Published in final edited form as:
Nature. 2013 January 10; 493(7431): 241–245. doi:10.1038/nature11781.

How insulin engages its primary binding site on the insulin receptor

John G. Menting[1,*], **Jonathan Whittaker**[2,*], **Mai B. Margetts**[1], **Linda J. Whittaker**[2], **Geoffrey K.-W. Kong**[1], **Brian J. Smith**[1,3], **Christopher J. Watson**[4], **Lenka Žáková**[5], **Emília Kletvíková**[5], **Jiří Jiráček**[5], **Shu Jin Chan**[6], **Donald F. Steiner**[6], **Guy G. Dodson**[4,‡], **Andrzej M. Brzozowski**[4], **Michael A. Weiss**[2], **Colin W. Ward**[1], and **Michael C. Lawrence**[1,7]

[1]Walter and Eliza Hall Institute of Medical Research, 1G Royal Parade, Parkville, Victoria 3052, Australia

[2]Department of Biochemistry, School of Medicine, Case Western Reserve University, Cleveland, Ohio 44106, USA

[3]Department of Chemistry, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, Victoria 3086, Australia

[4]York Structural Biology Laboratory, Department of Chemistry, University of York, York YO10 5DD, UK

[5]Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, v.v.i., 16610 Prague, Czech Republic

[6]Department of Medicine, University of Chicago, Chicago, Illinois 60637, USA

[7]Department of Medical Biology, University of Melbourne, Royal Parade, Parkville, Victoria 3010, Australia

Abstract

Insulin receptor signalling has a central role in mammalian biology, regulating cellular metabolism, growth, division, differentiation and survival[1,2]. Insulin resistance contributes to the pathogenesis of type 2 diabetes mellitus and the onset of Alzheimer's disease[3]; aberrant signalling occurs in diverse cancers, exacerbated by crosstalk with the homologous type 1 insulin-like growth factor receptor (IGF1R)[4]. Despite more than three decades of investigation, the three-dimensional structure of the insulin–insulin receptor complex has proved elusive, confounded by the complexity of producing the receptor protein. Here we present the first view, to our

Correspondence and requests for materials should be addressed to: M.C.L. (lawrence@wehi.edu.au) or M.A.W. (michael.weiss@case.edu).
*These authors contributed equally to this work.
‡Deceased.

Supplementary Information is available in the online version of the paper.

Author Contributions J.G.M. and J.W. contributed equally to the paper. J.G.M. purified and crystallized samples, collected data and performed the ITC study; J.W. and L.J.W. performed receptor photo-crosslinking experiments; M.B.M. performed molecular biology, cell culture and crystallization experiments; S.J.C. performed insulin photo-crosslinking experiments; G.K.-W.K. and C.J.W. performed crystallography experiments; B.J.S. performed calculations; E.K., L.Z. and J.J. prepared insulin analogues; C.W.W., M.A.W., J.W., D.F.S., S.J.C., J.G.M. and M.C.L. designed the experiments and analysed data. C.W.W., M.A.W., A.M.B., G.G.D. and M.C.L. wrote the paper. All authors discussed the results and commented on the manuscript.

Author Information Atomic coordinates and structure factors for complexes A,B,C and D have been deposited with the Protein Data Bank under accession codes 3W11, 3W12, 3W13 and 3W14, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper.

knowledge, of the interaction of insulin with its primary binding site on the insulin receptor, on the basis of four crystal structures of insulin bound to truncated insulin receptor constructs. The direct interaction of insulin with the first leucine-rich-repeat domain (L1) of insulin receptor is seen to be sparse, the hormone instead engaging the insulin receptor carboxy-terminal α-chain (αCT) segment, which is itself remodelled on the face of L1 upon insulin binding. Contact between insulin and L1 is restricted to insulin B-chain residues. The αCT segment displaces the B-chain C-terminal β-strand away from the hormone core, revealing the mechanism of a long-proposed conformational switch in insulin upon receptor engagement. This mode of hormone–receptor recognition is novel within the broader family of receptor tyrosine kinases[5]. We support these findings by photo-crosslinking data that place the suggested interactions into the context of the holoreceptor and by isothermal titration calorimetry data that dissect the hormone–insulin receptor interface. Together, our findings provide an explanation for a wealth of biochemical data from the insulin receptor and IGF1R systems relevant to the design of therapeutic insulin analogues.

