

N.E. Robinson
Z.W. Robinson
B.R. Robinson
A.L. Robinson
J.A. Robinson
M.L. Robinson
A.B. Robinson

Structure-dependent nonenzymatic deamidation of glutaminyl and asparaginyl pentapeptides

Authors' affiliations:

N.E. Robinson, Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA, USA.

Z.W. Robinson, B.R. Robinson, A.L. Robinson, J.A. Robinson, M.L. Robinson and A.B. Robinson, Oregon Institute of Science and Medicine, OR, USA.

Correspondence to:

N.E. Robinson
Division of Chemistry and Chemical Engineering
California Institute of Technology
Pasadena, CA 91125
USA
Tel.: 1-626-577-7849
Fax: 1-626-577-7849
E-mail: noahr@its.caltech.edu

Key words: deamidation; glutaminyl; asparaginyl; glutamine; asparagine; peptides; proteins; mass spectrometry

Abstract: Nonenzymatic deamidation rates for 52 glutaminyl and 52 asparaginyl pentapeptides in pH 7.4, 37.0 °C. 0.15 M Tris HCl buffer have been determined by direct injection mass spectrometry. These and the previously reported 306 asparaginyl rates have been combined in a self consistent model for peptide deamidation. This model depends quantitatively upon peptide structure and involves succinimide, glutarimide and hydrolysis mechanisms. The experimental values and suitable interpolated values have been combined to provide deamidation rate values in pH 7.4, 37.0 °C. 0.15 M Tris HCl buffer for the entire set of 648 single amide permutations of ordinary amino acid residues in GlyXxxAsnYyyGly and GlyXxxGlnYyyGly. Thus, knowledge about sequence dependent deamidation in peptides is extended to include very long deamidation half times in the range of 2–50 years.

Abbreviations: Asn, asparaginyl residue; Gln, glutaminyl residue; Tris, tris(hydroxymethyl)aminomethane.

It has been hypothesized that nonenzymatic deamidations of Asn and Gln residues serve as molecular clocks for the timing of biological processes (1–9). Since this hypothesis was first proposed, amide molecular clocks have been demonstrated in several biological systems. These include the timing of protein turnover in cytochrome c (10,11) and aldolase (12–14); the counting of individual enzyme molecule catalytic cycles in triosephosphate isomerase (15,16); and the resettable time dependent monitoring of DNA repair in Bcl X_L (17).

Nonenzymatic deamidation of asparaginyl (Asn) or glutaminyl (Gln) residues in vivo or in vitro under physiological solvent conditions has been reported for more than 200

Dates:

Received 10 October 2003
Revised 8 December 2003
Accepted 1 February 2004

To cite this article:

Robinson, N.E., Robinson, Z.W., Robinson, B.R., Robinson, A.L., Robinson, J.A., Robinson, M.L. & Robinson, A.B. Structure-dependent nonenzymatic deamidation of glutaminyl and asparaginyl pentapeptides.

J. Peptide Res., 2004, 63, 426–436.

Copyright Blackwell Munksgaard, 2004

types of naturally occurring peptides and proteins. Most of these molecules deamidate primarily at Asn residues, including those of special current interest such as ribonuclease (18) and phenylalanine hydroxylase (19). Deamidation rates of Gln residues are, on average, much slower than those of Asn residues. Consequently, nonenzymatic deamidation of Gln has been observed primarily in very long lived proteins such as the crystallins of the eye lens (20–23).

Deamidation of Asn in peptides and proteins is moderately well understood. At physiological pHs, this reaction apparently proceeds primarily through formation of a five member cyclic imide intermediate (24–28) except in AsnPro sequences. Isomeric products of this reaction include L Asp, L isoAsp, D Asp, and D isoAsp. At neutral pH, all of these products involve the introduction of a negative charge into the peptide or protein. At acidic pH or at neutral to basic pH in those instances when imide formation is structurally prevented, Asn deamidation apparently proceeds primarily through general acid and general base catalyzed hydrolysis.

The sequence dependence of Asn deamidation has been previously measured for 368 peptides in pH 7.4, 0.15 M Tris HCl buffer, at 37 °C (6,29), including 306 peptides of sequence GlyXxxAsnYyyGly, and in 65 peptides in phosphate buffer (3,30,31). Asn and Gln deamidation is strongly influenced by buffer anions, especially phosphate (3,32,33).

Sequence dependent rates of deamidation of Asn in Tris HCl, three dimensional protein structures, and qualitative reports of deamidation in proteins under a wide variety of solvent conditions have been combined to produce a reliable method of calculation for estimating the deamidation indexes and deamidation coefficients of Asn in all proteins for which the three dimensional structure is known (7–9). In proteins, these coefficients and indexes have been found to depend about 60% on primary structure and 40% on three dimensional structure. This method has proved reliable for predicting the relative deamidation rates of Asn in individual proteins and is also remarkably accurate in predicting absolute deamidation rates of proteins (8).

It has been found that protein deamidation rates are about the same as pentapeptide model rates for amides in unhindered regions of the protein. When protein structure interferes with deamidation, this interference can be quantitatively estimated from three dimensional protein structure. This estimate can be combined with the appropriate pentapeptide rate to give an accurate deamidation rate for each individual amide in the protein. These computations have been found to be correct for more than 95% of the protein amides for which deamidation measurements and three dimensional structures are available (7–9).

This method of estimating Asn deamidation rates is proving useful in individual protein investigations (17,19) and has also provided estimates of the prevalence of deamidation in biological systems. Calculations of 126 human proteins (8) indicate that more than 30% of these are at least one tenth singly deamidated in 5 days under physiological solvent conditions, and calculations of 10 369 proteins in the Brookhaven Protein Data Bank (9) indicate that more than 15% of these are at least one tenth singly deamidated in 5 days.

A similar treatment for Gln deamidation in proteins has not been possible because appropriate sequence dependent deamidation rates for Gln peptides were not available. These rates are reported herein. Deamidation rates for 30 Gln peptides have been reported (3,30), but these measurements were carried out in phosphate buffer, which substantially complicates their use. It has been found that the reported deamidation rates for Asn peptides in phosphate can be self consistently described (31), but phosphate effects on deamidation rates of Gln peptides are apparently larger than for Asn (3,32).

We now report, therefore, the rates of deamidation of 52 Gln pentapeptides and 52 additional Asn pentapeptides in pH 7.4, 37 °C, 0.15 M Tris HCl buffer. We show that these and the previously measured (6,29) 306 Asn rates are understandable through a combination of succinimide, glutarimide and hydrolysis mechanisms. The measured Gln rates are similar in magnitude to the Gln deamidation rates found for long lived proteins in vivo (20–23). When combined with three dimensional structures for long lived proteins, these sequence dependent rates may eventually permit quantitative understanding of Gln deamidation on a similar basis to that now available for Asn (7–9).

