13
240	T	4	2	0	2	0	8	14
241	L	0	0	0	0	0	0	12
242	Y	0	0	0	0	1	1	14
243	V	0	0	0	0	0	0	13

Position	Wildtype (WT)	Small Hydrophobes (7)	Polar Uncharged (5)	Polar Negative (2)	Polar Positive (3)	Large Hydrophobes (3)	Count Active (>40%, no errors)	Tested (w/o errors)
244	R	0	0	0	0	0	0	6
245	N	0	0	0	1	0	1	11
246	R	0	0	0	0	0	0	15
247	V	3	0	0	0	0	3	15
248	R	1	0	0	1	2	4	12
249	E	0	0	x	0	0	1	10
250	A	0	0	0	0	0	0	15
251	I	2	0	0	0	0	2	11
252	R	0	0	0	0	0	0	14
253	V	1	0	0	0	0	1	13
254	S	0	0	0	0	0	0	14
255	K	1	3	0	1	0	5	11
256	I	3	0	0	1	0	4	9
257	P	6	3	0	2	0	11	12
258	D	0	2	x	2	0	4	8
259	A	5	4	1	2	2	14	14
260	K	3	2	2	2	1	10	12
261	S	3	3	0	2	2	10	11
262	P	0	0	0	0	0	0	15
263	L	3	1	0	2	0	6	11
264	P	1	0	0	1	0	2	13
265	V	1	0	0	0	0	1	13
266	F	0	0	0	0	1	1	13
267	A	1	1	0	0	0	2	11
268	Y	0	0	0	0	0	0	14
269	T	1	2	1	0	0	4	12
270	R	1	3	0	0	0	4	13
271	I	4	1	0	0	1	6	13
272	V	0	2	1	1	0	4	10
273	F	0	1	0	1	1	3	14
274	T	1	1	0	0	1	3	11
275	D	0	0	x	0	0	0	9
276	Q	3	2	2	2	1	10	13
277	V	3	5	2	3	1	14	14
278	L	4	3	1	3	2	13	15
279	K	0	2	0	2	0	4	10
280	F	1	1	0	0	0	2	12
281	L	0	0	0	0	0	0	13
282	S	2	1	1	0	0	4	9
283	Q	1	2	1	1	0	5	10
284	D	4	3	1	1	0	9	12
285	E	2	2	x	1	2	7	12
286	L	0	1	0	1	0	2	12
287	V	1	1	0	0	0	2	13
288	Y	1	0	0	0	1	2	12
289	T	0	0	0	1	0	1	11
290	F	2	0	0	0	0	2	5
291	G	1	2	0	1	0	4	13
292	E	4	2	x	3	2	11	14
293	T	6	3	1	1	2	13	15
294	V	1	0	0	0	0	1	14
295	A	0	0	0	0	0	0	10
296	L	0	0	0	0	0	0	13
297	G	1	0	0	0	0	1	11
298	A	2	0	0	0	0	2	13
299	S	0	0	0	0	0	0	11
300	G	0	0	0	1	0	1	15
301	I	2	0	0	0	0	2	14
302	V	1	0	0	0	1	2	15
303	I	1	0	0	0	0	1	12
304	W	0	0	0	0	0	0	11
305	G	0	0	2	0	0	2	10
306	T	0	1	2	0	0	3	11
307	L	1	4	0	1	2	8	11
308	S	2	2	1	3	0	8	15
309	I	4	4	2	3	1	14	14
310	M	3	2	0	1	0	6	10
311	R	0	3	0	2	0	6	14
312	S	1	2	0	0	0	3	8
313	M	5	2	1	3	1	12	13
314	K	2	4	1	2	1	10	13
315	S	4	1	1	3	1	10	13
316	C	0	0	1	0	0	1	14
317	L	3	4	1	3	1	12	14
318	L	3	4	1	3	1	12	15
319	L	0	0	0	0	0	0	15
320	D	3	2	1	3	2	11	13
321	N	1	2	1	3	1	8	11
322	Y	0	0	0	0	0	0	13
323	M	2	0	0	0	1	3	14
324	E	2	2	1	2	0	7	13

Position	Wildtype (WT)	Small Hydrophobes (7)	Polar Uncharged (5)	Polar Negative (2)	Polar Positive (3)	Large Hydrophobes (3)	Count Active (>40%, no errors)	Tested (w/o errors)
325	T	4	3	2	2	1	12	14
326	I	2	1	0	1	1	5	10
327	L	1	0	0	0	0	1	14
328	N	5	4	0	3	1	13	13
329	P	0	0	0	0	0	0	15
330	Y	0	0	0	0	0	0	14
331	I	1	0	0	0	0	1	12
332	I	0	0	0	0	0	0	15
333	N	0	0	0	0	0	0	10
334	V	1	0	0	0	0	1	10
335	T	0	1	0	0	0	1	11
336	L	0	0	0	0	0	0	13
337	A	0	0	0	0	0	0	10
338	A	0	1	0	0	0	1	14
339	K	1	0	0	0	0	1	14
340	M	0	0	0	0	0	0	14
341	C	0	0	0	0	0	0	14
342	S	1	0	0	0	0	1	13
343	Q	1	1	0	0	0	2	8
344	V	0	0	0	0	0	0	13
345	L	0	0	0	0	0	0	11
346	C	0	0	0	0	0	0	14
347	Q	3	1	1	1	0	6	13
348	E	1	1	1	0	0	3	14
349	Q	2	2	1	2	0	7	14
350	G	0	0	0	0	0	0	15
351	V	1	1	0	0	0	2	10
352	C	0	0	0	0	0	0	15
353	I	1	1	0	0	0	2	13
354	R	0	0	0	0	0	0	15
355	K	0	0	0	0	0	0	15
356	N	1	1	1	1	0	4	13
357	W	1	3	0	0	0	4	12
358	N	1	2	0	0	0	3	11
359	S	2	1	2	2	0	7	13
360	S	0	1	0	0	0	1	14
361	D	0	0	x	0	0	0	12
362	Y	0	0	0	0	0	0	15
363	L	0	0	0	0	0	0	15
364	H	0	0	0	0	0	0	15
365	L	0	0	0	0	0	0	14
366	N	0	0	0	0	0	0	12
367	P	2	2	0	2	0	6	13
368	D	5	2	1	3	0	11	14
369	N	0	1	0	2	0	3	13
370	F	0	0	0	0	0	0	14
371	A	4	1	1	3	1	10	12
372	I	0	0	0	0	0	0	15
373	Q	4	1	1	2	1	9	12
374	L	5	3	0	2	1	11	13
375	E	4	3	x	2	0	9	14
376	K	3	3	2	1	1	10	13
377	G	1	2	2	3	0	8	12
378	G	0	1	0	2	0	3	13
379	K	1	2	0	2	0	5	13
380	F	3	0	0	0	2	5	12
381	T	1	3	0	3	0	7	12
382	V	0	0	0	0	0	0	14
383	R	5	3	1	2	0	11	13
384	G	0	0	0	0	0	0	6
385	K	3	4	0	2	0	9	15
386	P	0	0	0	0	0	0	15
387	T	0	1	0	0	0	1	13
388	L	3	1	0	2	3	9	14
389	E	5	3	x	3	1	12	14
390	D	0	0	x	0	0	0	15
391	L	0	1	0	0	0	1	15
392	E	5	3	x	2	2	12	14
393	Q	3	3	1	3	1	11	13
394	F	1	0	0	0	1	2	15
395	S	2	0	0	3	1	6	10
396	E	1	3	1	2	0	7	13
397	K	0	0	0	0	0	0	12
398	F	0	0	0	0	0	0	14
399	Y	3	5	1	2	1	12	14
400	C	0	0	0	0	0	0	15
401	S	1	1	1	0	0	3	10
402	C	0	0	0	0	0	0	14
403	Y	0	0	0	0	1	1	14
404	S	2	0	0	0	0	2	14
405	T	4	2	0	2	2	10	13

Position	Wildtype (WT)	Small Hydrophobes (7)	Polar Uncharged (5)	Polar Negative (2)	Polar Positive (3)	Large Hydrophobes (3)	Count Active (>40%, no errors)	Tested (w/o errors)
406	L	4	5	1	0	2	12	14
407	S	6	2	2	2	2	14	14
408	C	0	0	0	0	0	0	14
409	K	5	3	2	2	0	12	12
410	E	3	4	1	2	1	11	12
411	K	3	3	0	1	0	7	11
412	A	5	2	1	1	2	11	12
413	D	0	2	1	1	0	4	7
414	V	2	0	0	0	0	2	10
415	K	1	1	0	0	1	3	7
416	D	4	3	x	3	2	12	14
417	T	1	0	0	0	0	1	12
418	D	7	3	1	1	2	14	14
419	A	3	2	1	3	3	12	14
420	V	2	0	0	0	0	2	15
421	D	4	4	1	3	1	13	13
422	V	1	1	0	0	0	2	13
423	C	0	0	0	0	0	0	15
424	I	0	0	0	0	0	0	11
425	A	3	2	0	2	0	7	11
426	D	2	2	1	0	1	6	10
427	G	1	1	0	1	0	3	11
428	V	3	1	0	0	0	4	14
429	C	0	0	0	0	0	0	11
430	I	0	0	0	0	0	0	9
431	D	5	3	1	3	2	14	15
432	A	2	2	1	0	0	5	13
433	F	6	3	2	3	1	15	15
434	L	4	0	0	0	1	5	11
435	K	3	3	1	2	1	10	10
436	P	4	3	2	3	2	14	14
437	P	5	2	1	3	1	12	13
438	M	5	5	2	1	1	14	15
439	E	5	4	x	2	2	13	15
440	T	7	1	2	2	2	14	15
441	E	4	4	1	2	2	13	14
442	E	4	3	x	3	1	11	14
443	P	4	4	1	2	2	13	14
444	Q	3	1	2	3	3	12	12
445	I	6	4	0	2	2	14	14
446	F	6	3	2	3	1	15	15
447	Y	6	3	2	1	2	14	14

Appendix B-3

Count Comparison $PH20_{1-447}$ Residues

Source(s): EX1026 ('731 App.), Tables 3, 5, 9, 10; EX2166 (Petsko Patent Analysis Data); EX2068 (Simpson Dec.), EX2070 (Petsko Dec.)

# of Positions Across "Tolerance" Groups				
	Dr. Petsko	Dr. Petsko Corrected	Dr. Naismith	Prior Art / Patent Confirmed Essential & Dr. Park Essential Residues
	94	108	21	75
	120	118	0	0
	75	72	0	0
	115	106	12	0
	43	43	0	0
	0	0	414	0

Note:
Cols. B & C:
Red - residues for which no tested substitutions produced an "active" PH20 protein

Orange - residues for which 30% or less of the tested substitutions produced an "active" protein

Unshaded - residues for which between 30% and 60% of the tested substitutions produced an "active" protein

Light Green - residues fior which ≥ 60% of the tested substitutions yielded an "active" PH20 protein

Green - residues for which all of the tested substitutions were "tolerated"

Grey - residues for which their "tolerability" to modification is unknown

Col. D (Dr. Naismith):
Red - residues reported in the prior art and "confirmed" by the common disclosure as being essential to PH20 structure and/or activity

Col. E:
Red - total number of residues reported in the prior art and "confirmed" by the common disclosure and identified by Dr. Park as being essential to PH20 structure and/or activity

	Residues with the Identified Characterstic
Buried or partially buried residues with fSASA < 0.65	5, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 28, 29, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 60, 61, 62, 63, 64, 65, 66, 67, 68, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 95, 96, 97, 99, 100, 102, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 141, 144, 145, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 196, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 236, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 358, 359, 360, 361, 362, 363, 364, 365, 366, 369, 370, 371, 372, 373, 374, 375, 378, 379, 380, 381, 382, 383, 385, 386, 387, 390, 391, 393, 394, 395, 397, 398, 399, 400, 401, 402, 403
Stretches of Residues with low local QMEAN (<0.60)	3, 4, 5, 27, 28, 29, 30, 31, 32, 33, 34, 35, 44, 45, 46, 47, 48, 104, 105, 234, 235, 236, 237, 256, 257, 258, 259, 260, 261, 262, 274, 275, 276, 277, 278, 357, 358, 405, 406
Residues having ≥ 11 substitutions yielding active mutants (> 40% activity)	1, 24, 26, 27, 28, 29, 30, 31, 32, 46, 47, 48, 50, 58, 63, 69, 70, 73, 74, 75, 82, 84, 86, 87, 90, 93, 94, 97, 98, 102, 120, 127, 131, 132, 135, 136, 138, 139, 140, 141, 142, 145, 146, 147, 148, 150, 151, 155, 156, 159, 160, 162, 166, 167, 174, 179, 197, 198, 205, 213, 215, 219, 237, 257, 259, 277, 278, 292, 293, 309, 313, 317, 318, 320, 325, 328, 368, 374, 383, 389, 392, 393, 399, 406, 407, 409, 410, 412, 416, 418, 419, 421, 431, 433, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447
Residues with ≥ 11 substitutions yielding active mutants (> 40% activity) and sufficient physicochemical	1, 24, 26, 27, 28, 29, 30, 31, 46, 47, 50, 69, 73, 74, 84, 86, 87, 90, 93, 94, 97, 98, 102, 120, 127, 131, 132, 135, 136, 139, 140, 141, 146, 147, 148, 150, 151, 155, 156, 162, 166, 167, 197, 198, 205, 213, 215, 219, 237, 259, 277, 278, 292, 293, 309, 313, 317, 318, 320, 325, 389, 392, 393, 399, 407, 410, 412, 416, 418, 419, 421, 431, 433, 436, 437, 438, 439, 440, 441, 442, 443, 444, 446, 447
Residues with ≥ 11 substitutions yielding active mutants (> 40% activity) but insufficient physicochemical	32, 48, 58, 63, 70, 75, 82, 138, 142, 145, 159, 160, 174, 179, 257, 328, 368, 374, 383, 406, 409, 445
Residues with: 1.) ≥ 11 substitutions yielding active mutants (> 40% activity); 2.) at least one substitution from each major physicochemical class; 3.) solvent exposed (*i.e.*, surface residue, fSASA > 0.65); and 4.) had a corresponding position in PH20 homology model.	27, 30, 69, 94, 98, 140, 146, 147, 197, 237, 389, 392
Prior art / common disclosure "confirmed essential" residues	25, 111, 113, 176, 189, 200, 203, 246, 249, 252, 316, 333, 341, 346, 352, 358, 400, 402, 408, 423, 429

Appendix B-4

Solvent Accessibile Surface Area Data of PH20

Source(s): EX2176 (Park Analysis Spreadsheet)

Res #	Res Name	SASA (Å)	fSASA	SS
39	T	139.073	0.966	L
40	A	48.781	0.439	L
41	P	65.732	0.487	L
42	P	38.050	0.282	L
43	V	33.672	0.226	L
44	I	52.912	0.306	L
45	P	103.226	0.765	L
46	N	90.515	0.545	L
47	V	16.843	0.113	L
48	P	8.983	0.067	L
49	F	0.000	0.000	S
50	L	8.367	0.047	S
51	W	2.350	0.009	S
52	A	7.617	0.069	S
53	W	1.440	0.006	S
54	N	10.312	0.062	S
55	A	0.706	0.006	L
56	P	3.031	0.022	L
57	S	6.223	0.050	H
58	E	83.710	0.448	H
59	F	63.014	0.303	H
60	C	1.435	0.009	H
61	L	96.728	0.540	H
62	G	76.159	0.886	H
63	K	124.901	0.589	H
64	F	48.655	0.234	H
65	D	130.140	0.813	L
66	E	25.495	0.136	L
67	P	84.022	0.622	L
68	L	16.981	0.095	L
69	D	44.945	0.281	L
70	M	68.964	0.343	L
71	S	65.053	0.520	L
72	L	36.039	0.201	L
73	F	14.561	0.070	L
74	S	51.573	0.413	L
75	F	43.210	0.208	S
76	I	60.110	0.347	S
77	G	10.926	0.127	S
78	S	13.844	0.111	L
79	P	20.476	0.152	L
80	R	113.880	0.456	L
81	I	60.998	0.353	L
82	N	90.218	0.543	L
83	A	83.814	0.755	L
84	T	92.499	0.642	L
85	G	33.490	0.389	L
86	Q	109.269	0.563	L

Res #	Res Name	SASA (Å)	fSASA	SS
87	G	8.324	0.097	L
88	V	36.816	0.247	S
89	T	12.982	0.090	S
90	I	10.222	0.059	S
91	F	0.000	0.000	S
92	Y	64.770	0.285	L
93	V	34.213	0.230	L
94	D	112.167	0.701	L
95	R	93.608	0.374	L
96	L	0.000	0.000	L
97	G	9.936	0.116	L
98	Y	92.692	0.408	L
99	Y	4.748	0.021	L
100	P	0.720	0.005	L
101	Y	51.764	0.228	L
102	I	8.003	0.046	L
103	D	54.181	0.339	L
104	S	111.432	0.891	L
105	I	144.199	0.834	L
106	T	88.846	0.617	L
107	G	20.860	0.243	S
108	V	79.688	0.535	S
109	T	78.760	0.547	S
110	V	45.747	0.307	S
111	N	69.106	0.416	L
112	G	28.072	0.326	L
113	G	0.000	0.000	L
114	I	5.990	0.035	L
115	P	4.421	0.033	L
116	Q	46.361	0.239	L
117	K	117.628	0.555	L
118	I	34.146	0.197	L
119	S	47.481	0.380	L
120	L	46.614	0.260	H
121	Q	112.172	0.578	H
122	D	78.456	0.490	H
123	H	0.000	0.000	H
124	L	35.091	0.196	H
125	D	79.306	0.496	H
126	K	67.628	0.319	H
127	A	1.383	0.012	H
128	K	114.335	0.539	H
129	K	141.266	0.666	H
130	D	16.707	0.104	H
131	I	0.795	0.005	H
132	T	60.642	0.421	H
133	F	170.109	0.818	H
134	Y	46.946	0.207	H

Res #	Res Name	SASA (Å)	fSASA	SS
135	M	5.098	0.025	L
136	P	90.756	0.672	L
137	V	57.386	0.385	L
138	D	104.174	0.651	L
139	N	4.550	0.027	L
140	L	99.070	0.553	L
141	G	3.543	0.041	L
142	M	6.944	0.035	S
143	A	0.000	0.000	S
144	V	1.591	0.011	S
145	I	0.000	0.000	S
146	D	30.647	0.192	S
147	W	8.741	0.035	S
148	E	81.629	0.437	L
149	E	54.483	0.291	L
150	W	5.584	0.022	L
151	R	29.305	0.117	L
152	P	6.733	0.050	L
153	T	7.966	0.055	L
154	W	22.218	0.089	L
155	A	53.068	0.478	L
156	R	2.963	0.012	L
157	N	2.114	0.013	L
158	W	91.137	0.366	L
159	K	177.903	0.839	L
160	P	102.480	0.759	L
161	K	37.646	0.178	H
162	D	48.435	0.303	H
163	V	58.060	0.390	H
164	Y	3.493	0.015	H
165	K	50.824	0.240	H
166	N	86.923	0.524	H
167	R	93.616	0.374	H
168	S	0.000	0.000	H
169	I	38.703	0.224	H
170	E	83.130	0.445	H
171	L	41.446	0.232	H
172	V	13.182	0.088	H
173	Q	57.823	0.298	H
174	Q	119.753	0.617	H
175	Q	160.155	0.826	H
176	N	65.932	0.397	L
177	V	99.585	0.668	L
178	Q	165.074	0.851	L
179	L	32.641	0.182	L
180	S	68.577	0.549	L
181	L	122.629	0.685	H
182	T	98.809	0.686	H

Res #	Res Name	SASA (Å)	fSASA	SS
183	E	92.061	0.492	H
184	A	0.000	0.000	H
185	T	47.998	0.333	H
186	E	86.410	0.462	H
187	K	102.753	0.485	H
188	A	0.000	0.000	H
189	K	65.497	0.309	H
190	Q	111.552	0.575	H
191	E	73.070	0.391	H
192	F	3.035	0.015	H
193	E	43.559	0.233	H
194	K	139.479	0.658	H
195	A	10.651	0.096	H
196	G	0.000	0.000	H
197	K	79.805	0.376	H
198	D	55.396	0.346	H
199	F	3.866	0.019	H
200	L	5.098	0.028	H
201	V	25.229	0.169	H
202	E	61.558	0.329	H
203	T	0.000	0.000	H
204	I	0.720	0.004	H
205	K	89.258	0.421	H
206	L	29.926	0.167	H
207	G	0.000	0.000	H
208	K	84.793	0.400	H
209	L	121.372	0.678	H
210	L	46.659	0.261	H
211	R	39.598	0.158	L
212	P	74.735	0.554	L
213	N	79.973	0.482	L
214	H	2.085	0.011	L
215	L	13.986	0.078	S
216	W	0.000	0.000	S
217	G	0.000	0.000	S
218	Y	5.539	0.024	S
219	Y	17.785	0.078	L
220	L	27.644	0.154	L
221	F	12.206	0.059	L
222	P	4.158	0.031	L
223	D	9.692	0.061	L
224	C	2.092	0.013	L
225	Y	24.249	0.107	L
226	N	3.670	0.022	L
227	H	98.584	0.516	L
228	H	101.781	0.533	L
229	Y	68.699	0.303	L
230	K	163.551	0.771	L

Res #	Res Name	SASA (Å)	fSASA	SS
231	K	101.758	0.480	L
232	P	144.905	1.073	L
233	G	50.427	0.586	L
234	Y	38.340	0.169	L
235	N	101.580	0.612	L
236	G	0.720	0.008	L
237	S	65.334	0.523	L
238	C	13.271	0.085	L
239	F	98.579	0.474	L
240	N	98.259	0.592	L
241	V	57.311	0.385	H
242	E	24.338	0.130	H
243	I	64.061	0.370	H
244	K	107.619	0.508	H
245	R	71.659	0.287	H
246	N	1.184	0.007	H
247	D	52.008	0.325	H
248	D	90.794	0.567	H
249	L	5.023	0.028	H
250	S	24.295	0.194	H
251	W	38.551	0.155	H
252	L	2.311	0.013	H
253	W	3.197	0.013	H
254	N	71.240	0.429	H
255	E	41.651	0.223	H
256	S	8.290	0.066	L
257	T	25.292	0.176	L
258	A	0.639	0.006	L
259	L	2.235	0.012	L
260	Y	0.000	0.000	L
261	P	0.760	0.006	L
262	S	6.187	0.049	L
263	I	2.505	0.014	L
264	Y	35.171	0.155	L
265	L	3.400	0.019	L
266	N	69.570	0.419	L
267	T	56.716	0.394	L
268	Q	94.120	0.485	L
269	Q	10.225	0.053	L
270	S	102.880	0.823	L
271	P	58.304	0.432	L
272	V	135.527	0.910	L
273	A	12.639	0.114	H
274	A	2.243	0.020	H
275	T	24.367	0.169	H
276	L	57.938	0.324	H
277	Y	10.153	0.045	H
278	V	0.000	0.000	H

Res #	Res Name	SASA (Å)	fSASA	SS
279	R	68.249	0.273	H
280	N	3.179	0.019	H
281	R	7.459	0.030	H
282	V	0.000	0.000	H
283	R	89.071	0.356	H
284	E	2.065	0.011	H
285	A	0.000	0.000	H
286	I	38.094	0.220	H
287	R	58.364	0.233	H
288	V	0.000	0.000	H
289	S	2.200	0.018	H
290	K	117.817	0.556	H
291	I	54.806	0.317	L
292	P	84.209	0.624	L
293	D	92.576	0.579	L
294	A	21.353	0.192	L
295	K	175.390	0.827	L
296	S	68.849	0.551	L
297	P	57.941	0.429	L
298	L	25.418	0.142	L
299	P	16.735	0.124	L
300	V	6.141	0.041	L
301	F	2.350	0.011	L
302	A	0.760	0.007	L
303	Y	12.220	0.054	L
304	T	2.823	0.020	L
305	R	13.176	0.053	L
306	I	7.853	0.045	L
307	V	2.114	0.014	L
308	F	26.679	0.128	L
309	T	23.215	0.161	L
310	D	111.462	0.697	L
311	Q	46.527	0.240	L
312	V	53.292	0.358	L
313	L	110.288	0.616	L
314	K	105.162	0.496	L
315	F	57.997	0.279	L
316	L	5.232	0.029	L
317	S	53.343	0.427	L
318	Q	106.552	0.549	H
319	D	79.206	0.495	H
320	E	11.785	0.063	H
321	L	0.000	0.000	H
322	V	41.730	0.280	H
323	Y	63.463	0.280	H
324	T	3.830	0.027	H
325	F	0.000	0.000	H
326	G	3.146	0.037	H

Res #	Res Name	SASA (Å)	fSASA	SS
327	E	1.380	0.007	H
328	T	0.000	0.000	H
329	V	1.829	0.012	H
330	A	0.000	0.000	H
331	L	28.766	0.161	H
332	G	23.523	0.274	L
333	A	2.079	0.019	L
334	S	26.443	0.212	L
335	G	0.000	0.000	S
336	I	2.018	0.012	S
337	V	2.862	0.019	S
338	I	0.000	0.000	S
339	W	69.383	0.279	S
340	G	5.680	0.066	L
341	T	19.027	0.132	L
342	L	112.244	0.627	H
343	S	64.223	0.514	H
344	I	7.201	0.042	H
345	M	1.383	0.007	H
346	R	177.550	0.710	H
347	S	42.440	0.340	L
348	M	107.328	0.534	H
349	K	136.842	0.645	H
350	S	11.506	0.092	H
351	C	2.591	0.017	H
352	L	75.812	0.424	H
353	L	72.002	0.402	H
354	L	2.100	0.012	H
355	D	23.798	0.149	H
356	N	76.586	0.461	H
357	Y	10.085	0.044	H
358	M	3.715	0.018	H
359	E	89.034	0.476	H
360	T	90.864	0.631	H
361	I	43.424	0.251	H
362	L	0.000	0.000	H
363	N	0.652	0.004	H
364	P	32.519	0.241	H
365	Y	32.104	0.141	H
366	I	0.720	0.004	H
367	I	31.326	0.181	H
368	N	35.677	0.215	H
369	V	0.720	0.005	H
370	T	33.546	0.233	H
371	L	29.148	0.163	H
372	A	0.000	0.000	H
373	A	3.664	0.033	H
374	K	57.662	0.272	H

Res #	Res Name	SASA (Å)	fSASA	SS
375	M	33.489	0.167	H
376	C	0.000	0.000	H
377	S	5.448	0.044	H
378	Q	58.968	0.304	H
379	V	49.922	0.335	H
380	L	31.197	0.174	H
381	C	0.000	0.000	L
382	Q	86.748	0.447	L
383	E	101.491	0.543	L
384	Q	70.960	0.366	L
385	G	0.000	0.000	S
386	V	18.705	0.126	S
387	C	0.652	0.004	S
388	I	42.834	0.248	S
389	R	20.892	0.084	S
390	K	54.632	0.258	L
391	N	147.976	0.891	L
392	W	219.026	0.880	L
393	N	54.014	0.325	L
394	S	72.418	0.579	L
395	S	71.417	0.571	L
396	D	29.415	0.184	L
397	Y	39.880	0.176	L
398	L	6.819	0.038	L
399	H	91.537	0.479	L
400	L	11.873	0.066	L
401	N	46.049	0.277	L
402	P	105.954	0.785	L
403	D	140.876	0.880	L
404	N	25.624	0.154	L
405	F	20.643	0.099	S
406	A	35.895	0.323	S
407	I	44.651	0.258	S
408	Q	103.495	0.533	S
409	L	73.823	0.412	S
410	E	106.861	0.571	L
411	K	205.579	0.970	L
412	G	78.889	0.917	L
413	G	28.816	0.335	S
414	K	131.483	0.620	S
415	F	19.524	0.094	S
416	T	23.941	0.166	S
417	V	41.918	0.281	S
418	R	140.040	0.560	S
419	G	61.768	0.718	L
420	K	135.441	0.639	L
421	P	50.489	0.374	L
422	T	51.574	0.358	L

Res #	Res Name	SASA (Å)	fSASA	SS
423	L	132.293	0.739	H
424	E	134.042	0.717	H
425	D	19.943	0.125	H
426	L	55.602	0.311	H
427	E	122.480	0.655	H
428	Q	93.910	0.484	H
429	F	16.147	0.078	H
430	S	52.380	0.419	H
431	E	123.156	0.659	H
432	K	28.635	0.135	H
433	F	0.000	0.000	S
434	Y	119.359	0.526	S
435	C	20.202	0.129	S
436	S	55.944	0.448	S
437	C	40.505	0.258	S
438	Y	13.826	0.061	L

based on 587 structures

Analysis of accessible surface of residues in proteins

LAURENCE LINS, ANNICK THOMAS, AND ROBERT BRASSEUR

Centre de Biophysique Moléculaire Numérique (CBMN), Faculté des Sciences Agronomiques de Gembloux (FSAGx), 5030 Gembloux, Belgium

(RECEIVED January 31, 2003; FINAL REVISION March 26, 2003; ACCEPTED March 28, 2003)

residue	max sasa	med sasa	med sasa fraction
A	111	14	0.126
C	157	5	0.032
D	160	62	0.388
E	187	83	0.444
F	208	13	0.063
G	86	19	0.221
H	191	46	0.241
I	173	6	0.035
K	212	102	0.481
L	179	9	0.050
M	201	13	0.065
N	166	59	0.355
P	135	49	0.363
Q	194	74	0.381
R	250	87	0.348
S	125	35	0.280
T	144	37	0.257
V	149	8	0.054
W	249	25	0.100
Y	227	31	0.137

Note:
Information in this sheet was obtained from EX1004 (Park Decl.), Appx. D, Part 5: SASA, at 178.

