Three-Dimensional Structural Interactions of Insulin and its Receptor

Cecil C. Yip‡ and Peter Ottensmeyer§

From the Banting and Best Department of Medical Research‡, and the Ontario Cancer Institute and Department of Medical Biophysics§, University of Toronto, Toronto, Ontario, M5G 2M9, Canada.

‡To whom correspondence should be addressed: Banting and Best Department of Medical Research, University of Toronto, 112 College Street, Toronto, Ont., Canada, M5G 1L6

1 Abbreviations used are: IR, insulin receptor; IGF-1, insulin-like growth factor 1; IGF-1R, insulin-like growth factor 1 receptor; L1 and L2, large domain 1 and 2; CR, cys-rich region; LCL, L1-CR-L2; Fn, fibronectin III-like; ID, insert domain; TM, transmembrane; TK, tyrosine kinase; CT, C-terminal.

Acknowledgments - We thank Puja Malik, Daniel Beniac, Allan Fernandes, for insights, cross-checks, and creative and artistic technical assistance, and Ira Goldfine for stimulating discussions on IR and IGF-1R.
The insulin receptor (IR) belongs to the superfamily of transmembrane receptor tyrosine kinases (TKs) (reviewed in: 1). In contrast to other family members which are monomeric in their structure, IR and its homologue, insulin-like growth factor I receptor (IGF-1R), are intrinsic disulfide-linked dimers of heterodimeric disulfide-linked proteins of the form (αβ)₂. The 135-kDa α subunit of IR is extracellular, whereas the 95 kDa β subunit contains an extracellular portion, a single transmembrane sequence, and an intracellular TK domain. Fig. 1a depicts the major structural features of the αβ dimer (see review by Tavaré and Siddle, 2). Ligand-specific binding to the α subunits activates the TK, initiating a signal cascade that results in numerous cellular responses. Our understanding of the mechanics of this signal transduction process has been hampered by the unavailability of an atomic structure of the whole IR protein. However, the quaternary structure of the isolated complex of biologically active IR and insulin was recently solved by 3D reconstruction from low-dose scanning transmission electron micrographs (STEM) (3) (Fig. 1b1-1b3). Atomic structures of subdomains of IR or of highly analogous proteins were fitted into the complex (e.g. Fig. 1b4), creating the only available atomic structural model of IR. The model reveals structural details of the interaction of insulin with the receptor that lead to the activation of the intracellular TK (4). Here we review previous biochemical observations on IR binding of natural and modified insulins and of IGF-1 against this atomic structural model, and, in the light of recent structural data on the unbound receptor, we discuss the mechanics of a model of receptor activation arising from insulin binding.

3D Scanning Transmission Electron Microscopy (STEM)

While crystallography remains the technique of choice for atomic structure determination of large proteins, considerable advances in electron microscopy (EM) and in image processing have been made for 3D structure determination of proteins that are refractory to crystallization. For a growing number of proteins, including the plant light-harvesting complex (5), a human water channel (6), and aquaporin-1 (7), 3D structural information has been obtained at resolutions of several Angstrom units, revealing α-helices and other structural details.

In STEM, image acquisition is digital, as a 3 Å small electron beam probe scans across the specimen at low temperature. Signal intensity is directly proportional to the molecular mass of each molecule. Moreover, STEM in dark field mode can readily visualize clusters of heavy atoms on specifically marked biological molecules (3, 8). STEM images have been used to reconstruct 3D structures of several proteins at resolutions of 12 - 20 Å, such as SRP54, the Klenow fragment of DNA polymerase I, and IR (3, 9-12).

The STEM IR reconstruction at 20 Å was docked with atomic structures of insulin, and of subdomains of IR or highly similar proteins from crystallography or NMR spectroscopy (3, 4, 13-17). The resulting resolution is atomic within those domains, and as accurate between domains as the alignment of their centers-of-mass and their rotations. For instance, the centers-of-mass of the 17.5 kDa L1 regions can be aligned in their EM domains to within 3.0 Å, with larger domains located more accurately.
Rotational accuracy depends on recognizable asymmetries, and is as good as 5° for the L1-CR-L2 (LCL) region.