Experimental Procedures

Peptide synthesis

The peptides were synthesized by means of Merrifield solid phase peptide synthesis (34,35) as previously described (6,29). Fmoc derivatives were used with appropriate side chain blocking groups. The peptides were cleaved from the resin in trifluoroacetic acid, washed with methyl *t* butyl ether, and freeze dried. Yields were about 70% with the remaining peptide either not washed from the resin or dissolved in the ether washes.

The peptides had correct mass spectra, and impurities were indistinguishable from background in the mass

spectrometer. Each peptide was also analyzed by hydrolysis in 6 N HCl in H₂O at 95 °C for 48 h followed by amino acid analysis in a Thermoquest HPLC with a P4000 pump, AS3000 autosampler, and UV6000LP detector (Thermoquest Finnigan, San Jose, CA, USA) combined with a Pickering PCX5200 derivatizer (Pickering, Mountain View, CA, USA), column, ninhydrin, and lithium buffer system and showed correct amounts of amino acids.

Deamidation of peptides

Peptide deamidation was carried out in pH 7.4, 37 °C, 0.15 M Tris HCl buffer at 37.00 ± 0.02 °C in the apparatus and reaction tubes previously described (6,29). Peptide concentrations were 0.0010 M. The 18 polypropylene tubes used for each peptide were fitted with Teflon liners. Sealed sets of 100 µL samples were incubated for times between 0 and 300 days in a hydrated environment to prevent evaporation. Individual tubes were periodically removed from the reaction chamber and stored at −80 °C until analysis. At the end of the 300 day incubation period, the pH was checked in the 0th, 9th, and 18th tubes and found to be pH 7.4.

Measurement of deamidation by mass spectrometry

Measurements were made in a Thermoquest Model LCQ mass spectrometer (Thermoquest Finnigan) with a Thermoquest electrospray source, a Thermoquest AS3500 autosampler (Thermoquest Finnigan), and a special sample delivery system as previously described (6,29). Samples were alternated to correct for spectrometer drift and analyzed by direct loop injection into the mass spectrometer. This analytical system includes several special characteristics necessary for reliable, quantitative determination of deamidation as have been described in detail (29). Results were calculated by means of a customized procedure (6,29). Deamidation rates were assumed to be first order in peptide.

Results

The experimental results are summarized in Tables 1–3. The indicated experimental errors reflect statistical analysis. Small systematic errors are probably also present. For GlyXxxGlnGlyGly, where glutarimide formation would be expected to be favorable, the median deamidation half time is about 660 days. The median values then range upward to

Table 1. GlyXxxGlnYyyGly (pH 7.4, 37 °C, 0.15 M Tris-HCl)

Peptides	t 1/2 Days	Probable Error (%)	$k \times 10^9$ (s ^{−1})	ln(k) + 22
GlyCysGlnGlyGly	559	5	14.35	3.94
GlyAlaGlnGlyGly	605	1	13.26	3.86
GlyIleGlnGlyGly	621	1	12.92	3.84
GlyGlyGlnGlyGly	649	1	12.36	3.79
GlyArgGlnGlyGly	659	1	12.17	3.78
GlyThrGlnGlyGly	668	1	12.01	3.76
GlyLeuGlnGlyGly	668	3	12.01	3.76
GlyTyrGlnGlyGly	849	3	9.45	3.52
GlyHisGlnGlyGly	855	2	9.38	3.52
GlyArgGlnCysGly	1040	12	7.71	3.32
GlyAlaGlnCysGly	1910	3	4.20	2.71
GlyGlyGlnMetGly	4500	8	1.78	1.85
GlyIleGlnMetGly	5140	9	1.56	1.72
GlyAlaGlnThrGly	5070	11	1.58	1.74
GlyTrpGlnThrGly	5580	10	1.44	1.64
GlyGlyGlnSerGly	5710	9	1.40	1.62
GlyLysGlnAlaGly	4320	3	1.86	1.90
GlyMetGlnAlaGly	4400	11	1.82	1.88
GlyValGlnAlaGly	4460	4	1.80	1.86
GlyProGlnAlaGly	4730	3	1.70	1.81
GlyAlaGlnAlaGly	5300	8	1.51	1.69
GlyGlyGlnAlaGly	5880	6	1.36	1.59
GlyPheGlnAlaGly	6270	9	1.28	1.52
GlySerGlnAlaGly	6420	5	1.25	1.50
GlyAspGlnAlaGly	7090	6	1.13	1.40
GlyGluGlnAlaGly	7100	4	1.13	1.40
GlyAspGlnHisGly	2540	7	3.16	2.43
GlyLysGlnHisGly	5100	4	1.57	1.73
GlySerGlnHisGly	6520	5	1.23	1.48
GlyAspGlnLysGly	4600	4	1.74	1.83
GlyValGlnLysGly	5210	4	1.54	1.71
GlyGluGlnIleGly	4230	7	1.90	1.92
GlyValGlnIleGly	5570	4	1.44	1.64
GlySerGlnIleGly	5870	9	1.37	1.59
GlyAlaGlnIleGly	6170	14	1.30	1.54
GlyGlyGlnValGly	6470	3	1.24	1.49
GlyLysGlnArgGly	2300	2	3.49	2.53
GlyHisGlnArgGly	4530	21	1.77	1.85
GlyThrGlnArgGly	5110	5	1.57	1.73
GlyProGlnArgGly	6370	2	1.26	1.51
GlyAlaGlnArgGly	7190	4	1.12	1.39
GlyLysGlnGluGly	5350	5	1.50	1.68
GlyHisGlnAspGly	5630	5	1.42	1.63
GlyAlaGlnAspGly	7450	4	1.08	1.35
GlyTrpGlnAspGly	8780	6	0.91	1.19

Table 1. Continued

GlyGlyGlnPheGly	7610	13	1.05	1.33
GlyTrpGlnPheGly	8580	6	0.94	1.21
GlyGlyGlnProGly	10 040	8	0.80	1.05
GlyAlaGlnTyrGly	10 000	5	0.80	1.06
GlyLysGlnTrpGly	10 400	37	0.77	1.02
GlyGlyGlnTrpGly	15 300	17	0.52	0.63
GlyTyrGlnTrpGly	17 900	36	0.45	0.47

Table 2. GlyXxxAsnProGly (pH 7.4, 37 °C, 0.15 M Tris-HCl)