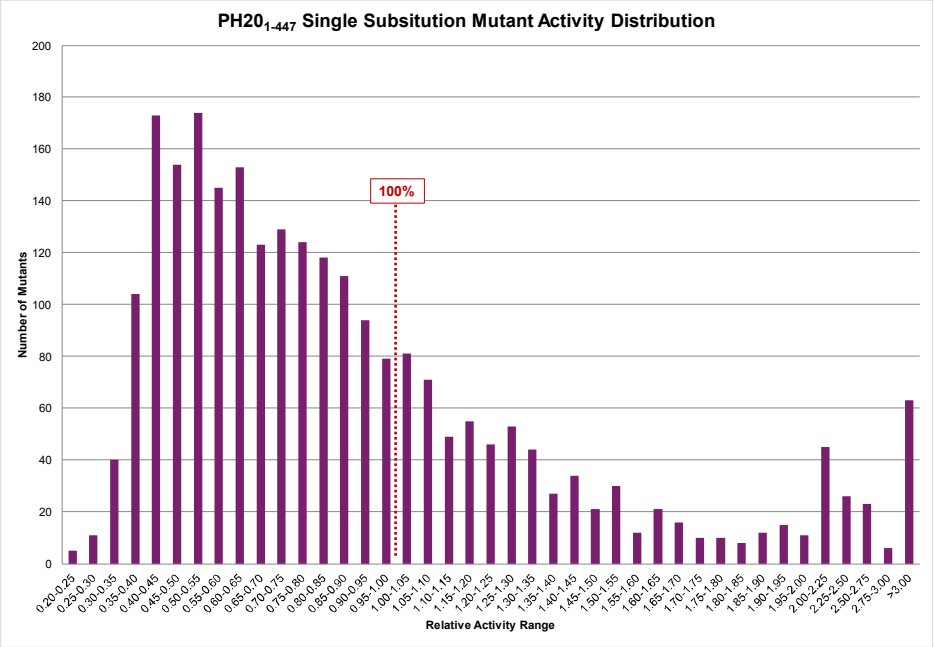
Information in this sheet was also obtained from EX2175, at 0062-0065.

Appendix B-5

Histogram of $PH20_{1-447}$ Activity Data

Source(s): EX1026 ('731 App.), at Table 9

Activity Range	Count
0.20-0.25	5
0.25-0.30	11
0.30-0.35	40
0.35-0.40	104
0.40-0.45	173
0.45-0.50	154
0.50-0.55	174
0.55-0.60	145
0.60-0.65	153
0.65-0.70	123
0.70-0.75	129
0.75-0.80	124
0.80-0.85	118
0.85-0.90	111
0.90-0.95	94
0.95-1.00	79
1.00-1.05	81
1.05-1.10	71
1.10-1.15	49
1.15-1.20	55
1.20-1.25	46
1.25-1.30	53
1.30-1.35	44
1.35-1.40	27
1.40-1.45	34
1.45-1.50	21
1.50-1.55	30
1.55-1.60	12
1.60-1.65	21
1.65-1.70	16
1.70-1.75	10
1.75-1.80	10
1.80-1.85	8
1.85-1.90	12
1.90-1.95	15
1.95-2.00	11
2.00-2.25	45
2.25-2.50	26
2.50-2.75	23
2.75-3.00	6
>3.00	63
TOTAL	**2526**



Note:

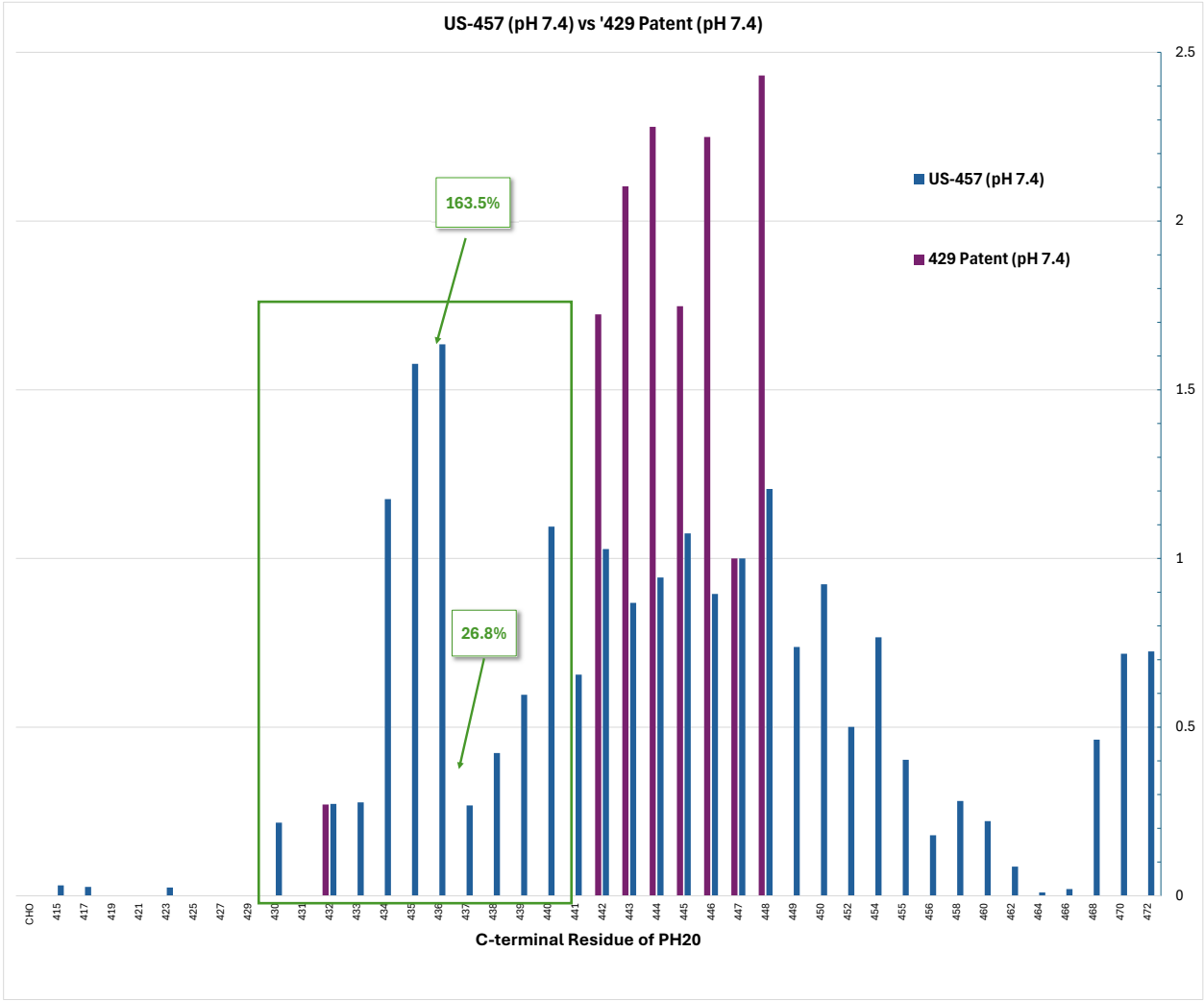
Information in this spreadsheet was obtained from the common disclosure (EX1026 ('731 App.), Table 9).

Appendix B-6

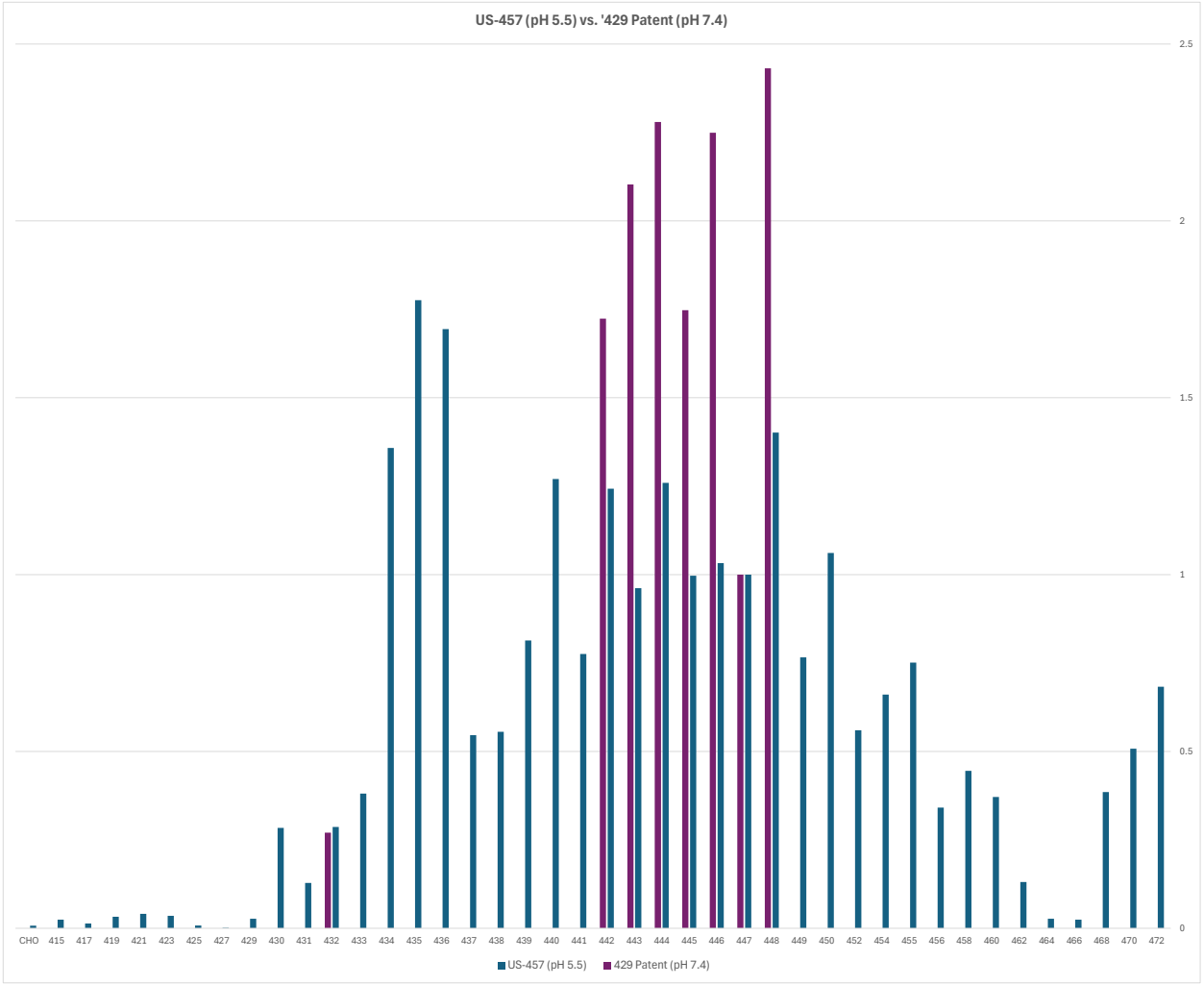
Activity Data Plot of C-terminal Truncation PH20 Mutants from US-457 and the '429 Patent

Source(s): EX2165 (US-457), EX1005 ('429 Patent)

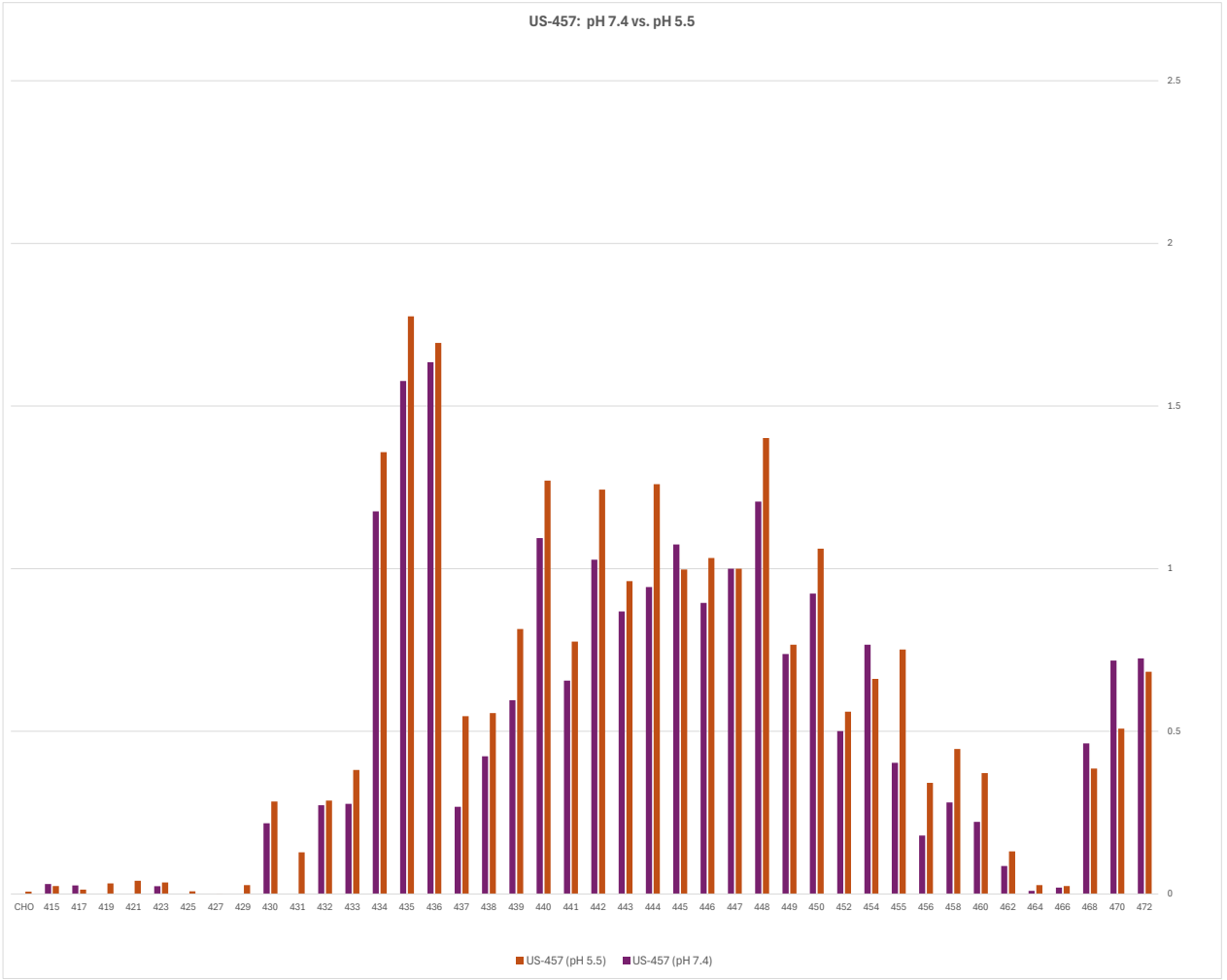
Deletion Mutant	Precursor	Mature	US-457 (pH 7.4)	429 Patent (pH 7.4)	US-457 pH7.4 (Units/ml)
SPAM1-SSVA	507	472	72.4%		1.4715
SPAM1-IISS	505	470	71.8%		1.458
SPAM1-FLII	503	468	46.3%		0.9405
SPAM1-ILFL	501	466	2.0%		0.0405
SPAM1-VSIL	499	464	1.0%		0.02025
SPAM1-FIVS	497	462	8.6%		0.1755
SPAM1-TMFI	495	460	22.1%		0.45
SPAM1-SATM	493	458	28.1%		0.5715
SPAM1-TLSA	491	456	17.9%		0.3645
SPAM1-STLS	490	455	40.3%		0.819
SPAM1-PSTL	489	454	76.6%		1.557
SPAM1-ASPS	487	452	50.1%		1.017
SPAM1-YNAS	485	450	92.4%		1.8765
SPAM1-FYNA	484	449	73.8%		1.4985
SPAM1-IFYN	483	448	120.6%	243.2%	2.45025
SPAM1-QIFY	482	447	100.0%	100.0%	2.03175
SPAM1-PQIF	481	446	89.5%	224.9%	1.818
SPAM1-EPQI	480	445	107.4%	174.8%	2.1825
SPAM1-EEPQ	479	444	94.4%	228.0%	1.917
SPAM1-TEEP	478	443	86.8%	210.3%	1.764
SPAM1-ETEE	477	442	102.8%	172.3%	2.088
SPAM1-METE	476	441	65.6%		1.332
SPAM1-PMET	475	440	109.4%		2.223
SPAM1-PPME	474	439	59.6%		1.2105
SPAM1-KPPM	473	438	42.3%		0.8595
SPAM1-LKPP	472	437	26.8%		0.5445
SPAM1-FLKP	471	436	163.5%		3.321
SPAM1-AFLK	470	435	157.7%		3.204
SPAM1-DAFL	469	434	117.6%		2.3895
SPAM1-IDAF	468	433	27.7%		0.5625
SPAM1-CIDA	467	432	27.2%	27.1%	0.5535
SPAM1-VCID	466	431	0.0%		0
SPAM1-GVCI	465	430	21.7%		0.441
SPAM1-DGVC	464	429	0.0%		0
SPAM1-IADG	462	427	0.0%		0
SPAM1-VCIA	460	425	0.0%		0
SPAM1-VDVC	458	423	2.4%		0.0495
SPAM1-DAVD	456	421	0.0%		0
SPAM1-DTDA	454	419	0.0%		0
SPAM1-VKDT	452	417	2.7%		0.054
SPAM1-ADVK	450	415	3.1%		0.063
CHO-S	CHO	CHO	0.0%		0



Deletion Mutant	Precursor	Mature	US-457 (pH 5.5)	429 Patent (pH 7.4)	US-457 (pH 5.5)(Units/ml)
SPAM1-SSVA	507	472	68.3%		1.125
SPAM1-IISS	505	470	50.8%		0.837
SPAM1-FLII	503	468	38.5%		0.6345
SPAM1-ILFL	501	466	2.5%		0.0405
SPAM1-VSIL	499	464	2.7%		0.045
SPAM1-FIVS	497	462	13.1%		0.216
SPAM1-TMFI	495	460	37.2%		0.612
SPAM1-SATM	493	458	44.5%		0.7335
SPAM1-TLSA	491	456	34.2%		0.5625
SPAM1-STLS	490	455	75.1%		1.2375
SPAM1-PSTL	489	454	66.1%		1.089
SPAM1-ASPS	487	452	56.0%		0.9225
SPAM1-YNAS	485	450	106.1%		1.74825
SPAM1-FYNA	484	449	76.6%		1.26225
SPAM1-IFYN	483	448	140.2%	243.2%	2.3085
SPAM1-QIFY	482	447	100.0%	100.0%	1.647
SPAM1-PQIF	481	446	103.3%	224.9%	1.701
SPAM1-EPQI	480	445	99.7%	174.8%	1.6425
SPAM1-EEPQ	479	444	126.0%	228.0%	2.0745
SPAM1-TEEP	478	443	96.2%	210.3%	1.584
SPAM1-ETEE	477	442	124.3%	172.3%	2.0475
SPAM1-METE	476	441	77.6%		1.278
SPAM1-PMET	475	440	127.0%		2.0925
SPAM1-PPME	474	439	81.4%		1.341
SPAM1-KPPM	473	438	55.6%		0.91575
SPAM1-LKPP	472	437	54.6%		0.9
SPAM1-FLKP	471	436	169.4%		2.79
SPAM1-AFLK	470	435	177.6%		2.925
SPAM1-DAFL	469	434	135.8%		2.2365
SPAM1-IDAF	468	433	38.1%		0.62775
SPAM1-CIDA	467	432	28.7%	27.1%	0.4725
SPAM1-VCID	466	431	12.8%		0.2115
SPAM1-GVCI	465	430	28.4%		0.468
SPAM1-DGVC	464	429	2.7%		0.045
SPAM1-IADG	462	427	0.1%		0.00225
SPAM1-VCIA	460	425	0.8%		0.0135
SPAM1-VDVC	458	423	3.6%		0.0585
SPAM1-DAVD	456	421	4.1%		0.0675
SPAM1-DTDA	454	419	3.3%		0.054
SPAM1-VKDT	452	417	1.4%		0.0225
SPAM1-ADVK	450	415	2.5%		0.0405
CHO-S	CHO	CHO	0.8%		0.012375



Deletion Mutant	Precursor	Mature	US-457 (pH 5.5)	US-457 (pH 7.4)	US 457 - 7.4 vs. 5.5
SPAM1-SSVA	507	472	68.3%	72.4%	130.8%
SPAM1-IISS	505	470	50.8%	71.8%	174.2%
SPAM1-FLII	503	468	38.5%	46.3%	148.2%
SPAM1-ILFL	501	466	2.5%	2.0%	100.0%
SPAM1-VSIL	499	464	2.7%	1.0%	45.0%
SPAM1-FIVS	497	462	13.1%	8.6%	81.3%
SPAM1-TMFI	495	460	37.2%	22.1%	73.5%
SPAM1-SATM	493	458	44.5%	28.1%	77.9%
SPAM1-TLSA	491	456	34.2%	17.9%	64.8%
SPAM1-STLS	490	455	75.1%	40.3%	66.2%
SPAM1-PSTL	489	454	66.1%	76.6%	143.0%
SPAM1-ASPS	487	452	56.0%	50.1%	110.2%
SPAM1-YNAS	485	450	106.1%	92.4%	107.3%
SPAM1-FYNA	484	449	76.6%	73.8%	118.7%
SPAM1-IFYN	483	448	140.2%	120.6%	106.1%
SPAM1-QIFY	482	447	100.0%	100.0%	123.4%
SPAM1-PQIF	481	446	103.3%	89.5%	106.9%
SPAM1-EPQI	480	445	99.7%	107.4%	132.9%
SPAM1-EEPQ	479	444	126.0%	94.4%	92.4%
SPAM1-TEEP	478	443	96.2%	86.8%	111.4%
SPAM1-ETEE	477	442	124.3%	102.8%	102.0%
SPAM1-METE	476	441	77.6%	65.6%	104.2%
SPAM1-PMET	475	440	127.0%	109.4%	106.2%
SPAM1-PPME	474	439	81.4%	59.6%	90.3%
SPAM1-KPPM	473	438	55.6%	42.3%	93.9%
SPAM1-LKPP	472	437	54.6%	26.8%	60.5%
SPAM1-FLKP	471	436	169.4%	163.5%	119.0%
SPAM1-AFLK	470	435	177.6%	157.7%	109.5%
SPAM1-DAFL	469	434	135.8%	117.6%	106.8%
SPAM1-IDAF	468	433	38.1%	27.7%	89.6%
SPAM1-CIDA	467	432	28.7%	27.2%	117.1%
SPAM1-VCID	466	431	12.8%	0.0%	0.0%
SPAM1-GVCI	465	430	28.4%	21.7%	94.2%
SPAM1-DGVC	464	429	2.7%	0.0%	0.0%
SPAM1-IADG	462	427	0.1%	0.0%	0.0%
SPAM1-VCIA	460	425	0.8%	0.0%	0.0%
SPAM1-VDVC	458	423	3.6%	2.4%	84.6%
SPAM1-DAVD	456	421	4.1%	0.0%	0.0%
SPAM1-DTDA	454	419	3.3%	0.0%	0.0%
SPAM1-VKDT	452	417	1.4%	2.7%	240.0%
SPAM1-ADVK	450	415	2.5%	3.1%	155.6%
CHO-S	CHO	CHO	0.8%	0.0%	0.0%



Note:
The information in this sheet was obtained from Published U.S. Patent Application 2010/0143457 ("US-457"), Example 5, ¶¶ 360-362 (Table 6).