**Ligand-Binding Tunnel in the Insulin-IR Complex**

In the fitted structure of IR, the two atomic LCL regions formed a diamond-shaped tunnel, the walls and the entrances of which were lined with most of the amino acids known to be involved in insulin binding. One insulin could be directly fitted into this tunnel as a rigid body to produce the side-chain pairings shown in Table 1. In this fit the insulin A chain interacts primarily electrostatically with the L1 and L2 domains of one α subunit, with no obvious hydrophobic components. The B chain interacts with the other α subunit, chiefly hydrophobically with L1, and electrostatically with CR. Thus the model points to multi-site interactions with IR, as expected from the $10^9$ M$^{-1}$ binding affinity of insulin.

The deduced interactions between insulin and LCL tunnel encompass insulin residues A4, A5, B12, B16, B17, B24, and B26, out of 11 that are known to be important in receptor binding (1, 18, 19). These residues interact to within 2.5 Å with residues 12, 34, 86, 89, 91, 247, and 249 among those in the LCL region with known involvement in binding (18, 20-24). Additional relevant receptor residues, such as 14, 250, and 323, are just beyond this distance (4), while 36, 39, and 64 (24) not interacting directly with insulin in the model, are on close facing regions on the two L1 domains. Other important residues are discussed below.

Alanine scanning mutagenesis in the receptor α-subunit segment 704-716 outside of the LCL domain greatly decreases the receptor affinity (25). In the schematic placement of the receptor subdomains (Fig. 1c), this C-terminal sequence emerges from the α-α disulfide bond at Cys682, 683 or 685 (26) on the 2-fold axis of the IR dimer. The α-terminal sequence 686-719 can easily reach the insulin-binding tunnel. Residues within 704-718 photocrosslinked to carboxyamidated azido-PheB25 on insulin (27), suggesting a juxtaposition between this sequence and PheB25. PheB25, facing into the entrance of the binding tunnel (Fig. 2), does not interact with the LCL regions in the atomic model (4), but is important in receptor binding. It is conceivable that it interacts hydrophobically with Tyr-Leu (708-709) as suggested by Kuros et al. (27), or with Val-Val-Phe-Val (712-715). Tyr-Leu interaction with PheB25, would provide a structural base for the suggestion (27), that the C-terminal(719) of the α subunit is near its N-terminal (Fig. 2). Deletion of insulin C-terminal B26-B30 did not decrease biological potency, implying that this sequence of the B chain may affect the interaction of insulin with IR through steric effects rather than binding (28). Such a steric or orientational role for this region is also suggested by the deduced electrostatic interaction of insulin LysB29 with Asp12 in L1 (4).

**Ligand Specificity and the Ligand-Binding Tunnel**

The primary amino acid sequence and the tertiary structure of all known insulins are highly conserved (reviewed in 29). Nevertheless, several insulins, including that of
guinea pig, the Atlantic hagfish, amphiuma, chicken and *Xenopus*, exhibit altered activity from porcine, bovine, or human insulin when assayed using mammalian insulin receptor. The altered bioactivities of guinea pig and amphiuma insulins were readily explained by the structural fitting of insulin in the binding tunnel (4). For chicken and *Xenopus* insulins (30, 31) the 2-3-fold greater activity was attributed to a common Thr(A8)His substitution, along with Ser(A9)Asn in chicken insulin. As also modeled for amphiuma, the replacement His residue can reach Asp59 in the L1 domain to potentiate binding. Moreover, AsnA9 in chicken insulin can readily bond with Gln328 of IR.

The binding affinity of hagfish ligand is about 25% of that of porcine insulin, and its biological potency less than 5% (32). Significant sequence changes are Phe(B1)Arg, Ser(B9)Lys, Thr(A8)His, Ser(A9)Lys, and Ile(A10)Arg. When hagfish insulin is fitted into the mammalian LCL tunnel, HisA8 can bond with Asp 59, and ArgA10 with Glu329. However, ArgB1 is positioned to repel LysB29 and compete for its interaction with Asp12, disrupting the proposed orientational function (28) and permitting LysB9 to bond Asp12 (on monomer B). A resulting misalignment of hagfish insulin between the two human LCLs could create the observed consequences on biological signaling.