Peptides	$t_{1/2}$ Days	Probable error (%)	$k \times 10^9$ (s^{-1})	$\ln(k) + 22$
GlyCysAsnProGly	1550	9	5.18	2.92
GlyLysAsnProGly	4940	8	1.62	1.76
GlyArgAsnProGly	5790	11	1.39	1.60
GlyThrAsnProGly	6290	9	1.28	1.52
GlyProAsnProGly	6590	8	1.22	1.47
GlySerAsnProGly	7060	11	1.14	1.40
GlyGlyAsnProGly	7170	10	1.12	1.39
GlyAlaAsnProGly	7370	8	1.09	1.36
GlyPheAsnProGly	7990	11	1.00	1.28
GlyValAsnProGly	8030	8	1.00	1.28
GlyHisAsnProGly	8440	13	0.95	1.23
GlyGluAsnProGly	8860	39	0.91	1.18
GlyMetAsnProGly	9300	21	0.86	1.13
GlyTyrAsnProGly	9830	24	0.82	1.07
GlyLeuAsnProGly	10 500	15	0.76	1.01
GlyIleAsnProGly	11 600	13	0.69	0.91
GlyAspAsnProGly	11 800	39	0.68	0.89
Median	8000		1.00	1.28
AlaAlaAlaAla NH ₂	6590	16	1.16	1.43
AlaAlaSerAsn	3440	15	2.23	2.08

about 10 000 days in GlyXxxGlnProGly where glutarimide formation is not possible and a little higher for Yyy of Tyr or Trp. The median value for GlyXxxAsnProGly, where succinimide formation is not possible is about 8000 days. These values are also in agreement with the value of 6600 days determined for AlaAlaAlaAla NH₂ and given in Table 2.

Sequence dependent effects are evident throughout the deamidation half times, with the effects of steric hindrance and acidic and basic side chains most pronounced. Reported experimental errors for the longer deamidation half times are higher than for the shorter largely because the experimental deamidation interval constituted a relatively smaller part of the half time.

The 410 experimental structure dependent deamidation pentapeptide rates were quantitatively organized, so that, by interpolation, those that have not been measured could be

Table 3. AcetamidomethylCys Peptides (pH 7.4, 37 °C, 0.15 M Tris-HCl)

Peptides	$t_{1/2}$ Days	Probable error (%)	$k \times 10^9$ (s^{-1})	$\ln(k) + 22$
GlyAmCysAsnGlyGly	1.14	1.0	7037	10.14
GlyAmCysAsnHisGly	10.9	0.5	736	7.88
GlyAmCysAsnSerGly	15.4	1.0	521	7.53
GlyAmCysAsnAlaGly	21.5	0.5	373	7.20
GlyAmCysAsnAspGly	32.9	1.0	244	6.77
GlyAmCysAsnAmCysGly	39.3	1.0	204	6.60
GlyAmCysAsnThrGly	41.7	0.5	192	6.54
GlyAmCysAsnLysGly	48.9	1.5	164	6.38
GlyAmCysAsnMetGly	56.5	0.5	142	6.23
GlyAmCysAsnGluGly	45.0	1.0	178	6.46
GlyAmCysAsnArgGly	58.8	1.0	136	6.19
GlyAmCysAsnPheGly	63.3	0.5	127	6.12
GlyAmCysAsnTyrGly	78.8	1.5	102	5.90
GlyAmCysAsnTrpGly	81.3	0.5	99	5.87
GlyAmCysAsnLeuGly	100	1.0	80	5.66
GlyAmCysAsnValGly	215	1.0	37	4.90
GlyAmCysAsnIleGly	250	0.5	32	4.75
GlyAmCysAsnProGly	3910	9.0	2.05	2.00
GlyGlyAsnAmCysGly	27.6	0.5	291	6.95
GlySerAsnAmCysGly	41.3	0.5	194	6.55
GlyThrAsnAmCysGly	34.1	0.5	235	6.74
GlyMetAsnAmCysGly	33.8	0.5	237	6.75
GlyPheAsnAmCysGly	29.8	0.5	269	6.87
GlyTyrAsnAmCysGly	33.3	1.0	241	6.76
GlyAspAsnAmCysGly	45.8	2.5	175	6.44
GlyGluAsnAmCysGly	32.1	1.0	250	6.80
GlyHisAsnAmCysGly	33.8	1.0	237	6.75
GlyLysAsnAmCysGly	36.5	0.5	220	6.67
GlyArgAsnAmCysGly	42.3	0.5	190	6.52
GlyAlaAsnAmCysGly	40.6	1.0	198	6.56
GlyLeuAsnAmCysGly	33.6	1.0	239	6.75
GlyValAsnAmCysGly	34.7	0.5	231	6.72
GlyIleAsnAmCysGly	33.0	0.5	243	6.77
GlyTrpAsnAmCysGly	42.9	0.5	187	6.51
GlyProAsnAmCysGly	43.7	0.5	184	6.49

estimated. This provides a complete set of sequence dependent rates as listed in Tables 4 and 5. This includes 648 ordinary peptide sequences with exclusion only of the 152 double amides for which no measurements have yet been reported. The reliability of the experimental and interpolated values should be sufficient for most protein applications.

First, the values were arranged in an ordered array in ascending order of deamidation half time with respect to

Table 4. First-order deamidation half-times of GlyXxxAsnYyyGly in days at pH 7.4, 37.0 °C, 0.15 M Tris HCl

