

Crystal Structure of the Ligand-bound Glucagon-like Peptide-1 Receptor Extracellular Domain^{*[5]}

Received for publication, October 22, 2007, and in revised form, January 24, 2008. Published, JBC Papers in Press, February 20, 2008, DOI 10.1074/jbc.M708740200

Steffen Runge^{‡1}, Henning Thøgersen[‡], Kjeld Madsen[§], Jesper Lau[§], and Rainer Rudolph[¶]

From the [‡]Department of Structure and Biophysical Chemistry and [§]Department of Protein and Peptide Chemistry, Novo Nordisk, 2760 Måløv, Denmark and [¶]Institute of Biotechnology, Martin Luther University Halle-Wittenberg, 06120 Halle (Saale), Germany

The glucagon-like peptide-1 receptor (GLP-1R) belongs to Family B1 of the seven-transmembrane G protein-coupled receptors, and its natural agonist ligand is the peptide hormone glucagon-like peptide-1 (GLP-1). GLP-1 is involved in glucose homeostasis, and activation of GLP-1R in the plasma membrane of pancreatic β -cells potentiates glucose-dependent insulin secretion. The N-terminal extracellular domain (nGLP-1R) is an important ligand binding domain that binds GLP-1 and the homologous peptide Exendin-4 with differential affinity. Exendin-4 has a C-terminal extension of nine amino acid residues known as the “Trp cage”, which is absent in GLP-1. The Trp cage was believed to interact with nGLP-1R and thereby explain the superior affinity of Exendin-4. However, the molecular details that govern ligand binding and specificity of nGLP-1R remain undefined. Here we report the crystal structure of human nGLP-1R in complex with the antagonist Exendin-4 (9–39) solved by the multiwavelength anomalous dispersion method to 2.2 Å resolution. The structure reveals that Exendin-4 (9–39) is an amphipathic α -helix forming both hydrophobic and hydrophilic interactions with nGLP-1R. The Trp cage of Exendin-4 is not involved in binding to nGLP-1R. The hydrophobic binding site of nGLP-1R is defined by discontinuous segments including primarily a well defined α -helix in the N terminus of nGLP-1R and a loop between two antiparallel β -strands. The structure provides for the first time detailed molecular insight into ligand binding of the human GLP-1 receptor, an established target for treatment of type 2 diabetes.

G protein-coupled receptors represent the largest protein family encoded by the human genome, and they are defined by the presence of seven transmembrane helices, an extracellular N terminus, ligand binding via the extracellular face, an intracellular C terminus, G protein coupling, and signaling via the intracellular face. Receptors of the B1 subfamily are specifically

characterized by three conserved disulfide bonds in the N-terminal extracellular domain (Nt-domain)² (1–4) and by their structurally related peptide hormone ligands, such as glucagon-like peptide-1 (GLP-1), glucagon-like peptide-2 (GLP-2), glucagon, glucose-dependent insulinotropic polypeptide (GIP), secretin, vasoactive intestinal polypeptide, pituitary adenylate cyclase-activating polypeptide (PACAP), growth hormone-releasing hormone, parathyroid hormone, calcitonin, and corticotropin-releasing factor (CRF) (5). Peptide hormone binding of cloned Family B1 receptors has been investigated for 15 years by pharmacological and biochemical approaches. The current binding model is a two-step mechanism where initially the C-terminal part of the peptide ligand interacts with the Nt-domain of the receptor (6–8). In the second step, the N-terminal part of the ligand interacts with the core domain of the receptor (transmembrane helices and connecting loops), which leads to activation and signal transduction (9–14). More recently, it was proposed that the Nt-domain of the secretin receptor was involved in the activation mechanism, but such an endogenous agonist mechanism has not been confirmed for GLP-1R (15).

The isolated soluble Nt-domains are able to bind their cognate ligands, although the affinity is often reduced compared with the full-length receptors (1, 3, 4, 6). The GLP-1R Nt-domain is important for both ligand binding and specificity, and it determines almost exclusively the ability of the full-length GLP-1R to discriminate between glucagon and GLP-1 (16). This is physiologically important, given the essentially opposite effects of glucagon and GLP-1 on blood glucose. Detailed structural information exists for ligand-bound forms of the Nt-domain of the mouse CRF receptor 2 β (nCRF-R2 β), the human PACAP receptor (nPAC1), and the human GIP receptor (nGIPR) (17–19). The structures confirmed both the existence of a common structural fold and the interaction with the C-terminal part of their cognate ligands. Surprisingly, the binding site and orientation of PACAP(6–38) was completely different from the binding site and orientation of astressin and GIP.

The soluble refolded nGLP-1R binds GLP-1 with lower affinity than the full-length GLP-1R in membranes, suggesting that high affinity binding of GLP-1 requires additional interaction with the core domain of GLP-1R (IC₅₀ of 398 and 1.0 nM, respectively) (6). Specifically, the N-terminal part of GLP-1 is

^{*} This work was supported by the European Membrane Protein Consortium (E-Mep), the federal state of Saxony-Anhalt (3324 A/0021 L), and the Deutsche Forschungsgemeinschaft (SFRB 610/TP A11). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 3C59) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

[5] The on-line version of this article (available at <http://www.jbc.org>) contains supplemental Figs. S1–S3.

¹ To whom correspondence should be addressed: Novo Nordisk Park G8.S.439, DK-2760 Måløv, Denmark. Tel.: 4544434431; Fax: +4544422284; E-mail: sfrr@novonordisk.com.

² The abbreviations used are: Nt-domain, N-terminal extracellular domain; GIP, glucose-dependent insulinotropic polypeptide; PACAP, pituitary adenylate cyclase-activating polypeptide; CRF, corticotropin-releasing factor; Ex4, Exendin-4; TFE, trifluoroethanol; SeMet, selenomethionine; Fmoc, N-(9-fluorenyl)methoxycarbonyl.

expected to interact with the first extracellular loop (ECL1) and the extracellular end of transmembrane helix 2 of GLP-1R (10, 13, 14). Exendin-4 (Ex4) is homologous to GLP-1, and it binds and activates GLP-1R with similar affinity and potency as GLP-1 (20). In contrast to GLP-1, Ex4 maintains high affinity for nGLP-1R, comparable with that of the full-length GLP-1R (1.0 and 1.6 nM, respectively) (6). Ex4 has a C-terminal extension of nine amino acid residues known as the Trp cage, which is absent in GLP-1. The NMR structure of Ex4 in aqueous trifluoroethanol (TFE) showed that the Trp cage folds back onto the central α -helical part of Ex4, forming the smallest known protein-like fold (21, 22). The Trp cage was suggested to interact with nGLP-1R, thereby explaining the superior affinity of Ex4 compared with GLP-1 (23). However, recent results suggest that the Trp cage plays only a minor role in receptor binding (24). Instead, the superior affinity of Ex4 was explained concomitantly by its superior α -helical propensity in solution and by superior interactions, due to specific divergent residues in the C-terminal part of GLP-1 and Ex4 (24). Therefore, the role of the Trp cage in receptor binding is not crystal clear.