The biological activity of insulin analogues with specific amino acid substitutions has been extensively studied. A His(B10)Asp change resulted in a superactive insulin (21), explainable by a stronger ionic interaction with Arg14 in the binding tunnel. On the other hand, the interchange of ProB28 and LysB29 in the Lys-Pro insulin produces no major alterations in binding (33). In the model, lysine in either position can bond with Asp12 on L1.

The invariant B23-B26 core (Gly-Phe-Phe-Tyr) has been probed by replacement of the aromatic residues with less hydrophobic amino acids. For both [LeuB24]- and [LeuB25]insulins, receptor-binding activity and biological was respectively 10% and 1% of porcine insulin (34). [LeuB24]insulin, but not [LeuB25]insulin, was a partial antagonist. In the model PheB24 lies flat in the surface of insulin while interacting with the side chain of Leu87 of L1. A bulkier and longer LeuB24 side chain would cause a slight separation between facing hydrophobic surfaces on insulin and L1 to account for lower binding, and could correspondingly separate and misalign the two LCLs just enough to lessen biological potency. Thus [LeuB24]insulin could act as a partial antagonist.

The role of PheB25 is more complex. Not only is PheB25 important for insulin binding, it also enhances the release of receptor-bound insulin, a kinetic phenomenon described as negative cooperativity (35, 36). Previous studies (27, 28, 34) placed PheB25 near the 704-719 tail of the IR α subunit. The accessible position of PheB25 in the IR model (Fig. 2) easily permits this. Low affinity insulin analogues, in which PheB25 was replaced by Ala, Ser, Leu or homophenylalanine (37), abrogated negative cooperativity, possibly due to a reduced interaction with the α tail. However, a truncated carboxy-amidated [PheB25](des(B26-30)insulin bound better than native insulin and retained negative cooperativity. Its azido derivative cross-linked to the α tail, suggesting that the non-azido form could also bind this region. Similarly, such truncated insulins with LeuB25, SerB25, AlaB25 or homoPheB25 all retained 50% or more negative
cooperativity compared to native insulin (28), suggesting that the binding of B25 to the α tail of IR still occurred, through the carboxyamide moiety. This should be testable. Thus negative cooperativity, which demands multiple sites of ligand-receptor interactions (35, 36), may for IR require also specific binding to the α-subunit C-terminal sequence.

Like PheB25, TyrA19 is important for receptor binding and biological potency (38), yet does not interact with LCL in the atomic model. Structurally it is sandwiched between AsnA18 and IleA2 and interacts edge-on with the plane of the phenyl ring of PheB25. Modifications or substitutions of TyrA19 can therefore directly affect the functional role of PheB25.

**IGF-1, IR, and IGF-1R**

Insulin-like growth factor 1 (IGF-1) and insulin are homologous peptide hormones that bind to homologous but different receptors. IGF-1 is a single-chain polypeptide with four sequential structural domains, B, C, A and D. The B and A domains are equivalent to the B and A chains of insulin, and their secondary structure can be largely superposed (39). The D domain extends 8 residues beyond the A chain. The C domain is equivalent to the C peptide of proinsulin. In spite of these strong similarities, insulin and IGF-1 bind only weakly to the alternative receptor. Extensive investigations indicate that this binding specificity for IR and IGF-1R resides in different regions of a common binding site (40, 41).