XxxYyy	G y	His	Ser	A a	Asp	AmCys	Thr	Cys	Lys	Met	G u	Arg	Phe	Tyr	Trp	Leu	Va	I e	Pro	Median*
G y	1.03	9.2	11.8	21.1	28.0	27.6	39.8	40.6	48.2	50.4	73.9	57.8	64.0	63.6	77.1	104	224	287	7170	50 4
Ser	0.96	9.0	15.1	24.1	30.3	41.3	45.7	60.2	55.5	54.9	59.7	59.7	52.2	64.7	76.8	110	233	285	7060	55 5
Thr	1.04	9.6	17.1	24.6	27.9	34.4	50.0	55.5	57.6	47.6	60.8	51.2	76.4	80.6	72.5	110	237	279	6290	55 5
Cys	1.14	10.8	19.0	26.4	30.6	38.3	48.7	46.0	46.6	64.5	48.3	83.1	73.9	83.9	111	119	229	304	1550	48 7
AmCys	1.14	10.9	15.4	21.5	32.9	39.3	41.7	46.6	48.9	56.5	45	58.8	63.3	78.8	81.3	100	215	250	3900	48 9
Met	1.04	10.2	15.2	22.1	26.4	33.8	43.6	49.6	60.4	56.9	72.4	58.8	61.9	74.0	92.7	113	211	275	9300	57 9
Phe	1.15	10.2	18.1	24.2	27.4	29.8	39.0	46.5	58.2	58.6	62.4	61.2	69.5	75.1	102	118	203	287	7990	58 6
Tyr	1.49	10.2	11.9	24.3	28.4	33.3	38.1	48.6	55.1	64.3	41.0	56.9	58.0	70.6	120	118	241	306	9830	51 8
Asp	1.53	9.7	17.0	24.0	29.4	45.8	52.4	54.1	75.9	57.3	46.8	87.2	70.1	70.4	80.3	111	241	298	11 800	55 7
G u	1.45	9.0	16.4	23.8	32.0	32.1	36.8	44.2	77.8	59.6	60.3	80.9	70.2	94.5	98.4	130	268	279	8860	59 9
H s	1.14	10.7	15.7	24.6	31.2	33.8	47.2	43.9	50.2	63.1	69.4	48.9	72.1	82.3	95.4	116	247	327	8440	50 2
Lys	1.02	10.5	15.6	23.6	34.0	36.5	58.1	49.0	53.5	60.9	72.5	57.4	70.1	96.7	98.1	119	246	313	4940	58 1
Arg	1.00	10.0	14.3	24.4	34.7	42.3	50.7	50.5	49.6	74.4	68.3	67.4	68.3	90.0	127	128	247	311	5790	67 4
A a	1.05	9.3	14.9	22.5	31.9	40.6	43.5	63.7	55.9	59.2	74.1	62.4	65.6	73.9	130	124	254	300	7370	62 4
Leu	1.08	10.7	16.7	25.1	32.1	33.6	46.1	53.5	60.1	62.6	56.7	62.1	72.4	75.7	74.5	155	294	391	10 500	60 1
Va	1.23	10.2	18.2	27.5	33.5	34.7	49.9	63.2	63.8	65.7	64.8	67.4	66.6	79.2	88.9	154	291	366	8030	64 8
e	1.26	11.5	14.5	25.9	33.8	33.0	46.3	52.7	64.4	58.8	58.6	66.4	61.5	79.3	86.7	154	295	384	11 600	58 8
Trp	1.75	11.3	15.5	30.7	43.6	42.9	38.9	83.1	59.4	64.2	75.7	73.9	71.1	92.6	135	133	226	286	12 000	67 6
Pro	1.18	12.8	18.9	31.8	48.6	43.7	63.1	60.0	67.8	78.4	92.0	72.9	100	114	122	181	364	455	6590	72 9
Mean	1.19	10.3	15.9	25.0	32.5	36.7	46.3	53.2	58.4	60.9	63.3	65.0	68.8	81.1	98	126	251	315	8300	60 9
SD	0.05	0.23	0.49	0.67	1.3	1.2	1.7	2.4	2.1	1.8	3.1	2.5	2.3	3.0	4.9	5.1	9.3	12.2	600	2 3
% SD	4.4	2.2	3.1	2.7	4.1	3.3	3.6	4.5	3.6	2.9	4.8	3.9	3.4	3.7	5.0	4.0	3.7	3.9	8 8	3 7
Median	1.14	10.2	15.6	24.4	31.9	34.7	46.1	50.5	57.6	59.6	62.4	62.1	69.5	79.2	95	119	241	300	8010	59 6

*Median does not include Yyy = AmCys
Values in bold are experimental

Table 5. First-order deamidation half-times of GlyXxxGlnYyyGly in days at pH 7.4, 37.0 °C, 0.15 M Tris-HCl

Xxx\Yyy	G y	Cys	Met	Thr	Ser	A a	His	Lys	Leu	I e	V a	Arg	G u	Asp	Phe	Pro	Tyr	Trp	Median
Cys	560	800	3200	3500	3800	4100	4200	4400	4800	4900	5000	5100	5600	6100	6500	7100	7900	9100	4800
Met	600	900	3500	3800	4100	4400	4400	4600	5000	5000	5000	5100	5800	6200	6600	7300	8200	9400	5000
Thr	670	1000	3700	4000	4200	4300	4500	4800	5200	5300	5100	5100	5900	6300	6800	7500	8400	9700	5100
Lys	650	1000	4000	4100	4200	4300	4900	4000	5300	5400	5700	2300	5400	5900	7000	7700	8800	10 000	5300
Arg	660	1300	4100	4200	4300	4400	4900	4000	5400	5500	5800	2300	5400	5900	7100	8100	9200	11 000	4900
Va	640	1300	4200	4300	4400	4500	5000	5200	5500	5600	5900	6100	6500	7000	7200	8500	9700	12 000	5500
Pro	630	1600	4500	4600	4600	4700	5200	5500	5800	6000	6200	6400	6800	7200	7300	8900	10 000	13 000	5800
A a	610	1900	4400	5100	5200	5300	5500	5700	6100	6200	6400	7200	7300	7400	7500	9300	10 000	14 000	6100
G y	650	1900	4500	5200	5700	5900	5900	6000	6200	6300	6500	7200	7300	7600	7600	10 000	12 000	15 000	6200
Leu	670	2000	4800	5300	5800	6000	6100	6100	6300	6500	6800	7200	7400	7800	8000	10 000	12 000	16 000	6300
e	620	2000	5100	5300	5800	6200	6100	6100	6300	6500	7100	7200	7700	8100	8100	10 000	12 000	16 000	6300
Phe	660	2000	5100	5300	5900	6300	6200	6200	6400	6400	7100	7200	8100	8200	8200	10 000	12 000	16 000	6400
Ser	700	2100	5100	5400	6000	6400	6500	6300	6100	5900	6800	7200	8100	8200	8300	10 000	13 000	17 000	6400
G u	750	2100	5200	5400	6100	7100	2500	4600	4300	4200	6400	5200	8200	8300	8400	10 000	13 000	17 000	5400
Asp	800	2100	5200	5400	6200	7100	2500	4600	6200	6400	6600	5200	8200	8400	8500	11 000	13 000	17 000	6200
H s	850	2200	5200	5500	6300	7200	7200	4000	6600	6700	6800	4500	5800	5600	8600	11 000	14 000	18 000	6300
Tyr	850	2200	5300	5600	6400	7300	7400	7500	7800	7900	8000	8100	8300	8600	8700	11 000	14 000	18 000	7800
Trp	850	2300	5300	5600	6500	7400	7500	7600	7900	8000	8200	8300	8500	8800	8600	11 000	14 000	19 000	7900
Mean	690	1700	4600	4900	5300	5700	5400	5400	6000	6000	6400	5900	7000	7300	7700	9400	11 200	14 300	6000
SD	22	129	163	169	228	296	352	272	226	233	221	423	273	259	180	329	521	809	246
% SD	3.2	7.6	3.5	3.4	4.3	5.2	6.5	5.0	3.8	3.9	3.4	7.2	3.9	3.5	2.3	3.5	4.7	5.7	4
Median*	660	1950	4650	5250	5750	5950	6000	6050	6250	6400	6650	7200	7350	7700	7800	10 000	12 000	15 500	6150

*Median without charged residues
Values in bold are experimental

the C terminal and N terminal residues. This ordering excluded those peptides with acidic or basic residues at both positions, as these doubly charged side chain molecules previously have been found to have unusual values in Asn peptides (6) and are also found to have unusual values in the Gln peptides reported here.