Exendin-4(9–39) (Ex4(9–39)) is a competitive antagonist of GLP-1R, displacing both GLP-1 and Ex4 from receptor binding due to its high affinity interaction with nGLP-1R (IC_{50} of 0.6 nM) (6, 25). To increase the knowledge about ligand binding of GLP-1R and Family B1 receptors in general, we solved the crystal structure of the complex between nGLP-1R and Ex4(9–39) to 2.1 Å resolution. The structure provides for the first time molecular details of the initial ligand binding step of GLP-1R, and it gives a structural explanation of the differential affinity of nGLP-1R toward GLP-1 and Ex4.

EXPERIMENTAL PROCEDURES

Protein and Peptide Preparation—The nGLP-1R was prepared as earlier described (24). Briefly, N-terminal His₆-tagged nGLP-1R was expressed in *Escherichia coli* inclusion bodies, isolated as inclusion body protein, solubilized in guanidine HCl and dithiothreitol, dialyzed against guanidine HCl to remove the dithiothreitol, and refolded using L-Arg and a 1:5 molar ratio of reduced and oxidized glutathione. Refolded nGLP-1R was purified by hydrophobic interaction chromatography and size exclusion chromatography. The His₆ tag was removed by thrombin cleavage.

Native Ex4(9–39)-amide was synthesized as previously described (16). [SeMet^{14,21}]Ex4(9–39)-amide was synthesized by incorporation of selenomethionine (SeMet) residues using the Fmoc strategy on Solid Phase Peptide Synthesis. Fmoc-SeMet-OH was prepared by dissolving L-(+)-selenomethionine (Acros Organics) in tetrahydrofuran/water with 1 equivalent of Na₂CO₃ and reacting with Fmoc-ONSu. The product was crystallized from ethyl acetate and n-heptan as off-white crystals with a slight odor of selenide.

Purification and Crystallization of Ligand-bound nGLP-1R—The purified nGLP-1R was concentrated to 1 mg/ml and mixed with 2-fold molar excess of [SeMet^{14,21}]Ex4(9–39) or Ex4(9–39). Ligand-bound nGLP-1R was purified by size exclusion chromatography using a Superdex75 column running in 10 mM Tris-HCl, pH 7.5 (supplemental Fig. S1). The eluted complex was characterized by SDS-PAGE and Trp fluores-

cence spectroscopy (data not shown). The complex was concentrated to 7 mg/ml and crystallized by hanging drop vapor diffusion. The crystallization conditions was identified initially using the Crystal Screen and Detergent Screen from Hampton Research and subsequently optimized to 0.1 M Tris-HCl, pH 8.5, 0.1 M MgCl₂, 0.4 M magnesium tartrate, and 9 mM n-decyl- β -D-thiomaltoside. Single crystals grew to a size of $\sim 0.5 \times 0.1 \times 0.1 \mu\text{m}$, and they were frozen in liquid N₂ using 25% glycerol in the cryo solution.

Data Collection and Structure Determination—Diffraction data were collected from a single crystal at three different wavelengths (peak, remote, inflection) at the MAX-lab beamline I911–3 (Lund, Sweden). The data were integrated and scaled using XDS (26); the crystals belonged to space group P3₁21, with the unit cell dimensions $a = b = 75.9 \text{ Å}$ and $c = 87.6 \text{ Å}$. The two selenium sites of [SeMet^{14,21}]Ex4(9–39), the phases and initial electron density map were calculated using autoSHARP (27). The initial electron density map was of excellent quality and allowed automated building of 65% of the structural model by ARP/wARP (28) (in the autoSHARP interface) with one complex in the asymmetric unit. The rest of the structural model was built manually without any ambiguity and refined (TLS and restrained) using COOT (29), REFMAC5 (30), and the CCP4 program suite (31). The final SeMet model has 124 residues in allowed and six residues in disallowed regions of the Ramachandran plot, a working R -factor of 20.4%, and a free R -factor of 25.0%. Data collection, phasing, and refinement statistics are summarized in Table 1. Coordinates and structure factors are deposited in the Protein Data Bank under accession code 3C59. Molecular graphics were prepared in PyMOL.

RESULTS AND DISCUSSION

Crystallization and Structure Determination—The nGLP-1R was expressed in *E. coli* inclusion bodies, refolded, and purified as described previously (24). We synthesized a selenium derivative of the ligand, where Met¹⁴ and Leu²¹ were substituted by SeMet: [SeMet^{14,21}]Ex4(9–39). The solution NMR structure of Ex4 showed previously that Met¹⁴ and Leu²¹ were on the same face of the α -helix of Ex4 (22). In addition, structure/activity studies of GLP-1 showed that these two positions were not important for receptor binding (32). Indeed, [SeMet^{14,21}]Ex4(9–39) displaced ¹²⁵I-Ex4(9–39) from binding to the refolded nGLP-1R, with the same IC_{50} value as native Ex4(9–39) ($IC_{50} = 5 \text{ nM}$, data not shown). The high affinity interaction allowed co-purification of nGLP-1R and [SeMet^{14,21}]Ex4(9–39) by size exclusion chromatography (supplemental Fig. S1). The intrinsic Trp fluorescence properties of the eluted protein suggested a 1:1 complex (data not shown), where the ligand provided the heavy atoms necessary for the experimental phasing. The complex was crystallized in hanging drops, and the structure was solved by the multi-wavelength anomalous dispersion method to 2.2 Å resolution, using the absorption edge of the two selenium atoms in [SeMet^{14,21}]Ex4(9–39). Subsequently, we crystallized and solved the structure of the complex with native Ex4(9–39) to 2.1 Å resolution by molecular replacement, using nGLP-1R from the initial SeMet structure as the search model (Protein Data Bank code 3C5T). The two structures were essentially

TABLE 1

Data and refinement statistics

Values in parentheses are for the highest resolution shell. Phasing statistics were extracted from the autoSHARP logfiles. Rmsd, root mean square deviation. FOM, figure of merit.

Crystal parameters		SeMet		Native	
Space group		P3 ₁ 21		P3 ₁ 21	
Unit cell dimensions (Å)		a = b = 75.9, c = 87.6		a = b = 76.0, c = 87.8	
Data collection		Peak	Inflection	Remote	
Wavelength (Å)		0.9782	0.9790	0.96	1.0
Resolution range (Å)		40–2.2 (2.3–2.2)	40–2.2 (2.2–2.1)	40–2.2 (2.2–2.1)	40–2.1 (2.2–2.1)
Unique reflections		28480	28468	28504	17273
Total reflections		165607	165747	165997	188373
Completeness (%)		99.4 (98.9)	99.4 (98.9)	99.4 (99.0)	98.3 (97.8)
I/σ(I)		19.0 (4.1)	18.7 (4.7)	15.9 (3.8)	27.1 (7.3)
R _{sym}		7.4 (56.1)	6.3 (42.3)	7.7 (53.3)	6.1 (39.5)
Phasing statistics		Peak	Inflection	Remote	
Number of selenium sites			2		
Phasing power anomalous		0.641	0.533	0.542	
Phasing power isomorphous (acentric/centric)		–	0.268/0.259	0.444/0.476	
R _{cullis} anomalous		0.902	0.933	0.937	
R _{cullis} isomorphous (acentric/centric)		–	0.606/0.597	0.731/0.700	
Figure of merit (acentric/centric)			0.295/0.227		
Refinement statistics		SeMet		Native	
Number of atoms (non-hydrogen)		1194		1257	
Resolution		2.2		2.1	
R _{work} /R _{free}		0.204/0.250		0.205/0.235	
Rmsd bond length (Å)		0.027		0.018	
Rmsd angles (°)		2.3		2.1	
Average B factors		40.1		39.3	

identical except for Met¹⁴ and Leu²¹ of Ex4(9–39). Data collection and refinement statistics are summarized in Table 1, and all the figures illustrate the structure with the native ligand.