These findings are corroborated by an examination of the atomic model in which insulin is replaced by IGF-1 (13; PDB: 2GF1), and the positions of the LCL regions in IR substituted by the exact crystal structure of the LCL 1 domains of IGF-1R (42; PDB: 1IGR). The pair-wise amino acid interactions found are shown in Table 1. For insulin on IR, 14 close interactions were obtained. For IGF-1 on IR, the ligand had to be rotated to relieve a serious steric hindrance between the D-domain of IGF-1 and the L2 region of IR. Then only 11 pair-wise interactions formed. Conversely, for IGF-1 on IGF-1R 14 pairs of interactions were obtained, while insulin in the IGF-1R tunnel had only 10 close interactions. The character of binding (hydrophobic, polar, etc.) also changed. Nevertheless, the atomic models based on the IR configuration indicate that the ten-fold weaker binding of each hormone to the homologous receptor is accompanied by a reduction of side-chain interactions.

The pattern of interaction of each hormone with its homologous receptor is also different (Table 1). Insulin binds predominantly to the L1 regions on IR, while IGF-1 has a preference for the CR regions and L2 on IGF-1R. Both hormones virtually eschew the CR regions of the homologous receptors, equally preferring the L1 and L2 domains. It is known that the CR region confers binding specificity in IGF-1R, and that L1 and L2 do so in IR (13, 18, 40, 41).

**The Mechanics of Transmembrane Signaling**

A fundamental question on the mechanism of transmembrane activation of receptors by extracellular ligand binding is how the binding action is transformed into
key intracellular biochemical reactions such as receptor autophosphorylation. Inactive monomeric TK receptors, such as the epidermal growth factor receptor and platelet-derived growth factor receptor, are dimerized and activated by ligand binding (see Review 43). Since IR is intrinsically dimeric without insulin, the distance between the cytoplasmic TKs must be too large for activation without ligand binding. In the 3D reconstruction of the insulin-bound IR complex the two TK domains are juxtaposed, and an extended flexible activation loop of one fitted atomic TK structure can just reach the catalytic loop of the other TK (3). In a simple mechanical paradigm of activation (4) the binding of insulin not only overcomes an energy barrier to binding, but holds the LCL domains together to permit a lateral shift and approach of the two transmembrane domains to allow trans-autophosphorylation of the tyrosine residues in the activation loop (4) (Fig. 3a,b).

The model presupposed an “open” configuration of the free IR. Such a configuration has been confirmed by the 3D reconstruction from STEM micrographs of insulin-free IR dimers (44) (Fig. 3d). Moreover, this structure, together with the insulin-bound 3D structure, makes it possible to reconcile virtually all structural forms seen previously in electron microscopic examinations of the IR, except for long parallel T-shaped structures occasionally observed (45, 46). This includes Y-, H-, V-, and X-shapes, and parallel bar-like images, observed by 2D electron microscopy of detergent solubilized IR, its ectodomain or vesicle-reconstituted insulin-IR complex (45-48). Such configurations can be formed by the projection of the 3D reconstructions at various orientations (Fig. 3e). Moreover, the 3D reconstruction can also explain the images of parallel rods and crosses (Fig. 3e, last column) produced by pairs of antibodies to Fn0, to the β-ID, and to the CR domain (49).

**Summary**

Electron microscopic imaging of the large insulin receptor protein in complex with insulin and in its free form, coupled with 3D reconstruction and fitting of atomic subdomains, has yielded a structural model that permits an understanding of insulin binding at the atomic level. The model corroborates and explains the biological activity and binding behavior of virtually all naturally occurring insulins and insulin analogues, as well as the effects of receptor mutations on insulin binding. The model also provides a basis to address the fundamental question on the mechanism of transmembrane activation of receptors by ligand binding.
Fig. 1. **Dimeric Insulin Receptor Structure.** (a) **Schematic.** IR represented in its \((\alpha\beta)_2\) form. Residues at the start of named domains, and disulfides (italics) are numbered. L, large domain; CR, cys-rich; Fn, fibronectin III-like; ID, insert domain; TM, transmembrane; TK, tyrosine kinase; CT, C-terminal; juxtamembrane domains not labelled.

(b) **3D reconstruction from EM.** (b1) Structure at full volume of 480 kD dimer (side view). (b2) Structure at 85% of full volume to show domain subdivision (side view). (b3) Top view of (b2); regions 1, 2, and 3 represent epitope locations for monoclonal Fab fragments 18-44, 83-14, and 83-7 (49), respectively (see text and Fig. 3e).