Secondly, the $\ln(k)$ values for the Gln array were interpolated by a three dimensional surface fitting calculation. The experimental values and the interpolated values are shown in Table 5. The resulting median deamidation half time values for Gly AsnYyyGly and Gly GlnYyyGly, excluding the doubly charged peptides, are given in Tables 6 and 7 and displayed in Fig. 1. These tables and figure summarize the effects of the C terminal nearest neighbor residues.

The Asn peptides deamidate primarily by formation of a succinimide ring (24–27) with the evident marked sequence dependence of formation of this ring. When formation of the succinimide is not possible as in the case of AsnPro peptides, the Asn rate decreases to the hydrolysis rate of the similar GlnPro peptides.

Imide formation in Asn peptides can involve either the peptide backbone nitrogen or the amide side chain nitrogen.

Table 6. Gly AsnYyyGly (pH 7.4, 37 °C, 0.15 M Tris-HCl)

Peptide*	Median $t_{1/2}$ Days	$k \times 10^9$ (s^{-1})	22 + $\ln(k)$	kJ/mole†
Gly AsnGlyGly	1.14	7037	10.14	22.8
Gly AsnHisGly	10.2	787	7.94	17.2
Gly AsnSerGly	15.7	511	7.51	16.1
Gly AsnAlaGly	24.5	327	7.07	14.9
Gly AsnAspGly	31.6	254	6.81	14.3
Gly AsnThrGly	46.2	174	6.43	13.3
Gly AsnCysGly	51.6	155	6.32	13.0
Gly AsnLysGly	57.9	139	6.21	12.7
Gly AsnMetGly	60.3	133	6.17	12.6
Gly AsnArgGly	62.3	129	6.13	12.5
Gly AsnGluGly	63.6	126	6.11	12.4
Gly AsnPheGly	69.8	115	6.02	12.2
Gly AsnTyrGly	79.3	101	5.89	11.9
Gly AsnTrpGly	97.0	83	5.69	11.4
Gly AsnLeuGly	119	67	5.49	10.8
Gly AsnValGly	243	33	4.77	9.0
Gly AsnIleGly	302	27	4.56	8.5
Gly AsnProGly‡	8000	1.00	1.28	0

*Median values.

†Arrhenius activation energy relative to Gly AsnProGly assuming constant Arrhenius frequency factor.

‡This work. Other values from reference (6).

Table 7. Gly GlnYyyGly (pH 7.4, 37 °C, 0.15 M Tris-HCl)

Peptide	Median $t_{1/2}$ Days	$k \times 10^9$ (s^{-1})	22 + $\ln(k)$	kJ/mole*
Gly GlnGlyGly	660	12.16	3.77	7.0
Gly GlnCysGly	1950	4.11	2.69	4.2
Gly GlnMetGly	4650	1.73	1.82	2.2
Gly GlnThrGly	5250	1.53	1.70	2.0
Gly GlnSerGly	5750	1.40	1.61	1.4
Gly GlnAlaGly	5950	1.35	1.58	1.3
Gly GlnHisGly	6000	1.34	1.57	1.3
Gly GlnLysGly	6050	1.33	1.56	1.3
Gly GlnLeuGly	6250	1.28	1.53	1.2
Gly GlnIleGly	6400	1.25	1.50	1.1
Gly GlnValGly	6650	1.21	1.46	1.0
Gly GlnArgGly	7200	1.11	1.38	0.8
Gly GlnGluGly	7350	1.09	1.36	0.8
Gly GlnAspGly	7700	1.04	1.32	0.7
Gly GlnPheGly	7800	1.03	1.30	0.6
Gly GlnProGly	10 000	0.80	1.06	0
Gly GlnTyrGly	12 000	0.67	0.87	0.5
Gly GlnTrpGly	15 500	0.52	0.62	1.1

*Arrhenius activation energy relative to Gly GlnProGly assuming constant Arrhenius frequency factor.

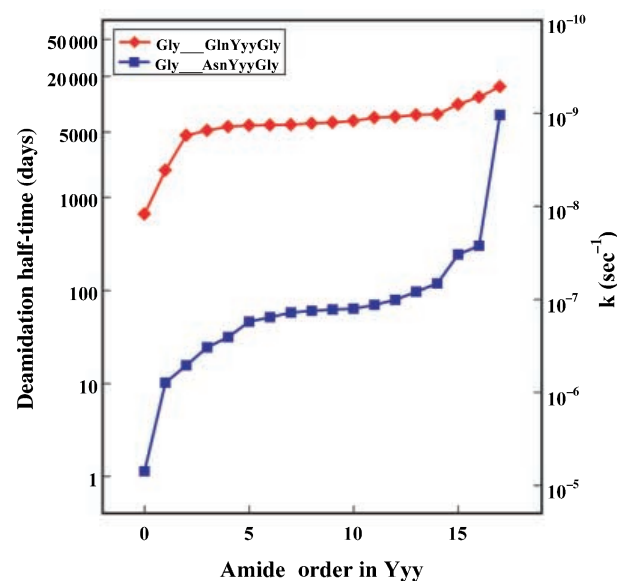


Figure 1. Deamidation rate constants and half times in pH 7.4, 37.0 °C, 0.15 M Tris HCl buffer for the Xxx medians of 324 GlyXxxAsnYyyGly peptides and 324 GlyXxxGlnYyyGly peptides referenced, measured, or computed as described in the text. The medians for the two sets of 18 peptides were calculated and then plotted in order of increasing deamidation half time. The plotted values are given in Tables 6 and 7.