Insulin comprises two chains (A and B) containing three α-helices (residues A1–A8, A12–A18 and B9–B19) constrained by one intra-and two interchain disulphide bonds[6] (Fig. 1a). The insulin receptor is a disulphide-linked $(\alpha\beta)_2$ homodimer; the extracellular portion of each αβ protomer contains six domains (L1, CR, L2, FnIII-1, FnIII-2 and FnIII-3) and an insert domain (ID) within FnIII-2[7] (Fig. 1b). The α-chain component of the ID is terminated by a segment termed αCT, spanning residues 704–719 (using the numbering of the insulin receptor exon 11 – isoform[8]). Two surfaces of insulin are understood to interact with the insulin receptor[9,10]. The first consists predominantly of hormone-dimerizing residues and contacts the primary binding site on the receptor (site 1; dissociation constant (K_d) ~6.4 nM) comprising the αCT segment from one insulin receptor α-chain and the central β-sheet of L1 (L1–β_2) of the other α-chain within the insulin receptor dimer[11–14] (Supplementary Fig. 1a). The second consists predominantly of hormone-hexamerizing residues and is proposed to interact with a secondary insulin receptor site (site 2; K_d ~400 nM) at the junction of FnIII-1 and FnIII-2 of the insulin receptor α-chain opposite to that contributing L1 to site 1 (Supplementary Fig. 1a)[7,9–11,13–15]. Structural study of the insulin receptor and IGF1R ectodomains is difficult; their large size, flexible multidomain morphology and extensive glycosylation and disulphide bonding posing challenges in production and crystallization. Only three ectodomain structures exist: those of the *apo*-insulin receptor[7,12] and the L1–CR–L2 fragments of insulin receptor[16] and IGF1R[17] (3.7 Å, 2.3 Å and 2.6 Å resolution, respectively); none is in complex with ligand.

We present here the first view of the insulin–site 1 interaction on the basis of four crystal structures (complexes A, B, C and D) varyingly using the L1–CR construct IR310.T (Fig. 1c) and the L1–CR–L2– (FnIII-1)–αCT(704–719) homodimeric construct IR593.αCT (Fig. 1d), which has αCT(704–719) linked to FnIII-1. Complex A (3.9 Å resolution) contains human-insulin-bound IR310.T in complex with a Fab fragment from monoclonal antibody 83-7 (ref. 7) and exogenous αCT(704–719) peptide. Complex B (4.3 Å resolution) is the same as complex A but uses the high-affinity, truncated insulin analogue [D- Pro^{B26}]-DTI-NH_2[18]. Complex C (4.4 Å resolution) is the same as complex B but uses the longer peptide αCT(697–719). Complex D (4.4 Å resolution) contains bovine insulin bound to IR593.αCT in complex with a Fab fragment from monoclonal antibody 83-14 (ref. 7).

The limited diffraction of these crystals necessitated the use of the following: (1) weak data at the diffraction limit (a process validated by the half-data set Pearson correlation coefficient $CC_{1/2}$ criterion[19,20] and ultimately by the evident quality (see below) of the resultant electron-density difference maps (EDDMs)); (2) B-factor-sharpened EDDMs in model building[21]; and (3) dynamic ‘reference-structure’ and ‘jelly-body’ restraints during individual atomic-coordinate refinement whereby we exploited the known higher-resolution

structures of the underlying fragments[22,23]. Molecular replacement located the insulin receptor and Fab domains and generated a clear solution for the insulin three-helix 'core' structure within each complex, positioning the core identically on the respective L1–β_2 surfaces. The strongest feature in the resultant EDDMs was a common, helical-like structure lying adjacent to the insulin B-helix on the L1–β_2 surface that, upon B-factor sharpening, exhibited side-chain density of distinctive sizes and at spacing consistent with an α-helix. In complex D, this feature was connected to the C-terminal strand of FnIII-1, allowing its unambiguous interpretation as αCT and restricting its prima facie sequence register to an ambiguity of at most one residue unit, with one alternative strongly favoured. We confirmed our register assignment objectively by devising a broad-window analysis of sequence fit both to the higher-resolution EDDM of complex A and to the profiled residue environments within that complex. Crystallographic processing and refinement statistics are provided in Table 1 and Supplementary Table 1, with further detail in Methods.