Appendix B-7

Comparison Table of Dr. Simpson's Activity Classification of C-terminal Truncation PH20 Mutants vs. US-457

Source(s): EX2068 (Simpson Dec.), EX2165 (US-457)

Terminus (precursor)	Terminus (mature)	pH 7.4	pH 5.5	Normalized Activity (pH 7.4)	Normalized Activity (pH 5.5)	Simpson Classification	US-457 Classification	Table
482	**447**	**2.03175**	**1.647**	**100%**	**100%**	**Active**	**Active**	**6**
465	**430**	0.441	0.468	22%	28%	Inactive	Active	6
466	**431**	0.000	0.212	0%	13%	Inactive	Active	6
467	**432**	0.545	0.900	27%	55%	Reduced Activity	Active	6
468	**433**	0.860	0.916	42%	56%	Reduced Activity	Active	6
472	**437**	0.545	0.900	27%	55%	Active	Active	6
490	**455**	0.819	1.238	40%	75%	Reduced Activity	Active	6
491	**456**	0.365	0.563	18%	34%	Reduced Activity	Active	6
491	**456**	0.851	0.729	42%	44%	Reduced Activity	Active	6A
492	**457**	n/t	n/t	n/a	n/a	Reduced Activity	-	none
493	**458**	0.689	0.851	34%	52%	Reduced Activity	Active	6
494	**459**	n/t	n/t	n/a	n/a	Reduced Activity	-	none
495	**460**	0.689	0.770	34%	47%	Reduced Activity	Active	6
496	**461**	n/t	n/t	n/a	n/a	Reduced Activity	-	none
497	**462**	0.176	0.216	9%	13%	Reduced Activity	Active	6
497	**462**	0.365	0.162	18%	10%	Reduced Activity	Active	6A
498	**463**	0.527	0.203	26%	12%	Inactive/Little Activity	Low expression	6A
499	**464**	0.020	0.045	1%	3%	Inactive/Little Activity	Low expression	6
499	**464**	0.486	0.122	24%	7%	Inactive/Little Activity	Low expression	6A
500	**465**	0.365	0.162	18%	10%	Inactive/Little Activity	Low expression	6A
501	**466**	0.446	0.122	22%	7%	None	Active	6
502	**467**	0.608	0.324	30%	20%	None	Active	6
503	**468**	1.094	0.648	54%	39%	None	Active	6

Note:
The information in this sheet is from Published U.S. Patent Application 2010/0143457 ("US-457"), Example 5, ¶¶ 360-362 (Table 6) and EX2068 (Simpson Decl.), ¶ 188.

Appendix B-8

Frequency Chart of Alternative Residues Identified in MSA and Number of MSA Alternatives Tested

Source(s): EX1004 (Park Dec.), EX2176 (Park Analysis Spreadsheet)

Position	WT	MSA Alternates	# of MSA Alternatives	Alternatives Assessed	# of MSA Alternatives Assessed	% of MSA Alternatives Assessed
1	L					
2	N					
3	F					
4	R	RTAQSKG	7		0	0%
5	A	ARQGTN	6		0	0%
6	P	PLAGDRWSH	9		0	0%
7	P	PTL	3		0	0%
8	V	IVLMTQNF	8	IL	2	25%
9	I	IFLYVGS	7	FLYV	4	57%
10	P	PTQSKGLRV	9	TQSK	4	44%
11	N	NRGEKDSHPA	10	RGED	4	40%
12	V	RKVQHTLSAICE	12	RK	2	17%
13	P	PATNSIG	7		0	0%
14	F		0		0	0%
15	L	ILVTSNAMQC	10	IVTS	4	40%
16	W	VATWISLGC	9	VAT	3	33%
17	A	AVILFG	6	VIL	3	50%
18	W		0		0	0%
19	N	NDA	3	D	1	33%
20	A	AVIGT	5	V	1	20%
21	P	PND	3	ND	2	67%
22	S	TSILV	5	T	1	20%
23	E	EQDAHSNLMYF	11	QDA	3	27%
24	F	DQWRLFSHNYPKTAE	15	DQWRL	5	33%
25	C		0		0	0%
26	L	LKGARQMPTVISEW	14	KGARQ	5	36%
27	G	PEIGKTQALRDMSNF	15	PEI	3	20%
28	K	RKNQMIHEGFD	11	R	1	9%
29	F	FYHLTVS	7	YHL	3	43%
30	D	NGKDSQYHR	9	NGK	3	33%
31	E	VLEIMQTHD	9	V	1	11%
32	P	PDRHTSAGNQKE	12	DRHS	4	33%
33	L	LVIPMFT	7	VI	2	29%
34	D	DNPQS	5	NP	2	40%
35	M	LVMFIT	6	LV	2	33%
36	S	SKNEDQRGA	9	KNEDQ	5	56%
37	L	VLAMIFQSDT	10	VAM	3	30%
38	F	FLY	3	L	1	33%
39	S	DQSGNPHTE	9	DQG	3	33%
40	F	VILFMA	6	VIL	3	50%
41	I	VIQTLNKFHAMSE	13	VQ	2	15%
42	G	AGVPHTSQLE	10	A	1	10%
43	S	SNT	3	NT	2	67%
44	P	PTHKRQCDLS	10	R	1	10%
45	R	NGRLFEQSHKA	11	NGL	3	27%
46	I	EQAKILGSTHN	11	EQAK	4	36%
47	N	TGKSRNDHEA	10	TGKS	4	40%
48	A	FALVQIDSWYKT	12	F	1	8%

Position	WT	MSA Alternates	# of MSA Alternatives	Alternatives Assessed	# of MSA Alternatives Assessed	% of MSA Alternatives Assessed
49	T	RVTSHQIMFAKN	12	RV	2	17%
50	G	GNDARK	6	N	1	17%
51	Q	QPSRNHD	7	PS	2	29%
52	G	NDGPTSEFKIQ	11	ND	2	18%
53	V	IVML	4	IML	3	75%
54	T	TVSA	4		0	0%
55	I	ILVT	4	L	1	25%
56	F		0		0	0%
57	Y		0		0	0%
58	V	VSKRAPTHQY	10	SKRAPH	6	60%
59	D	DNSETRKYAW	10	NSE	3	30%
60	R	RQKEHL	6	Q	1	17%
61	L	LFI	3		0	0%
62	G		0		0	0%
63	Y	YLTMFHSK	8	LTF	3	38%
64	Y		0		0	0%
65	P		0		0	0%
66	Y	YHWRSK	6	HWR	3	50%
67	I	YFILVHK	7	YF	2	29%
68	D	TDNGSEK	7	TNG	3	43%
69	S	SPEADTKQLH	10	PET	3	30%
70	I	QTARDSHEIVNKGM	14	QTAR	4	29%
71	T	TNS	3	S	1	33%
72	G	GENDQWMKHRS	11		0	0%
73	V	VEKTRADMIQL	11	EKTR	4	36%
74	T	PSATIENHVK	10	PSA	3	30%
75	V	VIHFLSYKT	9	IH	2	22%
76	N	NHFYSAK	7	HF	2	29%
77	G		0		0	0%
78	G		0		0	0%
79	I	LIV	3	LV	2	67%
80	P		0		0	0%
81	Q		0		0	0%
82	K	NLKARVQSE	9	NL	2	22%
83	I	AGVISTQCENLFM	13	AGV	3	23%
84	S	SNPYLD	6	N	1	17%
85	L	LIMRY	5		0	0%
86	Q	QDKSWIAREVTGNMC	15	DKSWI	5	33%
87	D	AVETKDRQSHCFNLM	15	AVE	3	20%
88	H		0		0	0%
89	L	LRIY	4		0	0%
90	D	EAKDRNGSITPFQ	13	EAK	3	23%
91	K	KRMHQLCVIY	10	RM	2	20%
92	A	ALTSVMC	7	LT	2	29%
93	K	KFADRQPEYSNVGT	14	FADRQ	5	36%
94	K	QEKDGHYSLVNTMA	14	QE	2	14%
95	D	DQGRHENSAP	10	Q	1	10%
96	I	IVLF	4	V	1	25%

Position	WT	MSA Alternates	# of MSA Alternatives	Alternatives Assessed	# of MSA Alternatives Assessed	% of MSA Alternatives Assessed
97	T	LENRKAQHDTGMSV	14	LEN	3	21%
98	F	HYAKFREDNTLQV	13	HYAK	4	31%
99	Y	YASTLFNRVKG	11	A	1	9%
100	M	IMLV	4	IL	2	50%
101	P	PRTSQGWKH	9	R	1	11%
102	V	SATDVIELWNMHYK	14	SAT	3	21%
103	D	DEGRNSKTAWI	11	E	1	9%
104	N	FSNKPYQVRA	10	FS	2	20%
105	L	SAEVLRTNQHKDPWFI	16	SA	2	13%
106	G		0		0	0%
107	M	LMIVP	5	L	1	20%
108	A	AGVS	4		0	0%
109	V		0		0	0%
110	I	ILVM	4		0	0%
111	D		0		0	0%
112	W		0		0	0%
113	E		0		0	0%
114	E	EAYDNSGKHF	10	AYDNS	5	50%
115	W		0		0	0%
116	R	RCKYESL	7		0	0%
117	P		0		0	0%
118	T	QLTVRIEK	8	QLVR	4	50%
119	W		0		0	0%
120	A	AVDIESMLTK	10	VDIE	4	40%
121	R	RFGWLC	6	F	1	17%
122	N		0		0	0%
123	W		0		0	0%
124	K	GDKQNARHE	9	GDQN	4	44%
125	P	PTSDARNEKLFG	12	TSD	3	25%
126	K	KRMG	4	R	1	25%
127	D	DRNIQALKEHMT	12		0	0%
128	V	VIAN	4	I	1	25%
129	Y		0		0	0%
130	K	RKQWLM	6	RQ	2	33%
131	N	QNRKEADIHLTM	12	QR	2	17%
132	R	KRALSHQNVEGW	12	KALSH	5	42%
133	S		0		0	0%
134	I	RILKWQMVEC	10	R	1	10%
135	E	EAQKTNRHVD	10	AQK	3	30%
136	L	LWFQIKY	7		0	0%
137	V	VIAT	4	IA	2	50%
138	Q	QRSAKLEF	8	RSAK	4	50%
139	Q	QSADEKHNWRILG	13	SADEK	5	38%
140	Q	QRMKHETVLANF	12	RMKHEV	6	50%
141	N	HNQFGDYKL	9	HQ	2	22%
142	V	PEVIRKQASLT	11	P	1	9%
143	Q	DNQTHGSYEFK	11	DNT	3	27%
144	L	WLVIMFAY	8	WV	2	25%

Position	WT	MSA Alternates	# of MSA Alternatives	Alternatives Assessed	# of MSA Alternatives Assessed	% of MSA Alternatives Assessed
145	S	SPNTDAQKL	9	PNTD	4	44%
146	L	PAELIHSVQDFRTKM	15	PAE	3	20%
147	T	TDPEAQSKNMW	11	DPEAQS	6	55%
148	E	EQDRKALWSY	10	QDRK	4	40%
149	A	VIAAQLS	6	VI	2	33%
150	T	ETVKLRDANIWHGS	14	EV	2	14%
151	E	KAYTEHRQDNSGLC	14	KAY	3	21%
152	K	LKVQEAIRSMT	11	LVQE	4	36%
153	A		0		0	0%
154	K	KQRVLIYAHE	10	QR	2	20%
155	Q	YDAQVETKRIFNLM	14	YDAVET	6	43%
156	E	EQTDGSAVR	9	QTDG	4	44%
157	F		0		0	0%
158	E	EQND	4	Q	1	25%
159	K	EKFNQGSTRLCAM	13	EFN	3	23%
160	A	ASCET	5	S	1	20%
161	G	AGS	3	A	1	33%
162	K	RKQEMSC	7	RQ	2	29%
163	D	AQSNDEKHTCRLVI	14	AQS	3	21%
164	F	FWLY	4	WL	2	50%
165	L	ML	2	M	1	50%
166	V	LEKANVQTMIS	11	LEKA	4	36%
167	E	EGDQKSTVIHRLNY	14	G	1	7%
168	T		0		0	0%
169	I	LI	2	L	1	50%
170	K	KRQETSAHWLN	11	RQE	3	27%
171	L	LYFVHKQW	8	Y	1	13%
172	G	GAV	3	AV	2	67%
173	K	KRQILEMS	8	RQIL	4	50%
174	L	AKSLENTRVFMIQG	14	AKS	3	21%
175	L	LSMVFRTYKAEN	12	SMVF	4	33%
176	R		0		0	0%
177	P		0		0	0%
178	N	RKNQHSGYDAE	11	RK	2	18%
179	H	GHQRYACS	8	GQ	2	25%
180	L		0		0	0%
181	W		0		0	0%
182	G		0		0	0%
183	Y	FYL	3	F	1	33%
184	Y		0		0	0%
185	L	LGRHN	5	G	1	20%
186	F	FY	2	Y	1	50%
187	P		0		0	0%
188	D	DANSCEV	7		0	0%
189	C		0		0	0%
190	Y	YHGF	4	H	1	25%
191	N		0		0	0%
192	H	YHNGTSD	7	Y	1	14%

Position	WT	MSA Alternates	# of MSA Alternatives	Alternatives Assessed	# of MSA Alternatives Assessed	% of MSA Alternatives Assessed
193	H	DNHWGKYQ	8	DN	2	25%
194	Y	YFVHRILWKMSG	12	FVH	3	25%
195	K	KVYLQNRHDSGTMA	14	VYLQR	5	36%
196	K	KSQADITGEHN	11	SQA	3	27%
197	P	EPSDQKALFNTH	12	ES	2	17%
198	G	NSGTFKQRHPID	12	NS	2	17%
199	Y		0		0	0%
200	N	TNKDRS	6	T	1	17%
201	G		0		0	0%
202	S	SRQHETKNAFDI	12	RQHE	4	33%
203	C		0		0	0%
204	F	PFHSLRTA	8	P	1	13%
205	N	DEPANLSKIVT	11	DEP	3	27%
206	V	VEGIDALSNKTFY	13	EGIDA	5	38%
207	E	EITAV	5	I	1	20%
208	I	VIKRLSHEMQPTCY	14	VKR	3	21%
209	K	ALSKQTRFDVEINHP	15	ALS	3	20%
210	R	RQLHKME	7	Q	1	14%
211	N		0		0	0%
212	D	DNTQ	4	N	1	25%
213	D	QEDAKHNLRG	10	QE	2	20%
214	L		0		0	0%
215	S	SAGLHQFKNDRMVIYTE	17	AGLH	4	24%
216	W		0		0	0%
217	L	LMI	3		0	0%
218	W		0		0	0%
219	N	NAKEGDTSRQ	10	AKEG	4	40%
220	E	EQSAKVDNIGHT	12	QSAK	4	33%
221	S		0		0	0%
222	T	TARSYQMNCL	10	ARS	3	30%
223	A		0		0	0%
224	L		0		0	0%
225	Y	YF	2	F	1	50%
226	P		0		0	0%
227	S	SDNA	4		0	0%
228	I	IVAT	4	V	1	25%
229	Y	YGSTVH	6	G	1	17%
230	L	LVMI	4	VM	2	50%
231	N	PDERNKSWGQT	11	PDER	4	36%
232	T	KAETPSLVQYRIC	13	KAEP	4	31%
233	Q	SARVQEGKM	9	SAREGK	6	67%
234	Q	LQMK	4	L	1	25%
235	S	AGSRN	5	AG	2	40%
236	P	SGDPNVL	7	SGD	3	43%
237	V	SNAVGHD	7	SNGH	4	57%
238	A	NKHYQASEVD	10	NKH	3	30%
239	A	AGSITVHRLE	10	GSITVHL	7	70%
240	T	LRQATVWPM	9	LRQA	4	44%

Position	WT	MSA Alternates	# of MSA Alternatives	Alternatives Assessed	# of MSA Alternatives Assessed	% of MSA Alternatives Assessed
241	L	LKMRNAHQPTE	11	KMRNA	5	45%
242	Y	FYM	3	F	1	33%
243	V	VSATC	5	S	1	20%
244	R	RQHSKWA	7	QHS	3	43%
245	N	FHNYSAE	7	FHY	3	43%
246	R		0		0	0%
247	V	VLI	3	L	1	33%
248	R	QRHKAELNMVTGWC	14	QH	2	14%
249	E		0		0	0%
250	A	ASG	3	S	1	33%
251	I	LMFIRV	6	LMF	3	50%
252	R		0		0	0%
253	V	VILMAT	6	I	1	17%
254	S	ASRDQ	5	A	1	20%
255	K	KTSRLEVQAHDIYF	14	TSRLEA	6	43%
256	I	MVATIQLGRKSEDF	14	MVATL	5	36%
257	P	AHTPGSVKMR	10	AHT	3	30%
258	D	HSRDNKG	7	HSR	3	43%
259	A	AHGKRPDETVNSCQ	14	HGKR	4	29%
260	K	DNKGSPEHQRALVI	14	DN	2	14%
261	S	YHPSNFDEGACIQ	13	YHP	3	23%
262	P	APSNTDVHKGRM	12	AS	2	17%
263	L	LVRAPTI	7		0	0%
264	P		0		0	0%
265	V	VIT	3	I	1	33%
266	F	FLYMV	5	LY	2	40%
267	A	VPAI	4	VP	2	50%
268	Y	YF	2	F	1	50%
269	T	TASVILMNF	9	ASV	3	33%
270	R	RQN	3	Q	1	33%
271	I	PLIV	4	PL	2	50%
272	V	TFVGAYDLS	9	TFG	3	33%
273	F	YFHL	4	YH	2	50%
274	T	TRSLIKGMN	9	RS	2	22%
275	D	DRNSETHLG	9	RNEH	4	44%
276	Q	TSQEKRAVNLMGD	13	TSEKRM	6	46%
277	V	PSTVGKIYNA	10	PST	3	30%
278	L	LNFSGDMITYEPK	13	NFS	3	23%
279	K	TREHFKLDYQNV	12	TREHF	5	42%
280	F	FYGLAVEQTP	10	YG	2	20%
281	L		0		0	0%
282	S	SPTENLQ	7	P	1	14%
283	Q	EQLKRTPAM	9	ELK	3	33%
284	D	DEMQIVSTKAPGH	13	EMQ	3	23%
285	E	DEH	3	D	1	33%
286	L		0		0	0%
287	V	VIEMA	5	IE	2	40%
288	Y	SHNQY	5	SHN	3	60%

Position	WT	MSA Alternates	# of MSA Alternatives	Alternatives Assessed	# of MSA Alternatives Assessed	% of MSA Alternatives Assessed
289	T	TSV	3	S	1	33%
290	F	ILFVS	5	IL	2	40%
291	G		0		0	0%
292	E	EVQT	4		0	0%
293	T	STIAC	5	S	1	20%
294	V	AVMI	4	A	1	25%
295	A	ASVP	4		0	0%
296	L	LQ	2	Q	1	50%
297	G		0		0	0%
298	A	AVSTP	5		0	0%
299	S	ASD	3	A	1	33%
300	G		0		0	0%
301	I	VIFATM	6	V	1	17%
302	V	VI	2	I	1	50%
303	I	LIMFV	5	LMF	3	60%
304	W		0		0	0%
305	G	GVDLI	5	V	1	20%
306	T	DSGTEAN	7	DS	2	29%
307	L	LMASWHVYIGTF	12	MASWHVYIGTF	11	92%
308	S	NESGQDATYIVL	12	NEG	3	25%
309	I	LYFNIDKMSCV	11	LYFNDKMSCV	10	91%
310	M	TSAMVY	6	TSA	3	50%
311	R	SRKTQLNHADE	11	SKT	3	27%
312	S	STNRD	5	TNRD	4	80%
313	M	KEMTARQYVNPL	12	KETARQYVNPL	11	92%
314	K	EKGAQRHTVPDN	12	E	1	8%
315	S	STNEAMGKQ	9	TNE	3	33%
316	C		0		0	0%
317	L	QTLWERMSPIA	11	QTWERMSPIA	10	91%
318	L	AKTNSYRLHIDMGEV	15	AKTN	4	27%
319	L	LVI	3	VI	2	67%
320	D	KDHRNQSGE	9	KHRNQSGE	8	89%
321	N	DQEKNSRTAYHG	12	DQEK	4	33%
322	Y	YFTS	4	F	1	25%
323	M	LVMI	4	LVI	3	75%
324	E	DTESVNKRLQHGA	13	DTSVNKRLQHGA	12	92%
325	T	STGRNEDKQVMH	12	SGRE	4	33%
326	I	TDLIPRSVKHEN	12	TDLP	4	33%
327	L		0		0	0%
328	N	GNVLTA	6	GV	2	33%
329	P	PHSRELYQNV	10	HS	2	20%
330	Y	YFS	3	F	1	33%
331	I	IVLA	4	VL	2	50%
332	I	VILATMKF	8	VL	2	25%
333	N		0		0	0%
334	V		0		0	0%
335	T	TSR	3	S	1	33%
336	L	SLRTWKAEDMG	11	SRTW	4	36%
337	A	AGS	3	G	1	33%
338	A	ATV	3	T	1	33%
339	K	KELQTMRDFVNHA	13	ELQM	4	31%
340	M	LMVYAIHREQN	11	LVY	3	27%
341	C		0		0	0%
342	S		0		0	0%
343	Q	QLRWHEKIT	9	LRWHEK	6	67%

Position	WT	MSA Alternates	# of MSA Alternatives	Alternatives Assessed	# of MSA Alternatives Assessed	% of MSA Alternatives Assessed
344	V	AVHTSQNRKEIDF	13	AHT	3	23%
345	L	LQRVHMT	7	Q	1	14%
346	C		0		0	0%
347	Q	HQSNRKGTEF	10	HSNR	4	40%
348	E	GNESKDRAF	9	GNS	3	33%
349	Q	HNQREKS	7	HN	2	29%
350	G		0		0	0%
351	V	RVILAMK	7	R	1	14%
352	C		0		0	0%
353	I	VIATLSRQKY	10	VA	2	20%
354	R	RWK	3		0	0%
355	K	KRQHN	5	R	1	20%
356	N	DNPTMHQVAKRSE	13	DP	2	15%
357	W	WPSNLEVID	9	PS	2	22%
358	N	NHSKGDEYTRQI	12	HSKGDE	6	50%
359	S	SAPQTEM	7	APQT	4	57%
360	S	SNDEPKHFALQGRT	14	NDEPK	5	36%
361	D	DATVFSHING	10	ATVS	4	40%
362	Y	YFLHR	5	FL	2	40%
363	L		0		0	0%
364	H	HLITVY	6	I	1	17%
365	L	LMI	3	M	1	33%
366	N	NSQHGDW	7	S	1	14%
367	P	PASTLN	6	AST	3	50%
368	D	ADSNRTGEKQMHL	13	AS	2	15%
369	N	SNTHFQMR	8	S	1	13%
370	F	FYHWLI	6	Y	1	17%
371	A	RSHAQKDYNVGT	12	RSH	3	25%
372	I	ILSV	4	L	1	25%
373	Q	QEVRHKILDAMGS	13	EVRK	4	31%
374	L	LPARFVTISHYKM	13	PA	2	15%
375	E	SGERATQNKIV	11	SGT	3	27%
376	K	GKEDQNSRPA	10	GEP	3	30%
377	G	GDENSATRKHY	11	DENS	4	36%
378	G	GPRFKNELQV	10	PR	2	20%
379	K	KPEQRLGSTDCWNH	14	PEQR	4	29%
380	F	FLYIPV	6	LY	2	33%
381	T	TRVSEIKWYHMAFLQ	15	RVSI	4	27%
382	V	VLPAIT	6	LPAI	4	67%
383	R	KRQESTHVPLNADI	14	KQE	3	21%
384	G	GSAPRE	6		0	0%
385	K	KAERNHTQSPYLW	13	AERNH	5	38%
386	P	LPAVMSNHFE	10	LA	2	20%
387	T	STKPNRGQW	9	S	1	11%
388	L	LDWAPHSKRQYVGEN	15	DWA	3	20%
389	E	EATKLDHRGQNWMS	14	ATK	3	21%
390	D	DGETASH	7		0	0%
391	L	LIRQKVGST	9	IRQ	3	33%
392	E	AEQKLDSHNMVTGI	14	AQ	2	14%
393	Q	QVYHRDANFELTIKG	15	VYHRD	5	33%
394	F	MLFW	4	ML	2	50%
395	S	ASQKERMVITG	11	AQK	3	27%
396	E	EKTVDMSNQIAH	12	KTVD	4	33%
397	K	KHENGSTRDQ	10	HEN	3	30%
398	F	F	1		0	0%

Position	WT	MSA Alternates	# of MSA Alternatives	Alternatives Assessed	# of MSA Alternatives Assessed	% of MSA Alternatives Assessed
399	Y	RSQKYTVMELHFD	13	RSQK	4	31%
400	C	C	1		0	0%
401	S	HQRSLV	6	HQR	3	50%
402	C	C	1		0	0%
403	Y	YFN	3		0	0%
404	S					
405	T					
406	L					
407	S					
408	C					
409	K					
410	E					
411	K					
412	A					
413	D					
414	V					
415	K					
416	D					
417	T					
418	D					
419	A					
420	V					
421	D					
422	V					
423	C					
424	I					
425	A					
426	D					
427	G					
428	V					
429	C					
430	I					
431	D					
432	A					
433	F					
434	L					
435	K					
436	P					
437	P					
438	M					
439	E					
440	T					
441	E					
442	E					
443	P					
444	Q					
445	I					
446	F					
447	Y					

Note:
Column A ("Positions"):
Dark Grey - residues not found in PH20 homology model

Red - essential residues identified by Dr. Park

Dark Green - positions where Dr. Park evaluated ≥ 66% of alternate residues identified in the MSA

Green - positions where Dr. Park evaluated betrween 33% and 66% of alternate residues identified in the MSA

Light Green - positions where Dr. Park evaluated < 33% of alternate residues identified in the MSA

Unshaded - positions where Dr. Park did not evaluate alternate residues

Columns C and E:
Color code corresponds to the frequency profile of amino acids at positions identified in EX1004 (Park), Appendix D Part I, the code below corresponds to rank order of frequency:

(1) Green (most frequent)
(2) Blue
(3) Light Blue
(4) Yellow
(5) Orange
(6+) Red (least frequent, remaining residues)

Underlined residues correspond to WT.

Note:
Information in this sheet was generate from EX1004 (Park) and EX2176, cols. D-G, at 1-50.