The Structure of nGLP-1R—The core structure of nGLP-1R is similar to that of nCRF-R2β, nPAC1, and nGIPR, showing the three conserved disulfide bonds, two regions of antiparallel β-sheets (β-strand 1–5), and centrally positioned conserved residues Asp⁶⁷, Trp⁷², Pro⁸⁶, Arg¹⁰², Gly¹⁰⁸, and Trp¹¹⁰ of nGLP-1R (Fig. 1, A and B) (18, 19). The N terminus of nGLP-1R contains an α-helix similar to that of nGIPR, and the tertiary structure is stabilized by the disulfide bridges and by multiple intramolecular interactions between the secondary structure elements. The α-helix is defined by residues Leu³²–Glu⁵² and includes Cys⁴⁶, which forms a disulfide bridge with Cys⁷¹ on β-strand 2 (β2). The α-helix is terminated by three successive Pro residues (Pro⁵⁴–Pro⁵⁶) positioned at the beginning of loop 1 (Pro⁵⁴–Phe⁶¹). The residues Cys⁶²–Asp⁶⁷ (β1) and Ala⁷⁰–Gly⁷⁵ (β2) constitute the first antiparallel β-sheet, although the backbone geometry is distorted around Arg⁶⁴ and Pro⁷³. β1 and β2 are separated by a short turn (turn 1, Glu⁶⁸ and Tyr⁶⁹). The side chain of Arg⁶⁴ interacts with the dipole of the α-helix through hydrogen bonds with the backbone of Leu⁵⁰ and Asp⁵³ (supplemental Fig. S2). The positioning of the Arg⁶⁴ side chain is further guided by a hydrogen bond with the backbone carbonyl of Pro⁵⁴ and by an ionic interaction with Asp⁷⁴ of β2. The residues Gly⁷⁸–Ser⁸⁴ (β3), His⁹⁹–Thr¹⁰⁵ (β4), Leu¹⁰⁹–Leu¹¹¹ (β5), and Asp¹²² (β-bridge with β4) constitute the second region of antiparallel β-sheets. β3 and β4 are separated by a long well defined loop (loop 2, Pro⁸⁶–Gly⁹⁸) that is important for ligand binding (see below). The disulfide bridge between Cys⁶² and Cys¹⁰⁴ connects the beginning of β1 and the end of β4. Cys⁸⁵ is positioned at the end of β3 and forms a disulfide bridge with Cys¹²⁶ in the C-terminal part of nGLP-1R. The segment from Leu¹⁰⁹–Asp¹²² resembles a β-strand with an insertion of a flexible loop

(loop 3, Gln¹¹²–Leu¹¹⁸). There was no detectable electron density for residues Arg²⁴–Gln²⁷ and Gly¹³²–Tyr¹⁴⁵.

The human nGLP-1R contains seven Trp residues, and substitution of Trp³⁹, Trp⁷², Trp⁹¹, Trp¹¹⁰, and Trp¹²⁰ by Ala resulted in complete loss of binding of the full-length rat GLP-1R, whereas substitution of Trp⁸⁷ had no apparent effect on binding or activation (33). The role of Trp³³, Trp³⁹, Trp⁷², Trp⁹¹, and Trp¹¹⁰ is discussed below. Trp¹²⁰ is not involved in ligand binding but appears to play a structural role by forming a well defined surface-exposed hydrophobic cluster together with Phe⁸⁰, Tyr¹⁰¹, Phe¹⁰³, and Leu¹¹¹.

Conserved Residues in the Nt-domain of Family B1 Receptors—The conserved Asp of nGLP-1R (Asp⁶⁷) is centrally positioned, forming intramolecular interactions, as also observed for nCRF-R2β, nPAC1, and nGIPR, although the molecular details are slightly different in our crystal structure. Here, the side chain of Asp⁶⁷ interacts indirectly via a water molecule with the side chain of Arg¹⁰² and directly with the side chain of Trp⁷² and Arg¹²¹ (Fig. 1B). In addition, the side chain and backbone of Asp⁶⁷ interact with Tyr⁶⁹ and Ala⁷⁰, stabilizing the turn between β-strand 1 and 2. Gly¹⁰⁸ has a structural function stabilizing the turn after β-strand 4 that allows the positioning of the Trp¹¹⁰ side chain below the Cys⁶²–Cys¹⁰⁴ disulfide bridge and above Arg¹⁰². Thereby, Arg¹⁰² is sandwiched between the side chains of Trp⁷² and Trp¹¹⁰ in a manner similar to nCRF-R2β. The importance of Asp⁶⁷, Trp⁷², and Trp¹¹⁰ for ligand binding was previously documented by receptor mutagenesis (33, 34). Pro⁸⁶ at the beginning of loop 2 plays a structurally important role for the formation of the ligand binding site of nGLP-1R (see below). The side chain of Pro⁸⁶ fills out a hydrophobic cavity formed by Tyr⁴² of the α-helix, Tyr⁶⁹ of turn 1, Ala⁷⁰ of β-strand 2, Val⁸³ of β-strand 3, Val¹⁰⁰ of β-strand 4, the Cys⁸⁵–Cys¹²⁶ disulfide bridge, and two residues of loop 2 itself, Tyr⁸⁸ and Leu⁸⁹ (supplemental Fig. S3).