The structures of the insulin-bound site 1 assembly within complexes A and D are shown in Fig. 1e, f, respectively; those within complexes B and C are effectively isomorphous with complex A (Fig. 1g). The Fab-complexed forms of the IR310.T and IR593.αCT monomers are shown in Supplementary Fig. 2a, b, respectively. As prepared and crystallized, complex D contains a dimer of IR593.αCT homodimers crosslinked by four insulins; further description of this assembly is provided in Supplementary Fig. 2c – e. Despite the non-native connection of αCT to FnIII-1 in IR593.αCT (Fig. 1d), the insulin-bound site 1 assembly within complex D is also superimposable on that of complex A (Fig. 1g). Refinement of these complexes provided EDDMs that revealed side-chain bulk for all components within the observed site 1 interface (Fig. 1h). The mode of insulin–site 1 interaction emerging from these structures is as follows. The insulin B-helix (B7–B21) engages the C-terminal end of the L1–β_2 strands; the insulin A-chain has no interaction with L1 (Fig. 1e, f). Both chains interact extensively with αCT. Notably, the αCT helix is, with respect to its *apo*-insulin receptor counterpart, both displaced on the L1–β_2 surface and C-terminally extended to include residues 711–715 (see Fig. 2a and further discussion below). Residues B22–B30 are not resolved in these structures, but, critically, the αCT helix occupies volume that would otherwise contain insulin residues B26–B30 if the latter retained their receptor-free conformation (see Fig. 2b and Supplementary Fig. 3a). Displacement of residues B26–B30 from the insulin core confirms a long-speculated[24] induced fit upon insulin receptor binding (Supplementary Fig. 1b). In contrast, the helical core of insulin seems, as far as can be discerned at the described resolution, closely similar to that of uncomplexed hormone (Supplementary Fig. 3b). The models suggest that the two most critical hormone-engaging residues in αCT are (1) His 710, which inserts into a pocket formed by invariant[10] insulin residues Val^{A3}, Gly^{B8}, Ser^{B9} and Val^{B12}; and (2) Phe 714, which occupies a hydrophobic crevice formed by invariant insulin residues Gly^{A1}, Ile^{A2}, Tyr^{A19}, Leu^{B11}, Val^{B12} and Leu^{B15} (Fig. 2c). Asn 711 is directed towards Gly^{A1}, Val^{A3} and Glu^{A4}. The hydrophobic face of αCT comprising Phe 705, Tyr 708, Leu 709, Val 712 and Val 713 engages a non-polar groove on L1-β_2 formed by Leu 36, Leu 37, Leu 62, Phe 64, Phe 88, Phe 89, Val 94 and Phe 96 (Fig. 2d). Direct interaction between the insulin B-chain B9–B19 helix and L1 seems to be limited: Val^{B12} is positioned between Phe 39, Phe 64 and Arg 65, whereas Tyr^{B16} adjoins Phe 39 (Fig. 2e). Detailed rotameric and atomic level detail of the above interactions is precluded by the limited diffraction.

Relocation of αCT upon hormone binding is supported by recent studies indicating that residues buried beneath αCT in the *apo*-ectodomain may directly contact insulin[25], and that peptide substitutions in the αCT segment upstream of Phe 705 that increase the affinity of the segment for L1 concomitantly impair insulin binding[12]. To investigate further, we used complex C to test whether the relocation observed in complexes A, B and D is a consequence of αCT truncation, as in the *apo*-insulin receptor the visible interaction of αCT

with L1–β_2 begins at Glu 697. It is not—the relocation persists in complex C (Fig. 1g) with no detectable ordering of residues 697–704.

We validated our interpretation of the αCT engagement with L1–β_2 and insulin further by isothermal titration calorimetry (ITC). First, Ala-substituted αCT(704–719) peptides were titrated against the L1– CR–L2 construct IR485 (ref. 16). Second, insulin was titrated against IR485 pre-saturated with Ala-substituted αCT(704–719) (Supplementary Tables 2 and 3 and Supplementary Fig. 4). Individual F705A and L709A substitutions significantly impaired αCT binding to IR485, supporting their modelled engagement with L1–β_2. In contrast, individual H710A, N711A and F714A substitutions significantly impaired insulin binding, supporting their modelled engagement with insulin. These data are summarized schematically in Fig. 3a. An apparent exception is V713A, which affects insulin but not αCT affinity, despite Val 713 being modelled as engaged with L1–β_2. However, this observation aligns with the disorder of Val 713 in the *apo*-insulin receptor ectodomain structure[12]; the interaction of Val 713 with L1–β_2 may thus instead be required to stabilize the critical[11] insulin-engaging residue Phe 714 at the terminus of the αCT helix.

The insulin–site 1 interactions described here are further supported by a 'photo-scan' of the *holo*-insulin receptor αCT segment, wherein individual αCT residues were successively substituted by the photo-activatable residue *para*-azido-Phe (Pap) and then exposed to ultraviolet light to probe covalent interaction with [^{125}I]-Tyr^{A14} insulin[25]. Strong crosslinking was observed for photo-probes at sites 711 and 714, medium at sites 707 and 710, and weak at sites 713 and 715; no crosslinking was observed at sites 705, 706, 708, 709 or 712 (Fig. 3b). This photo-crosslinking efficiency pattern in *holo*-insulin receptor exhibits a marked correlation with and supports the insulin–site 1 juxtaposition described above (Fig. 3c – e).

