Published in final edited form as:

Nature. 2018 April 05; 556(7699): 122–125. doi:10.1038/nature26153.

Structure of the Insulin Receptor-Insulin Complex by Single Particle CryoEM analysis

Giovanna Scapin[1,*], **Venkata P. Dandey**[2,*], **Zhening Zhang**[2], **Winifred Prosise**[1], **Alan Hruza**[1], **Theresa Kelly**[3], **Todd Mayhood**[3], **Corey Strickland**[1], **Clinton S. Potter**[2], and **Bridget Carragher**[2]

[1]Merck & Co., Inc., Department of Biochemical Engineering & Structure, 2000 Galloping Hill Rd. Kenilworth, NJ, 07033, USA

[2]Simons Electron Microscopy Center, National Resource for Automated Molecular Microscopy, New York Structural Biology Center, 89 Convent Avenue, New York, NY, 10027, USA

[3]Merck & Co., Inc., Department of Biophysics, NMR & Protein Products Characterization, 2000 Galloping Hill Rd. Kenilworth, NJ, 07033, USA

Summary

The insulin receptor (IR) is a dimeric protein that plays a crucial role in controlling glucose homeostasis, regulating lipid, protein and carbohydrate metabolism, and modulating brain neurotransmitter levels[1,2]. IR dysfunctions have been associated with many diseases, including diabetes, cancer, and Alzheimer's[1,3,4]. The primary sequence has been known since the 1980s[5], and is composed of an extracellular portion (ectodomain, ECD), a single transmembrane helix and an intracellular tyrosine kinase domain. Insulin binding to the dimeric ECD triggers kinase domain auto-phosphorylation and subsequent activation of downstream signaling molecules. Biochemical and mutagenesis data have identified two putative insulin binding sites (S1 and S2)[6]. While insulin bound to an ECD fragment containing S1 and the apo ectodomain have been characterized structurally[7,8], details of insulin binding to the full receptor and the signal propagation mechanism are still not understood. Here we report single particle cryoEM reconstructions for the 1:2 (4.3 Å) and 1:1 (7.4 Å) IR ECD dimer:Insulin complexes. The symmetric 4.3 Å structure shows two insulin molecules per dimer, each bound between the Leucine-rich sub domain L1 of one monomer and the first fibronectin-like domain (FnIII-1) of the other monomer, and making extensive interactions with the α subunit C-terminal helix (α-CT helix). The 7.4 Å structure has only one similarly bound insulin per receptor dimer. The structures confirm the S1 binding

Correspondence and requests for materials should be addressed to giovanna.scapin@merck.com.
*Authors contributed equally

Authors Contributions: GS collected and processed the data, interpreted the results and wrote the manuscript; VPD prepared the grids used in the studies, collected and processed the data, and assisted in writing the manuscript; ZZ prepared the grids used in the study; WP and AH obtained and prepared the receptor and insulin samples; TK produced the binding data; TM characterized the sample; ZZ, WP, AH, TK and TM also helped in writing the manuscript; CS, CSP and BC helped in planning the experiments, analyzing the data and writing and editing of the manuscript.

The authors declare they have no competing interests.

interactions and define the full S2 binding site. These insulin receptor states suggest that recruitment of the α-CT helix upon binding of the first insulin changes the relative sub domain orientations and triggers downstream signal propagation.

The insulin receptor (IR) is a dimer of heterodimers $((\alpha\beta)_2)$[5]. The α subunit and the β-chain N-terminal ~190 residues are located on the extracellular side of the plasma membrane (Figure 1a) and together constitute the full IR ECD; the remainder of the β-chain includes a transmembrane helix, the juxtamembrane domain and the intracellular tyrosine kinase domain. Structures of the individual sub-domains comprising the ECD (L1, CR, L2, and fibronectin type III domains FnIII-1, 2 and 3) are known, and the crystal structure of the apo IR ECD[8] provides one view of their quaternary organization. In the apo IR ECD structure, density for the insertion domain (120 residues located within FnIII-2, containing the cleavage site that generates the α and β chains, and the subunit α C-terminal helix, α-CT) was poor, and only a portion of α-CT helix was visible[9,10]. While only one (αβ) heterodimer is present in the asymmetric unit of the IR crystal structure, a symmetry generated tetramer $((\alpha\beta)_2)$ was proposed to represent the biologically relevant form of the ECD[8]. We will refer to each individual (αβ) heterodimer as IR-monomer, and to the $(\alpha\beta)_2$ tetramer as IR-dimer.

Current models of insulin binding to IR suggest two interaction sites on insulin, engaging with two binding sites (S1 and S2) on IR[11]. Binding is proposed to be in trans with each insulin interacting with S1 of one IR-monomer and S2 of the other. The bridging of these receptor binding sites constitutes the high affinity interactions, while low affinity interactions arise from single site occupancy. The high affinity interaction is characterized by a 6 -200 pM K_d for the solubilized and affinity purified full length receptor, while the low affinity interactions have K_ds between 6 (for S1) and 400 nM (for S2)[12]. Binding of insulin to full length receptor shows multiple binding events and dissociations and/or multiple rearrangements within low and high affinity binding poses[12]. While receptor activation is caused by high affinity binding, the activation curve is complex and shows anti-cooperative binding[13].

We report here the 4.3 Å resolution structure of the IR ECD in the presence of insulin obtained using single-particle cryoEM (Figure 1b,c). A pronounced preferred orientation of the complex in vitreous ice was overcome using a combination of tilted data collection[14] and fast plunge speeds using the Spotiton instrument[15,16]. Further details are provided in the methods section. Three-dimensional classification of the cryoEM data identified a dimeric structure with two populations, in a ~4:1 ratio (Figure 2b and Extended Data Figure 1). The largest, Class 1, is characterized by a symmetrical “head” and a poorly resolved, apparently flexible “stalk”. Refinement Class 1 after masking out flexible portions of the stalk and applying C2 symmetry provided a reconstruction to 4.3 Å (Extended Data Figure 1 and 2). The electron density map allowed unambiguous identification of the L1, CR, L2, and FnIII-1 receptor domains, as well as the clear positioning of the insulin molecule and the α-CT helix (Figure 1b and 1c). Refinement of the data using C1 symmetry produced a lower resolution map (4.7 Å; Extended Data Figure 3) in which additional density, assigned to FnIII-2 was visible for one monomer (Extended Data Figures 4a and 5b).

The Class 1 structure shows two insulin molecules symmetrically bound per dimeric receptor. Although asymmetric binding was predicted for the full length receptor[12], studies show that the soluble ECD binds two molecules of insulin each with a 3.5 nM K_d and a fast dissociation rate[17]. The symmetrical binding of both insulin molecules observed in the cryoEM map is consistent with these data. Each insulin is located between the L1 domain of one IR-monomer, the FnIII-1 domain of the other IR-monomer, and the α-CT helix (Figure 1c). Residues of the L1 subdomain and the α-CT helix, previously identified as essential for insulin binding by a variety of biochemical approaches[13], have been proposed to represent the IR S1 site[18]. Structures of "S1 microreceptors" (containing L1, CR and a portion of the α-CT) in complex with insulin and insulin mimetics[7] provided the first atomic details of some IR-insulin key interactions (Extended Data Table 1). In those structures insulin directly interacts with the IR L1 subdomain through only a small region of the B-helix, and most of the other IR residues that have been identified as essential for insulin binding (Extended Data Table 1) interact with the α-CT helix. Predicted interactions involving the C-terminal end of the insulin B-chain[6] could not be confirmed because those residues were not resolved in the electron density map[7]. The cryoEM structure confirms the engagement of S1, including the interactions involving the C-terminus of the insulin B-chain which is now resolved. The C-terminal region of the B-chain in the bound form assumes a different conformation from that observed in the hormone free form (for example, PDB entry 1ZNI, Extended Data Figure 4c), allowing the B-chain core residues (Gly8-Cys19) to interact with the α-CT helix, confirming the so-called detachment model of insulin binding[19]. Although from the current map it is not possible to determine from which IR monomer each α-CT helix originates, the fact that each insulin interacts with residues belonging to both monomers supports the trans binding mode hypothesis.

The cryoEM structure identifies the S2 location (Extended Data Table 1) as the area of interactions between the IR sequences Pro495-Arg498 and Arg539-Asn541 (located within the FnIII-1 domain) and chain A of insulin (residues Glu4-Thr8 and Ile10). Earlier work suggested that these insulin residues interact with the S2 binding site[6] but no IR partner residues were assigned. Interactions are observed between the α-CT helix residues Leu696-Lys703 and IR residues Gly346-Asn349 and Arg372-Tyr374 (L2 subdomain) and Arg498-Leu501 and Val570-Thr571 (FnIII-1 subdomain). Residues in L2 and FnIII-1 were correctly predicted as components of the IR S2 site[20]. Clinical mutations (S323L, F382V, K460E and N462S, all located within or near the L2 and FnIII-1 domain areas discussed above) result in impaired insulin binding or reduced signaling[21-24]. Further evidence that these regions represent S2 comes from Brandt et al.[25] in which a truncated receptor formed from the L1, CR, L2, and FnIII-1 subdomains fused to the α-CT helix is shown to have insulin binding properties identical to those of the full ECD. A second possible location for the S2 site (at the junction of the FnIII-1 and FnIII-2 domains) was hypothesized from the crystallographic dimer structure of the insulin-free IR ECD[8]. Mutagenesis data provided some support for this assignment[11]. However, in the un-symmetrized map (Extended Data Figure 4a and 5b) in which the FnIII-2 is visible, the bound insulin is more than 50 Å from this predicted S2 (Extended Data Figure 4b), and located on the opposite side of the receptor.