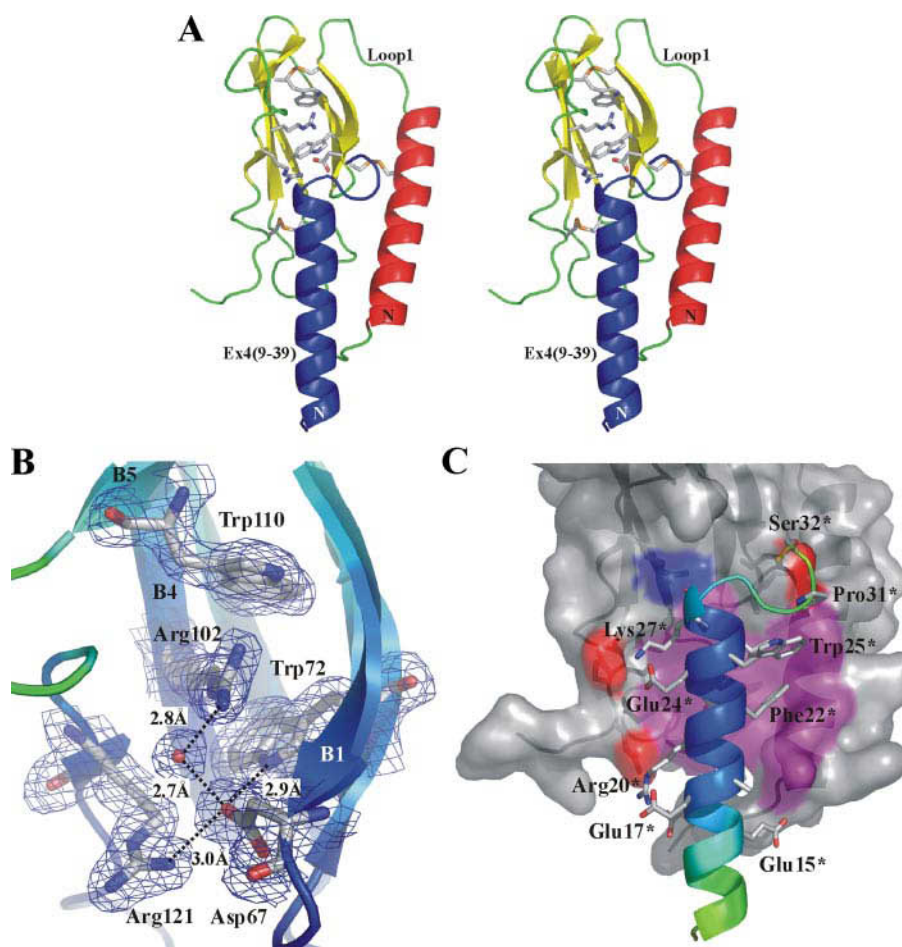


FIGURE 1. A, stereo ribbon diagram illustrating the structure of nGLP-1R (helix, red; β -strands, yellow; loops, green) in complex with the antagonist Exendin-4(9–39) shown in blue. The disulfide bridges and the residues Asp⁶⁷, Trp⁷², Arg¹⁰², Trp¹¹⁰, and Arg¹²¹ are shown in sticks colored by atoms. B, Asp⁶⁷, Trp⁷², Arg¹⁰², Trp¹¹⁰, and Arg¹²¹ of nGLP-1R and the coordinated water molecule are illustrated by stick representation and $2F_o - F_c$ electron density in blue, contoured at 1.0 σ level. Secondary structure elements are colored according to B-factors. C, surface representation of nGLP-1R in gray, highlighting the carboxyl group of Glu⁶⁸, Glu¹²⁷, and Glu¹²⁸ in red, the guanidinium group of Arg¹²¹ in blue, and the hydrophobic binding surface in magenta: Leu³², Thr³⁵, Val³⁶, Trp³⁹, Tyr⁶⁹, Tyr⁸⁸, Leu⁸⁹, Pro⁹⁰, Trp⁹¹, and Leu¹²³. The secondary structure representation of Ex4(9–39) is colored according to B-factors (blue = 20–25 and green-yellow = 45–60). The side chains of Glu^{15*}, Glu^{16*}, Glu^{17*}, Ala^{18*}, Arg^{20*}, Phe^{22*}, Glu^{24*}, Trp^{25*}, Leu^{26*}, Lys^{27*}, Pro^{31*}, and Ser^{32*} of Ex4(9–39) are shown by sticks colored according to atom type.

The Ligand and Its Interactions with nGLP-1R—Ex4(9–39) is a well defined α -helix from Leu^{10*} to Asn^{28*}, and the residues that interact with nGLP-1R lie within Glu^{15*} and Ser^{32*} (Fig. 1C) (throughout the text, ligand residues are designated with * and Ex4 is numbered 1–39). The residues Asp^{9*}–Met^{14*} do not interact with nGLP-1R. Glu^{15*} is conserved in GLP-1 and Ex4, and it interacts favorably with the dipole of the α -helix of nGLP-1R via a hydrogen bond with the backbone amide of Leu³², the first residue in the α -helix of nGLP-1R. The amphipathic nature of the Glu^{16*}–Lys^{27*} segment enables hydrophobic interactions with nGLP-1R via one face of the α -helix and hydrophilic interactions via the other face of the α -helix. The hydrophilic face is defined by Glu^{16*}, Glu^{17*}, Arg^{20*}, Glu^{24*}, and Lys^{27*}, of which only Arg^{20*} and Lys^{27*} interact directly with nGLP-1R (Fig. 2A). The side chain of Arg^{20*} binds to the side chain of Glu¹²⁸ of nGLP-1R, and the geometry is good for an ideal ionic interaction with two hydrogen bonds. Arg^{20*} is nicely positioned by Glu^{16*} and Glu^{17*} in

an arrangement that could stabilize the α -helical conformation of Ex4(9–39) itself, also in solution. In a similar arrangement, the side chain of Lys^{27*} appears to interact with both the side chain of Glu¹²⁷ of nGLP-1R and the side chain of Glu^{24*}. In addition, the backbone carbonyl of Lys^{27*} at the end of the α -helix forms a hydrogen bond with the side chain of Arg¹²¹ of nGLP-1R. After Lys^{27*}, the B-factors were increasing toward the C terminus of the ligand and truncation of Pro^{31*}–Ser^{39*} showed only a minor effect on binding to nGLP-1R, suggesting that the most critical interactions lie within Glu^{15*} and Lys^{27*} (24). Nevertheless, the structure suggests that the side chain of Ser^{32*} could form a hydrogen bond with Glu⁶⁸ of nGLP-1R.

The C-terminal tail after Ser^{32*} of receptor-bound Ex4(9–39) probably folds back on the α -helix in a manner similar to the Trp cage conformation of Ex4 in TFE solution (22). However, the electron density was very weak for the C terminus of Ex4(9–39) (Gly^{34*}–Ser^{39*}), indicating high mobility, and we were unable to build a representative structure of this part of the ligand. Accordingly, specific mutations in the C terminus of Ex4 were necessary to stabilize the Trp cage in aqueous solution, where it was otherwise only partially populated (21). Our results suggest that the specific Trp cage conformation is not stabilized in the receptor-bound state of Ex4(9–39) and that Ser^{33*}–Ser^{39*} does not interact specifically with nGLP-1R.

The hydrophobic face of Ex4(9–39) is defined by residues Ala^{18*}, Val^{19*}, Phe^{22*}, Ile^{23*}, Trp^{25*}, Leu^{26*}, and Pro^{31*} (Fig. 2B). Of these, Val^{19*}, Phe^{22*}, Ile^{23*}, and Leu^{26*} are buried by the hydrophobic interaction with nGLP-1R and they are the residues of the ligand with the lowest B-factors. Phe^{22*} of Ex4(9–39) is uniquely conserved in the glucagon peptide subfamily (Fig. 2D, GLP-1, Exendin-4, GLP-2, Glucagon, and GIP), and it interacts directly with Leu³², Thr³⁵, Val³⁶, and Trp³⁹ on the α -helix of nGLP-1R. The importance of Phe^{28*} and Trp³⁹ was demonstrated previously by alanine scanning of GLP-1 and by site-directed mutagenesis of GLP-1R (32, 35). The Phe^{28*}–Ala substitution of GLP-1 had by far the strongest impact in the Ala scan on the binding to the full-length GLP-1R. Ala substitution of Ile^{29*} and Leu^{32*} also reduced the affinity of GLP-1 significantly, although the effect was smaller than that of Phe^{28*}–Ala. Gly^{30*} and Pro^{31*} are stacking on Trp^{25*} and they are on the edge