The former leads to deamidation, while the latter leads to peptide cleavage (36). Formation of the imide that leads to cleavage is usually much slower than that of the imide that leads to deamidation, although cleavage products are detectable in the mass spectra of most of our peptide deamidation experiments. As long as cleavage is slow in comparison with deamidation or to the duration of the experiment, its effect on the measured deamidation rate is not significant. In AsnPro peptides, however, cleavage is faster than for other sequences. Also, the slow deamidation rate of AsnPro requires experiments of longer duration.

We have now undertaken a project to measure the cleavage rates in the samples utilized for all of our deamidation rate determinations. Preliminary results indicate that cleavage has not significantly affected our reported deamidation rates except in the case of AsnPro. In the case of AsnPro, the actual deamidation half times may be as much as 20% higher than reported in Table 2. We prefer not to make this correction until the actual cleavage rates are precisely known.

The Gln peptides show apparent glutarimide formation in the first few C terminal sequences, especially for GlnGly, but then establish a range wherein hydrolysis seems to predominate. The complete range of Asn and Gln C terminal side sequence dependence covers about four orders of magnitude varying from deamidation half times in pH 7.4, 37 °C, 0.15 M Tris HCl buffer of 1 day to more than 10 000 days. Apparently, this range includes successive succinimide, glutarimide, and hydrolysis mechanisms, with deamidation probably mildly catalyzed by Tris. Catalysis of long lived amides by phosphate ions can be much greater (3,32,33).

Table 2 and reference (6) provide experimental deamidation rates for 320 Asn sequences with the remaining four sequences interpolated. These Asn peptide rates are summarized in Table 4. Table 5 summarizes the experimental rates for 52 Gln sequences and interpolated values for the remainder. Figure 2 shows the cumulative distribution functions for the complete set of deamidation values in Tables 4 and 5, with the AmCys sequences excluded.

Using Gly AsnProGly as a standard for pure hydrolysis, median Gly AsnIleGly deamidation apparently proceeds approximately 96% through the imide and 4% through hydrolysis. This rises to more than 99% imide for the median of AsnPhe peptides with a deamidation half time of 70 days and for peptides with shorter half times.

Similarly, using Gly GlnProGly as a standard, the median of the Gly Gln Gly peptides is estimated to be about 40% glutarimide and 60% hydrolysis. Similar per

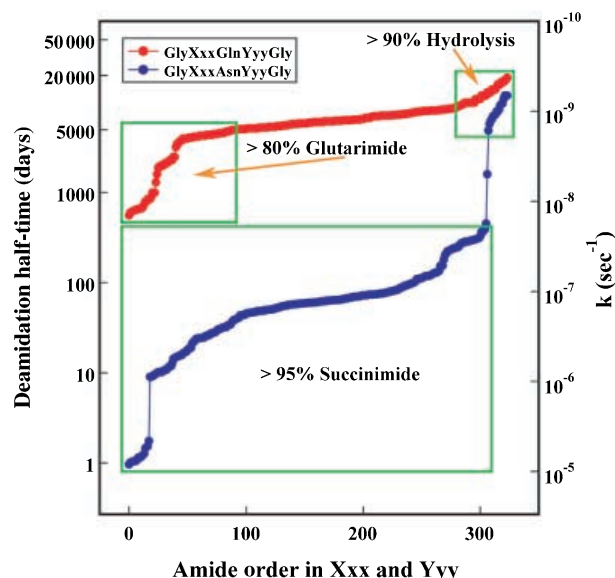


Figure 2. Deamidation rate constants and half times in pH 7.4, 37.0 °C, 0.15 M Tris HCl buffer for 324 GlyXxxAsnYyyGly peptides and 324 GlyXxxGlnYyyGly peptides referenced, measured, or computed as described in the text were plotted in order of increasing deamidation half time. The rates of deamidation can be explained by the succinimide, glutarimide, and hydrolysis mechanisms as indicated, although additional mechanisms may be present. The plotted values, which exclude AmCys peptides, are given in Tables 7 and 8. These rates are for peptides. In proteins, where the mechanisms of deamidation are selectively suppressed by three dimensional structure, the rates of deamidation are quantitatively different (7–9).

centages for specific Gln peptides, however, cannot be calculated precisely. The sequence dependence of the glutarimide and hydrolysis mechanisms is not quantitatively separable with these data, as the rates are of similar magnitude and the independent sequence dependence of hydrolysis is unknown.

As shown in Tables 6–8, comparative Arrhenius Activation Energies, assuming a constant Arrhenius Frequency Factor, for the AsnGly vs. AsnPro succinimide mechanism and the GlnGly vs. GlnPro glutarimide mechanism are 23 and 7 kJ/mole, respectively.

The values in Tables 6–8 were calculated as follows. First, for convenience, deamidation half times were expressed as $22 + \ln[k]$, where k is the assumed first order rate constant. This first order assumption is consistent, within experimental error, with the experimental data, but it is only an approximation for those peptides deamidating by two or more mechanisms. The activation energies, E_a were then estimated as $E_a = [\delta \ln[k]] (2.577) \text{ kJ/mole}$.

Deamidation of AcetylGlnGlyNHCH₃ in pH 7.4, 37 °C, 0.5 M phosphate buffer with ionic strength 0.5 has been

Table 8. Activation energy differences pH 7.4, 37 °C, 0.15 M Tris-HCl

Peptides	22 + ln(k)	$\delta \ln(k)$	kJ/mole*
Succinimide GlyGlyAsnProGly vs. GlyGlyAsnGlyGly	1.39 vs. 10.24	8.85	22.8
Succinimide Gly AsnProGly vs. Gly AsnGlyGly	1.28 vs. 10.14	8.86	22.8
Glutarimide GlyGlyGlnProGly vs. GlyGlyGlnGlyGly	1.05 vs. 3.80	2.75	7.1
Glutarimide Gly GlnProGly vs. Gly GlnGlyGly	1.06 vs. 3.77	2.71	7.0

* δ Arrhenius activation energy.

reported (37) with rate constants determined for the glutarimide and hydrolysis mechanisms. These give deamidation half times of 430 days for the glutarimide and 1800 days for hydrolysis. The combined deamidation half time is 350 days, with 82% proceeding via the glutarimide and 18% by hydrolysis (37).

In comparison, Table 7 gives median deamidation half times for Gly GlnGlyGly and Gly GlnProGly of 660 and 10 000 days, respectively. These measurements suggest that the deamidation of the average Gly GlnGlyGly peptide proceeds 93% through the glutarimide and 7% through hydrolysis. These values for similar but not identical peptide structures in Tris rather than phosphate buffer are in reasonably good agreement with those for Acetyl GlnGlyNHCH₃.