Refinement of the data for the smaller Class 2 produced a 7.4 Å map (Figure 2 and Extended Data Figures 1 and 6). In this map, half of the "head" has an overall conformation similar to

that observed in the Class 1 map, with visible α-CT helix and insulin density; the other half showed the L1-CR-L2 portion in a different conformation (Figure 2a), more open and characterized by less well-defined density. The IR model based on Class 1 map was rigid-body fitted into the Class 2 map, and only one insulin molecule and one α-CT helix were included in the final structure. The open side lacked both the insulin and the helix (Figure 2b). This conformation thus comprises both an insulin-bound monomer and an insulin-free monomer. Comparison of these structures suggests that binding of insulin induces both the closing of the top portion of the receptor (by rigid body motion of the L1-CR-L2 portion with respect to the FnIII-1 domain) and recruitment of the α-CT helix, contrary to a previously suggested mechanism in which the insulin molecule docks to a preformed harbor containing both the L1 and α-CT elements required for binding[26].

The IR-dimer identified in the cryoEM analysis does not resemble the crystallographic symmetry generated dimer[8]. The relative arrangement of the two monomers in the cryoEM and crystallographic dimers differ, and transitioning between them requires major conformational changes and disruption of extensive surface interfaces, specifically the interactions between L1 and FnIII-2′ and their dimeric symmetry mates, and between L2, FnIII-1′ and their dimeric symmetry mates. (Figure 3a and Extended Data Figure 7). However, the overall conformation of the crystallographic monomer (PDB 4ZXB) is similar to the unbound monomer observed in the second cryoEM class (Figure 3b), suggesting that it could be a biologically relevant representation of the unbound IR monomer. Attempts to obtain a cryoEM reconstruction for the ECD in the absence of insulin (apo) proved unsuccessful as the particles showed a high degree of heterogeneity and many were smaller than expected, suggesting that the ECD dimer may be unstable in the absence of ligand. Comparing the unbound monomer from the crystal structure to the insulin-bound monomer from the cryoEM structure (Extended data Figure 8), it appears that the conformational change that occurs upon insulin binding can be described as two rotations. The first is a 35° rotation of the L2-CR-L1 domains with respect to the linker connecting the L2 and FnIII-1 domains (residues Ala466- E469); the second is a 55° swing of the CR-L1 pair (as a rigid body) around the Gly306-Lys310 linker between the CR and L2 domains (not resolved in the current structure). These domain movements are similar to the rotation and translation observed for the related EGFR family upon ligand binding[27]. Based on the comparison between unbound and bound receptor, it appears that the α-CT helix moves between ~55 Å (if we consider the helix from the same monomer) and ~70 Å (if we consider the α-CT coming from the other monomer). Figure 3c shows a schematic diagram of a possible insulin receptor transduction mechanism. It is likely that following the binding of the first insulin molecule, the large shift in position of the α-CT helix causes conformational changes in the ID-α domain of both monomers. The ID-α domain is disulfide bonded to FnIII-3 (Cys647-Cys860) and this connection (so called “signaling bridge”[26]) may induce further conformational changes in the FnIII-3 domain, thus triggering the downstream signal propagation.

Methods

Recombinant human Insulin Receptor/CD220 was obtained from R&D systems (catalog # 1544-IR-50/CF, https://www.rndsystems.com/products/recombinant-human-insulin-r-cd220-

aa-28-944-protein-cf_1544-ir-cf), reconstituted in PBS (Phosphate buffered saline, pH=7.4) at 0.3 mg/ml (0.00293 mM), stored in 100 ul aliquots at -80, and used as is (without further purification). The R&D Systems protein was tested for insulin binding, using ^{125}I-insulin competition and Surface Plasmon Resonance (below and Extended Data table 2). The measured potency and binding affinities agreed very well with values from the literature[17,31]. A coomassie stained SDS PAGE gel (Extended Data Figure 5) run on the sample shows the expected bands for the α- and β-chains. For both chains the apparent molecular weight is higher than the predicted one (82.9 kDa for the α subunit and 22.9 kDa for the β subunit), and the β subunit appears as three discrete bands. This has been previously observed[32-34] and can be attributed to glycosylation (and to different states of glycosylation for the β subunit). The gel confirms the integrity of the sample used in the studies.

Recombinant human insulin was obtained from the laboratories of Merck & Co. RHI is a commercially available material; a generic description for the bioprocess used in preparing it can be found in the paper by Johnson[35]. A stock solution (5.6 mg/ml, 0.969 mM) was made in HBS (10 mM Hepes pH=7.4, 150 mM NaCl) and stored at 4°C.

IR ECD Competition Binding Assay

His tagged human insulin receptor ECD purified protein (R&D Systems) was bound to an anti-mouse IgG coated microplate via attachment to a mouse monoclonal anti-human IR capture antibody (R&D Systems). Varying concentrations of RHI were incubated with the IR ECD protein in the presence of [^{125}I]-insulin (PerkinElmer (catalogue NEX420050UC)) (0.1 nM) overnight at 4°C in binding buffer (100 mM HEPES, 100 mM NaCl, 10 mM MgCl2, 0.02% Triton X-100, pH 8). Plate washings with binding buffer were carried out after the antibody and protein capture steps and following ligand incubation. The amount of radio-ligand still bound to the IR ECD was determined with either a TopCount or a MicroBeta instrument using Microscint-40. Concentration response curves were generated and IC50 values calculated using a 4-parameter fit.

IR ECD Direct Binding (SPR; Biacore) Assay

His tagged human insulin receptor ECD purified protein (R&D Systems) was immobilized to a Biacore CM5 chip via an anti-His antibody kit (GE Healthcare Life Sciences) following manufacturer instructions. RHI binding to the immobilized human IR ECD protein was achieved by varying the compound concentration in running buffer (1X HBS-EP, 10 mM HEPES, 150 mM NaCl, 3 mM EDTA, 0.005 % P-20, pH 7.3) passed over the chip. General detection and data collection parameters for the Biacore T200 instrument were used to determine k_{on}, k_{off} and K_d values.

Grid preparation

Sample for grid preparation was generated by mixing 100 µl of IR solution and 3 µl of insulin stock (for a final ratio IR:insulin of 1:10) followed by incubating on ice for at least one hour before making vitrified grids. Grid reproducibility and obtaining suitable ice thickness were issues encountered when using manual grid preparation. To address these issues, the samples were vitrified using a semi-automated Spotiton V1.0 robot[15,16,36], a

novel device for preparing cryoEM samples that uses piezo-electric dispensing to apply small (50pL) drops of sample across a "self-blotting" nanowire grid as it flies past *en route* to plunge into liquid ethane. The nanowire grids, manufactured in house, used either lacey carbon or lacey gold supports[37]; the gold support helps to optimize image quality when acquiring data with the grid tilted relative to the electron beam[14,38]. Tilted data collection was necessary to address highly preferred particle orientation observed during initial data collections. For the tilted data sets, nanowire gold substrate grids were plasma cleaned for 10 secs (O_2 + H_2). The sample was dispensed onto these grids in 50pL drops for a total of ~5 nL of sample dispensed in a stripe from top to bottom across each grid, before the grid was plunged into liquid ethane. The time between sample application to the grid and plunge into ethane was typically ~300 msecs. Subsequently, we modified the SpotItOn instrument to further reduce the time between sample application and plunge to be ~170 msecs. This modification was undertaken in an attempt to reduce the time the sample has to interact with the air-water interfaces in the thin liquid film prior to vitrification. These interactions are assumed to be the cause of preferred particle orientations[39], and indeed the faster plunging speed resulted in grids that displayed markedly better particle distribution (paper in preparation), allowing for the collection of data at zero degrees tilt. Nanowire based, self-blotting grids with a lacey carbon supporting substrate were used for this purpose, following the same sample preparation methods described above.

Data Collection and processing

Tilted data were collected on a Thermo-Fisher Titan Krios equipped with an energy filter and a Gatan K2 Counting camera; the microscope was operated at 300 kV and a nominal magnification of 105,000×, with a calibrated pixel size of 1.1 Å. The defocus ranged from −1 to −2.5 μm. Images were collected at a tilt angle of −30° to address the issue of preferred orientation identified during initial data collection[14]; exposure was set to 10 sec (40 frames/image), for a total dose of ~68 $e^-/Å^2$. A total of 6805 images were collected in three sessions using Leginon[40]. 3056 images were deemed suitable for further data processing. Frames were aligned using MotionCor2[41]; global and per-particle CTF was calculated using gCTF[42]. Particle picking was done using Gautomatch (http://www.mrc-lmb.cam.ac.uk/kzhang/) and resulted in 1,206,222 particles. All subsequent data processing (from 2D classification to final reconstruction) was done using Cryosparc[43]. After several cycles of 3D classification, 151,409 particles in one class were used to generate a 4.6 Å map with C2 symmetry (and a 5.6 Å map with C1 symmetry). 48,315 particles went into a second class and generated a C1 symmetry map at 7.4 Å resolution (Extended Data Figure 1). We note that less than 20% of particles originally selected contribute to the final 3D maps. This arises firstly from very liberal initial picking criteria that result in a high rate of false positives that are rejected during initial 2D classification (approximately 60%). Secondly, there is a possibility that a substantial fraction of the receptor is not complexed, or alternatively that there is a fast equilibrium between bound and unbound species that shows up as heterogeneity limiting the contribution of these particle to the 3D structure.