Appendix B-9

PH20 Positional Alignment: Tested MSA Alternatives

Source(s): EX1004 (Park Dec.), EX2176 (Park Analysis Spreadsheet)

Position	WT	Position	WT	Position	WT	Position	WT	Position	WT	Position	WT	Position	WT
1	L	75	V	149	A	223	A	297	G	371	A	445	I
2	N	76	N	150	T	224	L	298	A	372	I	446	F
3	F	77	G	151	E	225	Y	299	S	373	Q	447	Y
4	R	78	G	152	K	226	P	300	G	374	L	447	Y
5	A	79	I	153	A	227	S	301	I	375	E		
6	P	80	P	154	K	228	I	302	V	376	K		
7	P	81	Q	155	Q	229	Y	303	I	377	G		
8	V	82	K	156	E	230	L	304	W	378	G		
9	I	83	I	157	F	231	N	305	G	379	K		
10	P	84	S	158	E	232	T	306	T	380	F		
11	N	85	L	159	K	233	Q	307	L	381	T		
12	V	86	Q	160	A	234	Q	308	S	382	V		
13	P	87	D	161	G	235	S	309	I	383	R		
14	F	88	H	162	K	236	P	310	M	384	G		
15	L	89	L	163	D	237	V	311	R	385	K		
16	W	90	D	164	F	238	A	312	S	386	P		
17	A	91	K	165	L	239	A	313	M	387	T		
18	W	92	A	166	V	240	T	314	K	388	L		
19	N	93	K	167	E	241	L	315	S	389	E		
20	A	94	K	168	T	242	Y	316	C	390	D		
21	P	95	D	169	I	243	V	317	L	391	L		
22	S	96	I	170	K	244	R	318	L	392	E		
23	E	97	T	171	L	245	N	319	L	393	Q		
24	F	98	F	172	G	246	R	320	D	394	F		
25	C	99	Y	173	K	247	V	321	N	395	S		
26	L	100	M	174	L	248	R	322	Y	396	E		
27	G	101	P	175	L	249	E	323	M	397	K		
28	K	102	V	176	R	250	A	324	E	398	F		
29	F	103	D	177	P	251	I	325	T	399	Y		
30	D	104	N	178	N	252	R	326	I	400	C		
31	E	105	L	179	H	253	V	327	L	401	S		
32	P	106	G	180	L	254	S	328	N	402	C		
33	L	107	M	181	W	255	K	329	P	403	Y		
34	D	108	A	182	G	256	I	330	Y	404	S		
35	M	109	V	183	Y	257	P	331	I	405	T		
36	S	110	I	184	Y	258	D	332	I	406	L		
37	L	111	D	185	L	259	A	333	N	407	S		
38	F	112	W	186	F	260	K	334	V	408	C		
39	S	113	E	187	P	261	S	335	T	409	K		
40	F	114	E	188	D	262	P	336	L	410	E		
41	I	115	W	189	C	263	L	337	A	411	K		
42	G	116	R	190	Y	264	P	338	A	412	A		
43	S	117	P	191	N	265	V	339	K	413	D		
44	P	118	T	192	H	266	F	340	M	414	V		
45	R	119	W	193	H	267	A	341	C	415	K		
46	I	120	A	194	Y	268	Y	342	S	416	D		
47	N	121	R	195	K	269	T	343	Q	417	T		
48	A	122	N	196	K	270	R	344	V	418	D		
49	T	123	W	197	P	271	I	345	L	419	A		
50	G	124	K	198	G	272	V	346	C	420	V		
51	Q	125	P	199	Y	273	F	347	Q	421	D		
52	G	126	K	200	N	274	T	348	E	422	V		
53	V	127	D	201	G	275	D	349	Q	423	C		
54	T	128	V	202	S	276	Q	350	G	424	I		
55	I	129	Y	203	C	277	V	351	V	425	A		
56	F	130	K	204	F	278	L	352	C	426	D		
57	Y	131	N	205	N	279	K	353	I	427	G		
58	V	132	R	206	V	280	F	354	R	428	V		
59	D	133	S	207	E	281	L	355	K	429	C		
60	R	134	I	208	I	282	S	356	N	430	I		
61	L	135	E	209	K	283	Q	357	W	431	D		
62	G	136	L	210	R	284	D	358	N	432	A		
63	Y	137	V	211	N	285	E	359	S	433	F		
64	Y	138	Q	212	D	286	L	360	S	434	L		
65	P	139	Q	213	D	287	V	361	D	435	K		
66	Y	140	Q	214	L	288	Y	362	Y	436	P		
67	I	141	N	215	S	289	T	363	L	437	P		
68	D	142	V	216	W	290	F	364	H	438	M		
69	S	143	Q	217	L	291	G	365	L	439	E		
70	I	144	L	218	W	292	E	366	N	440	T		
71	T	145	S	219	N	293	T	367	P	441	E		
72	G	146	L	220	E	294	V	368	D	442	E		
73	V	147	T	221	S	295	A	369	N	443	P		
74	T	148	E	222	T	296	L	370	F	444	Q		

Note:
Grey - residues not found in PH20 homology model

Red - essential residues identified by Dr. Park

Dark Green - positions where Dr. Park evaluated ≥ 66% of alternate residues identified in the MSA

Green - positions where Dr. Park evaluated between 33% and 66% of alternate residues identified in the MSA

Light Green - positions where Dr. Park evaluated < 33% of alternate residues identified in the MSA

Unshaded - positions where Dr. Park did not evaluate alternate substitutions

Information in this sheet was generate from EX1004 (Park), Appendix D Part 1, Appendix D Part 3, and EX2176, cols. D-G, at 1-50.

Appendix B-10

PH20 Positional Alignment: Domains, Motifs, and Prior Art

Source(s): EX1026 ('731 App.), EX1006 (Chao), EX1011 (Arming)

Precursor	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65
Mature	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
Residue	L	N	F	R	A	P	P	V	I	P	N	V	P	F	L	W	A	W	N	A	P	S	E	F	C	L	G	K	F	D

Precursor	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95
Mature	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**	**41**	**42**	**43**	**44**	**45**	**46**	**47**	**48**	**49**	**50**	**51**	**52**	**53**	**54**	**55**	**56**	**57**	**58**	**59**	**60**
Residue	E	P	L	D	M	S	L	F	S	F	I	G	S	P	R	I	N	A	T	G	Q	G	V	T	I	F	Y	V	D	R

Precursor	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125
Mature	**61**	**62**	**63**	**64**	**65**	**66**	**67**	**68**	**69**	**70**	**71**	**72**	**73**	**74**	**75**	**76**	**77**	**78**	**79**	**80**	**81**	**82**	**83**	**84**	**85**	**86**	**87**	**88**	**89**	**90**
Residue	L	G	Y	Y	P	Y	I	D	S	I	T	G	V	T	V	N	G	G	I	P	Q	K	I	S	L	Q	D	H	L	D

Precursor	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155
Mature	**91**	**92**	**93**	**94**	**95**	**96**	**97**	**98**	**99**	**100**	**101**	**102**	**103**	**104**	**105**	**106**	**107**	**108**	**109**	**110**	**111**	**112**	**113**	**114**	**115**	**116**	**117**	**118**	**119**	**120**
Residue	K	A	K	K	D	I	T	F	Y	M	P	V	D	N	L	G	M	A	V	I	D	W	E	E	W	R	P	T	W	A

Precursor	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185
Mature	**121**	**122**	**123**	**124**	**125**	**126**	**127**	**128**	**129**	**130**	**131**	**132**	**133**	**134**	**135**	**136**	**137**	**138**	**139**	**140**	**141**	**142**	**143**	**144**	**145**	**146**	**147**	**148**	**149**	**150**
Residue	R	N	W	K	P	K	D	V	Y	K	N	R	S	I	E	L	V	Q	Q	Q	N	V	Q	L	S	L	T	E	A	T

Precursor	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215
Mature	**151**	**152**	**153**	**154**	**155**	**156**	**157**	**158**	**159**	**160**	**161**	**162**	**163**	**164**	**165**	**166**	**167**	**168**	**169**	**170**	**171**	**172**	**173**	**174**	**175**	**176**	**177**	**178**	**179**	**180**
Residue	E	K	A	K	Q	E	F	E	K	A	G	K	D	F	L	V	E	T	I	K	L	G	K	L	L	R	P	N	H	L

Precursor	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245
Mature	**181**	**182**	**183**	**184**	**185**	**186**	**187**	**188**	**189**	**190**	**191**	**192**	**193**	**194**	**195**	**196**	**197**	**198**	**199**	**200**	**201**	**202**	**203**	**204**	**205**	**206**	**207**	**208**	**209**	**210**
Residue	W	G	Y	Y	L	F	P	D	C	Y	N	H	H	Y	K	K	P	G	Y	N	G	S	C	F	N	V	E	I	K	R

Precursor	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275
Mature	**211**	**212**	**213**	**214**	**215**	**216**	**217**	**218**	**219**	**220**	**221**	**222**	**223**	**224**	**225**	**226**	**227**	**228**	**229**	**230**	**231**	**232**	**233**	**234**	**235**	**236**	**237**	**238**	**239**	**240**
Residue	N	D	D	L	S	W	L	W	N	E	S	T	A	L	Y	P	S	I	Y	L	N	T	Q	Q	S	P	V	A	A	T

Precursor	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305
Mature	**241**	**242**	**243**	**244**	**245**	**246**	**247**	**248**	**249**	**250**	**251**	**252**	**253**	**254**	**255**	**256**	**257**	**258**	**259**	**260**	**261**	**262**	**263**	**264**	**265**	**266**	**267**	**268**	**269**	**270**
Residue	L	Y	V	R	N	R	V	R	E	A	I	R	V	S	K	I	P	D	A	K	S	P	L	P	V	F	A	Y	T	R

Precursor	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335
Mature	**271**	**272**	**273**	**274**	**275**	**276**	**277**	**278**	**279**	**280**	**281**	**282**	**283**	**284**	**285**	**286**	**287**	**288**	**289**	**290**	**291**	**292**	**293**	**294**	**295**	**296**	**297**	**298**	**299**	**300**
Residue	I	V	F	T	D	Q	V	L	K	F	L	S	Q	D	E	L	V	Y	T	F	G	E	T	V	A	L	G	A	S	G

Precursor	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365
Mature	**301**	**302**	**303**	**304**	**305**	**306**	**307**	**308**	**309**	**310**	**311**	**312**	**313**	**314**	**315**	**316**	**317**	**318**	**319**	**320**	**321**	**322**	**323**	**324**	**325**	**326**	**327**	**328**	**329**	**330**
Residue	I	V	I	W	G	T	L	S	I	M	R	S	M	K	S	C	L	L	L	D	N	Y	M	E	T	I	L	N	P	Y

Precursor	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395
Mature	**331**	**332**	**333**	**334**	**335**	**336**	**337**	**338**	**339**	**340**	**341**	**342**	**343**	**344**	**345**	**346**	**347**	**348**	**349**	**350**	**351**	**352**	**353**	**354**	**355**	**356**	**357**	**358**	**359**	**360**
Residue	I	I	N	V	T	L	A	A	K	M	C	S	Q	V	L	C	Q	E	Q	G	V	C	I	R	K	N	W	N	S	S

Precursor	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425
Mature	**361**	**362**	**363**	**364**	**365**	**366**	**367**	**368**	**369**	**370**	**371**	**372**	**373**	**374**	**375**	**376**	**377**	**378**	**379**	**380**	**381**	**382**	**383**	**384**	**385**	**386**	**387**	**388**	**389**	**390**
Residue	D	Y	L	H	L	N	P	D	N	F	A	I	Q	L	E	K	G	G	K	F	T	V	R	G	K	P	T	L	E	D

Precursor	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455
Mature	**391**	**392**	**393**	**394**	**395**	**396**	**397**	**398**	**399**	**400**	**401**	**402**	**403**	**404**	**405**	**406**	**407**	**408**	**409**	**410**	**411**	**412**	**413**	**414**	**415**	**416**	**417**	**418**	**419**	**420**
Residue	L	E	Q	F	S	E	K	F	Y	C	S	C	Y	S	T	L	S	C	K	E	K	A	D	V	K	D	T	D	A	V

Precursor	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482
Mature	**421**	**422**	**423**	**424**	**425**	**426**	**427**	**428**	**429**	**430**	**431**	**432**	**433**	**434**	**435**	**436**	**437**	**438**	**439**	**440**	**441**	**442**	**443**	**444**	**445**	**446**	**447**
Residue	D	V	C	I	A	D	G	V	C	I	D	A	F	L	K	P	P	M	E	T	E	E	P	Q	I	F	Y

Note

Dark Yellow - Hyal-EGF Domain (L336-D416)

Red - site of truncation reported in EX1011 (Arming), 813.

Light Blue (position numbers) - common hyaluronidase domain (3-339) (EX1026 ('731 App.), 79:20-31)

Purple - peptide 2 binding domain (170-200) (EX1026 ('731 App.), 81:7-14)

Appendix B-11

PH20 Residue Characterization

Source(s): EX1026 ('731 App.), EX2166 (Petsko Patent Analysis Data), EX1004 (Park Dec.)

Position	WT	Dr. Park Essential Residues	Prior Art / Confirmed Essential Residues
1	L		
2	N		
3	F		
4	R		
5	A		
6	P		
7	P		
8	V		
9	I		
10	P		
11	N		
12	V		
13	P		
14	F	X	
15	L		
16	W		
17	A		
18	W	X	
19	N		
20	A		
21	P		
22	S		
23	E		
24	F		
25	C	X	X
26	L		
27	G		
28	K		
29	F		
30	D		
31	E		
32	P		
33	L		
34	D		
35	M		
36	S		
37	L		
38	F		
39	S		
40	F		
41	I		
42	G		
43	S		
44	P		
45	R		
46	I		
47	N		
48	A		
49	T		
50	G		
51	Q		
52	G		
53	V		
54	T		
55	I		
56	F	X	
57	Y	X	
58	V		
59	D		
60	R		
61	L		
62	G	X	
63	Y		
64	Y	X	
65	P	X	
66	Y		
67	I		
68	D		
69	S		
70	I		
71	T		
72	G		
73	V		
74	T		
75	V		
76	N		
77	G	X	
78	G	X	
79	I		
80	P	X	
81	Q	X	
82	K		
83	I		
84	S		
85	L		
86	Q		
87	D		
88	H	X	
89	L		
90	D		
91	K		
92	A		
93	K		
94	K		
95	D		
96	I		
97	T		
98	F		
99	Y		
100	M		
101	P		
102	V		
103	D		
104	N		
105	L		
106	G	X	
107	M		
108	A		
109	V	X	
110	I		
111	D	X	X
112	W	X	
113	E	X	X
114	E		
115	W	X	
116	R		
117	P	X	
118	T		
119	W	X	
120	A		
121	R		
122	N	X	
123	W	X	
124	K		
125	P		
126	K		
127	D		
128	V		
129	Y	X	
130	K		
131	N		
132	R		
133	S	X	
134	I		
135	E		
136	L		
137	V		
138	Q		
139	Q		
140	Q		
141	N		
142	V		
143	Q		
144	L		
145	S		
146	L		
147	T		
148	E		
149	A		
150	T		
151	E		
152	K		
153	A	X	
154	K		
155	Q		
156	E		
157	F	X	
158	E		
159	K		
160	A		
161	G		
162	K		
163	D		
164	F		
165	L		
166	V		
167	E		
168	T	X	
169	I		
170	K		
171	L		
172	G		
173	K		
174	L		
175	L		
176	R	X	X
177	P	X	
178	N		
179	H		
180	L	X	
181	W	X	
182	G	X	
183	Y		
184	Y	X	
185	L		
186	F		
187	P	X	
188	D		
189	C	X	X
190	Y		
191	N	X	
192	H		
193	H		
194	Y		
195	K		
196	K		
197	P		
198	G		
199	Y	X	
200	N		X
201	G	X	
202	S		
203	C	X	X
204	F		
205	N		
206	V		
207	E		
208	I		
209	K		
210	R		
211	N	X	
212	D		
213	D		
214	L	X	
215	S		
216	W	X	
217	L		
218	W	X	
219	N		
220	E		
221	S	X	
222	T		
223	A	X	
224	L	X	
225	Y		
226	P	X	
227	S		
228	I		
229	Y		
230	L		
231	N		
232	T		
233	Q		
234	Q		
235	S		
236	P		
237	V		
238	A		
239	A		
240	T		
241	L		
242	Y		
243	V		
244	R		
245	N		
246	R	X	X
247	V		
248	R		
249	E	X	X
250	A		
251	I		
252	R	X	X
253	V		
254	S		
255	K		
256	I		
257	P		
258	D		
259	A		
260	K		
261	S		
262	P		
263	L		
264	P	X	
265	V		
266	F		
267	A		
268	Y		
269	T		
270	R		
271	I		
272	V		
273	F		
274	T		
275	D		
276	Q		
277	V		
278	L		
279	K		
280	F		
281	L	X	
282	S		
283	Q		
284	D		
285	E		
286	L	X	
287	V		
288	Y		
289	T		
290	F		
291	G	X	
292	E		
293	T		
294	V		
295	A		
296	L		
297	G	X	
298	A		
299	S		
300	G	X	
301	I		
302	V		
303	I		
304	W	X	
305	G		
306	T		
307	L		
308	S		
309	I		
310	M		
311	R		
312	S		
313	M		
314	K		
315	S		
316	C	X	X
317	L		
318	L		
319	L		
320	D		
321	N		
322	Y		
323	M		
324	E		
325	T		
326	I		
327	L	X	
328	N		
329	P		
330	Y		
331	I		
332	I		
333	N	X	X
334	V	X	
335	T		
336	L		
337	A		
338	A		
339	K		
340	M		
341	C	X	X
342	S	X	
343	Q		
344	V		

Position	WT	Dr. Park Essential Residues	Prior Art / Confirmed Essential Residues
345	L		
346	C	X	X
347	Q		
348	E		
349	Q		
350	G	X	
351	V		
352	C	X	X
353	I		
354	R		
355	K		
356	N		
357	W		
358	N		X
359	S		
360	S		
361	D		
362	Y		
363	L	X	
364	H		
365	L		
366	N		
367	P		
368	D		
369	N		
370	F		
371	A		
372	I		
373	Q		
374	L		
375	E		
376	K		
377	G		
378	G		
379	K		
380	F		
381	T		
382	V		
383	R		
384	G		
385	K		
386	P		
387	T		
388	L		
389	E		
390	D		
391	L		
392	E		
393	Q		
394	F		
395	S		
396	E		
397	K		
398	F		
399	Y		
400	C		X
401	S		
402	C		X
403	Y		
404	S		
405	T		
406	L		
407	S		
408	C		X
409	K		
410	E		
411	K		
412	A		
413	D		
414	V		
415	K		
416	D		
417	T		
418	D		
419	A		
420	V		
421	D		
422	V		
423	C		X
424	I		
425	A		
426	D		
427	G		
428	V		
429	C		X
430	I		
431	D		
432	A		
433	F		
434	L		

Position	WT	Dr. Park Essential Residues	Prior Art / Confirmed Essential Residues
435	K		
436	P		
437	P		
438	M		
439	E		
440	T		
441	E		
442	E		
443	P		
444	Q		
445	I		
446	F		
447	Y		

Note:
Col. A ("Position") reflects shading fof positions found from EX2166 (Petsko Patent Analysis Data):

Green - 100% of substitutions at the position yielded an "active" PH20 protein

Light Green - between 60% and 100% of substitutions at the position yielded an "active" PH20 protein

Grey - between 30% and 60% of substitutions at the position yielded an "active" PH20 protein

Orange - between 0% and 30% of substitutions at the position yielded an "active" PH20 protein

Red - no substitutions at the position yielded an "active" PH20 protein

An "X" in Cols. C-D denotes positions identified as either essential by Dr. Park (col. C), or identified in the prior art / confirmed essential by the common disclosure (col. D).

Other information from EX1026 ('731 App.), 80:19-35, and EX1004 (Park Dec.), Appendix D-3 (Essential Residues).

Appendix B-12

Compilation of Citations on Conventional Tools and Techniques from the Experts

Source(s): EX1003 (Hecht Dec.), EX1004 (Park Dec.), EX2068 (Simpson Dec.), EX2070 (Petsko Dec.), EX1131 (Petsko Dep.), EX1130 (Simpson Dep.), EX2076 (Hecht Dep.), EX2077 (Park Dep.)

POSA Knowledge in 2011	Supporting Expert Testimony
Sequence alignments and structural information found in the prior art, such as Chao, would guide the identification of residues suitable for modification.	Hecht: EX1003, ¶¶ 209-211, 213-216. Park: EX1004, ¶¶ 21-23, 92-93. Petsko: EX2070, ¶¶ 116, 126-127, 167, 290; EX1131, 73:14-74:17, 108:23-110:21, 177:6-13. Simpson: EX2068 ¶¶ 134-137, 299; EX1130, 75:3-14, 76:3-77:22, 92:15-93:5, 99:11-21.
POSAs would perform sequence alignments with homologous sequences to identify residues suitable for modification.	Hecht: EX1003, ¶¶ 49, 213-216. Park: EX1004, ¶¶ 27-32, 93. Petsko: EX2070, ¶¶ 87-92, 290; EX1131, 178:1-17, 179:10-21. Simpson: EX2068, ¶¶ 97, 130; EX1130, 101:21-103:9, 105:5-10.
POSAs would use BLAST and CLUSTAL Omega to generate MSAs.	Hecht: EX1003, ¶ 21. Park: EX1004, ¶¶ 24, 139. Petsko: EX2070, ¶¶ 93-94; EX1131, 168:8-169:16, 178:5-10. Simpson: EX2068, ¶¶ 98-99.
The crystal structure of Hyal-1 described in Chao was the best template to generate a PH20 structure homology model.	Hecht: EX1003, ¶ 225. Park: EX1004, ¶¶ 147-148, 152. Petsko: EX2070, ¶ 121, n.10; EX1131, 123:13-18. Simpson: EX2068, ¶¶ 119, 124, 134, 137; EX1130, 103:23-105:4.

POSA Knowledge in 2011	Supporting Expert Testimony
Homology models are important tools for evaluating substitutions.	Hecht: EX1003, ¶¶ 224, 226-228. Park: EX1004, ¶¶ 21, 156; EX2077, 176:2-5. Petsko: EX2070, ¶¶ 106-107, 118, 352; EX1131, 199:1-6. Simpson: EX2068, ¶¶ 108-109, 130, 138; EX1130, 125:14-126:6; EX1010.
POSAs would be able to generate reliable PH20 homology models using SWISS-MODEL and available sequence alignment information.	Hecht: EX1003, ¶ 233. Park: EX1004, ¶¶ 148, 151, 156. Petsko: EX2070, ¶¶ 104-105, 123-124, n.11. Simpson: EX2068, ¶¶ 109, 127, 245, 367; EX1130, 104:5-105:10, 107:12-108:21.
POSAs would examine residue interactions using, *e.g.*, PyMol, to evaluate substitutions.	Hecht: EX2076, 72:3-73:15. Park: EX1004, ¶¶ 56-63, 140, 163-64. Petsko: EX2070, ¶¶ 110-111; EX1131, 199:7-18. Simpson: EX1130, 34:1-10, 36:8-18.
POSAs would be interested in making substitutions that improve polypeptide characteristics.	Hecht: EX1003, ¶ 220. Park: EX1004, ¶¶ 68, 80-81. Petsko: EX2070, ¶ 288; EX1131, 105:11-23. Simpson: EX2068, ¶¶ 157, 215.

Appendix C

CV

JAMES HENDERSON NAISMITH
BSc, PhD, DSc, FRS, FRSE, FMedSci, MAE, FRSC, FinstP, FRSB, Member of EMBO, Fellow American Association Advancement of Science (AAAS).

Date of birth: 26th July 1968 Hamilton, Scotland, UK

Current role (from Oct 2023): Head of MPLS Division, Oxford University
As Head of Division I am responsible for teaching, research and administration across all science (except medicine) and engineering. I hold the budget of around £420 M pa, I agree with heads of the departments' academic strategy, I chair annual promotion panels, set the internal tone within Division and representing the division within the University. MPLS is ranked both in research and teaching as the leading Division in Oxford and the world.
Specific achievements include:

- Partnership with Ellison Institute around AI
- Partnered with Govt on AI & engineering initiatives
- Progress on broadening curriculum from ahistorical exclusive western lens
- Managing finances ahead of budget target
- Estates Strategy for science and engineering

Education
(1) Undergraduate Degree
University of Edinburgh 1985-1989 First Class Honours Bachelor of Science Degree in Chemistry
Awards: Crum Brown Medal 1986 (1st year Chemistry); Class Medal 1988 (3rd Year Chemistry); MacFarlane-Smith Prize 1989 (Honours Chemistry Medal); Stephenson Prize 1989 (joint) (Honours Project); Gillies Prize 1988 (Faculty of Science); SmithKline French Prize 1988 (joint) (U.K. Chemistry); Faculty of Science Scholarship 1988-1989 (joint); Dawson Trust Bursary 1986-1989; Carnegie Summer Scholarship 1988; I.C.I. Summer Studentship 1988; Class Medal Industrial Management 1988
(2) Postgraduate Degree (Carnegie Scholarship)
University of Manchester 1989-1992 Ph.D. "Structural studies of concanavalin A and a zinc aldolase"
Supervisors: Prof. J.R. Helliwell, Prof. C.D. Garner & Dr. W.N. Hunter (principal supervisor).
Award: Hibbert Prize 1992 (joint) (Thesis Prize).
University of Oxford, MA, Degree by Resolution 2023
(3) Higher Degree
University of St Andrews 2016 DSc "Chemical and structural studies upon proteins"
(4) Degree by resolution
University of Oxford 2023 MA

Previous Posts
(1) Post-doctoral, 2 years, 1993-1994
Howard Hughes Medical Institute, Dallas, (Postdoctoral Fellow)
Appointed SERC/NATO fellow (UK competition), in the laboratory of Professor Stephen R. Sprang.
(2) Principal Investigator St Andrews, 1995 - 2017
Biomedical Sciences Research Complex, University of St. Andrews (Lecturer - Professor), 1995-2017
Appointed Lecturer; Promoted Lecturer Grade B, October 1996; Reader, October 1998; Professor June 2001-17, Bishop Wardlaw Professor of Chemical Biology 2014-17
(3) Director of Research Complex Harwell 2017-2019
Appointed to turn around struggling institute, successfully lead the quinquennial review
(4) Founding Director and CEO Rosalind Franklin Institute 2017-2023

Specific achievements include:

- Operationalisation of a fully functioning independent research institute, during the covid19 pandemic (budgeting, charitable status, shareholder, governance)
- Establishment of an outstanding scientific and operational team, that was able to respond with resilience to the Covid pandemic and ensure delivery of the institutional objectives and provide a high level of support for all students and staff.
- An open diverse results-oriented culture, with career structures and initiatives that are designed to remove institutional barriers to groups that are often under-represented in UK science / research organisations.
- Delivery of a new building (of 5,500 m^2) 18 months faster than other comparable public sector buildings. The building secured a design award, is "Excellent" in terms of sustainability & >10 % more efficient in usable to gross space ratio than the next best public sector lab.
- A unique UK wide PhD program that will create a new generation of cross-disciplinary biomedical scientists (10 per year).
- Delivery of the national flagship program Electrifying Life Science, a £60 M revolution in electron cryo tomography
- New therapeutic agents which are progressing towards the clinic for treatment of covid19
- Establishing a unique cross disciplinary scientific environment

(5) Principal Investigator Oxford, 2017-

From June 2017 Professor Nuffield Department of Medicine with cross appointment to Chemistry, University of Oxford,

Previous Leadership

Deputy of Biomolecular Science Research Complex (BSRC) (then known as BMS) 2001-2009
Director of BSRC 2010-2017
I drove the vision, wrote the grant, secured University and Wellcome Trust funding for the BSRC and oversaw its design, construction and delivery.
Elected member of University Court St Andrews, 2012 - 2016

Major research recognition

1. Carbohydrate Chemistry Prize 2000, The Royal Society of Chemistry (Prize for most significant carbohydrate chemistry awarded to early career researcher open to UK citizen based in any country or UK based scientist from any country)
2. BBSRC Career Development Fellowship 200-2005
3. Leverhulme Trust Molecular Biology Prize 2001, The Leverhulme Trust (UK wide prize for most significant contribution of a person under 40 to molecular biology)
4. Corday-Morgan Medal 2004, The Royal Society of Chemistry (UK citizen from any country or UK based scientist, aged 36 or under, who has made the most significant contribution to chemistry)
5. Colworth Medal 2004, The Biochemical Society (UK based scientist identified by published independent work as the most promising Biochemist in the UK aged 36 or under)
6. Jeremy Knowles Medal 2009, The Royal Society of Chemistry (UK scientist under 55 who has made most meritorious contribution to Chemical Biology)
7. Rita & John Cornforth Award 2012, The Royal Society of Chemistry (Awarded to a team (CCP4), that has made the most significant contribution to life science)
8. Royal Society Wolfson Research Merit Award 2014
9. 1000 Talent Scholar (Chinese Academy of Science) 2015-2017
10. The Tilden Prize 2018, The Royal Society of Chemistry (Awarded to the scientist who has done the most to advance chemistry)
11. Horizon Prize 2022, The Royal Society of Chemistry (The nanobody team).