Complexes B and C probed whether the stable β-turn at B24–B26 in receptor-free [D-Pro^{B26}]-DTI-NH_2[18] might result in ordering of residues B22–B26 upon binding site 1—it did not, as these residues remained unresolved. We investigated the insulin B-chain C-terminal strand further by seeking to photo-crosslink individual Pap probes at B24 and B25 to the truncated L1–CR–L2 construct IR468[26] in the presence of an αCT peptide. In both *holo*- and ectodomain insulin receptor complexes, these probes respectively crosslink to L1 and αCT[27,28]. We observed photo-crosslinking for Pap^{A3}, Pap^{A8} and Pap^{B16} (as in *holo*-insulin receptor[27]) but not for Pap^{B24} or Pap^{B25} (Supplementary Fig. 5). We thus propose that full stabilization of B22–B30 upon detachment from the hormone core requires the presence of the FnIII domains.

Superposition of the site 1 complexed assembly onto the L1 domain in the *apo*-insulin receptor ectodomain structure (Supplementary Fig. 6) shows that the insulin surface proposed[10] to bind site 2 overlays the FnIII-1–FnIII-2 junction of the alternative αβ protomer, that is, the proposed site 2 itself. However, significant steric clash is present, indicating that insulin bridging of sites 1 and 2 requires relative displacement of the FnIII-1 and FnIII-2 domains away from the L1–CR module. Such change would alter the relative location of the membrane-proximal FnIII-3 domains, potentially initiating signal transduction[13].

Concordance of our structures with the extensive biochemical literature of this system is presented in Supplementary Discussion. Briefly, the structures are consistent with alanine scanning mutagenesis of the insulin receptor L1 domain, revealing now that of the 'hot-spot' residues, the set Gln 34, Leu 36, Leu 37, Phe 64, Phe 88, Phe 89 and Glu 120 contact the αCT segment whereas only Phe 39 contacts insulin. The structural engagement of insulin receptor residues His 710 and Phe 714 with insulin concords with (1) alanine scanning

mutagenesis of the insulin receptor ectodomain; (2) hydrogen/deuterium exchange studies of the closely related type 1 insulin-like growth factor receptor (IGF1R); and (3) the lack of hormone binding by insulin receptor chimaeras and IGF1R chimaeras containing the αCT segment from the insulin-receptor-related receptor, which has threonine at the residue equivalent to insulin receptor Phe 714 (IGF1R Phe 701). Our structures are also consistent with mutational analysis of insulin, showing now that the hotspot residues Gly^{A1}, Ile^{A2}, Val^{A3}, Tyr^{A19}, Gly^{B8} and Leu^{B11} make contact with the key αCT residues His 710 and Phe 714 (Fig. 2c). Although residues Gly^{B23}, Phe^{B24} and Phe^{B25} are unresolved here, evidence is presented for the likely repositioning of the B-chain C terminus into the volume between the αCT segment, the L1–β_2 surface and the adjacent CR domain; such repositioning rationalizes the enhanced affinity of analogues containing D-aminoacid substitutions at B24.

Our structural and biochemical findings provide groundbreaking insight into the primary insulin–insulin receptor interaction. Particularly important in binding is the αCT segment, which is without known counterpart in other mammalian receptor tyrosine kinases[5], although it is conserved among metazoan insulin receptor and IGF1R. The concomitant structural re-arrangements of hormone and receptor in our model rationalize long-standing anomalies in clinical insulin variants associated with diabetes[29] and define new structural targets for novel therapeutics targeting these receptors.

METHODS SUMMARY

IR310.T was produced by proteolysis of an engineered mini-receptor precursor cIR485 expressed from a stable Chinese hamster ovary Lec8 cell line and purified by ion-exchange and size-exclusion chromatography (SEC), followed by combination with 83-7, re-purification by SEC and finally combination with excess αCT peptide and insulin (or analogue). Crystallization conditions for IR310.T-derived complexes A, B and C were 0.9–1.1M tri-sodium citrate, 0.1 M imidazole-HCl (pH 8.0) + 0.02% sodium azide (NaN_3). IR593.αCT was expressed from a stable Chinese hamster ovary Lec8 cell line followed by insulin affinity chromatography and then SEC of the resultant bovine insulin complex. Samples were combined with 83-14 and re-purified by SEC. The crystallization condition for the IR593.αCT-derived complex D was 9% polyethylene glycol 3350, 200 mM proline, 100 mM HEPES-NaOH (pH 7.5). Diffraction data were collected at the Australian Synchrotron (beamline MX2) and Diamond Light Source (DLS; beamline I24). Structures were solved by molecular replacement. Introduction of Pap residues into the αCT segment used orthogonal tRNA/amber suppression technology. Mutant holoreceptors were expressed transiently using co-transfected 293PEAK rapid cells in Pap-containing medium. Cell detergent lysates were then subjected to wheat-germ agglutinin chromatography; eluates containing the receptor were concentrated before photo-crosslinking to ^{125}I-[Tyr^{A14}] insulin. Photo-crosslinked products were resolved by gel electrophoresis and auto-radiographed. IR468 was produced by transient expression in HEK293 cells. ^{125}I-photo-activatable insulin analogues were incubated with IR468 in the presence of 10^{-5} M αCT(703–719)– Myc and then photo-crosslinked. Samples were resolved by gel electrophoresis and auto-radiographed. For ITC, Ala-substituted αCT peptides (80 μM concentration) were titrated into a sample cell containing IR485 (10 μM concentration); and zinc-free porcine insulin (32–48 μM concentration) was titrated into a sample cell containing IR485 (4–6 μM concentration) pre-incubated with Ala-substituted αCT peptide (10× molar concentration). All samples were in Tris-buffered saline (pH 8.0) plus azide (TBSA).