Untilted data were collected on a Thermo-Fisher Titan Krios equipped with an energy filter and a Gatan K2 Counting camera; the microscope was operated at 300 kV and a nominal magnification of 105,000×, with a calibrated pixel size of 1.1 Å. The defocus ranged from

−1 to −2.5 μm. Exposure was set to 10 sec (40 frames/image), for a total dose of ~66 $e^{-}/Å^2$. A total of 2487 images were collected in three sessions using Leginon[40]. 1526 images were deemed suitable for further data processing. Frames were aligned using MotionCor2[41]; global and per-particle CTF was calculated using gCTF[42]. Particle picking was done using Gautomatch (http://www.mrc-lmb.cam.ac.uk/kzhang/) and produced 348,930 particles. All subsequent data processing (from 2D classification to final reconstruction) was done using Cryosparc[43]. After several cycles of 3D classification, a set of 40,390 particles was identified, which produced a 4.7 Å Class 1 map after symmetrized homogeneous refinement. This set was merged with the 151,409 particles from the tilted Class 1 set. After a round of heterogeneous refinement, 147,436 particles were used to generate a 4.3 Å map with C2 symmetry (and a 4.7 Å map with C1 symmetry)(Extended Data Figure 1). The resolution of each reconstruction was determined by the gold standard Fourier shell correlation criterion in Cryosparc[43] and Relion PostProcess[44]. The completeness of the maps was assessed by measuring the Euler angle orientation distribution (as calculated by Cryosparc[43]) and by evaluating each map 3D FSC[14]. (Extended Data Figures 2, 3, and 6). The quality of the map was confirmed by the ability to see separated β-strands and the presence of bulky side chains (Extended data figure 2f).

Structure Solution

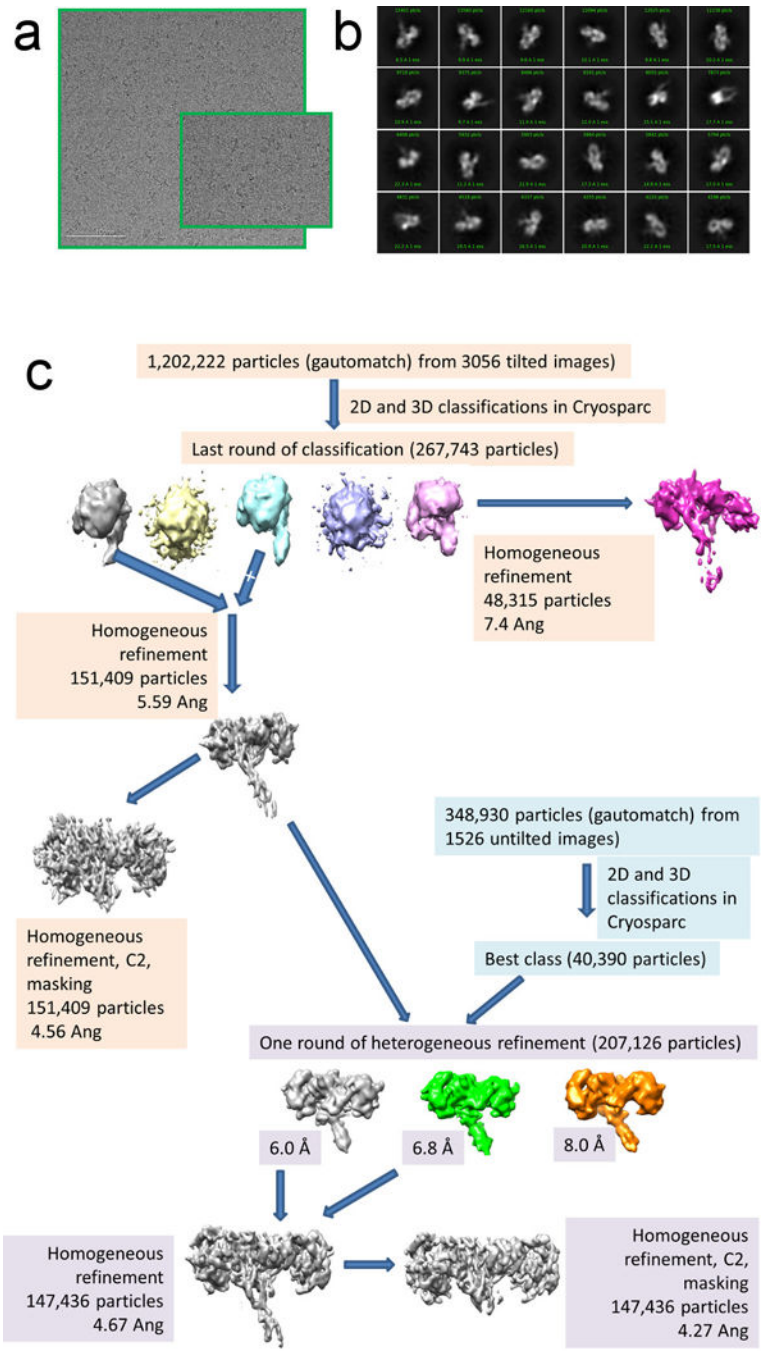
Individual domains of the Insulin receptor (L1, CR, L2, and FnIII-1, from PDB entry 4xzb) were positioned into the initial 4.6 Å resolution map using Molrep[45] and rigid-body refined using COOT[46]. The insulin and the α-CT helix were positioned by overlaying the microreceptor structure from PDB entry 3W11[9] onto the L1 domain of the complex and rigid body fitting them into the available density. The α-CT helix was manually extended, while the full insulin was built using as reference one of the monomers from PDB entry 1ZNI. The full structure was subjected to 5 cycles of global real-space refinement with NCS, rotamer, Ramachandran plot and C-beta deviations restraints enabled in Phenix[47]. This model was subsequently refined against the 4.3 Å map following the same real space refinement procedure in Phenix[47]. The resulting model is identical to the 4.6 Å model, but for the addition of the first 7 N-terminal residues (His1-Val7) to both receptor monomers, and contains residues His1-Asp591 and Lys691-Val720 (α-CT helix) for both of the IR monomers and residues Gly1-Lys21 of chain A and Phe1-Ala30 of chain B for both bound Insulins. Sugars were added to residues Asn16, Asn25, Asn111, Asn255, Asn397 and Asn418 of both monomers. The sugars were extracted from PDB entry 4ZXB. All subsequent, lower resolution models were generated by manually positioning the higher resolution structure onto the available density and rigid body refining the individual domains to their final position using COOT[46]. Subsequently, the structures were subjected to 5 cycles of global real-space refinement with rotamer, Ramachandran plot and C-beta deviations restraints enabled in Phenix[47]. The lack of definition for the FnIII-2 and especially FnIII-3 domains may be due to these regions intrinsic flexibility. While one FnIII-2 region was clearly visible in the un-symmetrized map, unstructured density that could be attributed to the remaining missing regions can be seen by contouring the maps at very low sigma (Extended Data Figure 5).

Figure 2A and extended Figures 1C, 2E, 3E and 6E were generated with Chimera[48]; all other figures, unless otherwise specified, were generated with PyMOL (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC).

Data Availability

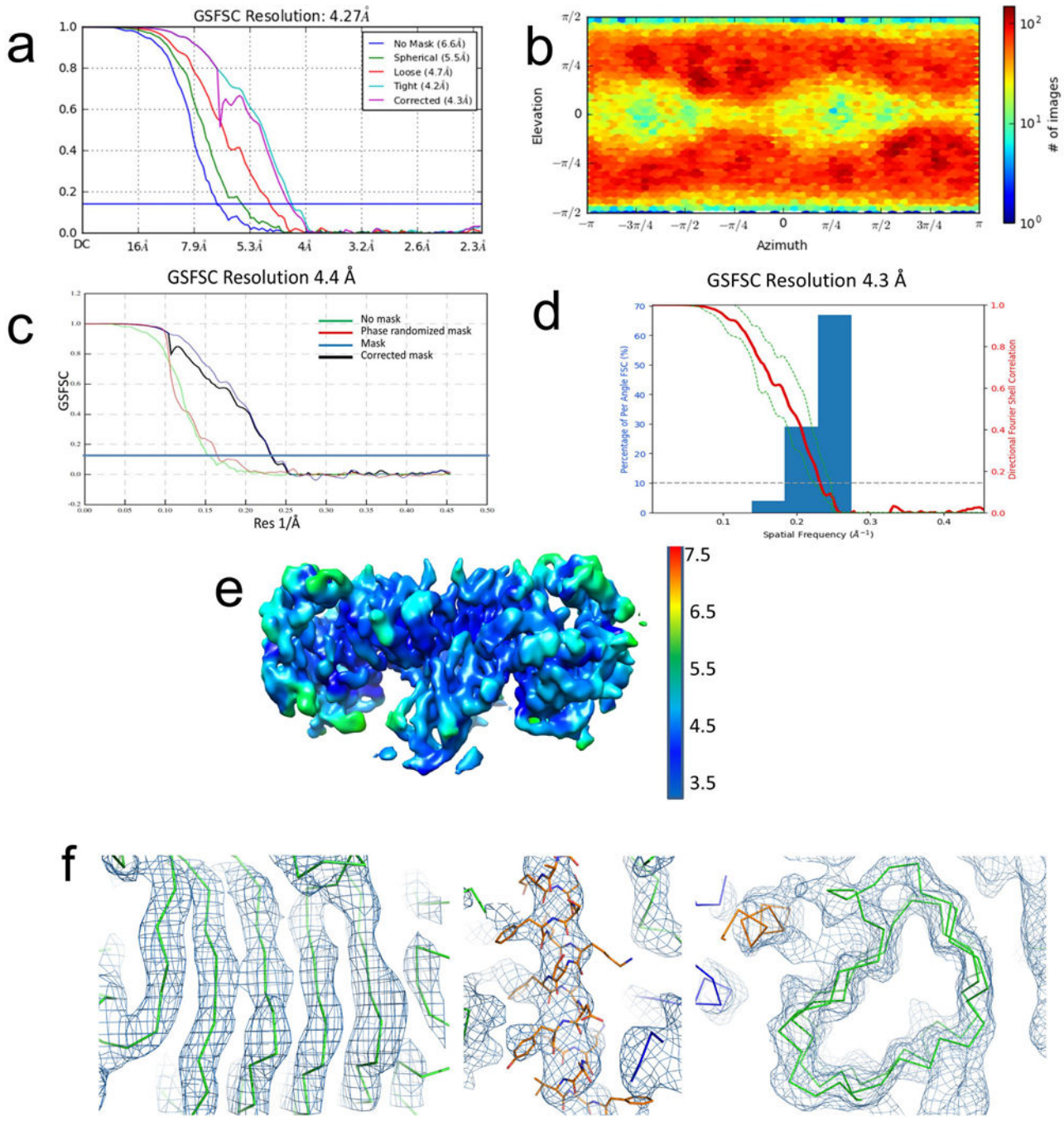
The maps and coordinates generated and analyzed during the current study have been deposited with the EMDB and PDB, accession codes EMD-7462 (PDB 6CE9), EMD-7463 (PDB 6CEB), EMD-7461 (PDB 6CE7) for the 4.3 Å, 4.7 Å and 7.4 Å maps, respectively.