Of special interest are the peptides containing CysAsn Pro, ArgGlnCys, and AlaGlnCys, which have deamidation half times of 1550, 1040, and 1910 days, respectively. These deamidation rates are markedly faster than expected for comparable peptides expected to deamidate primarily by hydrolysis. Table 3 gives experimentally determined deamidation rates for complete sets of XxxAsnAmCys and AmCysAsnYyy sequences wherein the Cys side chain has been blocked with an acetamidomethyl group. The blocked peptides have the expected rates of deamidation, although GlyAmCysAsnProGly with 3910 days is also somewhat faster than expected.

Conclusions

The measurements and calculations herein reported provide, for the first time, sequence dependent deamidation rates for Gln peptides and AsnPro peptides that are sufficiently accurate for use in efforts to understand the primary, secondary, tertiary, and quaternary structure aspects of deamidation of Gln and AsnPro in proteins.

For Asn deamidation, it has been found that the effects of protein three dimensional structure on deamidation can be

successfully discovered and quantitatively understood by measurement of the primary structure rates in model pentapeptides and comparison of these rates with those found in proteins for which the three dimensional structures are known (6,9,29). The pentapeptide deamidation rates and the three dimensional structures permit computation of the protein deamidation rates (7,9).

For an example of the potential application of this to Gln, it has been reported (21) that Gln(50)Ser in α A crystallin deamidates in human lenses with a half time of about 28 000 days beginning at age 25. We find that the median deamidation half time for GlyXxxGlnSerGly is 6000 days. In human growth hormone, the experimental deamidation half time for Asn(149)Ser is 60 days, while the median deamidation half time for GlyXxxAsnSerGly is 16 days (7). Thus, Gln(50) in vivo is fivefold slower than the peptide rate, and Asn(149) is fourfold slower. In a quantitative study of 30 proteins that deamidate at Asn, it has been found that an interfering hydrogen bond slows deamidation by about a factor of 2, so, like Asn(149), Gln(50) in vivo may have two such interfering hydrogen bonds. Alternatively, Gln(50) deamidation may be slowed by other aspects of three dimensional structure.

This example is hypothetical at present because an insufficient number of Gln protein deamidation rates have been measured in vivo to appropriately calibrate the various three dimensional structures that are involved. The Gln peptide deamidation rates reported herein will allow these calibrations to be made as Gln rates are measured in proteins. The fact that the rates for Gln peptides are similar in magnitude to the deamidation rates that have been observed so far in human lenses strongly suggests that Gln deamidation will be amenable to analysis by the same procedures that have been successful for Asn.

With the values herein and the previously reported deamidation rates for Asn peptides (6,29), complete sets of experimentally determined or suitably estimated sequence dependent deamidation rates are now available for 324 Asn and 324 Gln nearest neighbor sequences, which include all

except the double amides. It has been shown that a quantitative understanding of the deamidation rates of Asn in proteins can be achieved by combining pentapeptide deamidation rates with protein three dimensional structures (6,9,29). This may also prove true for Gln.

In pH 7.4, 37.0 °C, 0.15 M Tris HCl buffer, these nearest neighbor, sequence dependent deamidation half times of Asn and Gln range from about 1 day to more than 10 000 days. The reactions apparently involve succinimide, glutarimide, and hydrolysis mechanisms, with the proportions of each gradually changing to the next in that order as deamidation half times increase. Figure 1 shows these effects for the C terminal neighboring residues of Asn and Gln. Inclusion of N terminal neighbors provides specific values for each peptide as is illustrated in Fig. 2. Deamidation rates are valuable because peptides and proteins contain amide molecular clocks of substantial biological importance.

Of the 20 ordinary amino acid residues in peptides and proteins, Asn and Gln are uniquely unstable under physiological conditions. The nonenzymatic deamidation half times of these residues are genetically programmable throughout the range from less than a day to more than a century. Many are within the biological lifetimes of the proteins, organelles, and organisms of which they are a part. Deamidation of Asn or Gln introduces negative charges into proteins and, to a lesser but significant extent, isomerization as well.

Why do peptides and proteins include residues that are so unstable that they cause fundamental changes in the structures of peptides and proteins during the biological lifetimes of those molecules? These timed changes of

structure must be of substantial and positive biological importance. Otherwise, they would be unnecessarily disruptive to the integrity of the molecules essential to life.

We have suggested and continue to suggest (1-9) that this value resides in the ability of amide residues to serve as molecular clocks for the timing of biological processes.

This hypothesis originally presupposed (1-4) that there were not other undiscovered properties of Asn and Gln of such great biological importance that their presence in proteins is necessary regardless of instability. It is now known that, while a large percentage of proteins deaminate during their *in vivo* lifetimes, this deamidation occurs in a small percentage of amide structures (8,9). If Asn and Gln are needed for other purposes, more stable configurations are not only available but also make up most of the available configurations.

Amide residues serve as biological molecular clocks. A few examples of these clocks are now known (10-17). We suggest that amide clocks will be found to be ubiquitous in living things.

The sequence dependent deamidation rates reported herein are necessary for the understanding of deamidation rates of peptides and proteins and for the continuing experimental test of the molecular clock hypothesis.

Acknowledgements: We thank Prof. R.B. Merrifield, in whose laboratory and with whose help we synthesized the peptides, Prof. Brian Chait for his advice and counsel concerning the mass spectrometry, and Prof. H.B. Gray for advice and encouragement. A Caltech Special Institute Fellowship and grants to the Oregon Institute of Science and Medicine from the John Kinsman Foundation and the A. Reynolds Morse Foundation provided financial support.