Full Methods and any associated references are available in the online version of the paper.

METHODS

Cloning, production and purification of IR310.T

High yield and correct folding of IR310.T was achieved by design of a thrombin-cleavable version of IR485[16] (termed cIR485) for use as precursor, followed by production and purification protocols for cIR485 similar to those described for IR485[16]. A 1,741-nucleotide base fragment of insulin receptor was synthesized (DNA2.0) and inserted into the HindIII–XbaI sites of the vector pEE14 (Lonza). The fragment encoded a protein cIR485 ('cleavable' IR485) which consists, in order, of (1) the 27-residue insulin receptor native signal sequence; (2) residues 1–310 of the insulin receptor α-chain (the L1–CR module); (3) a thrombin cleavage site SSSLVPRGSSS; (4) residues 311–485 of the insulin receptor α-chain (the L2 domain); (5) an enterokinase cleavage site DDDDK; and (6) a c-Myc purification tag EQKLISEEDLN[31] (see Supplementary Fig. 7). A 106-nucleotide-base non-coding fragment 5′-GTCG ACGGTACCCCGGGGAATTAATTCCGGGGGCCGCCTCGGAGCATGACC CCCGCGGGCCAGCGCCGCGCGCTCTGATCCGAGGAGACCCCGCGCTC CCGCAGCC-3′ was included between the HindIII site and the region expressing the signal peptide; bases 29–106 of this fragment correspond to those immediately upstream of the human insulin receptor coding region. CHO Lec8 cells (CRL-1737; ATCC)[32] were transfected with the cIR485-containing expression vector using Lipofectamine 2000 (Life Technologies). Selection was undertaken in the presence of methionine sulphoximine (MSX), with cIR485 production monitored via western blot with monoclonal antibodies 83-7[33] and 9E10[34]. Production scale-up, performed under contract by CSIRO, was in spinner flasks (New Brunswick Scientific) containing DMEM/F12 medium (Life Technologies) supplemented with 10% fetal bovine serum (FBS; Life Technologies) plus 25 μM MSX. The medium was supplemented with 0.8 mM butyric acid (Sigma-Aldrich) during the last week of the production phase. A total of 50 l of culture supernatant was collected; harvests were filtered (0.2 μm cutoff) and then concentrated 10-fold (30 kDa cutoff) after addition of 0.1 mM phenylmethanesulphonylfluoride (PMSF; Sigma-Aldrich) and 0.02% sodium azide (NaN_3). Accumulated supernatant harvests were adjusted to pH 8 by addition of 3M Tris-HCl (pH 8.5). Initial purification of cIR485 was by 9E10 antibody affinity chromatography and size-exclusion chromatography (SEC) using procedures effectively identical to those described for purification of IR485[16]. Purified cIR485 in Tris-buffered saline (24.8 mM Tris-HCl pH 8.0, 137 mM NaCl, 2.7 mM KCl) plus 0.02% NaN_3 (TBSA) was then incubated overnight at 37 °C with 0.25 units human thrombin (Roche) per mg of cIR485 in the presence of 10 mM $CaCl_2$. Completion of pro-teolysis was assessed by SDS–PAGE, which revealed bands corresponding to the estimated molecular mass of the IR310.T fragment and the insulin receptor L2 domain (51 kDa and 30 kDa, respectively; Supplementary Fig. 8a). Western blotting with monoclonal antibody 83-7 confirmed that the CR domain was contained in the upper band alone. The sample was diluted 8-fold in buffer A (10 mM ethanolamine-HCl, pH 9.6 + 0.02% NaN_3), centrifuged for 5 min at 17,000g to remove particulates and then loaded onto a Mono Q 5/50 GL column (GE Healthcare). The sample was eluted with a 60 column-volume gradient of buffer A to buffer B (10 mM ethanolamine-HCl, 400 mM NaCl, pH 9.6 + 0.02% NaN_3; Supplementary Fig. 8b) and the fractions assessed by SDS–PAGE. Fractions containing IR310.T were pooled, concentrated and re-run in TBSA on a Superdex 200 10/300 column (GE Healthcare). The chromatogram exhibited three overlapping peaks, probably arising from multimerization (Supplementary Fig. 8c). SDS–PAGE of fractions revealed the presence of three closely spaced bands, which we attributed to varying glycosylation (Supplementary Fig. 8d). The final yield of IR301.T was ~0.5 mg ml^{-1} of cell culture.