Extended Data



Extended Data Figure 1. Data collection and processing

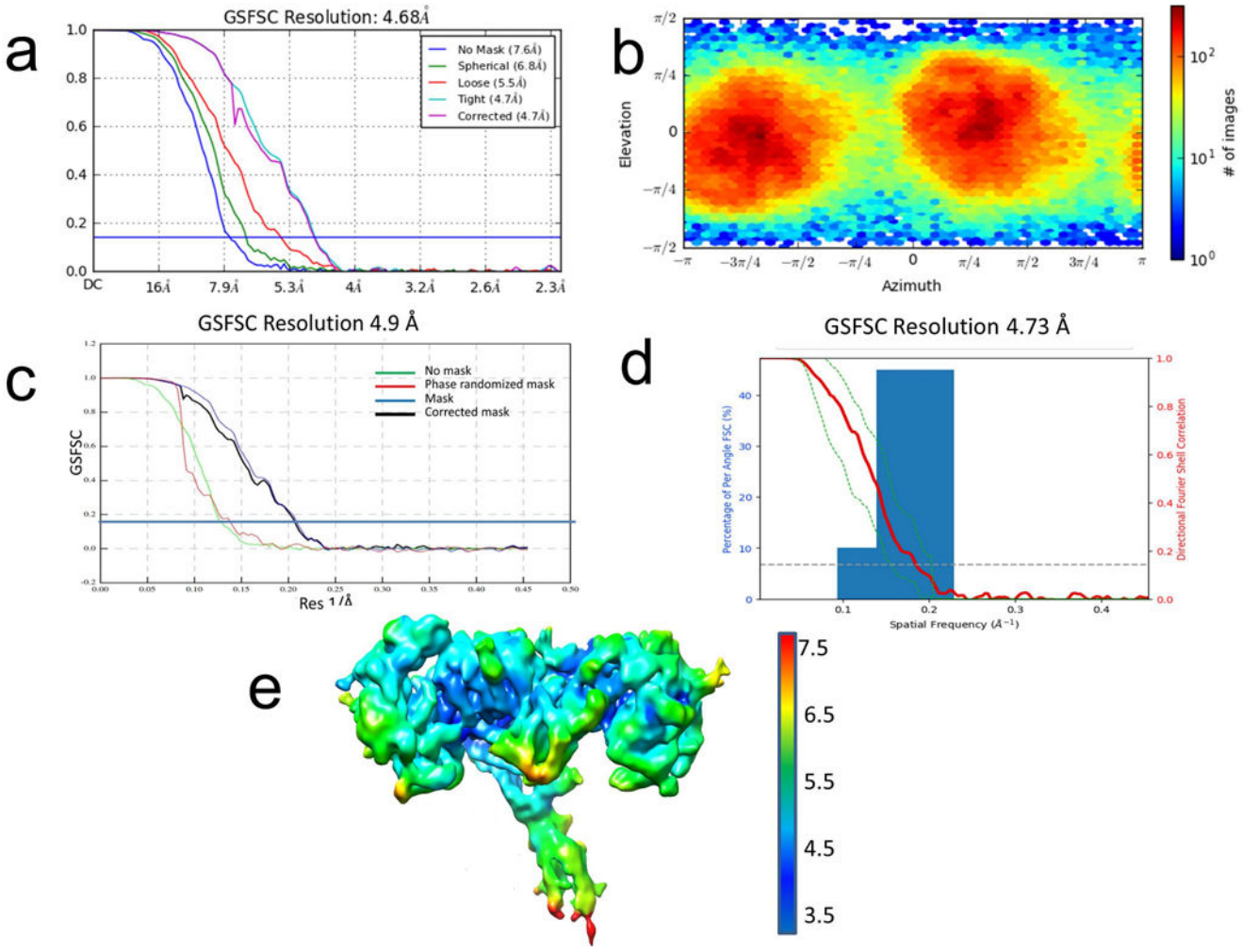
a, Representative micrograph for the IR:insulin complex. Images were collected on a Thermo-Fisher Krios equipped with an energy filter and a Gatan K2 Counting camera; the magnification was set to 105,000×, with a calibrated pixelsize of 1.10 Å. **b**, Representative 2D class averages as calculated with Cryosparc[41]. **c**, Schematic diagram of the data processing.



Extended Data Figure 2. Map of Class 1 reconstructed using C2 symmetry to a resolution of 4.3 Å

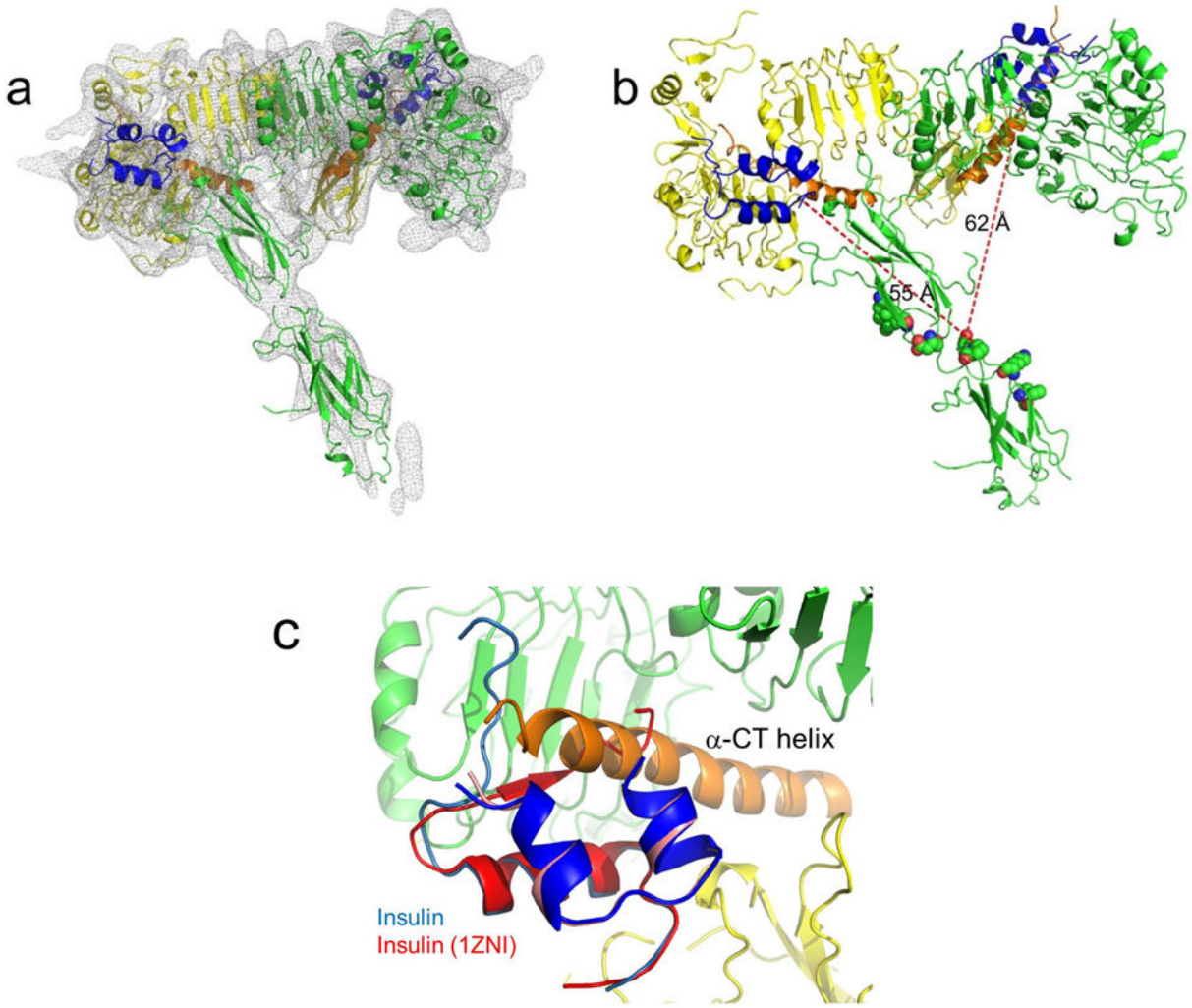
a, Gold standard FSC curve from Cryosparc[43]. **b**, Euler angle orientation distribution from Cryosparc[41]. **c**, Gold standard FSC curve as calculated in RELION[44]. **d**, Plot of the global half-map FSC (solid red line) and spread of directional resolution values (+/− 1σ from mean, green dotted lines; the blue bars are a histogram of 100 such values evenly sampled over the 3D FSC[14]. **e**, Local resolution distribution (as calculated in Cryosparc[43]). **f**, selected areas from the Class 1 C2 map at a resolution of 4.3 Å (map contoured at 12σ). Left, individual β-

strands in the L2 region are well separated; Center: bulky side chains are visible in the electron density; Right: density for the L1 β-barrel.