(b4) Ectodomain of EM structure (top view) fitted with atomic domains LCL (green), Fn0 (blue), \(\alpha\)Fn1-\(\alpha\)ID (red), \(\beta\)Fn1-Fn2 (violet); backbone ribbon representation. Second subunit in brown. Arrow indicates orientation of view of Fig. 2a.

(c) **Schematic of structural conformation in (b2).** One monomer labelled (colour), second monomer (white); In: insulin; A: activation loop and opposite catalytic site; X: C-terminal of \(\alpha\) subunit (white) and N-terminal of \(\beta\) subunit (black); 1, 2, and 3: disulfides at amino acids 524, 682-5, and between 647 and 872, respectively. Black bar represents 2-fold axis. Solid and open arrows point to insulin PheB25 and Glu6 of \(\alpha\) subunit.

Fig. 2 **Accessibility of Insulin PheB25.** (a) View of atomic insulin receptor model in direction of arrows in Fig. 1b4; insulin (purple) between the L1-CR domains of the two monomers (green and brown); (b) enlargement from (a); arrows at PheB25 and \(\alpha\)N-terminal Glu6 (start of IGF-1R crystal structure (13); L1(1-5) not modeled). (c) View perpendicular to (b), showing accessible locations of insulin PheB25 and L1 Glu6.

Fig. 3. **IR Signaling Mechanics and Conformational Changes.** (a) Schematic closed IR form, insulin-bound, permissive for activation by transphosphorylation. **Set 1, set 2:** amino acid interactions with monomer 1, monomer 2, respectively; C: catalytic site; A-loop: activation loop; 1, 2: \(\alpha\alpha\) disulfide bonds; cam: structural feature (bulge) on CR.

(b) Schematic open IR form, non-permissive for transphosphorylation, cam blocks approach of A-loop to C site. (c) 3D reconstruction of insulin-bound IR. (d) 3D reconstruction of insulin-free IR. (e) Projection images at different orientations of structures in (c) and (d) showing Y-, H-, V-, X-, and parallel bar shapes like those seen in prior work (45-48). Sixth column: modeled structures based on Fig. 1b4 of Fab/IR images by Tullock et al. (49).

Table 1. **Receptor-Ligand Side-Chain Interactions:** Simultaneous close amino acid pairings with observed side-chain distances within 2.5 Å for insulin or for IGF-1 in each of the modeled IR or IGF-1R ligand-binding tunnels. Color scheme: dark yellow: hydrophobic; green: polar; red: negative; blue: positive.
References

Table 1. **Receptor-Ligand Side-Chain Interactions**: Simultaneous close amino acid pairings with observed side-chain distances within 2.5 Å for insulin or for IGF-1 in each of the modeled IR or IGF-1R ligand-binding tunnels. Color scheme: dark yellow: hydrophobic; green: polar; red: negative; blue: positive.

<table>
<thead>
<tr>
<th>Insulin Receptor Interactions</th>
<th>IGF-1R Interactions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Insulin</strong></td>
<td><strong>IR</strong></td>
</tr>
<tr>
<td>Mon A</td>
<td>L1</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>CR</td>
<td>Glu287</td>
</tr>
<tr>
<td></td>
<td>ArgA50</td>
</tr>
<tr>
<td></td>
<td>ThrC41</td>
</tr>
<tr>
<td></td>
<td>ArgA56</td>
</tr>
<tr>
<td>L1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Mon B</td>
<td>L1</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>CR</td>
<td>GinB21</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>L2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* alternative pairing
Three-dimensional structural interactions of insulin and its receptor

Cecil C. Yip and Peter Ottensmeyer

J. Biol. Chem. published online May 22, 2003

Access the most updated version of this article at doi: 10.1074/jbc.R300021200

Alerts:
  • When this article is cited
  • When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2003/07/10/R300021200.DC1

This article cites 0 references, 0 of which can be accessed free at
http://www.jbc.org/content/early/2003/05/22/jbc.R300021200.citation.full.html#ref-list-1