References

- Robinson, A.B., McKerrow, J.H. & Cary, P. (1970) Controlled deamidation of peptides and proteins: an experimental hazard and a possible biological timer. *Proc. Natl. Acad. Sci. USA* **66**, 753-757.
- Robinson, A.B. (1974) Evolution and the distribution of glutamyl and asparaginyl residues in proteins. *Proc. Natl. Acad. Sci. USA* **71**, 885-888.
- Robinson, A.B. & Rudd, C.J. (1974) Deamidation of glutamyl and asparaginyl residues in peptides and proteins. *Curr. Topics Cell. Reg.* **8**, 247-295.
- Robinson, A.B. & Scotchler, J.W. (1974) Sequence-dependent deamidation rates for model peptides of histone IV. *Int. J. Peptide Protein Res.* **6**, 279-282.
- Robinson, A.B. & Robinson, L.R. (1991) Distribution of glutamine and asparagine residues and their near neighbors in peptides and proteins. *Proc. Natl. Acad. Sci. USA* **88**, 8880-8884.
- Robinson, N.E. & Robinson, A.B. (2001) Molecular clocks. *Proc. Natl. Acad. Sci. USA* **98**, 944-949.
- Robinson, N.E. & Robinson, A.B. (2001) Prediction of protein deamidation rates from primary and three-dimensional structure. *Proc. Natl. Acad. Sci. USA* **98**, 4367-4372.
- Robinson, N.E. & Robinson, A.B. (2001) Deamidation of human proteins. *Proc. Natl. Acad. Sci. USA* **98**, 12409-12413.
- Robinson, N.E. (2002) Protein deamidation. *Proc. Natl. Acad. Sci. USA* **99**, 5283-5288.
- Flatmark, T. & Sletten, K. (1968) Multiple forms of cytochrome c in the rat. *J. Biol. Chem.* **243**, 1623-1629.
- Robinson, A.B., McKerrow, J.H. & Legaz, M. (1974) Sequence-dependent deamidation rates for model peptides of cytochrome c. *Int. J. Peptide Protein Res.* **6**, 31-35.
- Midelfort, C.F. & Mehler, A.H. (1972) Deamidation *in vivo* of an asparagine residue of rabbit muscle aldolase. *Proc. Natl. Acad. Sci. USA* **69**, 1816-1819.

13. McKerrow, J.H. & Robinson, A.B. (1974) Primary sequence-dependence of the deamidation of rabbit muscle aldolase. *Science* **183**, 85.
14. Lai, C.Y., Chen, C. & Horecker, B.L. (1970) Primary structure of two COOH-terminal hexapeptides from rabbit muscle aldolase: a difference in the structure of the α and the β subunits. *Biochem. Biophys. Res. Comm.* **40**, 461–468.
15. Sun, A., Yüksel, K.U. & Gracy, R.W. (1995) Terminal marking of triosephosphate isomerase: consequences of deamidation. *Arch. Biochem. Biophys.* **322**, 361–368.
16. Gracy, R.W., Talent, J.M. & Zvaigzne, A.I. (1998) Molecular wear and tear leads to terminal marking and the unstable isoforms of aging. *J. Exp. Zoology* **282**, 18–27.
17. Deverman, B.E., Cook, B.L., Manson, S.R., Niederhoff, R.A., Langer, E.M., Rosová, I., Kulans, L.A., Fu, X., Weinberg, J.S., Heinecke, J.W., Roth, K.A. & Weintraub, S.J. (2002) Bcl-X_L deamidation is a critical switch in the regulation of the response to DNA damage. *Cell* **111**, 51–62.
18. Capasso, S. & Salvadori, S. (1999) Effect of the three-dimensional structure on the deamidation reaction of ribonuclease A. *J. Peptide Res.* **54**, 377–382.
19. Solstad, T. & Flatmark, T. (2000) Microheterogeneity of recombinant human phenylalanine hydroxylase as a result of nonenzymatic deamidations of labile amide containing amino acids. *Eur. J. Biochem.* **267**, 6302–6310.
20. Palmer, W.G. & Papaconstantinou, J. (1969) Aging of α -crystallins during development of the lens. *Proc. Natl. Acad. Sci. USA* **64**, 404–410.
21. Takamoto, L. & Boyle, D. (1998) Deamidation of specific glutamine residues from α A-crystallin during aging of the human lens. *Biochemistry* **37**, 13681–13685.
22. Takamoto, L. (1998) Quantitation of asparagine-101 deamidation from α A-crystallin during aging of the human lens. *Current Eye Res.* **17**, 247–250.
23. Lapko, V.N., Purkiss, A.G., Smith, D.L. & Smith, J.B. (2002) Deamidation in human γ S-crystallin from cataractous lenses is influenced by surface exposure. *Biochemistry* **41**, 8638–8648.
24. Sondheimer, E. & Holley, R.W. (1954) Imides from asparagine and glutamine. *J. Am. Chem. Soc.* **76**, 2467–2470.
25. Meinwald, Y.C., Stimson, E.R. & Scheraga, H.A. (1986) Deamidation of the asparaginyl-glycyl sequence. *Int. J. Pept. Protein Res.* **28**, 79–84.
26. Geiger, T. & Clarke, S. (1987) Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. *J. Biol. Chem.* **262**, 785–794.
27. Tam, J.P., Riemen, M.W. & Merrifield, R.B. (1988) Mechanisms of aspartamide formation: the effects of protecting groups, acid, base, temperature, and time. *Peptide Res.* **1**, 6–18.
28. Capasso, S., Mazzarella, L. & Zagari, A. (1991) Deamidation via cyclic imide of asparaginyl peptides: dependence on salts, buffers, and organic solvents. *Peptide Res.* **4**, 234–238.
29. Robinson, N.E., Robinson, A.B. & Merrifield, R.B. (2001) Mass spectrometric evaluation of synthetic peptides as primary structure models for peptide and protein deamidation. *J. Peptide Res.* **57**, 483–493.
30. Robinson, A.B., Scotchler, J.W. & McKerrow, J. (1973) Rates of non-enzymatic deamidation of glutaminyl and asparaginyl residues in pentapeptides. *J. Am. Chem. Soc.* **95**, 8156–8159.
31. Capasso, S. (2000) Estimation of the deamidation rate of asparaginyl side chains. *J. Peptide Res.* **55**, 224–229.
32. Scotchler, J.W. & Robinson, A.B. (1971) Deamidation of glutaminyl residues: dependence on pH, temperature, and ionic strength. *Anal. Biochem.* **59**, 319–322.
33. McKerrow, J.H. & Robinson, A.B. (1971) Deamidation of asparaginyl residues as a hazard in experimental protein and peptide procedures. *Anal. Biochem.* **42**, 565–568.
34. Merrifield, R.B. (1963) Solid phase peptide synthesis I. Synthesis of a tetrapeptide. *J. Am. Chem. Soc.* **85**, 2149–2154.
35. Merrifield, R.B. (1995) Solid phase peptide synthesis. In: *Peptides: Synthesis, Structures, and Applications* (Gutte, B., ed.), pp. 94–169. Academic Press, New York.
36. Capasso, S., Mazzarella, L., Sorrentino, G., Balboni, G. & Kirby, A.J. (1996) Kinetics and mechanism of the cleavage of the peptide bond next to asparagine. *Peptides* **17**, 1075–1077.
37. Capasso, S., Mazzarella, L., Sica, F. & Zagari, A. (1991) First evidence of spontaneous deamidation of glutamine residue via cyclic imide to α - and γ -glutamic residue under physiological conditions. *J. Chem. Soc. Chem. Commun.* **23**, 1667–1668.