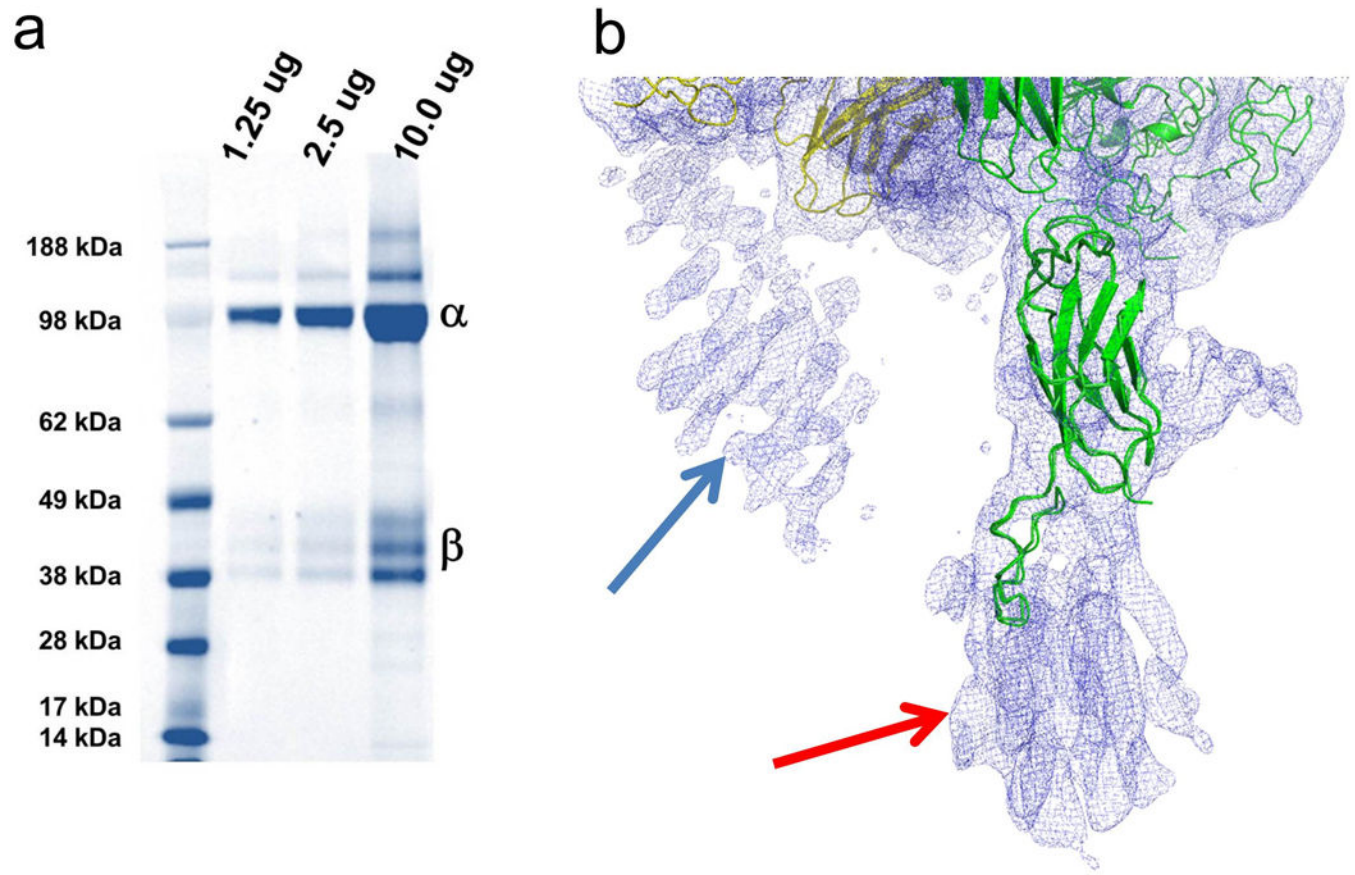
Extended Data Figure 3. Map of Class 1 reconstructed using C1 symmetry to a resolution of 4.7 Å

a, Gold standard FSC curve from Cryosparc[43]. **b**, Euler angle orientation distribution from Cryosparc[41]. **c**, Gold standard FSC curve as calculated in RELION[44]. **d**, Plot of the global half-map FSC (solid red line) and spread of directional resolution values (+/− 1σ from mean, green dotted lines; the blue bars are a histogram of 100 such values evenly sampled over the 3D FSC[14]. **e**, Local resolution distribution (as calculated in Cryosparc[43]).



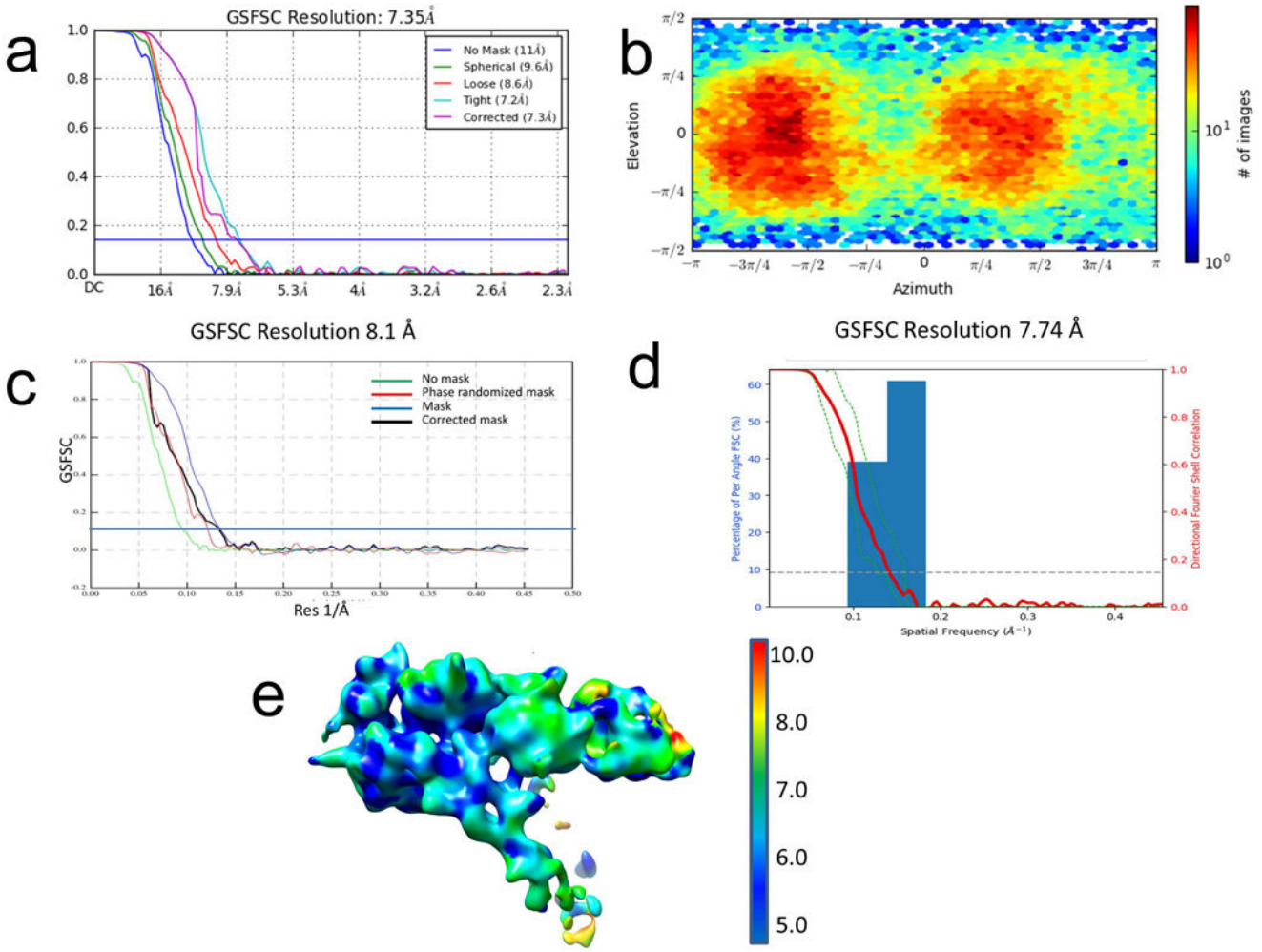
Extended Data Figure 4. Insulin binding and crystallographic S2 site

a, Map of the Class 1 structure obtained without applying C2 symmetry, with the IR model shown as a cartoon. The map is asymmetric and only one of the FnIII-2 sub domains is clearly visible. **b**, Positioning of the FnIII-2 sub domain allows for analyzing the relative position of the bound insulin (blue) and the proposed S2 site (shown as spheres, residues selected according to[26]: the insulin is between 55 and 62 Å away, and on the opposite side of the proposed S2. **c**, Overlay of free insulin (1ZNI) to the insulin bound to IR.