Elections to learned societies – research recognition

1. Fellowship of the Royal Society of Edinburgh (2005) (Scotland's 230 year old National Academy of Arts, Business & Science)
2. Membership of EMBO (2009) (European wide academy of Life Science)
3. Fellowship of the Academy Medical Sciences (2012) (The Academy is one the four UK National Academies)
4. Fellowship of the Royal Society (2014) (The UK's 350+ year old National Academy of Science)
5. Fellow of American Association for Advancement of Science 2016 (AAAS)
6. Member of Academia Europa 2022 (The European Academy of Arts and Science).

Professional societies

1. Fellow of the Royal Society of Chemistry (FRSC)
2. Fellow of the Royal Society of Biology (FRSB)
3. Fellow of the Institute of Physics (FInstP)
4. Member American Chemical Society

Positions of Trust

1. Non-executive Director of Diamond Light Source PLC 2008 - 2016
2. Member of Court (the University governing and corporate body) and Trustee of the University of St Andrews 2012-2016
3. Selection panel Vice-Chancellor of the University of St Andrews 2014
4. Director of the Rosalind Franklin Institute, (Charity Trustee), 2018-
5. UK government nominee on the Council of the European X-ray Free Electron Laser (equivalent to non-Exec Director) 2019-2022
6. Non Executive Director GyreOx, Spin out company from my research, 2020-
7. Council of the Royal Society, Nov 2020 – Nov 2022 (Charity Trustee)
8. Committee for nomination of the Vice-Chancellor Oxford 2021-22
9. Elected Vice-Chair of Council (*super partes*) of the European X-ray Free Electron Laser, 2022-2024
10. Council of Academy of Medical Sciences (Charity Trustee) Dec 2022 -2025
11. Vice-President (non-clinical) Academy of Medical Sciences Dec 2022 -2025

Other significant Responsibilities

1. Chair of SC6 at the Royal Society (2016-2018) (two cycles)
2. Served on Fellowship election committees at the Royal Society, Royal Society Edinburgh and Academy of Medical Science
3. Grant panels Wellcome, RCUK and Royal Society
4. Review panels for major Institutes
5. Royal Society group on Covid evidence assessment
6. Chair of financial sustainability working group on European X-ray Free Electron Laser

Consultancies

Hoescht Marrion Roussel 2000-2 complete
Limatech 2017 complete
Sidley Austin LLP 2017 complete, Expert Witness in Immunex vs Sandoz

Publication & metrics

Full list https://scholar.google.com/citations?user=fLmKKQMAAAAJ&hl=en

Outreach

I have lectured at both general science events (British Association for Advancement of Science) and in schools. I have carried out practical demonstrations for primary school children involving them in hands on experiments (Iodine clock, flame colours, splitting of water). During covid19 I have provided over 100 written comments on the progress of the pandemic through the Science Media Centre. https://www.sciencemediacentre.org/?s=naismith&cat=. Many of these comments were quoted by international (USA, Canada, Brazil, Australia, Germany, France, Japan, South Korea) & UK newspapers (Times, Telegraphy, Mail, Guardian, Financial Times, Sun, Express); often I have extensive background information for the articles (including for the Marr show). I have been interviewed on national TV (Sky News, BBC News, ITV news) and interviewed on international (USA, Australia) & UK radio (BBC R4, BBC Scotland, BBC Oxford, BBC West Midlands). I wrote a commentary for the Guardian and was given a small profile on Radio 5.

Service to community

For over 12 years, 8 years as Chairman I was heavily involved in running CCP4 (finishing in 2010). CCP4 is the *de facto* standard software for protein crystallography. I twice wrote and successfully renewed funding for the project. I managed both the strategic and operational aspects of the entire suite. I organised and taught on specific workshops (refinement) in addition to the regular every two years graduate course at St Andrews. I have led the UK structural biology into the European X-FEL. I was the elected head of the International User Consortium that oversaw the construction of the SFX/SPB beamline from outset to completion.

Undergraduate teaching

At St Andrews I have taught undergraduate chemistry and biochemistry every year since 1995. My teaching load in terms of lectures and tutorials has been on average been in the top 1/3 of academics. In 2017 I gave lectures on the origin of life, enzyme kinetics, chemistry in biology, new medicines and structural biology. I organise the final year module on structural biology. I have supervised labs for second year students. I was voted Chemistry teacher of the year.

Career progression of PhD students supervised to date

1. **Davina Moothoo** 1995-1998 BBSRC. "Protein carbohydrate interactions". Now medical doctor.
2. **Stephan McMahon** 1995-1998 BBSRC. "Protein carbohydrate recognition". Postdoc at UCSD, currently PDRA in Naismith lab.
3. **Louise Buchanan** 1997-2000 BBSRC. "Structural investigations of E. Coli KDPG aldolase a carbohydrate synthesising protein". Took up post developing and managing Unix computing for Ford UK.
4. **Jennifer Harrison** 1995-1998 BBSRC. "Investigation of the substrate specificity of recombinant Trypanosoma cruzi trans-sialidase". Joint supervision with Rob Field (main supervisor). Post-doc in Dundee.
5. **Simon Allard** 1998-2001 BBSRC. "Carbohydrate processing pathogenic bacteria". Post-doc in Madison Wisconsin then moved into scientific publishing.
6. **Michael Tatham** 1998-2001 BBSRC. "Conjugation and deconjugation of the small ubiquitin-like modifiers SUMO-1, SUMO-2 and SUMO-3". Joint with Ron Hay (main supervisor). Now senior researcher in Hay lab.
7. **Changjiang Dong** 1999-2002 ORS "RmlC structure, mechanism and inhibitor design". PDRA with JHN, now a WT career development fellow and then Professor at UEA
8. **Iain Kerr** 1999-2003 "Molecular recognition in proteins". PDRA at St Judes, facility manager at UCSF, now Biotech Salt lake City.
9. **Stephen Fullerton** 1999-2003 "Structural and biochemical studies of carbohydrate binding proteins". PDRA in Leicester, then staff scientist at EMBL. Steve is now working in the biotech sector in Germany.
10. **Kostas Beis** 2001-2004 BBSRC "Structural studies of carbohydrate biosynthesis and translocation across membranes". PDRA in Scripps, now a RCUK lecturer in MPL (Diamond / Imperial)
11. **Alexandra Merkel** 2001-2004 "Enzyme-catalysed carbohydrate biosynthesis and modification ". Now director of a German technology company.
12. **Andrew McEwan** 2002-2005 BBSRC "Structures and function of 5'-fluorodeoxyadenosine synthase". Andy took a post-doc in Aberdeen where he is now a senior researcher.
13. **Walter De Laurentis** 2003-2005 BBSRC "PyrH and PrnB crystal structures ". Returned to Italy, then into grant management in Spain, now post-doc in USA.
14. **Gareth Williams** 2001-2004 Wellcome Trust "Structural studies of ArnA, a bifunctional enzyme from E. Coli and an archaeal heterotrimeric sliding clamp from Sl sulfataricus". Post-doc in LBL, Berkley USA. Gareth is a senior scientist in at LBL synchrotron in the Tainer lab.
15. **Thomas Kowatz** 2005-2009 SFC "Mechanisms of silicate polymerisation, carbohydrate epimerisation and metalloprotease inhibition". Thomas is working as a PDRA in the USA.
16. **Stefan Schmelz** 2003-2006. SFC "Adenylate forming enzymes involved in NRPR-independent siderophore biosynthesis". Stefan has a post-doc in Hemlholz Insitute for infection.
17. **Hubing Lou** SFC 2006-2010. "Structural and functional studies of bacterial outer membrane proteins". Postdoc in Beijing University and is now a PDRA in the USA
18. **Xuan Yan** (self funded and University) 2007-2011. "Structural analysis of the potential therapeutic targets from specific genes in methicillin resistant Staphylococcus aureus" Xuan has a prosperous career in Biotech China.
19. **Wassila Abdelli Boulkeroura** (Algerian scholarship) 2009-2013. "The application of the fragment-based screening approach to RmlA protein and PA1645 structure" Wassila is now a senior pharmacist in the NHS (Milton Keynes Trust)
20. **Emma Brannigan** (BBSRC) "Spin labeling of MscS" PDRA in Dundee in Ron Hay's lab

21. **Judith A Reekes** (BBSRC), 2009-2013 "Structural Studies of CRISPR-associated proteins" PDRA in Newcastle with Martin Noble.
22. **Hexian Huang** (Self funded & part time) 2008 – 2013 "Regulation of export and chain length of extracellular bacterial polysaccharides" PDRA in Naismith lab.
23. **Gregor Wallat (University)** 2010-2014 I took over Gregor's supervision when his supervisor left St Andrews. "Structure determination of the MOMO from C. jejuni & structural and functional studies of the endonuclease from lassa virus" Works for a company in Germany
24. **Andrew F Bent (University)** 2011-2015 "Structural biology of patellamide biosynthesis" PDRA in RIPPTIDE PHARMA; deceased.
25. **Greg Mann (St And)** 2012-2016 "Cyanobactins: structural and chemical " PDRA Switzerland
26. **Laura Woodward (BBSRC)** 2013-2017 "Structural and Biochemical Characterisation of Enzymes Required for Gram-Negative Capsule Biosynthesis" PDRA Germany, now editor of Angewantie
27. **Luana Ferrara (EU)** 2013-2017 "Structural and functional studies of porins from pathogenic bacteria" Post-doc Italy
28. **Ying Ge (ERC)** 2014-2018 "Mechanistic structural biology" PDRA Imperial College
29. **Hannes Ludwig (BBSRC)** 2016-2020 "Structural and functional characterisation of enzymes promising biocatalytic prowess" Post-doc USA
30. **Sisi Gao (ERC)** 2016-2020 "Enzyme mechanism" PDRA LM-MRC
31. **Micah Lee (Self)** 2018-2021 "Carbohydrate processing" Legal profession
32. **Wenbo Li (Self)** 2020-2023 "C-C bond forming enzymes" PDRA
33. **Helena Watson (EPRSC)** 2021-2024 "Structural Biology and CryoET" PDRA
34. **William Bowles (BBSRC)** 2021-2024 "Electron Damage to biological samples" PDRA

*One student under my supervision (2001-2006) Michael Hutchings did not submit a thesis.

Destination of post-docs

1. Marie-France Giraud (Faculty position CNRS)
2. Wulf Blankenfelt (Professor Germany)
3. David Sanders (Faculty position Canada)
4. Haitham Idriss (Medical Student)
5. Miryam Asuncion (Permanent job in Industry)
6. Adam Staines (post-doc Dundee, now BBSRC)
7. Lesley Thompson (Permanent job in NHS)
8. Genevive Temple (Permanent job in NHS)
9. Irina Grishkovska (post-doc in Germany)
10. Vicki Bamford (Lectureship Reading University)
11. Changjiang Dong (Independent Wellcome Fellowship, Professor UEA)
12. Lester G. Carter (Post-doc in Paul Adams lab, California)
13. Ken Johnson (Scientist, Tromso Sweden)
14. Kerstin Steiner (Faculty position Graz, Austria)
15. Muse Oke (scientist, KAUST; now Faculty University of Lagos)
16. Gregor Hageluken (Faculty position, Bonn)
17. Simon Bushell (PDRA Oxford)
18. Jesko Koehnke (Faculty position Saarbruken)
19. Christos Pliotas (Faculty St Andrews)
20. Haigang Song (Senior Post-doc Oxford)
21. Lucile Moynie (Junior Investigator Franklin)
22. Jiwei Liu (1000 Talent Scholar & Investigator Sichuan University)

23. Yun Yang (Lecturer Sichuan University)

First destination of technicians & MSc students

1. Didier Danillon (permanent job in Industry)
2. Helen McMiken (permanent job in Industry)
3. Amanda Crawford (emigrated to Australia)
4. Helen Powers (resigned from employment to raise family)
5. Mark Dorward (permanent post Dundee)
6. Helen Falconer (post-doc at Edinburgh)
7. Michal Zawadski (job with Syngenta)
8. Fraser Duthie (science job in Bonn)
9. Phil Ward (retired reaching normal retirement age)

Naismith publications

2025 (6 publications)

292. Velazco, A., T. Glen, S. Klumpe, A. Pennington, J. Zhang, J. L. R. Smith, C. Glynn, W. Bowles, M. Kobylynska, R. A. Fleck, J.H. Naismith, J. S. Kim, M. C. Darrow, M. Grange, A. I Kirkland, M. Dumoux 2025 Reduction of SEM charging artefacts in native cryogenic biological samples *Nat Comm* 16, 5204 (doi 10.1038/s41467-025-60545-3)
291. Matabaro, E., H. Song, L. Sonderegger, F. Gherlone, A. M. Giltrap, S. Liver, A. D. Gossert, M. Künzler, J. H. Naismith 2025 Macrocyclization of backbone N-methylated peptides by a prolyl oligopeptidase with a distinctive substrate recognition mechanism *Chem Sci* 16, 14196-14206 (doi 10.1039/D5SC03723A)
290. Yang, Y., M. Batista, B. R. Clarke, M. R. Agyare-Tabbi, H. Song, N. M. Kuehfuss, A. Le Bas, C. V. Robinson, C. Whitfield, J. H. Naismith, J. Liu 2025 Molecular basis for the phosphorylation of bacterial tyrosine kinase Wzc. *Nat Comm* 16, 3437 (doi 10.1038/s41467-025-58693-7)
289. Berger, C., H. Watson, J.H. Naismith, M. Dumoux, M. Grange 2025 Xenon plasma focused ion beam lamella fabrication on high-pressure frozen specimens for structural cell biology *Nat Comm* 16, 2286 (doi 10.1038/s41467-025-57493-3)
288. Hussain, N., H. Mikolajek, P. J. Harrison, N. Paterson, M. W. Akhtar, S. Sadaf, J. H. Naismith 2025 Structural and functional snapshots of a broad-specificity endoglucanase from Thermogutta terrifontis for biomass saccharification Arch Bioch Biophy, 764, 110274 (doi 0.1016/j.abb.2024.110274)
287. Le Bas A, et al 2025 Structure of WzxE the lipid III flippase for Enterobacterial Common Antigen polysaccharide *Open Biol*15, 240310 (doi 10.1098/rsob.240310)

2024 (6 publications)

286. Cornish, K. et al. 2024 Structural and functional characterization of nanobodies that neutralize Omicron variants of SARS-CoV-2 *Open Biol* 14, 230252 (doi 10.1098/rsob.230252)
285. Richards, N. G. J., J. H Naismith 2024 The chemistry of Formycin biosynthesis *Front. Chem. Biol*. 3, 1428646 (doi 10.3389/fchbi.2024.1428646)
284. Deimel, L.P., et al. 2024 Covalent penicillin-protein conjugates elicit anti-drug antibodies that are clonally and functionally restricted. *Nat Comm* 15, 6851 (2024). (doi 10.1038/s41467-024-51138-7
283. Parkhurst, J. M., T. Varslot, M. Dumoux, C.A. Siebert, M. Darrow, M. Basham, A. Kirkland, M. Grange, G. Evans, J.H. Naismith 2024 Pillar data-acquisition strategies for cryo-electron tomography of beam-sensitive biological samples Acta Cryst D80, 421-438 (doi 10.1107/S2059798324004546)
282. Parkhurst, J. M., A. Cavalleri, M. Dumoux, D. Clare, C.A. Siebert, G. Evans, J.H. Naismith, A. Kirkland, J. W. Essex 2024 Computational models of amorphous ice for accurate simulation of cryo-EM images of biological samples Ultramicroscopy, 256, 113882 (doi 10.1016/j.ultramic.2023.113882)
281. Couturier, C. 2024. et al. Studies of antibacterial activity (in vitro and in vivo) and mode of action for des-acyl tridecaptins (DATs) *Eur. J. Med. Chem* 266, 116087 (doi 10.1016/j.ejmech.2023.116097)

2023 (8 publications)

280. Madhusudanan, A., C. Iddon, M. Cevik, J. H. Naismith, S. Fitzgerald 2023 Non-pharmaceutical interventions for COVID-19: a systematic review on environmental control measures *Phil Trans R Soc* A381 20230130 (doi 10.1098/rsta.2023.0130)
279. Gao, X., J. W. Thrush, J. Gor, J. H. Naismith, R. J. Owens, S. J. Perkins 2023 The solution structure of the heavy chain–only C5-Fc nanobody reveals exposed variable regions that are optimal for COVID-19 antigen interactions *J. Biol. Chem*. 299, 105337 (doi 10.1016/j.jbc.2023.105337)

278. Parkhurst, J. M. et al. 2023 Investigation of the milling characteristics of different focused-ion-beam sources assessed by three-dimensional electron diffraction from crystal lamellae. *IUCrJ* (doi 10.1107/S2052252523001902).
278. Dumoux, M., J. L. R. Smith, T. Glen, M. Grange, M. C. Darrow, & J. H. Naismith. 2023 A protocol for cryogenic volumetric imaging using serial plasma FIB/SEM. In *Methods in Cell Biology* (eds. pp. Elsevier) (doi: 10.1016/bs.mcb.2023.01.015)
277. Weckener, M., L. S. Woodward, B.R. Clarke, H. Liu, P. N. Ward, A. Le Bas, D. Bhella, C. Whitfield, J.H. Naismith 2023 The lipid linked oligosaccharide polymerase Wzy and its regulating co-polymerase, Wzz, from enterobacterial common antigen biosynthesis form a complex *Roy Soc Open Bio* (doi 10.1098/rsob.220373)
276. Berger, C., M. Dumoux, T. Glen, N. B. Yee, J. M. Mitchels, Z. Patáková, M. C. Darrow, J. H. Naismith, & M. Grange. 2023 Plasma FIB milling for the determination of structures *in situ*. *Nat Commun* **14**: 629 (doi 10.1038/s41467-023-36372-9).
275. Li, W., G. C. Girt, A. Radadiya, J. J. P. Stewart, N. G. J. Richards, & J. H. Naismith. 2023 Experimental and computational snapshots of C-C bond formation in a C-nucleoside synthase. *Open Biol* **13**: 220287 (doi 10.1098/rsob.220287).
274. Dumoux, M. et al. 2023 Cryo-plasma FIB/SEM volume imaging of biological specimens. *Elife* **12**: e83623 (doi 10.7554/eLife.83623).

2022 (6 publications)

273. Mikolajek, H. M. Weckener, Z.F. Brotzakis, J. Huo, E.V Dalietou, A. Le Bas, P. Sormanni, P.J. Harrison, P.N. Ward, S. Truong, L. Moynie, D. Clare, M. Dumoux, J. Dormon, C. Norman, N. Hussain, V. Vogirala, R. J. Owens, M. Vendruscolo & J. H. Naismith (2022) Correlation between the binding affinity and the conformational entropy of nanobodies targeting the SARS-CoV-2 spike protein. Proc Natl Acad Sci USA 118, e2205412119 (doi 10.1073/pnas.2205412119)
272. Le Bas, A., H. Mikolajek, J. Huo, C. Norman, J. Dormon, J. H. Naismith & R. J. Owens Production and Crystallization of Nanobodies in Complex with the Receptor Binding Domain of the SARS-CoV-2 Spike Protein 12, 10.21769/BioProtoc.4406
271. Buchanan, C.J., B. Gaunt, P. J. Harrison, Y. Yang, J. Liu, A. Khan, A. M. Giltrap, A. Le Bas, P. N. Ward, K. Gupta, M. Dumoux, T. K. Tan, L. Schimaski, S. Daga, N. Picchiotti, M. Baldassarri, E. Benetti, C. Fallerini, F. Fava, A. Giliberti, P. I. Koukos, M. J. Davy, A. Lakshminarayanan, X. Xue, G. Papadakis, L. P. Deimel, V. Casablancas-Antràs, T. D. W. Claridge, A. M .J .J. Bonvin, Q. J. Sattentau, S.Furini, M. Gori, J. Huo, R. J. Owens, C. Schaffitzel, I. Berger, A. Renieri, GEN-COVID Multicenter Study, Naismith, J.H., Baldwin, A.J. & B. G. Davis (2022) Cryptic pathogen-sugar interactions revealed by universal saturation transfer analysis. Science, 337, 385-401 (doi 10.1126/science.abm3125)
270. Levitz, T.S, Weckener, M. I. Fong, Naismith, J. H., Drennan, C. L., Brignole, E.J., Clare, D.K, & Darrow, M.C. (2022) Approaches to using the chameleon: robust, automated, fast-plunge cryoEM specimen preparation. Frontiers Molecular Biosciences. (doi 10.3389/fmolb.2022.903148)
269. Moynié, L., Hoegy, F., Milenkovic, S., Munier, M., Paulen, A., Gasser, V., Faucon, A.L., Zill, N., Naismith, J.H., Ceccarelli, M., Schalk, I.J., Mislin, G.L.A. (2022) Hijacking of the enterobactin pathway by a synthetic catechol vectors designed for antibiotic delivery in pseudomonas aeruginosa. ACS Infectious Disease (doi 10.1021/acsinfecdis.2c00202).
268. Le Bas, A., H. Mikolajek, J. Huo, C. Norman, J. Dormon, J. H. Naismith & R. J. Owens Production and Crystallization of Nanobodies in Complex with the Receptor Binding Domain of the SARS-CoVCoV-2 Spike Protein. Bio-protocol 2022 May 5; 12(09): e4406. (doi: 10.21769/BioProtoc.4406)

2021 (14 publications)

267. Tang, Q., R. J. Owens, & J. H. Naismith. 2021 Structural Biology of Nanobodies against the Spike Protein of SARS-CoV-2. *Viruses* **13**: 2214 (doi 10.3390/v13112214).

266. Zhang, J. et al. 2021 Catalytic flexibility of rice glycosyltransferase OsUGT91C1 for the production of palatable steviol glycosides. *Nat Commun* **12**: 7030 (doi 10.1038/s41467-021-27144-4).
265. Naismith, J. H. 2021 Iain Donald Campbell. 24 April 1941—5 March 2014. *Biographical Memoirs of Fellows of the Royal Society* **71**: 59-77 (doi 10.1098/rsbm.2021.0003).
264. Parkhurst, J. M., M. Dumoux, M. Basham, D. Clare, C. A. Siebert, T. Varslot, A. Kirkland, J. H. Naismith, & G. Evans. 2021 Parakeet: a digital twin software pipeline to assess the impact of experimental parameters on tomographic reconstructions for cryo-electron tomography. *Open Biology* **11**: (doi 10.1098/rsob.210160).
263. Ganyuan X., G., M.S. Alphey, F. Tran, L. Pirrie, P. Milbeo, Y. Zhou, J.K. Bickel, O. Kempf, K. Kempf, J.H. Naismith, N.J. Westwood 2021 Next generation Glucose-1-phosphate thymidylyltransferase (RmlA) inhibitors: An extended SAR study to direct future design. Bioorg. Med. Chem. ***In press*** (doi: 10.1016/j.bmc.2021.116477)
262. Girt, G. C., A. Lakshminarayanan, J. Huo, J. Dormon, C. Norman, B. Afrough, A. Harding, W. James, R. J. Owens, & J. H. Naismith. 2021 The use of nanobodies in a sensitive ELISA test for SARS-CoV-2 Spike 1 protein. *Royal Society Open Science* **8**: (doi 10.1098/rsos.211016).
261. Yang, Y., J. Liu, B. R. Clarke, L. Seidel, J. R. Bolla, P. N. Ward, P. Zhang, C. V. Robinson, C. Whitfield, & J. H. Naismith. 2021 The molecular basis of regulation of bacterial capsule assembly by Wzc. *Nat Commun* **12**: 4349 (doi 10.1038/s41467-021-24652-1).
260. Barnawi, H., L. Woodward, N. Fava, M. Roubakha, S. D. Shaw, C. Kubinec, J. H. Naismith, & C. Creuzenet. 2021 Structure-function studies of the C3/C5 epimerases and C4 reductases of the *Campylobacter jejuni* capsular heptose modification pathways. *J Biol Chem* 100352 (doi 10.1016/j.jbc.2021.100352).
259. Phillips, P., J. M. Parkhurst, I. Kounatidis, C. Okolo, T. M. Fish, J. H. Naismith, M. A. Walsh, M. Harkiolaki, & M. Dumoux. 2021 Single Cell Cryo-Soft X-ray Tomography Shows That Each *Chlamydia Trachomatis* Inclusion Is a Unique Community of Bacteria. *Life (Basel)* **11**: 842 (doi 10.3390/life11080842).
258. Uchański, T. et al. 2021 Megabodies expand the nanobody toolkit for protein structure determination by single-particle cryo-EM. *Nat Methods* **18**: 60-68 (doi 10.1038/s41592-020-01001-6).
257. Matabaro, E., H. Song, C. Chepkirui, H. Kaspar, L. Witte, J. H. Naismith, M. F. Freeman, & M. Künzler. 2021 Enzyme-mediated backbone N-methylation in ribosomally encoded peptides. *Methods Enzymol* **656**: 429-458 (doi 10.1016/bs.mie.2021.04.014).
256. Doble, M. V., L. Obrecht, H.-J. Joosten, M. Lee, H. J. Rozeboom, E. Branigan, J. H. Naismith, D. B. Janssen, A. G. Jarvis, & P. C. J. Kamer. 2021 Engineering Thermostability in Artificial Metalloenzymes to Increase Catalytic Activity. *ACS Catalysis* **11**: 3620-3627.
255. Song, H., A. J. Burton, S. L. Shirran, J. Fahrig-Kamarauskaitė, H. Kaspar, T. W. Muir, & M. Künzler & J.H. Naismith 2021 Engineering of a peptide α-N-methyltransferase to methylate non-proteinogenic amino acids. *Angew Chem Int Ed Engl* (doi 10.1002/anie.202100818).
254. Huo, J. et al. 2021 A potent SARS-CoV-2 neutralising nanobody shows therapeutic efficacy in the Syrian golden hamster model of COVID-19. *Nat Commun* **12**: 5469 (doi 10.1038/s41467-021-25480-z).

2020 (8 publications)

253. Gao, S., A. Radadiya, W. Li, H. Liu, W. Zhu, V. de Crécy-Lagard, N. G. J. Richards, & J. H. Naismith. 2020 Uncovering the chemistry of C-C bond formation in C-nucleoside biosynthesis: crystal structure of a C-glycoside synthase/PRPP complex. *Chem Commun (Camb)* **56**: 7617-7620 (doi 10.1039/d0cc02834g).
252. Bell, A., E. Severi, M. O. Lee, S. Monaco, D. Latousakis, J. Angulo, G. H. Thomas, J. Naismith, & N. Juge. 2020 Uncovering a novel molecular mechanism for scavenging sialic acids in bacteria. *J Biol Chem* (doi 10.1074/jbc.RA120.014454).