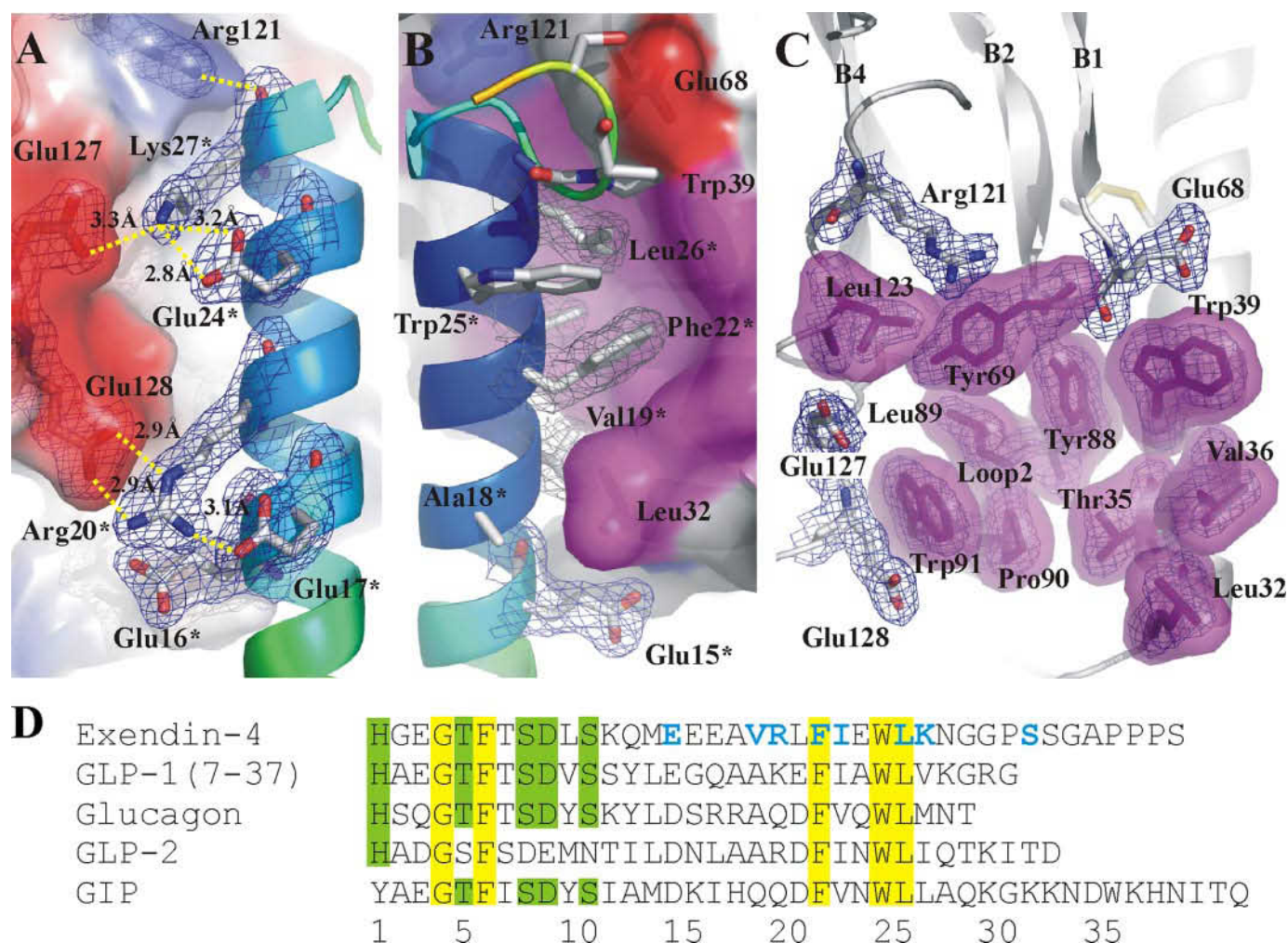


FIGURE 2. Hydrophilic and hydrophobic interactions between Ex4(9–39) and nGLP-1R. The ribbon diagram of Ex4(9–39) is colored according to B-factors. A, the residues Glu^{16*}, Glu^{17*}, Arg^{20*}, Glu^{24*}, and Lys^{27*} are illustrated as sticks, and the electron density is shown in blue. The residues Arg¹²¹, Glu¹²⁷, and Glu¹²⁸ of nGLP-1R are illustrated as sticks, showing also the electron density and charge density surface. B, the side chains of Glu^{15*}, Ala^{18*}, Val^{19*}, Phe^{22*}, Ile^{23*}, Trp^{25*}, Leu^{26*}, Pro^{31*}, and Ser^{32*} are illustrated as sticks. The electron density is shown of Glu^{15*} in blue and of Val^{19*}, Phe^{22*}, and Leu^{26*} in gray. The surface of the hydrophobic binding cavity of nGLP-1R is illustrated in magenta. The surface of Glu⁶⁸ and Arg¹²¹ is illustrated in red and blue, respectively. C, all the residues of nGLP-1R directly involved in binding of Ex4(9–39) are illustrated as sticks with electron density. In addition, the hydrophobic residues are highlighted by a magenta surface representation. Leu³², Thr³⁵, Val³⁶, and Trp³⁹ belong to the α -helix of nGLP-1R and Tyr⁸⁸, Leu⁸⁹, Pro⁹⁰, and Trp⁹¹ belong to loop 2 of nGLP-1R. The β -strands are labeled B1–B5. D, sequence alignment of the human glucagon peptide subfamily. The residues of Ex4(9–39) that interact with nGLP-1R are colored blue. Conserved residues are colored yellow. Positions where only four residues are conserved are colored green. All electron density is shown as $2F_o - F_c$ contoured at the 1.0 σ level. Stick representations are colored according to atom type.

of the hydrophobic interface with nGLP-1R. Trp^{25*} is not directly involved in binding to nGLP-1R, and the Trp^{31*}-Ala substitution of GLP-1 gave only a minor loss of binding affinity toward the full-length GLP-1R accordingly (32).

We previously used the intrinsic Trp fluorescence to investigate ligand binding of nGLP-1R (24). The results suggested that one or more Trp residues were buried in a hydrophobic environment upon binding of Ex4(9–39) to nGLP-1R. The hydrophobic binding cavity of nGLP-1R is formed by residues of four discontinuous receptor segments: Leu³², Thr³⁵, Val³⁶, and Trp³⁹ of the α -helix; Tyr⁶⁹ of turn 1; Tyr⁸⁸-Trp⁹¹ of loop 2; and Leu¹²³ just before the Cys⁸⁵-Cys¹²⁶ disulfide bridge (Fig. 2C). Trp³⁹ and Trp⁹¹ are directly at the hydrophobic interface with Ex4(9–39) and clearly protected from solvent exposure by ligand binding. The hydrophobic interaction between nGLP-1R and Ex4(9–39) defines an area of the structure where water molecules are absent, suggesting a strong hydrophobic

effect. The hydrophobic effect was also a strong contributor to the interaction between astressin and nCRF-R2 β (18). Specifically, loop 2 of unbound nCRF-R2 β was very flexible and adopted a well defined conformation upon ligand binding (18). Superposition of the ligand-bound nGLP-1R and unbound nCRF-R2 β suggests that loop 2 of nGLP-1R is pulled toward Ex4(9–39), engaging in the hydrophobic interaction by a similar induced fit mechanism. The conserved Pro⁸⁶ of nGLP-1R is probably important for this binding mechanism and for the specific conformation of loop 2. In the complex with Ex4(9–39), the side chain of Pro⁸⁶ fills out a hydrophobic cavity on the backside of the hydrophobic ligand binding site. Specifically, Tyr⁸⁸ of loop 2 is positioned between Leu^{26*} of Ex4(9–39) and Pro⁸⁶ of nGLP-1R. The conserved proline holds a similar position in ligand-bound nCRF-R2 β and nGIPR. The ligand-induced conformational change of loop 2 might alter the fluorescence properties of Trp⁸⁷ of nGLP-1R, although it is rather