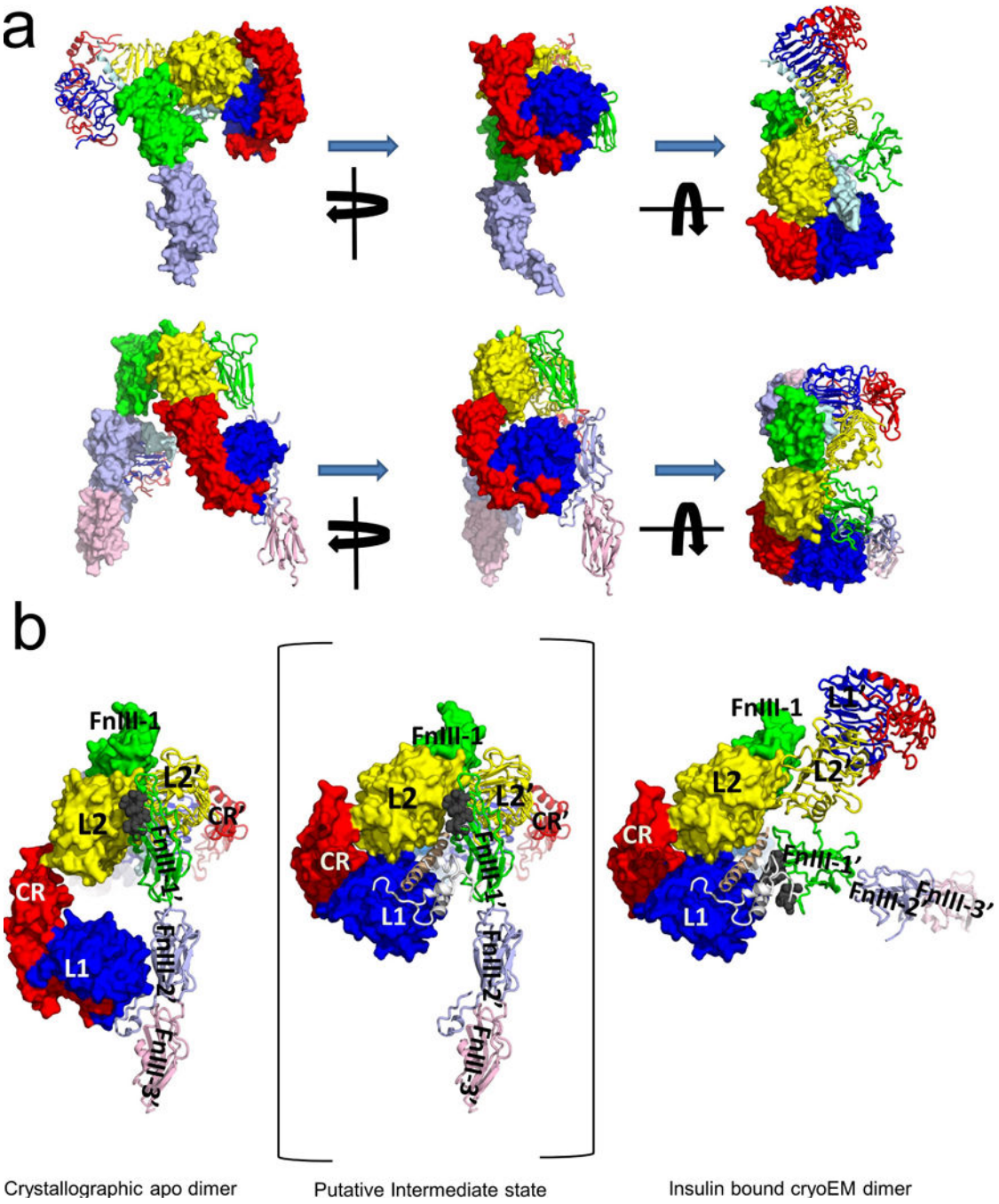
Extended Data Figure 5. IR sample characterization

a, Coomassie stained SDS-PAGE gel. The protein was solubilized in PBS at pH=7.2 and run on a 4-12% Bis-Tris gel in 1× MOPS. Molecular weight markers are labelled. Bands corresponding to the α- and β-chains are indicated. This experiment was only run once to confirm sample quality was as reported by R&D Systems. For gel source data, see Supplementary Figure 1. **b**, CryoEM density for the un-symmetrized map (4.7 Ang resolution) countered at 6σ. Density that can be attributed to the second FnIII-2 domain (blue arrow) as well as one of the FnIII-3 domains (red arrow) is visible, but it is not of sufficient quality for building the model.



Extended Data Figure 6. Map of Class 2 reconstructed using C1 symmetry to a resolution of 7.4 Å

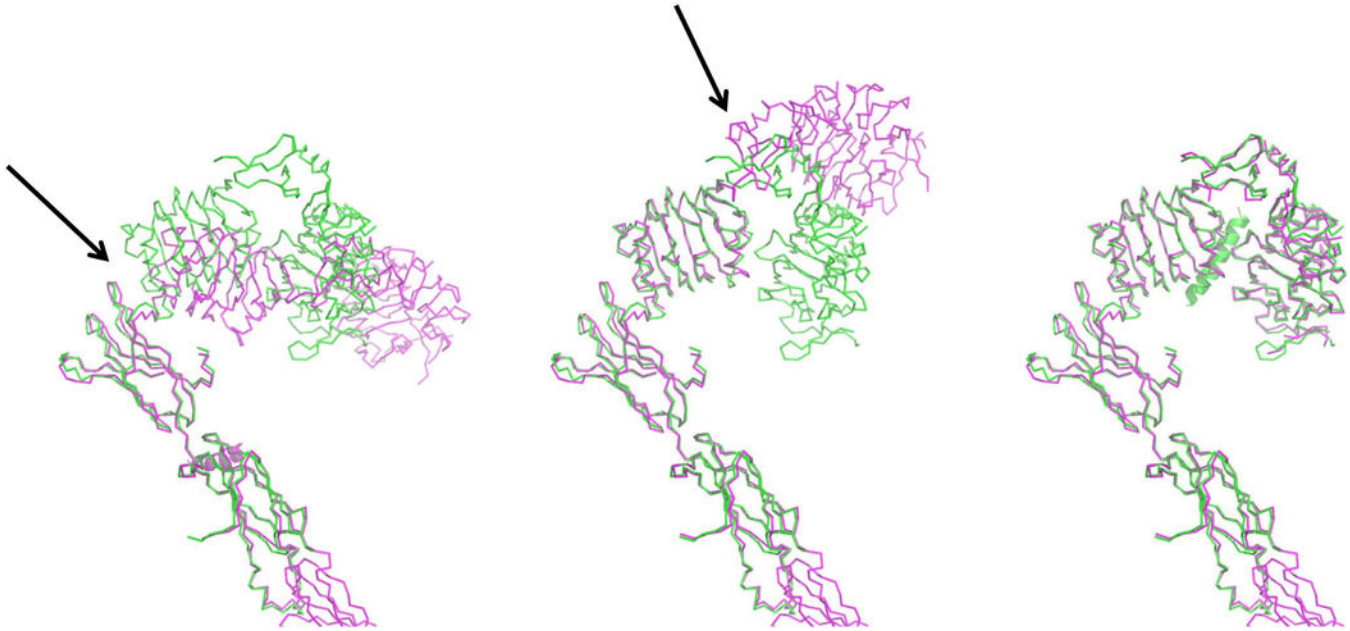
a, Gold standard FSC curve from Cryosparc[43]. **b**, Euler angle orientation distribution from Cryosparc[41]. **c**, Gold standard FSC curve as calculated in RELION[44]. **d**, Plot of the global half-map FSC (solid red line) and spread of directional resolution values (+/− 1σ from mean, green dotted lines; the blue bars are a histogram of 100 such values evenly sampled over the 3D FSC[14]. **e**, Local resolution distribution (as calculated in Cryosparc[43]).



Extended Data Figure 7. Comparison between the crystallographic and the cryoEM dimer

a, The top panel shows three different views of the cryoEM dimer, related by the 90° rotation shown in the figure. This is the surface (for one monomer) and cartoon (for the other) representation of the coordinates fitted to the un-symmetrized, 4.7 Å map: in this model one of the FnIII-2 subdomains is visible. Both monomers are colored according to the subdomains (L1, blue; CR, red; L2, yellow; FnIII-1, green, FnIII-2, light blue). The bottom panel shows three corresponding views for the crystallographic dimer. One monomer is shown as surface, the other as a cartoon; both are colored according to the subdomains (as

for the cryoEM dimer; the additional FnIII-3 is colored in light purple). The relative arrangement of the two monomers in the cryoEM and crystallographic dimers are completely different. **b**, If the crystallographic IR dimer (left panel) represents the biological apo IR dimer, the transition to the insulin bound dimer (right panel) would imply a putative intermediate state shown in the middle panel: in this state the large conformational changes required to accommodate the αCT helix (orange ribbon) and the insulin (white ribbon), and to allow them to engage the S2 site (black spheres) would require disruption of the extensive surface interface between L1 and FnIII-2′/3′ and L2-FnIII-1′ interactions. In the middle and right panel the αCT helix and the insulin are from the cryo-EM structure.



Extended data Figure 8. Comparison between the bound and unbound IR monomer

Transition from the unliganded form (magenta, 4ZXB monomer) to the bound form (green, cryo-EM structure) for the IR monomer; this transition can be described as two rotations with respect to the linker regions identified by the black arrows. The α-CT helix is shown as cartoon.

Extended Data Table 1

Summary of the interactions between the insulin receptor and bound insulin.

Insulin residues	IR L1 (IR monomer A)	IR α-CT helix	IR L2 (IR monomer A)	IR FnIII-1′ (IR monomer B)
	L62, F64, F88, F89, Y91, V94, F96, R188	F701, F705		
	R14, Q34, L36, F88	Y708		
	R118	E698		

Insulin residues	IR L1 (IR monomer A)	IR α-CT helix	IR L2 (IR monomer A)	IR Fnlll-1′ (IR monomer B)
	E120	R702		
	R37, F64	L709		
V3A, G8B, L11B, V12B		D707, H710		
G1A, I2A, Y19A, L15B		F714		
V12B	F39, F64, R65			
Y16B	F39, K40			
F24B	R14, N15, L37	V713		
F25B		P716, S717, S719		
Y26B	D12, R14	P716		
Chain A, C7-T8; Chain B, E4-H10 *				P495-R498, R539-N541, E575
		L696-K703	G346-N349, R372-Y374	
		L696-K703		D496, R498-L501, V570-T571

Interactions are shown along the rows. Light blue: interactions structurally characterized[7]. Red: interactions predicted by biochemical analysis[6], but not seen in the microreceptor structures[7]. Green: new interactions identified in the cryoEM map.