Large-scale production and purification of Fab 83-7

A hybridomal cell line expressing monoclonal antibody 83-7[33] was a gift from K. Siddle. Large-scale production and purification of monoclonal antibody 83-7 and preparation of Fab 83-7 from monoclonal antibody 83-7 were based on protocols described previously[7]. In particular, monoclonal antibody 83-7 was proteolysed with dithiothreitol-activated papain (Sigma-Aldrich) at 37 °C. Proteolysis was stopped by adding iodoacetamide (IAA; Sigma-Aldrich) and the reaction mixture then passed down a Superdex 200 26/60 column (GE Healthcare). Fractions containing 83-7 $F(ab')_2$ were isolated and reduced with mercaptoethylamine (Sigma-Aldrich) and then alkylated with IAA, followed by further SEC and cation-exchange chromatography on Mono S (GE Healthcare).

Production of Fab 83-7-complexed IR310.T

IR310.T was mixed with a slight molar excess of Fab 83-7 and the complex purified by SEC using a Superdex S200 10/300 column (GE Healthcare) in TBSA buffer. Fractions containing the complex of IR310.T/Fab 83-7 were then pooled, concentrated and exchanged into 10 mM HEPES-NaOH (pH 7.5) + 0.02% NaN_3 for crystallization screens.

Cloning, production and purification of IR593.αCT

CHO Lec8 cells stably expressing IR593.αCT were obtained from CSIRO. These cells were derived by co-transfection with both the linearized expression plasmid pEE14 (Lonza) and a pEF puromycin plasmid (Springer Labs). The pEE14 expression plasmid contained, inserted between its HindIII–XbaI sites, a fragment encoding (in order) the 27-amino-acid residue insulin receptor signal sequence, insulin receptor residues 1–593 and insulin receptor residues 704–719 (see Supplementary Fig. 9). The 106-nucleotide-base non-coding fragment described above was also included here between the HindIII site and the start codon. Selection was undertaken in the presence of puromycin, with IR593.αCT production monitored via europium-labelled insulin binding assays[35]. Production scale-up, performed under contract by CSIRO, was in spinner flasks (New Brunswick Scientific) containing DMEM/ F12 medium (Life Technologies) supplemented with 10% FBS plus 10 μg ml^{-1} puromycin. Accumulated supernatant batches were adjusted to pH8 by addition of 3M Tris-HCl (pH 8.5). PMSF was added to 0.1 mM concentration and NaN_3 to 0.02%. The culture supernatant was filtered (0.2 μm cutoff), concentrated 10-fold (10 kDa cutoff) and pumped through a Sephacryl S100 column (GE Healthcare) onto an in-house manufactured *bis*-Boc-insulin MiniLeak medium resin column (Kem-En-Tec)[36]. The latter was eluted with bovine insulin in TBSA. PMSF was again added to 0.1 mM concentration and the eluates concentrated for Superdex S200 26/60 or S200 10/300 (GE Healthcare) SEC. The SEC profiles showed a single major peak corresponding to the insulin-complexed IR593.αCT homodimer and a further peak corresponding to excess bovine insulin (Supplementary Fig. 8g). At higher concentrations, an additional peak appeared corresponding to a dimer of IR593.αCT homodimers (Supplementary Fig. 8h). Contents of fractions were analysed by SDS–PAGE (data not shown). The final yield of IR593.αCT was ~0.4 mg l^{-1} of cell culture supernatant.

Large-scale production and purification of Fab 83-14

A hybridomal cell line expressing Mab 83-14[33] was a gift from K. Siddle. Cells were grown in 950 cm^2 roller bottles at 37 °C in 5% CO_2 (balanced with air) at 10 r.p.m., using approximately 300 ml H-SFM medium supplemented with 5% fetal calf serum (FCS). Typically, 2–3 l of cell culture at a viable cell density of 2–3 × 10^6 cells ml^{-1} was achieved. Cells were then pelleted (5 min, 350g) and re-suspended in serum-free medium. The culture was collected when cell viability dropped below 30% (after 5–7 days). Monoclonal antibody 83-14 from collected supernatant was captured using a ProSep-vA column (Millipore).