251. Song, H., C. Rao, Z. Deng, Y. Yi, & J. H. Naismith. 2020 The biosynthetic pathway of the benzoxazole in nataxazole proceeds via an unstable ester and has synthetic utility. *Angew Chem Int Ed Engl* (doi 10.1002/anie.201915685).
250. Song, H., J. R. Fahrig-Kamarauskaitè, E. Matabaro, H. Kaspar, S. L. Shirran, C. Zach, A. Pace, B. A. Stefanov, J. H. Naismith, & M. Künzler. 2020 Substrate Plasticity of a Fungal Peptide α-N-Methyltransferase. *ACS Chem Biol* **15**: 1901-1912 (doi 10.1021/acschembio.0c00237).
249. Zhou, D. et al. 2020 Structural basis for the neutralization of SARS-CoV-2 by an antibody from a convalescent patient. *Nat Struct Mol Biol* (doi 10.1038/s41594-020-0480-y).
248. Huo, J. et al. & J. H. Naismith 2020 Neutralizing nanobodies bind SARS-CoV-2 spike RBD and block interaction with ACE2. *Nat Struct Mol Biol* **27**: 846-854 (doi 10.1038/s41594-020-0469-6).
247. Song, H. & J. H. Naismith. 2020 Enzymatic methylation of the amide bond. *Curr Opin Struct Biol* **65**: 79-88 (doi 10.1016/j.sbi.2020.06.004).
246. Naismith, J. H. & E. Parker. 2020 Editorial overview: Catalysis and regulation. *Curr Opin Struct Biol* **65**: iii-iv (doi 10.1016/j.sbi.2020.10.003).

2019 (8 publications)

245. Harrison, P. J., K. Gable, N. Somashekarappa, V. Kelly, D. J. Clarke, J. H. Naismith, T. M. Dunn, & D. J. Campopiano. 2019 Use of isotopically labeled substrates reveals kinetic differences between human and bacterial serine palmitoyltransferase. *J Lipid Res* **60**: 953-962 (doi 10.1194/jlr.M089367).
244. Moynié, L. et al. 2019 The complex of ferric-enterobactin with its transporter from Pseudomonas aeruginosa suggests a two-site model. *Nat Commun* **10**: 3673 (doi 10.1038/s41467-019-11508-y).
243. Vergalli, J. et al. 2019 Porins and small-molecule translocation across the outer membrane of Gram-negative bacteria. *Nat Rev Microbiol* (doi 10.1038/s41579-019-0294-2).
242. Gao, S., H. Liu, V. de Crécy-Lagard, W. Zhu, N. G. J. Richards, & J. H. Naismith. 2019 PMP-diketopiperazine adducts form at the active site of a PLP dependent enzyme involved in formycin biosynthesis. *Chem Commun (Camb)* **55**: 14502-14505 (doi 10.1039/c9cc06975e).
241. Ge, Y., C. M. Czekster, O. K. Miller, C. H. Botting, U. Schwarz-Linek, & J. H. Naismith. 2019 Insights into the Mechanism of the Cyanobactin Heterocyclase Enzyme. *Biochemistry* **58**: 2125-2132 (doi 10.1021/acs.biochem.9b00084).
240. Yang, T. et al. 2019 Hydrophobic recognition allows the glycosyltransferase UGT76G1 to catalyze its substrate in two orientations. *Nat Commun* **10**: 3214 (doi 10.1038/s41467-019-11154-4).
239. Scorciapino, M. A., G. Malloci, I. Serra, S. Milenkovic, L. Moynié, J. H. Naismith, E. Desarbre, M. G. P. Page, & M. Ceccarelli. 2019 Complexes formed by the siderophore-based monosulfactam antibiotic BAL30072 and their interaction with the outer membrane receptor PiuA of P. aeruginosa. *Biometals* **32**: 155-170 (doi 10.1007/s10534-018-00166-0).
238. Gkotsi, D. S. et al. 2019 A marine viral halogenase that iodinates diverse substrates. *Nat Chem* **11**: 1091-1097 (doi 10.1038/s41557-019-0349-z).

2018 (10 publications))

237. Luscher, A., L. Moynié, P. S. Auguste, D. Bumann, L. Mazza, D. Pletzer, J. H. Naismith, & T. Köhler. 2018 TonB-Dependent Receptor Repertoire of Pseudomonas aeruginosa for Uptake of Siderophore-Drug Conjugates. *Antimicrob Agents Chemother* **62**: (doi 10.1128/AAC.00097-18).
236. Ojo, O. S. et al. 2018 Synthesis of the natural product descurainolide and cyclic peptides from lignin-derived aromatics. *Org Biomol Chem* **16**: 266-273 (doi 10.1039/c7ob02697h).
235. Moynié, L., I. Serra, M. A. Scorciapino, E. Oueis, M. G. Page, M. Ceccarelli, & J. H. Naismith. 2018 Preacinetobactin not acinetobactin is essential for iron uptake by the BauA transporter of the pathogen Acinetobacter baumannii. *Elife* **7**: (doi 10.7554/eLife.42270).
234. Liston, S. D., S. A. McMahon, A. Le Bas, M. D. L. Suits, J. H. Naismith, & C. Whitfield. 2018 Periplasmic depolymerase provides insight into ABC transporter-dependent secretion of

bacterial capsular polysaccharides. *Proc Natl Acad Sci U S A* **115**: E4870-E4879 (doi 10.1073/pnas.1801336115).

233. Gao, S., Y. Ge, A. F. Bent, U. Schwarz-Linek, & J. H. Naismith. 2018 Oxidation of the Cyanobactin Precursor Peptide Is Independent of the Leader Peptide and Operates in a Defined Order. *Biochemistry* **57**: 5996-6002 (doi 10.1021/acs.biochem.8b00835).
232. Acosta-Gutiérrez, S. et al. 2018 Getting Drugs into Gram-Negative Bacteria: Rational Rules for Permeation through General Porins. *ACS Infect Dis* **4**: 1487-1498 (doi 10.1021/acsinfecdis.8b00108).
231. Ronan, J. L., N. Kadi, S. A. McMahon, J. H. Naismith, L. M. Alkhalaf, & G. L. Challis. 2018 Desferrioxamine biosynthesis: diverse hydroxamate assembly by substrate-tolerant acyl transferase DesC. *Philos Trans R Soc Lond B Biol Sci* **373**: (doi 10.1098/rstb.2017.0068).
230. Ludewig, H., C. M. Czekster, E. Oueis, E. S. Munday, M. Arshad, S. A. Synowsky, A. F. Bent, & J. H. Naismith. 2018 Characterization of the Fast and Promiscuous Macrocyclase from Plant PCY1 Enables the Use of Simple Substrates. *ACS Chem Biol* **13**: 801-811 (doi 10.1021/acschembio.8b00050).
229. Song, H. et al. 2018 A molecular mechanism for the enzymatic methylation of nitrogen atoms within peptide bonds. *Sci Adv* **4**: eaat2720 (doi 10.1126/sciadv.aat2720).
228. Perraud, Q., L. Moynié, V. Gasser, M. Munier, J. Godet, F. Hoegy, Y. Mély, G. L. A. Mislin, J. H. Naismith, & I. J. Schalk. 2018 A Key Role for the Periplasmic PfeE Esterase in Iron Acquisition via the Siderophore Enterobactin in Pseudomonas aeruginosa. *ACS Chem Biol* **13**: 2603-2614 (doi 10.1021/acschembio.8b00543).

2017 (14 publications)

227. Burkhart, B. J., C. J. Schwalen, G. Mann, J. H. Naismith, & D. A. Mitchell. 2017 YcaO-Dependent Posttranslational Amide Activation: Biosynthesis, Structure, and Function. *Chem Rev* **117**: 5389-5456 (doi 10.1021/acs.chemrev.6b00623).
226. Wang, M., L. Moynié, P. J. Harrison, V. Kelly, A. Piper, J. H. Naismith, & D. J. Campopiano. 2017 Using the pimeloyl-CoA synthetase adenylation fold to synthesize fatty acid thioesters. *Nat Chem Biol* **13**: 660-667 (doi 10.1038/nchembio.2361).
225. Carroll, C. S., C. L. Grieve, I. Murugathasan, A. J. Bennet, C. M. Czekster, H. Liu, J. Naismith, & M. M. Moore. 2017 The rhizoferrin biosynthetic gene in the fungal pathogen Rhizopus delemar is a novel member of the NIS gene family. *Int J Biochem Cell Biol* **89**: 136-146 (doi 10.1016/j.biocel.2017.06.005).
224. Oueis, E., B. Nardone, M. Jaspars, N. J. Westwood, & J. H. Naismith. 2017 Synthesis of Hybrid Cyclopeptides through Enzymatic Macrocyclization. *ChemistryOpen* **6**: 11-14 (doi 10.1002/open.201600134).
223. Moynié, L., A. Luscher, D. Rolo, D. Pletzer, A. Tortajada, H. Weingart, Y. Braun, M. G. Page, J. H. Naismith, & T. Köhler. 2017 Structure and Function of the PiuA and PirA Siderophore-Drug Receptors from Pseudomonas aeruginosa and Acinetobacter baumannii. *Antimicrob Agents Chemother* **61**: (doi 10.1128/AAC.02531-16).
222. Ackermann, K., C. Pliotas, S. Valera, J. H. Naismith, & B. E. Bode. 2017 Sparse Labeling PELDOR Spectroscopy on Multimeric Mechanosensitive Membrane Channels. *Biophys J* **113**: 1968-1978 (doi 10.1016/j.bpj.2017.09.005).
221. Barski, M., B. Brennan, O. K. Miller, J. A. Potter, S. Vijayakrishnan, D. Bhella, J. H. Naismith, R. M. Elliott, & U. Schwarz-Linek. 2017 Rift Valley fever phlebovirus NSs protein core domain structure suggests molecular basis for nuclear filaments. *Elife* **6**: (doi 10.7554/eLife.29236).
220. Michailidou, F., C. W. Chung, M. J. B. Brown, A. F. Bent, J. H. Naismith, W. J. Leavens, S. M. Lynn, S. V. Sharma, & R. J. M. Goss. 2017 Pac13 is a Small, Monomeric Dehydratase that Mediates the Formation of the 3'-Deoxy Nucleoside of Pacidamycins. *Angew Chem Int Ed Engl* **56**: 12492-12497 (doi 10.1002/anie.201705639).
219. Stroek, R. et al. 2017 Kinetics and Structure of a Cold-Adapted Hetero-Octameric ATP Phosphoribosyltransferase. *Biochemistry* **56**: 793-803 (doi 10.1021/acs.biochem.6b01138).

218. Czekster, C. M. & J. H. Naismith. 2017 Kinetic Landscape of a Peptide Bond-Forming Prolyl Oligopeptidase. *Biochemistry* **56**: 2086-2095 (doi 10.1021/acs.biochem.7b00012).
217. Asadi, J., S. Ferguson, H. Raja, C. Hacker, P. Marius, R. Ward, C. Pliotas, J. Naismith, & J. Lucocq. 2017 Enhanced imaging of lipid rich nanoparticles embedded in methylcellulose films for transmission electron microscopy using mixtures of heavy metals. *Micron* **99**: 40-48 (doi 10.1016/j.micron.2017.03.019).
216. Czekster, C. M., H. Ludewig, S. A. McMahon, & J. H. Naismith. 2017 Characterization of a dual function macrocyclase enables design and use of efficient macrocyclization substrates. *Nat Commun* **8**: 1045 (doi 10.1038/s41467-017-00862-4).
215. Oueis, E., H. Stevenson, M. Jaspars, N. J. Westwood, & J. H. Naismith. 2017 Bypassing the proline/thiazoline requirement of the macrocyclase PatG. *Chem Commun (Camb)* **53**: 12274-12277 (doi 10.1039/c7cc06550g).
214. Pliotas, C. et al. 2017 Adenosine Monophosphate Binding Stabilizes the KTN Domain of the *Shewanella denitrificans* Kef Potassium Efflux System. *Biochemistry* **56**: 4219-4234 (doi 10.1021/acs.biochem.7b00300).

2016 (17 publications)

213. Brás, N. F., P. Ferreira, A. R. Calixto, M. Jaspars, W. Houssen, J. H. Naismith, P. A. Fernandes, & M. J. Ramos. 2016 The Catalytic Mechanism of the Marine-Derived Macrocyclase PatGmac. *Chemistry* **22**: 13089-13097 (doi 10.1002/chem.201601670).
212. Constantinescu-Aruxandei, D., B. Petrovic-Stojanovska, O. Schiemann, J. H. Naismith, & M. F. White. 2016 Taking a molecular motor for a spin: helicase mechanism studied by spin labeling and PELDOR. *Nucleic Acids Res* **44**: 954-968 (doi 10.1093/nar/gkv1373).
211. Bent, A. F., G. Mann, W. E. Houssen, V. Mykhaylyk, R. Duman, L. Thomas, M. Jaspars, A. Wagner, & J. H. Naismith. 2016 Structure of the cyanobactin oxidase ThcOx from Cyanothece sp. PCC 7425, the first structure to be solved at Diamond Light Source beamline I23 by means of S-SAD. *Acta Crystallogr D Struct Biol* **72**: 1174-1180 (doi 10.1107/S2059798316015850).
210. Pliotas, C. & J. H. Naismith. 2016 Spectator no more, the role of the membrane in regulating ion channel function. *Curr Opin Struct Biol* **45**: 59-66 (doi 10.1016/j.sbi.2016.10.017).
209. Yang, J. et al. 2016 Pironetin reacts covalently with cysteine-316 of α-tubulin to destabilize microtubule. *Nat Commun* **7**: 12103 (doi 10.1038/ncomms12103).
208. Hacker, C., J. Asadi, C. Pliotas, S. Ferguson, L. Sherry, P. Marius, J. Tello, D. Jackson, J. Naismith, & J. M. Lucocq. 2016 Nanoparticle suspensions enclosed in methylcellulose: a new approach for quantifying nanoparticles in transmission electron microscopy. *Sci Rep* **6**: 25275 (doi 10.1038/srep25275).
207. Ferrara, L. G. M. et al. 2016 MOMP from Campylobacter jejuni Is a Trimer of 18-Stranded β-Barrel Monomers with a Ca^{2+} Ion Bound at the Constriction Zone. *J Mol Biol* **428**: 4528-4543 (doi 10.1016/j.jmb.2016.09.021).
206. Bajaj, H., M. A. Scorciapino, L. Moynié, M. G. Page, J. H. Naismith, M. Ceccarelli, & M. Winterhalter. 2016 Molecular Basis of Filtering Carbapenems by Porins from β-Lactam-resistant Clinical Strains of Escherichia coli. *J Biol Chem* **291**: 2837-2847 (doi 10.1074/jbc.M115.690156).
205. Czekster, C. M., Y. Ge, & J. H. Naismith. 2016 Mechanisms of cyanobactin biosynthesis. *Curr Opin Chem Biol* **35**: 80-88 (doi 10.1016/j.cbpa.2016.08.029).
204. Constantinescu-Aruxandei, D., B. Petrovic-Stojanovska, J. C. Penedo, M. F. White, & J. H. Naismith. 2016 Mechanism of DNA loading by the DNA repair helicase XPD. *Nucleic Acids Res* (doi 10.1093/nar/gkw102).
203. Thompson, S., S. A. McMahon, J. H. Naismith, & D. O'Hagan. 2016 Exploration of a potential difluoromethyl-nucleoside substrate with the fluorinase enzyme. *Bioorg Chem* **64**: 37-41 (doi 10.1016/j.bioorg.2015.11.003).
202. Oueis, E., M. Jaspars, N. J. Westwood, & J. H. Naismith. 2016 Enzymatic Macrocyclization of 1,2,3-Triazole Peptide Mimetics. *Angew Chem Int Ed Engl* **55**: 5842-5845 (doi 10.1002/anie.201601564).

201. Woodward, L. & J. H. Naismith. 2016 Bacterial polysaccharide synthesis and export. *Curr Opin Struct Biol* **40**: 81-88 (doi 10.1016/j.sbi.2016.07.016).
200. Adaba, R. I., G. Mann, A. Raab, W. E. Houssen, A. R. McEwan, L. Thomas, J. Tabudravu, J. H. Naismith, & M. Jaspars. 2016 Accurate quantification of modified cyclic peptides without the need for authentic standards. *Tetrahedron* **72**: 8603-8609 (doi 10.1016/j.tet.2016.11.040).
199. Valera, S., K. Ackermann, C. Pliotas, H. Huang, J. H. Naismith, & B. E. Bode. 2016 Accurate Extraction of Nanometer Distances in Multimers by Pulse EPR. *Chemistry* (doi 10.1002/chem.201505143).
198. Parajuli, A. et al. 2016 A Unique Tryptophan C-Prenyltransferase from the Kawaguchipeptin Biosynthetic Pathway. *Angew Chem Int Ed Engl* (doi 10.1002/anie.201509920).
197. Moynié, L., A. G. Hope, K. Finzel, J. Schmidberger, S. M. Leckie, G. Schneider, M. D. Burkart, A. D. Smith, D. W. Gray, & J. H. Naismith. 2016 A Substrate Mimic Allows High-Throughput Assay of the FabA Protein and Consequently the Identification of a Novel Inhibitor of Pseudomonas aeruginosa FabA. *J Mol Biol* **428**: 108-120 (doi 10.1016/j.jmb.2015.10.027).

2015 (6 publications)

196. Juárez, J. F., H. Liu, M. T. Zamarro, S. McMahon, H. Liu, J. H. Naismith, C. Eberlein, M. Boll, M. Carmona, & E. Díaz. 2015 Unraveling the specific regulation of the central pathway for anaerobic degradation of 3-methylbenzoate. *J Biol Chem* **290**: 12165-12183 (doi 10.1074/jbc.M115.637074).
195. Pliotas, C. A E. Dahl, T. Rasmussen, K. R. Mahendran, T. K. Smith, P. Marius, J. Gault, T. Banda, A. Rasmussen, S. Miller, C.V. Robinson , H. Bayley , M. S. P. Sansom, I.R. Booth, J. H. Naismith. 2015 The role of lipids in mechanosensation. *Nat Struct Mol Biol* **22**: 991-998 (doi 10.1038/nsmb.3120).
194. Branigan, E., A. Plechanovová, E. G. Jaffray, J. H. Naismith, & R. T. Hay. 2015 Structural basis for the RING-catalyzed synthesis of K63-linked ubiquitin chains. *Nat Struct Mol Biol* **22**: 597-602 (doi 10.1038/nsmb.3052).
193. Koehnke, J., G. Mann, A. F. Bent, H. Ludewig, S. Shirran, C. Botting, T. Lebl, W. E. Houssen, M. Jaspars, & J. H. Naismith. 2015 Structural analysis of leader peptide binding enables leader-free cyanobactin processing. *Nat Chem Biol* **11**: 558-563 (doi 10.1038/nchembio.1841).
192. Oueis, E., C. Adamson, G. Mann, H. Ludewig, P. Redpath, M. Migaud, N. J. Westwood, & J. H. Naismith. 2015 Derivatisable Cyanobactin Analogues: A Semisynthetic Approach. *Chembiochem* **16**: 2646-2650 (doi 10.1002/cbic.201500494).
191. Hagelueken, G., B. R. Clarke, H. Huang, A. Tuukkanen, I. Danciu, D. I. Svergun, R. Hussain, H. Liu, C. Whitfield, & J. H. Naismith. 2015 A coiled-coil domain acts as a molecular ruler to regulate O-antigen chain length in lipopolysaccharide. *Nat Struct Mol Biol* **22**: 50-56 (doi 10.1038/nsmb.2935).

2014 (9 publications)

190. Nickerson, N. N., I. L. Mainprize, L. Hampton, M. L. Jones, J. H. Naismith, & C. Whitfield. 2014 Trapped translocation intermediates establish the route for export of capsular polysaccharides across Escherichia coli outer membranes. *Proc Natl Acad Sci U S A* **111**: 8203-8208 (doi 10.1073/pnas.1400341111).
189. Mann, G., J. Koehnke, A. F. Bent, R. Graham, W. Houssen, M. Jaspars, U. Schwarz-Linek, & J. H. Naismith. 2014 The structure of the cyanobactin domain of unknown function from PatG in the patellamide gene cluster. *Acta Crystallogr F Struct Biol Commun* **70**: 1597-1603 (doi 10.1107/S2053230X1402425X).
188. Koehnke, J., A. F. Bent, W. E. Houssen, G. Mann, M. Jaspars, & J. H. Naismith. 2014 The structural biology of patellamide biosynthesis. *Curr Opin Struct Biol* **29**: 112-121 (doi 10.1016/j.sbi.2014.10.006).
187. Ward, R. et al. 2014 Probing the structure of the mechanosensitive channel of small conductance in lipid bilayers with pulsed electron-electron double resonance. *Biophys J* **106**: 834-842 (doi 10.1016/j.bpj.2014.01.008).

186. Deng, H., L. Ma, N. Bandaranayaka, Z. Qin, G. Mann, K. Kyeremeh, Y. Yu, T. Shepherd, J. H. Naismith, & D. O'Hagan. 2014 Identification of fluorinases from Streptomyces sp MA37, Norcardia brasiliensis, and Actinoplanes sp N902-109 by genome mining. *Chembiochem* **15**: 364-368 (doi 10.1002/cbic.201300732).
185. Naismith, J. H. & E. J. Parker. 2014 Editorial overview: Catalysis and regulation: enzyme catalysis, biosynthetic pathways and regulation. *Curr Opin Struct Biol* **29**: iv-v (doi 10.1016/j.sbi.2014.11.004).
184. Kassaar, O., S. A. McMahon, R. Thompson, C. H. Botting, J. H. Naismith, & A. J. Stewart. 2014 Crystal structure of histidine-rich glycoprotein N2 domain reveals redox activity at an interdomain disulfide bridge: implications for angiogenic regulation. *Blood* **123**: 1948-1955 (doi 10.1182/blood-2013-11-535963).
183. Houssen, W. E. et al. 2014 An efficient method for the in vitro production of azol(in)e-based cyclic peptides. *Angew Chem Int Ed Engl* **53**: 14171-14174 (doi 10.1002/anie.201408082).
182. Thompson, S., Q. Zhang, M. Onega, S. McMahon, I. Fleming, S. Ashworth, J. H. Naismith, J. Passchier, & D. O'Hagan. 2014 A Localized Tolerance in the Substrate Specificity of the Fluorinase Enzyme enables "Last-Step" (18) F Fluorination of a RGD Peptide under Ambient Aqueous Conditions. *Angew Chem Int Ed Engl* **53**: 8913-8918 (doi 10.1002/anie.201403345).

2013 (14 publications)

181. Bushell, S. R., I. L. Mainprize, M. A. Wear, H. Lou, C. Whitfield, & J. H. Naismith. 2013 Wzi is an outer membrane lectin that underpins group 1 capsule assembly in Escherichia coli. *Structure* **21**: 844-853 (doi 10.1016/j.str.2013.03.010).
180. Warelow, T. P. et al. 2013 The respiratory arsenite oxidase: structure and the role of residues surrounding the rieske cluster. *PLoS One* **8**: e72535 (doi 10.1371/journal.pone.0072535).
179. Koehnke, J. et al. 2013 The cyanobactin heterocyclase enzyme: a processive adenylase that operates with a defined order of reaction. *Angew Chem Int Ed Engl* **52**: 13991-13996 (doi 10.1002/anie.201306302).
178. Wadsworth, J. M., D. J. Clarke, S. A. McMahon, J. P. Lowther, A. E. Beattie, P. R. Langridge-Smith, H. B. Broughton, T. M. Dunn, J. H. Naismith, & D. J. Campopiano. 2013 The chemical basis of serine palmitoyltransferase inhibition by myriocin. *J Am Chem Soc* **135**: 14276-14285 (doi 10.1021/ja4059876).
177. Moynie, L. et al. 2013 The AEROPATH project targeting Pseudomonas aeruginosa: crystallographic studies for assessment of potential targets in early-stage drug discovery. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **69**: 25-34 (doi 10.1107/S1744309112044739).
176. Reeks, J., S. Graham, L. Anderson, H. Liu, M. F. White, & J. H. Naismith. 2013 Structure of the archaeal Cascade subunit Csa5: relating the small subunits of CRISPR effector complexes. *RNA Biol* **10**: 762-769 (doi 10.4161/rna.23854).
175. Bent, A. F., J. Koehnke, W. E. Houssen, M. C. Smith, M. Jaspars, & J. H. Naismith. 2013 Structure of PatF from Prochloron didemni. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **69**: 618-623 (doi 10.1107/S1744309113012931).
174. Reeks, J., R. D. Sokolowski, S. Graham, H. Liu, J. H. Naismith, & M. F. White. 2013 Structure of a dimeric crenarchaeal Cas6 enzyme with an atypical active site for CRISPR RNA processing. *Biochem J* **452**: 223-230 (doi 10.1042/BJ20130269).
173. Moynie, L., S. M. Leckie, S. A. McMahon, F. G. Duthie, A. Koehnke, J. W. Taylor, M. S. Alphey, R. Brenk, A. D. Smith, & J. H. Naismith. 2013 Structural insights into the mechanism and inhibition of the beta-hydroxydecanoyl-acyl carrier protein dehydratase from Pseudomonas aeruginosa. *J Mol Biol* **425**: 365-377 (doi 10.1016/j.jmb.2012.11.017).
172. Branigan, E., C. Pliotas, G. Hagelueken, & J. H. Naismith. 2013 Quantification of free cysteines in membrane and soluble proteins using a fluorescent dye and thermal unfolding. *Nat Protoc* **8**: 2090-2097 (doi 10.1038/nprot.2013.128).
171. Giannoulis, A., R. Ward, E. Branigan, J. H. Naismith, & B. E. Bode. 2013 PELDOR in rotationally symmetric homo-oligomers. *Mol Phys* **111**: 2845-2854 (doi 10.1080/00268976.2013.798697).

170. Cukier, C. D. et al. 2013 Discovery of an allosteric inhibitor binding site in 3-Oxo-acyl-ACP reductase from Pseudomonas aeruginosa. *ACS Chem Biol* **8**: 2518-2527 (doi 10.1021/cb4005063).
169. Reeks, J., J. H. Naismith, & M. F. White. 2013 CRISPR interference: a structural perspective. *Biochem J* **453**: 155-166 (doi 10.1042/BJ20130316).
168. Koehnke, J. et al. 2013 An enzymatic route to selenazolines. *Chembiochem* **14**: 564-567 (doi 10.1002/cbic.201300037).