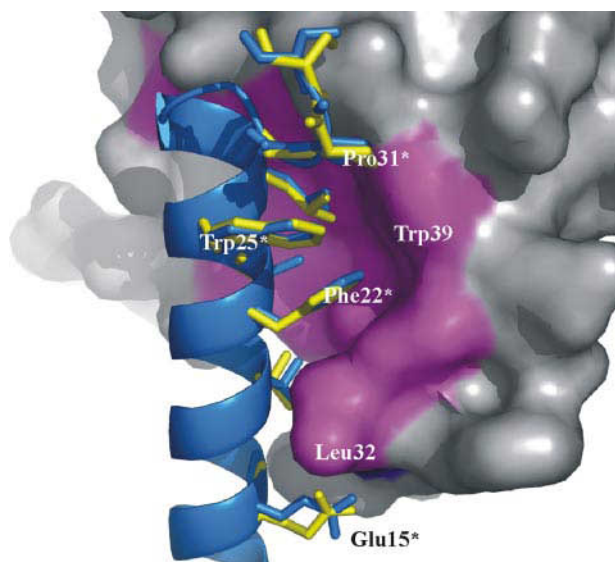


FIGURE 3. Superposition of the solution NMR structure of Ex4 in TFE (Protein Data Bank code 1JRJ) and the receptor-bound crystal structure of Ex4(9–39). The surface of nGLP-1R is illustrated in gray, highlighting the hydrophobic binding cavity in magenta. The secondary structure of receptor-bound Ex4(9–39) is illustrated in blue. The side chains of Glu^{15*}, Val^{19*}, Phe^{22*}, Ile^{23*}, Trp^{25*}, Leu^{26*}, Pro^{31*}, and Ser^{32*}, of receptor-bound Ex4(9–39) and Ex4 in solution are shown in blue and yellow sticks, respectively.

surface-exposed also in the ligand-bound conformation shown here. The extent to which Trp^{25*} of Ex4(9–39) itself changes its fluorescence properties upon receptor binding is difficult to evaluate. However, Trp^{25*} is on the edge of the hydrophobic interface with nGLP-1R and it is protected from direct surface exposure only if the Trp cage is populated. C-terminal truncation of Ex4 altered the fluorescence properties of ligand-bound nGLP-1R (24). It seems that the Trp cage influences the fluorescence properties of Ex4 itself while bound to nGLP-1R, supporting the hypothesis that the Trp cage is partially populated in the complex between Ex4(9–39) and nGLP-1R.

Crystal Packing—Four Arg residues on the α -helix of nGLP-1R played a major role in crystal packing, forming ionic interactions and hydrogen bonds with a symmetry-related receptor molecule. In addition, the N-terminal part of Ex4(9–39) packed against a symmetry-related receptor molecule. We co-purified, crystallized, and solved the structure of nGLP-1R bound to Ex4(5–39) by molecular replacement to 2.1 Å resolution, using nGLP-1R from the initial SeMet structure as the search model (data not shown). The four additional residues at the N terminus of the ligand altered the packing; as a result, the complex crystallized in space group P2₁2₁2₁, with two complexes in the asymmetric unit. Nevertheless, the overall structure and the ligand-receptor interactions were not influenced by the new crystal packing.

Structural Properties and Differential Affinity of GLP-1 and Exendin-4—Superposition of unbound Ex4 in TFE and receptor-bound Ex4(9–39) in crystals illustrates a striking similarity (Fig. 3). The conformation of the hydrophobic residues Val^{19*}, Phe^{22*}, Trp^{25*}, and Leu^{26*} of Ex4 in TFE (Protein Data Bank code 1JRJ) is well defined and almost identical with their receptor-bound conformation shown here (22). The similarity extends further to the structure around Gly^{29*}, Gly^{30*}, and

Pro^{31*} that allows formation of the Trp cage in TFE solution. However, the absence of the Trp cage in our crystal structure strongly suggests that the Trp cage does not bind to nGLP-1R, and this is supported by the ligand binding properties of nGLP-1R (24). The hydrophobic residues Phe^{22*}, Ile^{23*}, Trp^{25*}, and Leu^{26*} are conserved in GLP-1 and Ex4, and the competitive nature of their binding to nGLP-1R and GLP-1R indicates that the two peptides share the same hydrophobic binding site of nGLP-1R. We previously showed a correlation between the α -helical propensity of peptide ligands in solution and their affinity for nGLP-1R (24). Therefore, the differential affinity of Ex4 and GLP-1 for nGLP-1R is best explained by divergent residues in the central part of the peptides, Glu^{15*}–Gly^{30*} in Ex4 and Glu^{21*}–Arg^{36*} in GLP-1. Specifically, Arg^{20*} and Lys^{27*} of Ex4 have a dual function by binding directly to nGLP-1R and by stabilizing the α -helical conformation of Ex4 itself via Glu^{16*}, Glu^{17*}, and Glu^{24*}. In addition, Val^{19*} on the hydrophobic face of Ex4(9–39) forms a hydrophobic interaction with Leu³², Thr³⁵, and Pro⁹⁰ of nGLP-1R. The corresponding divergent residues in GLP-1 (Gly^{22*}, Gln^{23*}, Ala^{25*}, Lys^{26*}, Ala^{30*}, and Val^{33*}) most likely explain its lower α -helical propensity in solution as described by NMR and CD spectroscopy (22, 36) and its lower affinity for nGLP-1R.

In conclusion, the structural characterization of GLP-1 and Ex4 in solution and the receptor-bound crystal structure of Ex4(9–39) show that the superior affinity of nGLP-1R for Ex4 is determined concomitantly by superior hydrophilic and hydrophobic interactions and by a superior α -helical propensity of Ex4 in solution, thus lowering the entropic cost of receptor binding, compared with GLP-1. This explanation also accounts for the ability of Ex4(9–39) to maintain high affinity for GLP-1R in contrast to N-terminal-truncated GLP-1 analogs, *i.e.* Ex4(9–39) is a potent antagonist and GLP-1(9–36) is not (37).