* Insulin residues predicted to belong to S2, with no IR partner assigned[6].

Extended Data Table 2

Human IR ECD binding and kinetic data. Top: Human IR ECD binding data for recombinant human insulin. Data is reported as the geometric mean for 23 replicates; the geometric standard deviation is shown in parentheses. The potency measured with the ^{125}I-insulin competition assay agrees very well with values from the literature[17].

Bottom: Human IR ECD Binding Kinetic Data for RHI. Data is reported as the geometric mean for three replicates; the geometric standard deviation is shown in parentheses. The SPR data is very consistent with values from the literature[31].

[^{125}I]-Insulin Competition Binding: IC_{50} (nM)	Biacore Binding: High Affinity Site $K_{d\ (nM)}$	Biacore Binding: Low Affinity Site $K_{d\ (nM)}$
3.0	**2.7**	**220**
(2.6)	(1.2)	(1.7)

Biacore Binding: High Affinity Site k_{on} ($nM^{-1}\ s^{-1}$)	Biacore Binding: High Affinity Site k_{off} (s^{-1})	Biacore Binding: Low Affinity Site k_{on} ($nM^{-1}\ s^{-1}$)	Biacore Binding: Low Affinity Site k_{off} (s^{-1})
0.0016	**0.0044**	**0.00061**	**0.13**
(1.1)	(1.1)	(1.3)	(1.3)

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The work presented here was conducted at the National Resource for Automated Molecular Microscopy located at the New York Structural Biology Center, supported by grants from the NIH (GM103310, OD019994) and the Simons Foundation (349247).

The authors would like to acknowledge the entire staff of the Simons Electron Microscopy Center at the New York Structural Biology Center (New York, NY) for continuous help and technical support and George Boykow, Aimie Ogawa and Lei Zhang (Merck In vitro Pharmacology group (Kenilworth, NJ)) for providing assay support.

References

1. Saltiel A, Kahn C. Insulin signalling and the regulation of glucose and lipid metabolism. Nature. 2001; 414:799–806. [PubMed: 11742412]
2. Adamo M, Raizada M, LeRoith D. Insulin and insulin-like growth factor receptors in the nervous system. Mol Neurobiol. 1989; 3:71–100. [PubMed: 2553069]
3. Frasca F, et al. The role of insulin receptors and IGF-I receptors in cancer and other diseases. Arch Physiol Biochem. 2008; 114:23–37. [PubMed: 18465356]
4. Craft S. Alzheimer disease: Insulin resistance and AD–extending the translational path. Nat Rev Neurol. 2012; 8:360–362. [PubMed: 22710630]
5. Seino S, Seino M, Nishi S, Bell G. Structure of the human insulin receptor gene and characterization of its promoter. Proc Natl Acad Sci U S A. 1989; 86:114–118. [PubMed: 2911561]
6. De Meyts P. Insulin/Receptor binding: the last piece of the puzzle? Bioessays. 2015; 37:389–397. [PubMed: 25630923]
7. Menting J, et al. How insulin engages its primary binding site on the insulin receptor. Nature. 2013; 493:241–245. [PubMed: 23302862]
8. McKern NM, et al. Structure of the insulin receptor ectodomain reveals a folded over conformation. Nature. 2006; 443:218–221. [PubMed: 16957736]
9. Smith B, et al. Structural Resolution of a tandem hormone-bindng element in the insulin receptor and its implications for design of peptide agonists. Proc Natl Acad Sci USA. 2010; 107:6771–6776. [PubMed: 20348418]
10. Croll T, et al. Higher-Resolution Structure of the Human Insulin Receptor Ectodomain: Multi-Modal Inclusion of the Insert Domain. Structure. 2016; 24:469–476. [PubMed: 26853939]
11. Whittaker L, Hao C, Fu W, Whittaker J. High affinity insulin binding: insulin interacts with two receptor ligand binding sites. Biochemistry. 2008; 47:12900–12909. [PubMed: 18991400]
12. Tatulian SA. Structural dynamics of insulin Receptor and transmembarne signaling. Biochemistry. 2015; 54:5523–5532. [PubMed: 26322622]
13. De Meyts P, Whittaker J. Structural biology of insulin and IGF1 receptors: implications for drug design. Nat Rev Drug Discov. 2002; 1:769–783. [PubMed: 12360255]
14. Tan Y, Baldwin P, Davis J, Williamson J, Potter C, Carragher B, Lyumkis D. Addressing preferred specimen orientation in single-particle cryo-EM through tilting. Nat Methods. 2017; 14(8):793–796. [PubMed: 28671674]
15. Jain T, Sheehan P, Crum J, Carragher B, Potter C. Spotiton: a prototype for an integrated inkjet dispense and vitrification system for cryo-TEM. J Struct Biol. 2012; 179:68–75. [PubMed: 22569522]
16. Razinkov I, et al. A new method for vitrifying samples for cryoEM. J Struct Biol. 2016; 195:190–198. [PubMed: 27288865]
17. Markussen J, Halstrøm J, Wiberg F, Shaffer L. Immobilized insulin for high capacity affinity chromatography of insulin receptors. J Biol Chem. 1991; 266:18814–18818. [PubMed: 1918001]
18. Whittaker J, Whittaker L. Characterization of the Functional Insulin Binding Epitopes of the Full-length Insulin Receptor. J Biol Chem. 2005; 280:20932–20936. [PubMed: 15799978]
19. Ward CW, Menting J, Lawrence M. The insulin receptor changes conformation in unforeseen ways on ligand binding: sharpening the picture of insulin receptor activation. Bioessays. 2013; 35:945–954. [PubMed: 24037759]

20. Hao C, Whittaker L, Whittaker J. Characterization of a second ligand binding site of the insulin receptor. Biochem Biophys Res Commun. 2006; 347:334–339. [PubMed: 16814253]
21. Roach P, et al. A novel human insulin receptor gene mutation uniquely inhibits insulin binding without impairing posttranslational processing. Diabetes. 1994; 43:1096–1102. [PubMed: 8070609]
22. Accili D, Mosthaf L, Ullrich A, Taylor S. A mutation in the extracellular domain of the insulin receptor impairs the ability of insulin to stimulate receptor autophosphorylation. J Biol Chem. 1991; 266:434–439. [PubMed: 1845971]
23. Lebrun C, et al. Antibodies to the extracellular receptor domain restore the hormone-insensitive kinase and conformation of the mutant insulin receptor valine 382. J Biol Chem. 1993; 268:11272–11277. [PubMed: 8388389]
24. Kadowaki H, et al. Mutagenesis of lysine 460 in the human insulin receptor. Effects upon receptor recycling and cooperative interactions among binding sites. J Biol Chem. 1990; 265:21285–21296. [PubMed: 2123490]
25. Brandt J, Andersen A, Kristensen C. Dimeric fragment of the insulin receptor alpha-subunit binds insulin with full holoreceptor affinity. J Biol Chem. 2001; 276:12378–12384. [PubMed: 11278498]
26. Ye L, et al. Structure and dynamics of the insulin receptor: implications for receptor activation and drug discovery. Drug Discov Today. 2017; 22:1092–1102. [PubMed: 28476537]
27. Burgess A, et al. An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors. Mol Cell. 2003; 12:541–552. [PubMed: 14527402]
28. Schäffer L, Ljungqvist L. Identification of a disulfide bridge connecting the alpha-subunits of the extracellular domain of the insulin receptor. Biochem Biophys Res Commun. 1992; 189:650–653. [PubMed: 1472036]
29. Sparrow L, et al. The disulfide bonds in the C-terminal domains of the human insulin receptor ectodomain. J Biol Chem. 1997; 272:29460–29467. [PubMed: 9368005]
30. Gutmann, T., Kim, KH., Grzybek, M., Walz, T., Coskun, U. Visualization of ligand-induced transmembrane signalling in the full-length human insulin receptor. BioRxiv. 2017. https://doi.org/10.1101/207928
31. Subramanian K, Fee C, Fredericks R, Stubbs R, Hayes M. Insulin receptor-insulin interaction kinetics using multiplex surface plasmon resonance. J Mol Recognit. 2013; 26:643–652. [PubMed: 24277609]
32. Johnson J, Wong M, Rutter W. Properties of the insulin receptor ectodomain. Proc Natl Acad Sci USA. 1988; 85:7516–7520. [PubMed: 3050997]
33. Markussen J, Halstrhom J, Wiberg F, Schäffer L. Immobilized Insulin for High Capacity Affinity Chromatography of Insulin Receptors. J Biol Chem. 1991; 266:18814–18818. [PubMed: 1918001]
34. Cosgrove L, et al. Purification and properties of insulin receptor ectodomain from large-scale mammalian cell culture. Protein Expr Purif. 1995; 6:789–198. [PubMed: 8746631]
35. Johnson I. Human Insulin from recombinant DNA technology. Science. 1983; 219:632–637. [PubMed: 6337396]
36. Dandey VP, et al. Spotiton: New features and applications. J Struct Biol. 2018 [Epub ahead of print].
37. Wei H, et al. Optimizing "self-wicking" nanowire grids. J Struct Biol. 2018 [Epub ahead of print].
38. Russo C, Passmore L. Ultrastable gold substrates for electron cryomicroscopy. Science. 2014; 346:1377–1380. [PubMed: 25504723]
39. Glaeser R, Han B. Opinion: hazards faced by macromolecules when confined to thin aqueous films. Biophys Rep. 2017; 3:1–7. [PubMed: 28781996]
40. Suloway C, et al. Automated Molecular Microscopy: The New Leginon System. J Struct Biol. 2005; 151:41–50. [PubMed: 15890530]
41. Zheng S, et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods. 2017; 14:331–332. [PubMed: 28250466]
42. Zhang K. Gctf: real-time CTF determination and correction. J Struct Biol. 2016; 193:1–12. [PubMed: 26592709]