2012 (15 publications)

167. Naismith, J. H. 2012 Tryptophan oxygenation: mechanistic considerations. *Biochem Soc Trans* **40**: 509-514 (doi 10.1042/BST20120073).
166. Koehnke, J. et al. 2012 The mechanism of patellamide macrocyclization revealed by the characterization of the PatG macrocyclase domain. *Nat Struct Mol Biol* **19**: 767-772 (doi 10.1038/nsmb.2340).
165. Houssen, W. E., J. Koehnke, D. Zollman, J. Vendome, A. Raab, M. C. Smith, J. H. Naismith, & M. Jaspars. 2012 The discovery of new cyanobactins from Cyanothece PCC 7425 defines a new signature for processing of patellamides. *Chembiochem* **13**: 2683-2689 (doi 10.1002/cbic.201200661).
164. Houssen, W. E., J. Koehnke, D. Zollman, J. Vendome, A. Raab, M. C. M. Smith, J. H. Naismith, & M. Jaspars. 2012 The Discovery of New Cyanobactins from Cyanothece PCC 7425 Defines a New Signature for Processing of Patellamides. *ChemBiochem* **13**: 1-13.
163. Hagelueken, G., H. Huang, B. R. Clarke, T. Lebl, C. Whitfield, & J. H. Naismith. 2012 Structure of WbdD: a bifunctional kinase and methyltransferase that regulates the chain length of the O antigen in Escherichia coli O9a. *Mol Microbiol* **86**: 730-742 (doi 10.1111/mmi.12014).
162. Plechanovova, A., E. G. Jaffray, M. H. Tatham, J. H. Naismith, & R. T. Hay. 2012 Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis. *Nature* **489**: 115-120 (doi 10.1038/nature11376).
161. Zhang, J. et al. 2012 Structure and Mechanism of the CMR Complex for CRISPR-Mediated Antiviral Immunity. *Mol Cell* (doi 10.1016/j.molcel.2011.12.013).
160. Lowther, J., J. H. Naismith, T. M. Dunn, & D. J. Campopiano. 2012 Structural, mechanistic and regulatory studies of serine palmitoyltransferase. *Biochem Soc Trans* **40**: 547-554 (doi 10.1042/BST20110769).
159. Wang, W., T. Rasmussen, A. J. Harding, N. A. Booth, I. R. Booth, & J. H. Naismith. 2012 Salt bridges regulate both dimer formation and monomeric flexibility in HdeB and may have a role in periplasmic chaperone function. *J Mol Biol* **415**: 538-546 (doi 10.1016/j.jmb.2011.11.026).
158. Hagelueken, G., R. Ward, J. H. Naismith, & O. Schiemann. 2012 MtsslWizard: In Silico Spin-Labeling and Generation of Distance Distributions in PyMOL. *Appl Magn Reson* **42**: 377-391 (doi 10.1007/s00723-012-0314-0).
157. Hagelueken, G., H. Huang, K. Harlos, B. R. Clarke, C. Whitfield, & J. H. Naismith. 2012 Crystallization, dehydration and experimental phasing of WbdD, a bifunctional kinase and methyltransferase from Escherichia coli O9a. *Acta Crystallogr D Biol Crystallogr* **68**: 1371-1379 (doi 10.1107/S0907444912029599).
156. Pliotas, C., R. Ward, E. Branigan, A. Rasmussen, G. Hagelueken, H. Huang, S. S. Black, I. R. Booth, O. Schiemann, & J. H. Naismith. 2012 Conformational state of the MscS mechanosensitive channel in solution revealed by pulsed electron-electron double resonance (PELDOR) spectroscopy. *Proc Natl Acad Sci U S A* **109**: E2675-82 (doi 10.1073/pnas.1202286109).
155. Naismith, J. H. & I. R. Booth. 2012 Bacterial mechanosensitive channels--MscS: evolution's solution to creating sensitivity in function. *Annu Rev Biophys* **41**: 157-177 (doi 10.1146/annurev-biophys-101211-113227).
154. Alphey, M. S. et al. 2012 Allosteric Competitive Inhibitors of the Glucose-1-Phosphate Thymidylyltransferase (RmlA) from Pseudomonas aeruginosa. *ACS Chem Biol* (doi 10.1021/cb300426u).

153. Zhu, X., X. Yan, L. G. Carter, H. Liu, S. Graham, P. J. Coote, & J. Naismith. 2012 A model for 3-methyladenine recognition by 3-methyladenine DNA glycosylase I (TAG) from Staphylococcus aureus. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **68**: 610-615 (doi 10.1107/S1744309112016363).

2011 (10 publications)

152. Morris, C. et al. 2011 The Protein Information Management System (PiMS): a generic tool for any structural biology research laboratory. *Acta Crystallogr D Biol Crystallogr* **67**: 249-260 (doi 10.1107/S0907444911007943).
151. Schmelz, S., C. H. Botting, L. Song, N. F. Kadi, G. L. Challis, & J. H. Naismith. 2011 Structural basis for acyl acceptor specificity in the achromobactin biosynthetic enzyme AcsD. *J Mol Biol* **412**: 495-504 (doi 10.1016/j.jmb.2011.07.059).
150. Lintner, N. G. et al. 2011 Structural and functional characterization of an archaeal clustered regularly interspaced short palindromic repeat (CRISPR)-associated complex for antiviral defense (CASCADE). *J Biol Chem* **286**: 21643-21656 (doi 10.1074/jbc.M111.238485).
149. Harrison, J. A., K. P. Kartha, E. J. Fournier, T. L. Lowary, C. Malet, U. J. Nilsson, O. Hindsgaul, S. Schenkman, J. H. Naismith, & R. A. Field. 2011 Probing the acceptor substrate binding site of Trypanosoma cruzi trans-sialidase with systematically modified substrates and glycoside libraries. *Org Biomol Chem* **9**: 1653-1660 (doi 10.1039/c0ob00826e).
148. Plechanovova, A., E. G. Jaffray, S. A. McMahon, K. A. Johnson, I. Navratilova, J. H. Naismith, & R. T. Hay. 2011 Mechanism of ubiquitylation by dimeric RING ligase RNF4. *Nat Struct Mol Biol* **18**: 1052-1059 (doi 10.1038/nsmb.2108).
147. Compton, E. L., E. Karinou, J. H. Naismith, F. Gabel, & A. Javelle. 2011 Low resolution structure of a bacterial SLC26 transporter reveals dimeric stoichiometry and mobile intracellular domains. *J Biol Chem* **286**: 27058-27067 (doi 10.1074/jbc.M111.244533).
146. Paytubi, S., S. A. McMahon, S. Graham, H. Liu, C. H. Botting, K. S. Makarova, E. V. Koonin, J. H. Naismith, & M. F. White. 2011 Displacement of the canonical single-stranded DNA-binding protein in the Thermoproteales. *Proc Natl Acad Sci U S A* (doi 10.1073/pnas.1113277108).
145. Lang, A., S. Polnick, T. Nicke, P. William, E. P. Patallo, J. H. Naismith, & K. H. van Pee. 2011 Changing the regioselectivity of the tryptophan 7-halogenase PrnA by site-directed mutagenesis. *Angew Chem Int Ed Engl* **50**: 2951-2953 (doi 10.1002/anie.201007896).
144. Lou, H. et al. 2011 Altered antibiotic transport in OmpC mutants isolated from a series of clinical strains of multi-drug resistant E. coli. *PLoS One* **6**: e25825 (doi 10.1371/journal.pone.0025825).
143. Wilfong, E. M., Y. Kogiso, S. Muthukrishnan, T. Kowatz, Y. Du, A. Bowie, J. H. Naismith, C. M. Hadad, E. J. Toone, & T. L. Gustafson. 2011 A multidisciplinary approach to probing enthalpy-entropy compensation and the interfacial mobility model. *J Am Chem Soc* **133**: 11515-11523 (doi 10.1021/ja1098287).

2010 (10 publications)

142. Zhu, X., K. H. van Pee, & J. H. Naismith. 2010 The ternary complex of PrnB (the second enzyme in the pyrrolnitrin biosynthesis pathway), tryptophan, and cyanide yields new mechanistic insights into the indolamine dioxygenase superfamily. *J Biol Chem* **285**: 21126-21133 (doi 10.1074/jbc.M110.120485).
141. Raman, M. C., K. A. Johnson, D. J. Clarke, J. H. Naismith, & D. J. Campopiano. 2010 The serine palmitoyltransferase from Sphingomonas wittichii RW1: An interesting link to an unusual acyl carrier protein. *Biopolymers* **93**: 811-822 (doi 10.1002/bip.21482).
140. Oke, M. et al. 2010 The Scottish Structural Proteomics Facility: targets, methods and outputs. *J Struct Funct Genomics* **11**: 167-180 (doi 10.1007/s10969-010-9090-y).
139. Kowatz, T., J. P. Morrison, M. E. Tanner, & J. H. Naismith. 2010 The crystal structure of the Y140F mutant of ADP-L-glycero-D-manno-heptose 6-epimerase bound to ADP-beta-D-mannose suggests a one base mechanism. *Protein Sci* **19**: 1337-1343 (doi 10.1002/pro.410).

138. Steiner, K., G. Hagelueken, P. Messner, C. Schaffer, & J. H. Naismith. 2010 Structural basis of substrate binding in WsaF, a rhamnosyltransferase from Geobacillus stearothermophilus. *J Mol Biol* **397**: 436-447 (doi 10.1016/j.jmb.2010.01.035).
137. Hagan, R. M., R. Bjornsson, S. A. McMahon, B. Schomburg, V. Braithwaite, M. Buhl, J. H. Naismith, & U. Schwarz-Linek. 2010 NMR spectroscopic and theoretical analysis of a spontaneously formed Lys-Asp isopeptide bond. *Angew Chem Int Ed Engl* **49**: 8421-8425 (doi 10.1002/anie.201004340).
136. Lowther, J. et al. 2010 Inhibition of the PLP-dependent enzyme serine palmitoyltransferase by cycloserine: evidence for a novel decarboxylative mechanism of inactivation. *Mol Biosyst* **6**: 1682-1693 (doi 10.1039/c003743e).
135. Schmidberger, J. W., A. B. James, R. Edwards, J. H. Naismith, & D. O'Hagan. 2010 Halomethane biosynthesis: structure of a SAM-dependent halide methyltransferase from Arabidopsis thaliana. *Angew Chem Int Ed Engl* **49**: 3646-3648 (doi 10.1002/anie.201000119).
134. Bushell, S. R., H. Lou, G. D. Wallat, K. Beis, C. Whitfield, & J. H. Naismith. 2010 Crystallization and preliminary diffraction analysis of Wzi, a member of the capsule export and assembly pathway in Escherichia coli. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **66**: 1621-1625 (doi 10.1107/S1744309110040546).
133. Oke, M., M. Kerou, H. Liu, X. Peng, R. A. Garrett, D. Prangishvili, J. H. Naismith, & M. F. White. 2010 A dimeric Rep protein initiates replication of a linear archaeal virus genome: implications for Rep mechanism and viral replication. *J Virol* (doi 10.1128/JVI.01467-10).

2009 (15 publications)

132. Raman, M. C., K. A. Johnson, B. A. Yard, J. Lowther, L. G. Carter, J. H. Naismith, & D. J. Campopiano. 2009 The external aldimine form of serine palmitoyltransferase: structural, kinetic, and spectroscopic analysis of the wild-type enzyme and HSAN1 mutant mimics. *J Biol Chem* **284**: 17328-17339 (doi 10.1074/jbc.M109.008680).
131. Errey, J. C., M. C. Mann, S. A. Fairhurst, L. Hill, M. R. McNeil, J. H. Naismith, J. M. Percy, C. Whitfield, & R. A. Field. 2009 Sugar nucleotide recognition by Klebsiella pneumoniae UDP-D-galactopyranose mutase: fluorinated substrates, kinetics and equilibria. *Org Biomol Chem* **7**: 1009-1016 (doi 10.1039/b815549f).
130. Ford, R. C., A. L. Brunkan-LaMontagne, R. F. Collins, B. R. Clarke, R. Harris, J. H. Naismith, & C. Whitfield. 2009 Structure-function relationships of the outer membrane translocon Wza investigated by cryo-electron microscopy and mutagenesis. *J Struct Biol* **166**: 172-182 (doi 10.1016/j.jsb.2009.02.005).
129. Zhu, X., W. De Laurentis, K. Leang, J. Herrmann, K. Ihlefeld, K. H. van Pee, & J. H. Naismith. 2009 Structural insights into regioselectivity in the enzymatic chlorination of tryptophan. *J Mol Biol* **391**: 74-85 (doi 10.1016/j.jmb.2009.06.008).
128. McRobbie, A. M., L. G. Carter, M. Kerou, H. Liu, S. A. McMahon, K. A. Johnson, M. Oke, J. H. Naismith, & M. F. White. 2009 Structural and functional characterisation of a conserved archaeal RadA paralog with antirecombinase activity. *J Mol Biol* **389**: 661-673 (doi 10.1016/j.jmb.2009.04.060).
127. Cuthbertson, L., I. L. Mainprize, J. H. Naismith, & C. Whitfield. 2009 Pivotal roles of the outer membrane polysaccharide export and polysaccharide copolymerase protein families in export of extracellular polysaccharides in gram-negative bacteria. *Microbiol Mol Biol Rev* **73**: 155-177 (doi 10.1128/MMBR.00024-08).
126. Hagelueken, G., W. J. Ingledew, H. Huang, B. Petrovic-Stojanovska, C. Whitfield, H. ElMkami, O. Schiemann, & J. H. Naismith. 2009 PELDOR spectroscopy distance fingerprinting of the octameric outer-membrane protein Wza from Escherichia coli. *Angew Chem Int Ed Engl* **48**: 2904-2906 (doi 10.1002/anie.200805758).
125. Deng, H., S. A. McMahon, A. S. Eustaquio, B. S. Moore, J. H. Naismith, & D. O'Hagan. 2009 Mechanistic insights into water activation in SAM hydroxide adenosyltransferase (duf-62). *Chembiochem* **10**: 2455-2459 (doi 10.1002/cbic.200900369).

124. McMahon, S. A. et al. 2009 Extensive DNA mimicry by the ArdA anti-restriction protein and its role in the spread of antibiotic resistance. *Nucleic Acids Res* **37**: 4887-4897 (doi 10.1093/nar/gkp478).
123. Huang, H., G. Hagelueken, C. Whitfield, & J. H. Naismith. 2009 Crystallization and preliminary crystallographic analysis of the bacterial capsule assembly-regulating tyrosine phosphatases Wzb of Escherichia coli and Cps4B of Streptococcus pneumoniae. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **65**: 770-772 (doi 10.1107/S1744309109023914).
122. Hagelueken, G., H. Huang, I. L. Mainprize, C. Whitfield, & J. H. Naismith. 2009 Crystal structures of Wzb of Escherichia coli and CpsB of Streptococcus pneumoniae, representatives of two families of tyrosine phosphatases that regulate capsule assembly. *J Mol Biol* **392**: 678-688 (doi 10.1016/j.jmb.2009.07.026).
121. Rejzek, M., V. Sri Kannathasan, C. Wing, A. Preston, E. L. Westman, J. S. Lam, J. H. Naismith, D. J. Maskell, & R. A. Field. 2009 Chemical synthesis of UDP-Glc-2,3-diNAcA, a key intermediate in cell surface polysaccharide biosynthesis in the human respiratory pathogens B. pertussis and P. aeruginosa. *Org Biomol Chem* **7**: 1203-1210 (doi 10.1039/b819607a).
120. Schmelz, S. & J. H. Naismith. 2009 Adenylate-forming enzymes. *Curr Opin Struct Biol* **19**: 666-671 (doi 10.1016/j.sbi.2009.09.004).
119. Schmelz, S. et al. 2009 AcsD catalyzes enantioselective citrate desymmetrization in siderophore biosynthesis. *Nat Chem Biol* **5**: 174-182 (doi 10.1038/nchembio.145).
118. Liu, H. & J. H. Naismith. 2009 A simple and efficient expression and purification system using two newly constructed vectors. *Protein Expr Purif* **63**: 102-111 (doi 10.1016/j.pep.2008.09.008).

2008 (15 publications)

117. Oke, M. et al. 2008 Unusual chromophore and cross-links in ranasmurfin: a blue protein from the foam nests of a tropical frog. *Angew Chem Int Ed Engl* **47**: 7853-7856 (doi 10.1002/anie.200802901).
116. Wang, W., S. S. Black, M. D. Edwards, S. Miller, E. L. Morrison, W. Bartlett, C. Dong, J. H. Naismith, & I. R. Booth. 2008 The structure of an open form of an E. coli mechanosensitive channel at 3.45 A resolution. *Science* **321**: 1179-1183 (doi 10.1126/science.1159262).
115. Wang, W., C. Dong, M. McNeil, D. Kaur, S. Mahapatra, D. C. Crick, & J. H. Naismith. 2008 The structural basis of chain length control in Rv1086. *J Mol Biol* **381**: 129-140 (doi 10.1016/j.jmb.2008.05.060).
114. Overton, I. M. et al. 2008 TarO: a target optimisation system for structural biology. *Nucleic Acids Res* **36**: W190-6 (doi 10.1093/nar/gkn141).
113. Liu, H., J. Rudolf, K. A. Johnson, S. A. McMahon, M. Oke, L. Carter, A. M. McRobbie, S. E. Brown, J. H. Naismith, & M. F. White. 2008 Structure of the DNA repair helicase XPD. *Cell* **133**: 801-812 (doi 10.1016/j.cell.2008.04.029).
112. Richards, J. D., K. A. Johnson, H. Liu, A. M. McRobbie, S. McMahon, M. Oke, L. Carter, J. H. Naismith, & M. F. White. 2008 Structure of the DNA repair helicase hel308 reveals DNA binding and autoinhibitory domains. *J Biol Chem* **283**: 5118-5126 (doi 10.1074/jbc.M707548200).
111. Steiner, K., A. Wojciechowska, C. Schaffer, & J. H. Naismith. 2008 Purification, crystallization and preliminary crystallographic analysis of WsaF, an essential rhamnosyltransferase from Geobacillus stearothermophilus. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **64**: 1163-1165 (doi 10.1107/S1744309108036762).
110. McMahon, S. A., M. Oke, H. Liu, K. A. Johnson, L. Carter, N. Kadi, M. F. White, G. L. Challis, & J. H. Naismith. 2008 Purification, crystallization and data collection of Pectobacterium chrysanthemi AcsD, a type A siderophore synthetase. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **64**: 1052-1055 (doi 10.1107/S1744309108032132).
109. McMahon, S. A., M. Oke, H. Liu, K. A. Johnson, L. Carter, N. Kadi, M. F. White, G. L. Challis, & J. H. Naismith. 2008 Purification, crystallization and data collection of Pectobacterium

chrysanthemi AcsD, a type A siderophore synthetase. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **64**: 1052-1055 (doi 10.1107/S1744309108032132).

108. Whitfield, C. & J. H. Naismith. 2008 Periplasmic export machines for outer membrane assembly. *Curr Opin Struct Biol* **18**: 466-474 (doi 10.1016/j.sbi.2008.04.001).
107. Flecks, S., E. P. Patallo, X. Zhu, A. J. Ernyei, G. Seifert, A. Schneider, C. Dong, J. H. Naismith, & K. H. van Pee. 2008 New insights into the mechanism of enzymatic chlorination of tryptophan. *Angew Chem Int Ed Engl* **47**: 9533-9536 (doi 10.1002/anie.200802466).
106. Fairhead, M., K. A. Johnson, T. Kowatz, S. A. McMahon, L. G. Carter, M. Oke, H. Liu, J. H. Naismith, & C. F. van der Walle. 2008 Crystal structure and silica condensing activities of silicatein alpha-cathepsin L chimeras. *Chem Commun (Camb)* 1765-1767 (doi 10.1039/b718264c).
105. Davies, G. J. & J. H. Naismith. 2008 Chemistry's prodigal child: enzyme mechanism. *Curr Opin Chem Biol* **12**: 529-531 (doi 10.1016/j.cbpa.2008.08.018).
104. Walters, M. J., V. Srikannathasan, A. R. McEwan, J. H. Naismith, C. A. Fierke, & E. J. Toone. 2008 Characterization and crystal structure of Escherichia coli KDPGal aldolase. *Bioorg Med Chem* **16**: 710-720 (doi 10.1016/j.bmc.2007.10.043).
103. Liu, H. & J. H. Naismith. 2008 An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol* **8**: 91 (doi 10.1186/1472-6750-8-91).

2007 (8 publications)

102. Yard, B. A. et al. 2007 The structure of serine palmitoyltransferase; gateway to sphingolipid biosynthesis. *J Mol Biol* **370**: 870-886 (doi 10.1016/j.jmb.2007.04.086).
101. De Laurentis, W., L. Khim, J. L. Anderson, A. Adam, K. A. Johnson, R. S. Phillips, S. K. Chapman, K. H. van Pee, & J. H. Naismith. 2007 The second enzyme in pyrrolnitrin biosynthetic pathway is related to the heme-dependent dioxygenase superfamily. *Biochemistry* **46**: 12393-12404 (doi 10.1021/bi7012189).
100. Collins, R. F., K. Beis, C. Dong, C. H. Botting, C. McDonnell, R. C. Ford, B. R. Clarke, C. Whitfield, & J. H. Naismith. 2007 The 3D structure of a periplasm-spanning platform required for assembly of group 1 capsular polysaccharides in *Escherichia coli*. *Proceedings of the National Academy of Sciences USA* **104**: 2390-2395 (doi 10.1073/pnas.0607763104).
99. Liu, H., K. Woznica, G. Catton, A. Crawford, N. Botting, & J. H. Naismith. 2007 Structural and kinetic characterization of quinolinate phosphoribosyltransferase (hQPRTase) from homo sapiens. *J Mol Biol* **373**: 755-763 (doi 10.1016/j.jmb.2007.08.043).
98. Dong, C. et al. 2007 RmlC, a C3' and C5' carbohydrate epimerase, appears to operate via an intermediate with an unusual twist boat conformation. *J Mol Biol* **365**: 146-159 (doi 10.1016/j.jmb.2006.09.063).
97. Seetharamappa, J. et al. 2007 Purification, crystallization and data collection of methicillin-resistant Staphylococcus aureus Sar2676, a pantothenate synthetase. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **63**: 488-491 (doi 10.1107/S1744309107020362).
96. Zhu, X., D. A. Robinson, A. R. McEwan, D. O'Hagan, & J. H. Naismith. 2007 Mechanism of enzymatic fluorination in Streptomyces cattleya. *J Am Chem Soc* **129**: 14597-14604 (doi 10.1021/ja0731569).
95. Seetharamappa, J. et al. 2007 Expression, purification, crystallization, data collection and preliminary biochemical characterization of methicillin-resistant Staphylococcus aureus Sar2028, an aspartate/tyrosine/phenylalanine pyridoxal-5'-phosphate-dependent aminotransferase. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **63**: 452-456 (doi 10.1107/S1744309107019562).

2006 (15 publications)

94. Dong, C. J., K. Beis, J. Nesper, A. L. Brunkan-LaMontagne, B. R. Clarke, C. Whitfield, & J. H. Naismith. 2006 Wza the translocon for E-coli capsular polysaccharides defines a new class of membrane protein. *Nature* **444**: 226-229 (doi 10.1038/nature05267).

93. Beis, K., C. Whitfield, I. R. Booth, & J. H. Naismith. 2006 Two-step purification of outer membrane proteins. *International Journal of Biological Macromolecules* **39**: 10-14 (doi 10.1016/j.ijbiomac.2005.12.008).
92. Shen, L. N., C. Dong, H. Liu, J. H. Naismith, & R. T. Hay. 2006 The structure of SENP1-SUMO-2 complex suggests a structural basis for discrimination between SUMO paralogues during processing. *Biochem J* **397**: 279-288 (doi 10.1042/BJ20052030).
91. Deng, H., S. L. Cobb, A. R. McEwan, R. P. McGlinchey, J. H. Naismith, D. O'Hagan, D. A. Robinson, & J. B. Spencer. 2006 The fluorinase from Streptomyces cattleya is also a chlorinase. *Angewandte Chemie-International Edition* **45**: 759-762 (doi 10.1002/anie.200503582).
90. Shen, L., M. H. Tatham, C. Dong, A. Zagorska, J. H. Naismith, & R. T. Hay. 2006 SUMO protease SENP1 induces isomerization of the scissile peptide bond. *Nat Struct Mol Biol* **13**: 1069-1077 (doi 10.1038/nsmb1172).
89. Cobb, S. L., H. Deng, A. R. McEwan, J. H. Naismith, D. O'Hagan, & D. A. Robinson. 2006 Substrate specificity in enzymatic fluorination. The fluorinase from Streptomyces cattleya accepts 2'-deoxyadenosine substrates. *Org Biomol Chem* **4**: 1458-1460 (doi 10.1039/b600574h).
88. Cobb, S. L., H. Deng, A. R. McEwan, J. H. Naismith, D. O'Hagan, & D. A. Robinson. 2006 Substrate specificity in enzymatic fluorination. The fluorinase from Streptomyces cattleya accepts 2 '-deoxyadenosine substrates. *Organic & Biomolecular Chemistry* **4**: 1458-1460. (doi 10.1039/B600574H)
87. Williams, G. J., K. Johnson, J. Rudolf, S. A. McMahon, L. Carter, M. Oke, H. T. Liu, G. L. Taylor, M. F. White, & J. H. Naismith. 2006 Structure of the heterotrimeric PCNA from Sulfolobus solfataricus. *Acta Crystallographica Section F-Structural Biology and Crystallization Communications* **62**: 944-948 (doi 10.1107/S1744309106034075).
86. De Laurentis, W., K. Leang, K. Hahn, B. Podemski, A. Adam, S. Kroschwald, L. G. Carter, K. H. van Pee, & J. H. Naismith. 2006 Preliminary crystallographic characterization of PrnB, the second enzyme in the pyrrolnitrin biosynthetic pathway. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **62**: 1134-1137 (doi 10.1107/S1744309106041649).
85. Collins, R. F., K. Beis, B. R. Clarke, R. C. Ford, M. Hulley, J. H. Naismith, & C. Whitfield. 2006 Periplasmic protein-protein contacts in the inner membrane protein Wzc form a tetrameric complex required for the assembly of Escherichia coli group 1 capsules. *Journal of Biological Chemistry* **281**: 2144-2150 (doi 10.1074/jbc.M508078200).
84. McEwan, A. R., H. Liu, M. Oke, L. Carter, H. Powers, M. Dorward, S. A. McMahon, M. F. White, & J. H. Naismith. 2006 Overexpression, purification, crystallization and data collection of Sulfolobus solfataricus Sso6206, a novel highly conserved protein. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **62**: 228-230 (doi 10.1107/S1744309106003654).
83. Fullerton, S. W., J. S. Griffiths, A. B. Merkel, M. Cheriyan, N. J. Wymer, M. J. Hutchins, C. A. Fierke, E. J. Toone, & J. H. Naismith. 2006 Mechanism of the Class I KDPG aldolase. *Bioorg Med Chem* **14**: 3002-3010 (doi 10.1016/j.bmc.2005.12.022).
82. Naismith, J. H. 2006 Inferring the chemical mechanism from structures of enzymes. *Chem Soc Rev* **35**: 763-770 (doi 10.1039/b602313d).
81. McMahon, S. A. et al. 2006 Crystallization of Ranasmurfin, a blue-coloured protein from Polypedates leucomystax. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **62**: 1124-1126 (doi 10.1107/S1744309106040036).
80. van Pee, K. H., C. Dong, S. Flecks, J. Naismith, E. P. Patallo, & T. Wage. 2006 Biological halogenation has moved far beyond haloperoxidases. *Adv Appl Microbiol* **59**: 127-157 (doi 10.1016/S0065-2164(06)59005-7).

2005 (6 publications)

79. Dong, C., S. Flecks, S. Unversucht, C. Haupt, K. H. van Pee, & J. H. Naismith. 2005 Tryptophan 7-halogenase (PrnA) structure suggests a mechanism for regioselective chlorination. *Science* **309**: 2216-2219 (doi 10.1126/science.1116510).