Monoclonal antibody 83-14 was proteolysed with dithiothreitol-activated papain (Sigma-Aldrich) at 37 °C. Proteolysis was stopped by adding IAA and the reaction mixture containing Fab 83-14 was then passed down a Superdex 200 26/60 column. Final purification of Fab 83-14 was by cation-exchange chromatography on a MonoS column (GE Healthcare).

Purification of the IR593.αCT +Fab 83-14 + insulin complex

The IR593.αCT protein was combined with an excess of bovine insulin and Fab 83-14 and then re-subjected to SEC. The SEC profiles (Supplementary Fig. 8i) showed a major peak with a molecular mass corresponding to that of a dimer of Fab-complexed IR593. αCT homodimers and a shoulder of molecular weight corresponding to that of an Fab-complexed IR593. αCT homodimer, as well as peaks corresponding to excess Fab 83-14 and insulin. Fractions corresponding to the higher molecular mass component of the major peak were pooled and then concentrated in a 10 mM HEPES-NaOH (pH 7.5)10.02% NaN_3 buffer for crystallization.

Production of IR468-containing media

A DNA fragment encoding the signal peptide plus residues 1–468 of human insulin receptor was inserted into the HindIII–XbaI sites of the vector pcDNA3.1zeo(+) (Life Technologies) and then transiently transfected into HEK293 cells. Cells were grown in Dulbecco's modified Eagle's medium (Life Technologies) supplemented with 10% FCS for 72h after transfection; media containing IR468 were then collected and stored at −20 °C until used.

Crystallization and diffraction data collection of complexes A to D

Complexes A, B, C and D are as defined in Table 1. Complexes A, B and C were prepared for crystallization by combining the Fab 83-7 bound IR310.T (final concentration 3.5 mg ml^{-1}) with a 1.5× molar ratio of αCT peptide (Genscript) and 3× molar ratio of human insulin (Sigma-Aldrich) or [$_D$-Pro^{B26}]-DTI-NH_2 in 10 mM HEPES-NaOH (pH 7.5) +0.02% sodium azide (NaN_3). [D-Pro^{B26}]-DTI-NH_2 was prepared as described previously[37]. A single crystallization condition for complex A was detected using a sparse-matrix 792-condition screen (CSIRO Collaborative Crystallization Centre) and optimized manually to 0.9–1.1M tri-sodium citrate, 0.1M imidazole-HCl (pH 8.0)+ 0.02% NaN_3 for all complexes A, B and C. An initial crystallization condition for complex D at 2.0 mg ml^{-1} was detected in a similar 792-condition screen and optimized manually to 9% PEG3350, 200 mM proline, 100 mM HEPES-NaOH (pH 7.5). For data collection, the complex A and B crystals were cryo-protected by transfer to paraffin oil HR403 (Hampton Research), the complex C crystal by progressive increase of the precipitant concentration to saturated tri-sodium citrate, and complex D crystal by addition of 20% glycerol to the crystallant solution. All crystals were cryo-cooled by plunging into liquid nitrogen and maintained at ~100 K during data collection. Data set 1 for the complex A crystal and the data set for the complex D crystal were collected at beamline MX2 at the Australian Synchrotron[38] (λ = 0.95369 Å). The remaining data sets were collected at beamline I24 at the Diamond Light Source (λ = 0.96860 Å). Diffraction data were processed with XDS[39] and CCP4[40]. Resolution limits were assessed on the basis of the significance of $CC_{1/2}$ at the P = 0.001 level[19]. Data processing statistics are provided in Table 1.

Structure solution and refinement of complex A

PHASER[41] located within the asymmetric unit a single copy each of the L1–CR module and the variable module of 83-7 (using data set 1). The 83-7 constant module was not located and ultimately presumed disordered. TLS parameters and individual restrained isotropic B-factors were then refined using autoBUSTER[42], followed by atomic coordinate-only