43. Punjani A, Rubinstein J, Fleet D, Brubaker M. CryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods. 2017; 14:290–296. [PubMed: 28165473]

44. Scheres S. A Bayesian View on Cryo-EM Structure Determination. J Mol Biol. 2012; 415:406–418. [PubMed: 22100448]

45. Vagin A, Teplyakov AI. MOLREP: an automated program for molecular replacement. J Appl Cryst. 1997; 30:1022–1025.

46. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004; 60:2126–2132. [PubMed: 15572765]

47. Afonine P, Headd J, Terwilliger T, Adams P. Computational Crystallography Newsletter. 2013; 4:43–44.

48. Pettersen E, et al. UCSF Chimera–a visualization system for exploratory research and analysis. J Comput Chem. 2004; 25:1605–1612. [PubMed: 15264254]

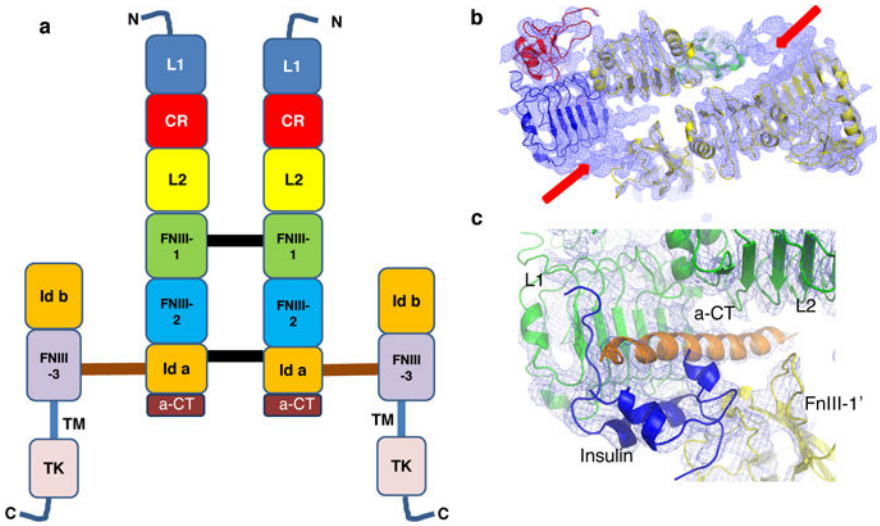


Figure 1. Structure of the Insulin Receptor dimer

a, Domain organization of the full length IR. The inter-monomer disulfide bonds[28,29] are shown as black lines; the intra-monomer "signaling bridge"[26] is shown as an orange line. **b**, CryoEM density map for Class 1 with the IR sub-domains fitted to the density; one monomer is yellow, the other is color coded as in **a**. Density identified by the red arrows can be attributed to the insulin and αCT helix. **c**, Close-up of density with insulin and α-CT helix fitted.

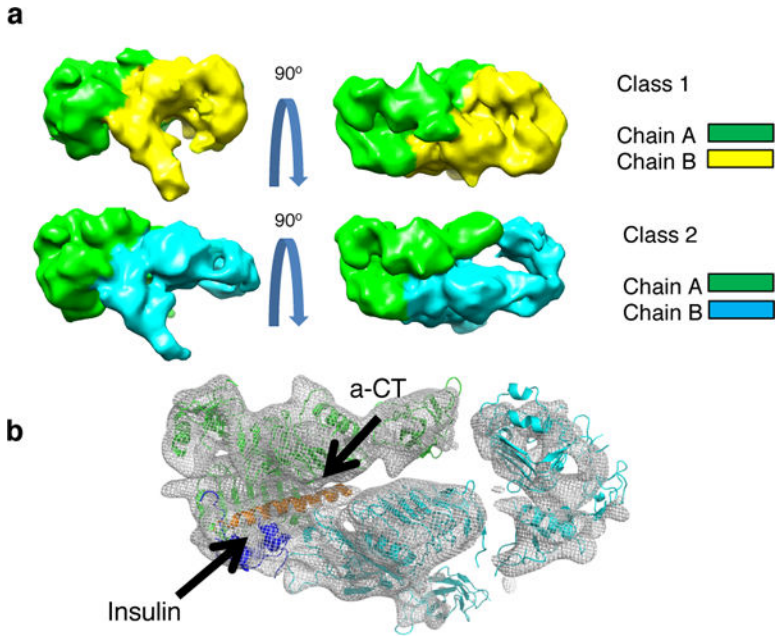


Figure 2. Class 1 and Class 2 dimers

a, Side and top view of the two classes identified during 3D reconstruction. **b** CryoEM map for the Class 2 showing density for the α-CT helix and insulin only in one monomer.

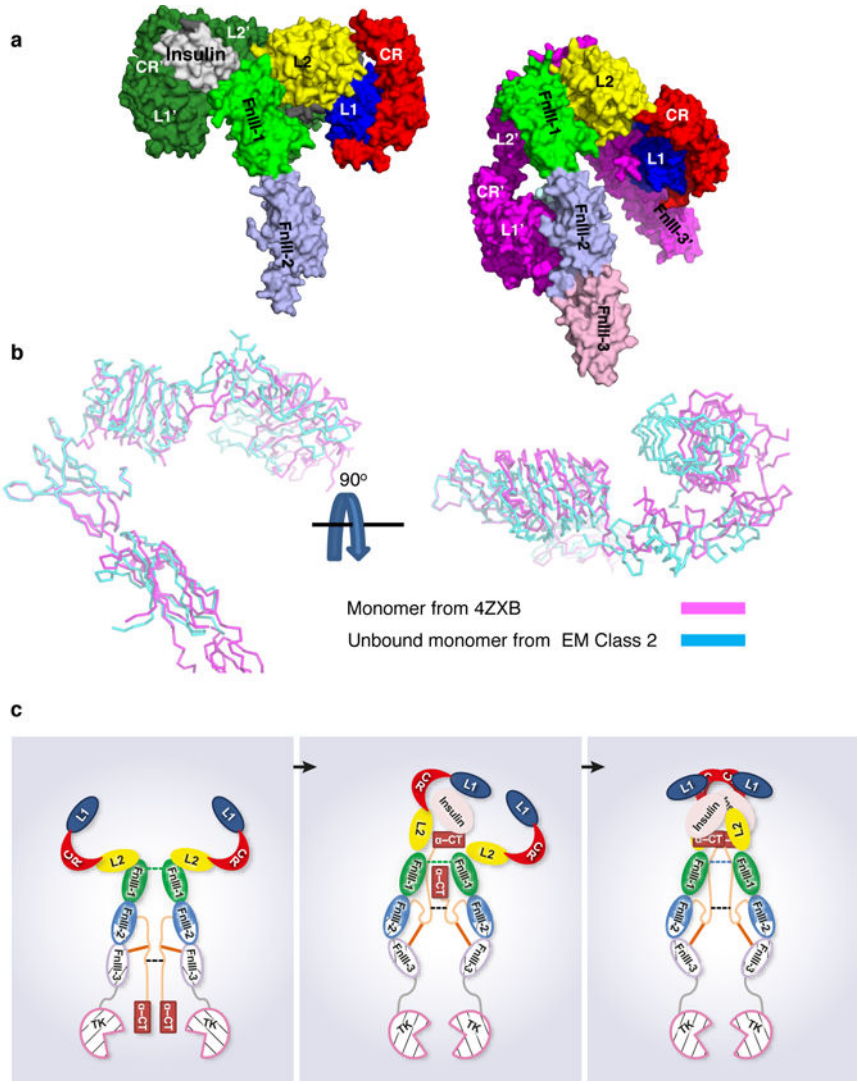


Figure 3. Proposed transduction mechanism

a, Comparison between the cryoEM (insulin bound) and the crystallographic (apo) dimers. Both dimers are shown as surface representation of the coordinates. Similar conformational differences between unbound and insulin bound IR have been observed in Gutman et al.[30]. **b**, Side and top view of the 4ZXB monomer overlaid onto the cryo EM "open" monomer. The overlay was done using the FnIII-1 domain. **c**, Schematic diagram of a possible activation mechanism for IR. The IR sub-domains color coded as in Figure 1a (solid color for the α-chain and light colors/thicker outer lines for the β-chain); the inter-monomer disulfide bonds are shown as dotted lines, the "signaling bridge"[26] as solid orange line. Binding of one insulin molecule to the apo receptor (left to middle panel) causes the L1-CR-L2 subdomains of one monomer, the FnIII-1 subdomain of the other and the α-CT helix to move to generate the binding site. The motion of the α-CT helix and the attached ID-α causes, through the "signaling bridge", a conformational change in the FnIII-3 domain. Since the two ID-α regions are also disulfide bonded, the motion of one is likely to be transmitted to the other, inducing a similar conformational change in the other FnIII-3

domain. These changes propagate through the transmembrane helix to the TK domains, inducing autophosphorylation and activation of the signaling pathway. This state is seen in the cryo-EM Class 2 map. Right: binding of a second insulin molecule recruits the second α-CT (cryo-EM Class1 map) and may fully stabilize the activated complex. Although the diagram suggests that the α-CT helix involved in insulin binding is the one from the same monomer (cis-interaction), there are no currently experimental evidences ruling out a trans-interaction.