The Glucagon Receptor Branch—The GLP-1R, GLP-2R, GluR, and GIPR define the glucagon receptor branch within Family B1. The crystal structures of ligand-bound nGLP-1R and nGIPR are very similar (root mean square deviation = 0.84 Å for C α atoms of 92 aligned residues of the receptors alone), and the hydrophobic interaction involves conserved residues of both the ligands and the receptors (17). Tyr⁴², Phe⁶⁶, Pro⁹⁰, and Trp⁹¹ of GLP-1R are exclusively conserved in the glucagon receptor branch and thus likely to define specificity for the ligands of the glucagon peptide subfamily. Pro⁹⁰ and Trp⁹¹ are directly at the hydrophobic interface with Val^{19*} and Ile^{23*} of Ex4(9–39). Tyr⁴² and Phe⁶⁶ are not directly involved in ligand binding, and they could instead modulate the binding site of the Nt-domain by intramolecular interactions in order to accommodate specifically the glucagon peptide subfamily. The role of Tyr⁴², Phe⁶⁶, Pro⁹⁰, and Trp⁹¹ for ligand binding and specificity of GLP-1R remains to be characterized by receptor mutagenesis.

Within the glucagon receptor branch, divergent residues determine specificity for the ligands. The ligand specificity of GLP-1R was shifted 50-fold in favor of glucagon by substitution of the residues Thr²⁹–Leu³² with the corresponding residues of the glucagon receptor (38). The ability of the Thr²⁹–Leu³² segment to establish ligand specificity of GLP-1R is corroborated by the crystal structure shown here, where specifically Leu³² of

Crystal Structure of the GLP-1R Extracellular Domain

nGLP-1R is directly involved in the hydrophobic interaction with Ex4(9–39).

The Full-length Family B1 Receptors—Assuming that the two-step model for ligand binding of Family B1 receptors is correct, the conserved N terminus of GLP-1 and Ex4 binds to the core domain of GLP-1R, which leads to receptor activation. Specifically, Glu^{9*} of GLP-1 (Glu^{3*} of Ex4) is probably accommodated by positively charged residues in the extracellular end of transmembrane helix 2 of GLP-1R (10). Glu^{15*} of Ex4 (Glu^{21*} of GLP-1) could define the boundary between residues that interact with the core domain of GLP-1R and residues that interact with nGLP-1R as previously suggested (24). The α -helices of nGLP-1R and Ex4(9–39) run in the same direction, suggesting that both the N-termini approach the core domain of GLP-1R. Dual cross-linking of the secretin N terminus to the secretin receptor also seems to agree with a model in which the N terminus of both the ligand and the receptor are in proximity with the core domain of the receptor (39). Intriguingly, the superior affinity of Ex4 for nGLP-1R is not translated into a superior affinity, potency, or efficacy at the full-length GLP-1R compared with GLP-1. Clearly, ligand binding of GLP-1R is more complex than that of the isolated nGLP-1R, and the second binding event of GLP-1R somehow compensates for the low affinity of nGLP-1R for GLP-1 relative to Ex4.

Trp³³ of the human GLP-1R was previously identified as a determinant of species-selective small molecule antagonist binding (40). The small molecule displaced ¹²⁵I-Ex4(9–39) binding of GLP-1R and reduced the efficacy of GLP-1, suggesting a non-competitive mechanism. Trp³³ is positioned in the N-terminal part of the α -helix of nGLP-1R, on the opposite site of Trp³⁹ but only 10 Å away from the hydrophobic ligand-receptor interface. Direct binding of the small molecule antagonist to Trp³³ of nGLP-1R could affect receptor activation either by interfering with the hydrophobic ligand-receptor interaction or by disrupting the interface between the receptor subdomains (the core domain and the α -helix of the Nt-domain). In both cases, binding of the small molecule antagonist to Trp³³ of nGLP-1R could drive the N terminus of GLP-1 and Ex4 into a suboptimal position for receptor activation, resulting in the non-competitive effect. However, it is unclear if the small molecule antagonist binds directly to Trp³³, and the orientation of nGLP-1R with respect to the core domain is not known.

One of the major differences between the ligand-bound structures of Family B1 receptor Nt-domains is the binding site and orientation of the ligand (17–19). The binding site of astressin, GIP, and Ex4(9–39) is on the same surface of the Nt-domain, and the ligand orientation is the same (Fig. 4). However, the position of Ex4(9–39) and GIP is shifted relative to astressin, because they interact with the α -helix of their receptor. Surprisingly, PACAP(6–38) binds to a completely different site of nPAC1 and has its termini upside-down compared with the other ligands. The N terminus of Ex4, GIP, CRF, and PACAP is important for receptor activation, and it is expected to interact with the receptor core domain. To accommodate this interaction, the orientation of nPAC1 relative to the receptor core domain (and membrane surface) must be rather different from the orientation of nGLP-1R, nGIPR, and nCRF-R2B.

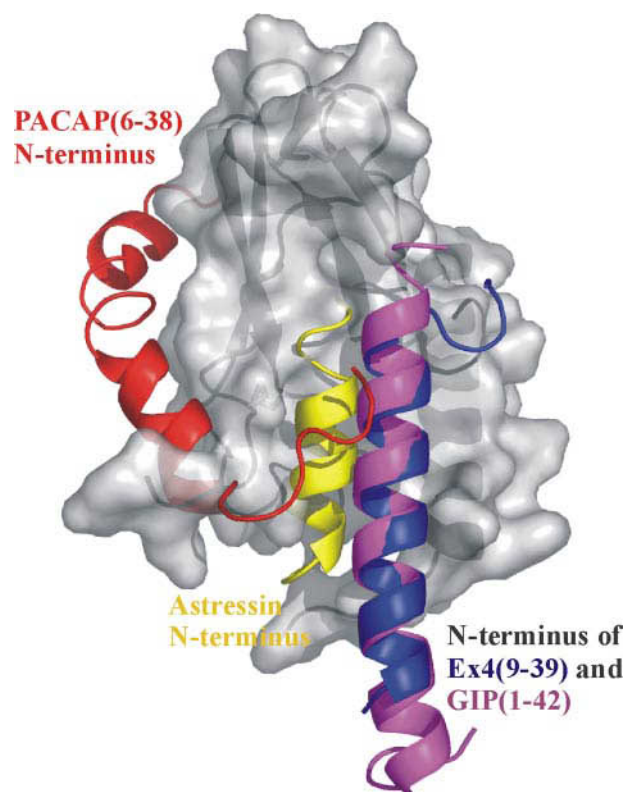


FIGURE 4. Superposition of ligand-bound nPAC1, nCRF-R2B, nGIPR, and nGLP-1R. The conserved core structures of the Nt-domains were aligned by superposition of the ligand-bound structures using PyMOL. A surface representation of nGLP-1R is shown in gray, and the ligands are illustrated by ribbon diagrams, Ex4(9–39) in blue, GIP(1–42) in magenta, astressin in yellow, and PACAP(6–38) in red.

The crystal structure reported here provides detailed molecular information about the first step of peptide ligand binding of the human GLP-1 receptor. It strongly supports the structure-based drug design for treatment of type 2 diabetes targeting the GLP-1 receptor. Structural characterization of the full-length receptor is necessary to determine the orientation of the Nt-domains and the intimacy of receptor subdomains and ultimately to understand the activation mechanism.

Acknowledgments—We thank Mathias Norrman, Gerd Schluckebier, and Anders Svensson for valuable help with data collection and crystallographic computation. We thank Carsten S. Stenvang for skillful laboratory assistance and the staff at the MAX-lab beamline I911-3 for technical help during data collection.