78. Major, L. L., B. A. Wolucka, & J. H. Naismith. 2005 Structure and function of GDP-mannose-3 ‘,5 ‘-epimerase: An enzyme which performs three chemical reactions at the same active site. *Journal of the American Chemical Society* **127**: 18309-18320 (doi 10.1021/ja056490i).
77. Williams, G. J., S. D. Breazeale, C. R. H. Raetz, & J. H. Naismith. 2005 Structure and function of both domains of ArnA, a dual function decarboxylase and a formyltransferase, involved in 4-amino-4deoxy-L-arabinose biosynthesis. *The Journal of Biological Chemistry* **280**: 23000-23008 (doi 10.1074/jbc.M501534200).
76. Williams, G. J., S. D. Breazeale, C. R. H. Raetz, & J. H. Naismith. 2005 Structure and function of both domains of ArnA, a dual function decarboxylase and a formyltransferase, involved in 4-amino-4-deoxy-L-arabinose biosynthesis. *Journal of Biological Chemistry* **280**: 23000-23008 (doi 10.1074/jbc.M501534200).
75. Shen, L. N., H. Liu, C. Dong, D. Xirodimas, J. H. Naismith, & R. T. Hay. 2005 Structural basis of NEDD8 ubiquitin discrimination by the deNEDDylating enzyme NEDP1. *EMBO J* **24**: 1341-1351 (doi 10.1038/sj.emboj.7600628).
74. Beis, K., V. Srikannathasan, H. Liu, S. W. Fullerton, V. A. Bamford, D. A. Sanders, C. Whitfield, M. R. McNeil, & J. H. Naismith. 2005 Crystal structures of Mycobacteria tuberculosis and Klebsiella pneumoniae UDP-galactopyranose mutase in the oxidised state and Klebsiella pneumoniae UDP-galactopyranose mutase in the (active) reduced state. *J Mol Biol* **348**: 971-982 (doi 10.1016/j.jmb.2005.02.057).

2004(7 publications)

73. Beis, K., R. F. Collins, R. C. Ford, A. B. Kamis, C. Whitfield, & J. H. Naismith. 2004 Three-dimensional structure of Wza, the protein required for translocation of group 1 capsular polysaccharide across the outer membrane of Escherichia coli. *Journal of Biological Chemistry* **279**: 28227-28232 (doi 10.1074/jbc.M402913200).
72. Merkel, A. B., L. L. Major, J. C. Errey, M. D. Burkart, R. A. Field, C. T. Walsh, & J. N. Naismith. 2004 The position of a key tyrosine in dTDP-4-Keto-6-deoxy-D-glucose-5-epimerase (EvaD) alters the substrate profile for this RmlC-like enzyme. *Journal of Biological Chemistry* **279**: 32684-32691 (doi 10.1074/jbc.M404091200).
71. Challis, G. L. & J. H. Naismith. 2004 Structural aspects of non-ribosomal peptide biosynthesis. *Curr Opin Struct Biol* **14**: 748-756 (doi 10.1016/j.sbi.2004.10.005).
70. Dong, C., A. Kotzsch, M. Dorward, K. H. van Pee, & J. H. Naismith. 2004 Crystallization and X-ray diffraction of a halogenating enzyme, tryptophan 7-halogenase, from Pseudomonas fluorescens. *Acta Crystallogr D Biol Crystallogr* **60**: 1438-1440 (doi 10.1107/S0907444904012521).
69. Beis, K., J. Nesper, C. Whitfield, & J. H. Naismith. 2004 Crystallization and preliminary X-ray diffraction analysis of Wza outer-membrane lipoprotein from Escherichia coli serotype O9a:K30. *Acta Crystallogr D Biol Crystallogr* **60**: 558-560 (doi 10.1107/S0907444903029494).
68. Dong, C. J., F. L. Huang, H. Deng, C. Schaffrath, J. B. Spencer, D. O’Hagan, & J. H. Naismith. 2004 Crystal structure and mechanism of a bacterial fluorinating enzyme. *Nature* **427**: 561-565 (doi 10.1038/nature02280).
67. Naismith, J. H. 2004 Chemical insights from structural studies of enzymes. *Biochem Soc Trans* **32**: 647-654 (doi 10.1042/BST0320647).

2003 (14 publications)

66. Nesper, J., C. M. Hill, A. Paiment, G. Harauz, K. Beis, J. H. Naismith, & C. Whitfield. 2003 Translocation of group 1 capsular polysaccharide in Escherichia coli serotype K30. Structural and functional analysis of the outer membrane lipoprotein Wza. *J Biol Chem* **278**: 49763-49772 (doi 10.1074/jbc.M308775200).
65. Beis, K., S. T. Allard, A. D. Hegeman, G. Murshudov, D. Philp, & J. H. Naismith. 2003 The structure of NADH in the enzyme dTDP-d-glucose dehydratase (RmlB). *J Am Chem Soc* **125**: 11872-11878 (doi 10.1021/ja035796r).
64. Field, R. A. & J. H. Naismith. 2003 Structural and mechanistic basis of bacterial sugar nucleotide-modifying enzymes. *Biochemistry* **42**: 7637-7647 (doi 10.1021/bi0345079).

63. Fullerton, S. W., S. Daff, D. A. Sanders, W. J. Ingledew, C. Whitfield, S. K. Chapman, & J. H. Naismith. 2003 Potentiometric analysis of UDP-galactopyranose mutase: stabilization of the flavosemiquinone by substrate. *Biochemistry* **42**: 2104-2109 (doi 10.1021/bi027077f).
62. Babaoglu, K., M. A. Page, V. C. Jones, M. R. McNeil, C. Dong, J. H. Naismith, & R. E. Lee. 2003 Novel inhibitors of an emerging target in Mycobacterium tuberculosis; substituted thiazolidinones as inhibitors of dTDP-rhamnose synthesis. *Bioorg Med Chem Lett* **13**: 3227-3230. (doi 10.1016/s0960-894x(03)00673-5)
61. Babaoglu, K., M. A. Page, V. C. Jones, M. R. McNeil, C. J. Dong, J. H. Naismith, & R. E. Lee. 2003 Novel inhibitors of an emerging target in Mycobacterium tuberculosis; Substituted thiazolidinones as inhibitors of dTDP-rhamnose synthesis. *Bioorganic & Medicinal Chemistry Letters* **13**: 3227-3230.
60. Kerr, I. D., R. I. M. Wadsworth, L. Cubbedu, W. Blankenfeldt, J. H. Naismith, & M. F. White. 2003 Insights into ssDNA recognition by the OB fold from a structural and thermodynamic study of Sulfolobus SSB. *Embo Journal* **22**: 2561-2570 (doi 10.1093/emboj/cdg272).
59. Dong, C., L. L. Major, A. Allen, W. Blankenfeldt, D. Maskell, & J. H. Naismith. 2003 High-resolution structures of RmlC from Streptococcus suis in complex with substrate analogs locate the active site of this class of enzyme. *Structure* **11**: 715-723.
58. Dong, C., H. Deng, M. Dorward, C. Schaffrath, D. O'Hagan, & J. H. Naismith. 2003 Crystallization and X-ray diffraction of 5'-fluoro-5'-deoxyadenosine synthase, a fluorination enzyme from Streptomyces cattleya. *Acta Crystallogr D Biol Crystallogr* **59**: 2292-2293.
57. Dong, C. J., H. Deng, M. Dorward, C. Schaffrath, D. O'Hagan, & J. H. Naismith. 2003 Crystallization and X-ray diffraction of 5 '-fluoro-5 '-deoxyadenosine synthase, a fluorination enzyme from Streptomyces cattleya. *Acta Crystallographica Section D-Biological Crystallography* **59**: 2292-2293 (doi 10.1107/S0907444903019826).
56. Liu, H., J. H. Naismith, & R. T. Hay. 2003 Adenovirus DNA replication. *Curr Top Microbiol Immunol* **272**: 131-164.
55. Dong, C., K. Beis, M. F. Giraud, W. Blankenfeldt, S. Allard, L. L. Major, I. D. Kerr, C. Whitfield, & J. H. Naismith. 2003 A structural perspective on the enzymes that convert dTDP-D-glucose into dTDP-L-rhamnose. *Biochemical Society Transactions* **31**: 532-536 (doi 10.1042/BST0310532).
54. Ramsay, R. R. & J. H. Naismith. 2003 A snapshot of carnitine acetyltransferase. *Trends Biochem Sci* **28**: 343-346 (doi 10.1016/S0968-0004(03)00137-3).
53. Rahn, A., K. Beis, J. H. Naismith, & C. Whitfield. 2003 A novel outer membrane protein, Wzi, is involved in surface assembly of the Escherichia coli K30 group 1 capsule. *J Bacteriol* **185**: 5882-5890 (doi 10.1128/JB.185.19.5882-5890.2003).

2002 (6 publications)

52. Blankenfeldt, W., I. D. Kerr, M.-F. Giraud, H. J. McMiken, G. Leonard, C. Whitfield, P. Messner, M. Graninger, & J. H. Naismith. 2002 Variation on a theme of SDR: dTDP-6-Deoxy-L-*lyxo*-4-hexulose reductase (RmlD) shows a new Mg^{2+}-dependent dimerization mode. *Structure* **10**: 773-786 (doi 10.1016/S0969-2126(02)00770-0).
51. Blankenfeldt, W., I. D. Kerr, M. F. Giraud, H. J. McMiken, G. Leonard, C. Whitfield, P. Messner, M. Graninger, & J. H. Naismith. 2002 Variation on a theme of SDR. dTDP-6-deoxy-L- lyxo-4-hexulose reductase (RmlD) shows a new Mg2+-dependent dimerization mode. *Structure (Camb)* **10**: 773-786 (doi 10.1016/S0969-2126(02)00770-0).
50. Allard, S. T. et al. 2002 Toward a structural understanding of the dehydratase mechanism. *Structure (Camb)* **10**: 81-92 (doi 10.1016/S0969-2126(01)00694-3).
49. Asuncion, M., W. Blankenfeldt, J. N. Barlow, D. Gani, & J. H. Naismith. 2002 The structure of 3-methylaspartase from Clostridium tetanomorphum functions via the common enolase chemical step. *J Biol Chem* **277**: 8306-8311 (doi 10.1074/jbc.M111180200).
48. Merkel, A. B., G. K. Temple, M. D. Burkart, H. C. Losey, K. Beis, C. T. Walsh, & J. H. Naismith. 2002 Purification, crystallization and preliminary structural studies of dTDP-4-keto-6-deoxy-glucose-5-epimerase (EvaD) from Amycolatopsis orientalis, the fourth enzyme in the

dTDP-L-epivancosamine biosynthetic pathway. *Acta Crystallogr D Biol Crystallogr* **58**: 1226-1228 (doi 10.1107/S0907444902007382).

47. Kannathasan, V. S., C. J. Dong, C. Wing, A. Preston, D. J. Maskell, J. H. Naismith, & R. A. Field. 2002 Biochemical and Structural Analysis of the Role of the Wlb Gene Locus in Bordetella Pertussis Lipopolysaccharide Biosynthesis. *ScientificWorldJournal* **2**: 55-56 (doi 10.1100/tsw.2002.28).

2001 (11 publications)

46. Sanders, D. A., A. G. Staines, S. A. McMahon, M. R. McNeil, C. Whitfield, & J. H. Naismith. 2001 UDP-galactopyranose mutase has a novel structure and mechanism. *Nat Struct Biol* **8**: 858-863 (doi 10.1038/nsb1001-858).
45. Allard, S. T., M. F. Giraud, C. Whitfield, M. Graninger, P. Messner, & J. H. Naismith. 2001 The crystal structure of dTDP-D-Glucose 4,6-dehydratase (RmlB) from Salmonella enterica serovar Typhimurium, the second enzyme in the dTDP-l-rhamnose pathway. *J Mol Biol* **307**: 283-295 (doi 10.1006/jmbi.2000.4470).
44. Sanders, D. A., D. N. Moothoo, J. Raftery, A. J. Howard, J. R. Helliwell, & J. H. Naismith. 2001 The 1.2 A resolution structure of the Con A-dimannose complex. *J Mol Biol* **310**: 875-884 (doi 10.1006/jmbi.2001.4806).
43. Tatham, M. H., E. Jaffray, O. A. Vaughan, J. M. Desterro, C. H. Botting, J. H. Naismith, & R. T. Hay. 2001 Polymeric chains of SUMO-2 and SUMO-3 are conjugated to protein substrates by SAE1/SAE2 and Ubc9. *J Biol Chem* **276**: 35368-35374 (doi 10.1074/jbc.M104214200).
42. Sri Kannathasan, V., A. G. Staines, C. J. Dong, R. A. Field, A. G. Preston, D. J. Maskell, & J. H. Naismith. 2001 Overexpression, purification, crystallization and data collection on the Bordetella pertussis wlbD gene product, a putative UDP-GlcNAc 2'-epimerase. *Acta Crystallogr D Biol Crystallogr* **57**: 1310-1312 (doi 10.1107/S0907444901010885).
41. Kerr, I. D., R. I. Wadsworth, W. Blankenfeldt, A. G. Staines, M. F. White, & J. H. Naismith. 2001 Overexpression, purification, crystallization and data collection of a single-stranded DNA-binding protein from Sulfolobus solfataricus. *Acta Crystallogr D Biol Crystallogr* **57**: 1290-1292 (doi 10.1107/S0907444901011374).
40. Asuncion, M., J. N. Barlow, J. Pollard, A. G. Staines, S. A. McMahon, W. Blankenfeldt, D. Gani, & J. H. Naismith. 2001 Overexpression, purification, crystallization and data collection of 3-methylaspartase from Clostridium tetanomorphum. *Acta Crystallogr D Biol Crystallogr* **57**: 731-733 (doi 10.1107/S0907444901003225).
39. Sanders, D. A., S. A. McMahon, G. L. Leonard, & J. H. Naismith. 2001 Molecular placement of experimental electron density: a case study on UDP-galactopyranose mutase. *Acta Crystallogr D Biol Crystallogr* **57**: 1415-1420 (doi 10.1107/S0907444901009829).
38. Harrison, J. A., K. P. Kartha, W. B. Turnbull, S. L. Scheuerl, J. H. Naismith, S. Schenkman, & R. A. Field. 2001 Hydrolase and sialyltransferase activities of trypanosoma cruzi trans-sialidase towards NeuAc-alpha-2,3-gal-Gal-beta-O-PNP. *Bioorg Med Chem Lett* **11**: 141-144 (doi 10.1016/S0960-894X(00)00611-9).
37. Allard, S. T., M. F. Giraud, & J. H. Naismith. 2001 Epimerases: structure, function and mechanism. *Cell Mol Life Sci* **58**: 1650-1665.
36. Wymer, N., L. V. Buchanan, D. Henderson, N. Mehta, C. H. Botting, L. Pocivavsek, C. A. Fierke, E. J. Toone, & J. H. Naismith. 2001 Directed evolution of a new catalytic site in 2-keto-3-deoxy-6-phosphogluconate aldolase from Escherichia coli. *Structure (Camb)* **9**: 1-9 (doi 10.1016/S0969-2126(00)00555-4).

2000 (8 publications)

35. Bryce, R. A., I. H. Hillier, & J. H. Naismith. 2001 Carbohydrate-protein recognition: molecular dynamics simulations and free energy analysis of oligosaccharide binding to concanavalin A. *Biophys J* **81**: 1373-1388 (doi 10.1016/S0006-3495(01)75793-1).
34. Idriss, H. T. & J. H. Naismith. 2000 TNF alpha and the TNF receptor superfamily: structure-function relationship(s). *Microsc Res Tech* **50**: 184-195.

33. Blankenfeldt, W., M. Asuncion, J. S. Lam, & J. H. Naismith. 2000 The structural basis of the catalytic mechanism and regulation of glucose-1-phosphate thymidylyltransferase (RmlA). *Embo J* **19**: 6652-6663 (doi 10.1093/emboj/19.24.6652).
32. Giraud, M. F. & J. H. Naismith. 2000 The rhamnose pathway. *Curr Opin Struct Biol* **10**: 687-696 (doi 10.1016/S0959-440X(00)00145-7).
31. Allard, S. T., M. F. Giraud, C. Whitfield, P. Messner, & J. H. Naismith. 2000 The purification, crystallization and structural elucidation of dTDP-D-glucose 4,6-dehydratase (RmlB), the second enzyme of the dTDP-L-rhamnose synthesis pathway from Salmonella enterica serovar typhimurium. *Acta Crystallogr D Biol Crystallogr* **56**: 222-225 (doi 10.1107/S0907444999016200).
30. Blankenfeldt, W., M. F. Giraud, G. Leonard, R. Rahim, C. Creuzenet, J. S. Lam, & J. H. Naismith. 2000 The purification, crystallization and preliminary structural characterization of glucose-1-phosphate thymidylyltransferase (RmlA), the first enzyme of the dTDP-L-rhamnose synthesis pathway from Pseudomonas aeruginosa. *Acta Crystallogr D Biol Crystallogr* **56**: 1501-1504 (doi 10.1107/S0907444900010040).
29. Giraud, M. F., G. A. Leonard, R. A. Field, C. Bernlind, & J. H. Naismith. 2000 RmlC, the third enzyme of dTDP-L-rhamnose pathway, is a new class of epimerase. *Nat Struct Biol* **7**: 398-402 (doi 10.1038/75178).
28. Liu, H., J. H. Naismith, & R. T. Hay. 2000 Identification of conserved residues contributing to the activities of adenovirus DNA polymerase. *J Virol* **74**: 11681-11689 (doi 10.1128/JVI.74.24.11681-11689.2000).

1999 (6 publications)

27. Giraud, M. F., F. M. Gordon, C. Whitfield, P. Messner, S. A. McMahon, & J. H. Naismith. 1999 Purification, crystallization and preliminary structural studies of dTDP-6-deoxy-D-xylo-4-hexulose 3,5-epimerase (RmlC), the third enzyme of the dTDP-L-rhamnose synthesis pathway, from Salmonella enterica serovar typhimurium. *Acta Crystallogr D Biol Crystallogr* **55**: 706-708 (doi 10.1107/S0907444998015042).
26. Giraud, M. F., H. J. McMiken, G. A. Leonard, P. Messner, C. Whitfield, & J. H. Naismith. 1999 Overexpression, purification, crystallization and preliminary structural study of dTDP-6-deoxy-L-lyxo-4-hexulose reductase (RmlD), the fourth enzyme of the dTDP-L-rhamnose synthesis pathway, from Salmonella enterica serovar Typhimurium. *Acta Crystallogr D Biol Crystallogr* **55**: 2043-2046 (doi 10.1107/S0907444999012251).
25. Moothoo, D. N., B. Canan, R. A. Field, & J. H. Naismith. 1999 Man alpha1-2 Man alpha-OMe-concanavalin A complex reveals a balance of forces involved in carbohydrate recognition. *Glycobiology* **9**: 539-545 (doi 10.1093/glycob/9.6.539).
24. Buchanan, L. V., N. Mehta, L. Pocivavsek, S. Niranjanakumari, E. J. Toone, & J. H. Naismith. 1999 Initiating a structural study of 2-keto-3-deoxy-6-phosphogluconate aldolase from Escherichia coli. *Acta Crystallogr D Biol Crystallogr* **55**: 1946-1948 (doi 10.1107/S0907444999011166).
23. McMahon, S. A., G. A. Leonard, L. V. Buchanan, M. F. Giraud, & J. H. Naismith. 1999 Initiating a crystallographic study of UDP-galactopyranose mutase from Escherichia coli. *Acta Crystallogr D Biol Crystallogr* **55**: 399-402 (doi 10.1107/S0907444998010877).
22. Moothoo, D. N. & J. H. Naismith. 1999 A general method for co-crystallization of concanavalin A with carbohydrates. *Acta Crystallogr D Biol Crystallogr* **55**: 353-355 (doi 10.1107/S0907444998008919).

1998 (7 publications)

21. Singh, J., E. Garber, H. Van Vlijmen, M. Karpusas, Y. M. Hsu, Z. Zheng, J. H. Naismith, & D. Thomas. 1998 The role of polar interactions in the molecular recognition of CD40L with its receptor CD40. *Protein Sci* **7**: 1124-1135 (doi 10.1002/pro.5560070506)
20. Giraud, M. F., J. M. Desterro, & J. H. Naismith. 1998 Structure of ubiquitin-conjugating enzyme 9 displays significant differences with other ubiquitin-conjugating enzymes which may reflect

its specificity for sumo rather than ubiquitin. *Acta Crystallogr D Biol Crystallogr* **54**: 891-898 (doi 10.1107/S0907444998002480).

19. Naismith, J. H. & S. R. Sprang. 1998 Modularity in the TNF-receptor family. *Trends Biochem Sci* **23**: 74-79 (doi 10.1016/S0968-0004(97)01164-X).
18. Moothoo, D. N., S. A. McMahon, S. M. Dimick, E. J. Toone, & J. H. Naismith. 1998 Crystallization of succinylated concanavalin A bound to a synthetic bivalent ligand and preliminary structural analysis. *Acta Crystallogr D Biol Crystallogr* **54**: 1023-1025 (doi 10.1107/S0907444998003965).
17. Moothoo, D. N. & J. H. Naismith. 1998 Concanavalin A distorts the beta-GlcNAc-(1-->2)-Man linkage of beta-GlcNAc-(1-->2)-alpha-Man-(1-->3)-[beta-GlcNAc-(1-->2)-alpha-Man- (1-->6)]-Man upon binding. *Glycobiology* **8**: 173-181 (doi 10.1093/glycob/8.2.173).
16. Moothoo, D. N. & J. H. Naismith. 1998 Concanavalin A distorts the b-GlcNAc-(1-2)-Man linkage of b-GlcNAc-(1-2)-a-Man-(1-3)-[b-GLcNAc-(1-2)-a-Man-(1-6)]-Man upon binding. *Glycobiology* **8**: 173-181 (doi 10.1093/glycob/8.2.173).
15. Moothoo, D. N. & J. H. Naismith. 1998 A general method for co-crystallisation of concanavalin A with carbohydrates. *Acta Crystallographica Section D-Biological Crystallography*. **D55**, 353-5 (doi: 10.1107/S0907444998008919.)

1997(1 publication)

14. Harrison, J. A., K. P. Kartha, S. L. Smith, J. H. Naismith, S. Schenkman, & R. A. Field. 1997 Development of a high through-put spectrophotometric assay to monitor Trypanosoma cruzi trans-sialidase. *Biochem Soc Trans* **25**: 424S. (doi 10.1042/bst025424s)

1996 (4 publications)

13. Cooper, S. J., G. A. Leonard, S. M. McSweeney, A. W. Thompson, J. H. Naismith, S. Qamar, A. Plater, A. Berry, & W. N. Hunter. 1996 The Crystal-Structure of a Class-II Fructose-1,6-Bisphosphate Aldolase Shows a Novel Binuclear Metal-Binding Active-Site Embedded in a Familiar Fold. *Structure* **4**: 1303-1315 (doi 10.1016/S0969-2126(96)00138-4).
12. Naismith, J. H., T. Q. Devine, T. Kohno, & S. R. Sprang. 1996 Structures of the extracellular domain of the type I tumor necrosis factor receptor. *Structure* **4**: 1251-1262 (doi 10.1016/S0969-2126(96)00134-7).
11. Naismith, J. H. & R. A. Field. 1996 Structural basis of trimannoside recognition by concanavalin A. *J Biol Chem* **271**: 972-976 (doi 10.1074/jbc.271.2.972).
10. Naismith, J. H., B. J. Brandhuber, T. Q. Devine, & S. R. Sprang. 1996 Seeing double: crystal structures of the type I TNF receptor. *J Mol Recognit* **9**: 113-117. (doi 10.1002/(SICI)1099-1352(199603)9:2<113::AID-JMR253>3.0.CO;2-H)

1995 (2 publications)

9. Naismith, J. H. & S. R. Sprang. 1995 Tumor necrosis factor receptor superfamily. *J Inflamm* **47**: 1-7. (PMID: 8913924)
8. Naismith, J. H., T. Q. Devine, B. J. Brandhuber, & S. R. Sprang. 1995 Crystallographic evidence for dimerization of unliganded tumor necrosis factor receptor. *J Biol Chem* **270**: 13303-13307 (doi 10.1074/jbc.270.22.13303).

1994 (4 publications)

7. Naismith, J. H., C. Emmerich, J. Habash, S. J. Harrop, J. R. Helliwell, W. N. Hunter, J. Raftery, A. J. Kalb, & J. Yariv. 1994 Refined Structure of Concanavalin-a Complexed With Methyl Alpha-D- Mannopyranoside At 2.0 Angstrom Resolution and Comparison With the Saccharide-Free Structure. *Acta Crystallographica D***50**: 847-858 (doi 10.1107/S0907444994005287).
6. Naismith, J. H., J. Habash, S. Harrop, J. R. Helliwell, W. N. Hunter, T. C. M. Wan, S. Weisgerber, A. J. Kalb, & J. Yariv. 1994 Refined Structure of Cadmium-Substituted Concanavalin-a At 2.0 Angstrom Resolution *Acta Crystallographica D* **50**: 236-236 (doi 10.1107/S090744499301399X).
5. Emmerich, C., J. R. Helliwell, M. Redshaw, J. H. Naismith, S. J. Harrop, J. Raftery, A. J. Kalb, J. Yariv, Z. Dauter, & K. S. Wilson. 1994 High-Resolution Structures of Single-Metal-

Substituted Concanavalin-a - the Co,Ca-Protein At 1.6-Angstrom and the Ni,Ca-Protein At 2.0-Angstrom. *Acta Crystallographica Section D* **50**: 749-756 (doi 10.1107/S0907444994002143).

4. Rodseth, L. E., B. Brandhuber, T. Q. Devine, M. J. Eck, K. Hale, J. H. Naismith, & S. R. Sprang. 1994 2 Crystal Forms of the Extracellular Domain of Type-I Tumor-Necrosis- Factor Receptor. *J Mol Biol* **239**: 332-335 (doi 10.1006/jmbi.1994.1371).

1993 (2 publications)

3. Zhang, Y. et al. 1993 Trypanosoma cruzi trypanothione reductase. Crystallization, unit cell dimensions and structure solution. *J Mol Biol* **232**: 1217-1220 (doi 10.1006/jmbi.1993.1475).
2. Naismith, J. H., J. Habash, S. Harrop, J. R. Helliwell, W. N. Hunter, T. C. Wan, S. Weisgerber, A. J. Kalb, & J. Yariv. 1993 Refined structure of cadmium-substituted concanavalin A at 2.0 A resolution. *Acta Crystallogr D* **49**: 561-571 (doi 10.1107/S0907444993006390).

1992 (1 publication)

1. Naismith, J. H., J. D. Ferrara, S. Bailey, K. Marshall, Z. Dauter, K. S. Wilson, J. Habash, S. J. Harrop, A. J. Berry, & W. N. Hunter. 1992 Initiating a crystallographic study of a class II fructose-1,6-bisphosphate aldolase. *J Mol Biol* **225**: 1137-1141 (doi 10.1016/0022-2836(92)90113-X).