REFERENCES

1. Perrin, M. H., Fischer, W. H., Kunitake, K. S., Craig, A. G., Koerber, S. C., Cervini, L. A., Rivier, J. E., Groppe, J. C., Greenwald, J., Moller, N. S., and Vale, W. W. (2001) *J. Biol. Chem.* **276**, 31528–31534
2. Lisenbee, C. S., Dong, M., and Miller, L. J. (2005) *J. Biol. Chem.* **280**, 12330–12338
3. Grauschoepf, U., Lilie, H., Honold, K., Wozny, M., Reusch, D., Esswein, A., Schafer, W., Rucknagel, K. P., and Rudolph, R. (2000) *Biochemistry* **39**, 8878–8887
4. Bazarsuren, A., Grauschoepf, U., Wozny, M., Reusch, D., Hoffmann, E., Schaefer, W., Panzner, S., and Rudolph, R. (2002) *Biophys. Chem.* **96**, 305–318
5. Foord, S. M., Bonner, T. I., Neubig, R. R., Rosser, E. M., Pin, J. P., Daven-

- port, A. P., Spedding, M., and Harmar, A. J. (2005) *Pharmacol. Rev.* **57**, 279–288
6. Lopez de Maturana, R., Willshaw, A., Kuntzsch, A., Rudolph, R., and Donnelly, D. (2003) *J. Biol. Chem.* **278**, 10195–10200
 7. Dong, M., Li, Z., Zang, M., Pinon, D. I., Lybrand, T. P., and Miller, L. J. (2003) *J. Biol. Chem.* **278**, 48300–48312
 8. Holtmann, M. H., Hadac, E. M., and Miller, L. J. (1995) *J. Biol. Chem.* **270**, 14394–14398
 9. Solano, R. M., Langer, I., Perret, J., Vertongen, P., Juarranz, M. G., Robberecht, P., and Waelbroeck, M. (2001) *J. Biol. Chem.* **276**, 1084–1088
 10. Runge, S., Gram, C., Brauner-Osborne, H., Madsen, K., Knudsen, L. B., and Wulff, B. S. (2003) *J. Biol. Chem.* **278**, 28005–28010
 11. Dong, M., Li, Z., Pinon, D. I., Lybrand, T. P., and Miller, L. J. (2004) *J. Biol. Chem.* **279**, 2894–2903
 12. Wittelsberger, A., Corich, M., Thomas, B. E., Lee, B. K., Barazza, A., Czodrowski, P., Mierke, D. F., Chorev, M., and Rosenblatt, M. (2006) *Biochemistry* **45**, 2027–2034
 13. Al-Sabah, S., and Donnelly, D. (2003) *FEBS Lett.* **553**, 342–346
 14. Lopez de Maturana, R., Treece-Birch, J., Abidi, F., Findlay, J. B., and Donnelly, D. (2004) *Protein Pept. Lett.* **11**, 15–22
 15. Dong, M., Pinon, D. I., Asmann, Y. W., and Miller, L. J. (2006) *Mol. Pharmacol.* **70**, 206–213
 16. Runge, S., Wulff, B. S., Madsen, K., Brauner-Osborne, H., and Knudsen, L. B. (2003) *Br. J. Pharmacol.* **138**, 787–794
 17. Parthier, C., Kleinschmidt, M., Neumann, P., Rudolph, R., Manhart, S., Schlentzig, D., Fanghanel, J., Rahfeld, J. U., Demuth, H. U., and Stubbs, M. T. (2007) *Proc. Natl. Acad. Sci. U. S. A.* **104**, 0706404104
 18. Grace, C. R., Perrin, M. H., Gulyas, J., DiGruccio, M. R., Cantle, J. P., Rivier, J. E., Vale, W. W., and Riek, R. (2007) *Proc. Natl. Acad. Sci. U. S. A.* **104**, 4858–4863
 19. Sun, C., Song, D., vis-Taber, R. A., Barrett, L. W., Scott, V. E., Richardson, P. L., Pereda-Lopez, A., Uchic, M. E., Solomon, L. R., Lake, M. R., Walter, K. A., Hajduk, P. J., and Olejniczak, E. T. (2007) *Proc. Natl. Acad. Sci. U. S. A.* **104**, 7875–7880
 20. Goke, R., Fehmann, H. C., Linn, T., Schmidt, H., Krause, M., Eng, J., and Goke, B. (1993) *J. Biol. Chem.* **268**, 19650–19655
 21. Neidigh, J. W., Fesinmeyer, R. M., and Andersen, N. H. (2002) *Nat. Struct. Biol.* **9**, 425–430
 22. Neidigh, J. W., Fesinmeyer, R. M., Prickett, K. S., and Andersen, N. H. (2001) *Biochemistry* **40**, 13188–13200
 23. Al-Sabah, S., and Donnelly, D. (2003) *Br. J. Pharmacol.* **140**, 339–346
 24. Runge, S., Schimmer, S., Oschmann, J., Schiodt, C. B., Knudsen, S. M., Jeppesen, C. B., Madsen, K., Lau, J., Thogersen, H., and Rudolph, R. (2007) *Biochemistry* **46**, 5830–5840
 25. Thorens, B., Porret, A., Buhler, L., Deng, S. P., Morel, P., and Widmann, C. (1993) *Diabetes* **42**, 1678–1682
 26. Kabsch, W. (1993) *J. Appl. Crystallogr.* **26**, 795–800
 27. Vornrhein, C., Blanc, E., Roversi, P., and Bricogne, G. (2006) *Methods Mol. Biol.* **364**, 215–230
 28. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) *Nat. Struct. Biol.* **6**, 458–463
 29. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2126–2132
 30. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr.* **53**, 240–255
 31. Collaborative Computational Project 4 (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50**, 760–763
 32. Adelhorst, K., Hedegaard, B. B., Knudsen, L. B., and Kirk, O. (1994) *J. Biol. Chem.* **269**, 6275–6278
 33. Wilmen, A., Van, E. B., Goke, B., and Goke, R. (1997) *Peptides* **18**, 301–305
 34. Carruthers, C. J., Unson, C. G., Kim, H. N., and Sakmar, T. P. (1994) *J. Biol. Chem.* **269**, 29321–29328
 35. Van, E. B., Goke, B., Wilmen, A., and Goke, R. (1996) *Peptides* **17**, 565–570
 36. Andersen, N. H., Brodsky, Y., Neidigh, J. W., and Prickett, K. S. (2002) *Bioorg. Med. Chem.* **10**, 79–85
 37. Montrose-Rafizadeh, C., Yang, H., Rodgers, B. D., Beday, A., Pritchette, L. A., and Eng, J. (1997) *J. Biol. Chem.* **272**, 21201–21206
 38. Graziano, M. P., Hey, P. J., and Strader, C. D. (1996) *Receptors Channels* **4**, 9–17
 39. Dong, M., Pinon, D. I., and Miller, L. J. (2005) *Mol. Endocrinol.* **19**, 1821–1836
 40. Tibaduiza, E. C., Chen, C., and Beinborn, M. (2001) *J. Biol. Chem.* **276**, 37787–37793