refinement, yielding $R^{xpct}_{work}/R^{xpct}_{free}$ = 0.368/0.363 (where R^{xpct} is the crystallographic *R*-factor reported by autoBUSTER, computed using the expectation value of F_{calc} instead of the value itself)[43]. The difference density maps contained four helix-like features encompassing the first, third, fourth, fifth and seventh highest peaks in the map (the second highest peak being glycan conjugated to Asn 111; Supplementary Fig. 10a). These features lay adjacent to the L1–β_2 surface, and PHASER then positioned the three insulin helices as a rigid entity onto three of these (TFZ = 11.4). Correctness of this solution was verified by exhaustive six-dimensional real-space search using ESSENS[44] (Supplementary Fig. 10b). The structure was then refined further, yielding $R^{xpct}_{work}/R^{xpct}_{free}$ = 0.320/0.339. The remaining helix-like feature in the above difference map (the first and fifth highest peaks) revealed clear side-chain protrusions upon negative B-factor map sharpening[21], spaced consistent with an underlying α-helix (Supplementary Fig. 10c). We concluded that it arose from the αCT peptide as there were no further sources of peptide in the crystallization sample. A ten-residue poly-alanine α-helix was then docked into this feature using ESSENS; the best fit was adequately discriminated from lower-scoring fits and visual inspection confirmed the direction of the helix as judged by protruding side-chain density (Supplementary Fig. 10c). A procedure was then devised to assign, in an objective fashion, the sequence register to this 10-residue segment, based on compatibility of individual residue side chains both with sharpened difference electron density and with their surrounding protein environment. (1) Fit to difference density used a method designed not only to assess the fit of atoms within the density, but also to penalize the existence of volumes of positive difference density into which no atoms had been placed. Briefly, voxels corresponding to the αCT feature were excised from the B-factor-sharpened[21] map using CHIMERA[45] (cutoff level = 0.16 $e^{-}Å^{-3}$) and placed within a rectangular grid volume large enough to allow a >8 Å buffer zone around the feature. All grid points outside of the feature were set to 0 $e^{-}Å^{-3}$, with the resultant map being termed M_{obs}. Coordinates for residues 1, 2 and 10 of the fitted poly-alanine helix were then deleted from the model, as the density associated with them displayed poorer α-helical geometry. Ten alternatives, corresponding to all possible alignments of sub-sequences of αCT(704–719) with the heptameric poly-alanine structure, were then assessed. Rotamers for each residue within each of these ten models were then assigned manually using COOT[46] based on visual inspection of the fit of the trial rotamer to the density at the corresponding site inM_{obs}. The ten individual models were then 'real-space-refined' within COOT to achieve optimal fit to M_{obs}, maintaining tight helical restraints. An electron density map M_{calc}, on the same grid as M_{obs}, was then generated (using SFALL within CCP4) for each heptamer model in isolation, with the B-factors of all main-chain atoms being set to 10 $Å^2$ and of all side-chain atoms set to 20 $Å^2$ to allow for subsequent comparison with a sharpened map. All grid points within M_{calc} with ρ < 0.45 e^{-} $Å^{-3}$ were set to 0$e^{-}Å^{-3}$. Correlation coefficients CC = $(<xy>-<x><y>)/\sqrt{[(<x^2>-<x>^2)(<y^2>-<y>^2)]}$ between M_{obs} and each M_{calc} were then calculated using MAPMAN[47]. Each CC was then 'normalized' by dividing its value by the CC value calculated for the underlying poly-alanine heptamer, and the quotient termed the trial sequence's 'density score'. The heptamer with trial register 705–711 was seen to have the highest density score (Supplementary Table 4). (2) Compatibility of residue side-chain environment within the L1–αCT–insulin interface was assessed as follows. For each of the ten trial heptamer (real-space refined) models described above, an 'environment score' was generated using VERIFY3D[48] to assess compatibility with the surrounding L1–β_2 and insulin surfaces. Trial register 705–711 again scored highest (Supplementary Table 4). A 'combined score' was then computed as the product (density score)×(environment score) to assist assessment of lower-ranked trial registers (Supplementary Table 4). The next highest combined score was for trial register 709–715, which is related to trial register 705–711 by a one-turn translation along the helix, effectively maintaining hydrophobic-to-hydrophobic docking with the L1 surface. However, register 709–715 was judged to be most unlikely, as it would bring Pro 716 and Pro 718 into the remaining C-terminal region of the helix. We

concluded that register assignment 705–711 was correct, given that it ranked highest on all criteria. The assignment aligned with that assessed directly from the EDDMs for complex D, wherein the αCT segment is directly attached to the C-terminus of FnIII-1 and the ambiguity of assignment strictly limited. Nine sugar residues were also included at N-linked sites 16, 25, 111, 225 and 255[16,49]. Final refinement was against data set 2, using local structural similarity restraints (LSSRs)[22] against Protein Data Bank (PDB) structures 2HR7-A (L1-CR), 1IL1-B (83-7 light chain) and 1FNS-H (83-7 heavy chain), yielding $R^{xpct}_{work}/R^{xpct}_{free}$ = 0.264/0.292 (see Supplementary Table 1). Ramachandran statistics are: favoured region 91%, acceptable region 5% and disallowed region 4%.

Structure solution and refinement of complexes B and C

These structures were solved independently. MOLREP[50] was used to locate the L1–CR/83-7 fragment and the structure then refined with REFMAC5[23]. [D-Pro^{B26}]-DTI-NH_2 (PDB entry 2WRW) was then located using MOLREP, its position being effectively identical to that in the native insulin complex (complex A). The αCT segment from the wild-type insulin complex was fitted in the final electron density maps using COOT. Final refinement gave R_{work}/R_{free} = 0.289/0.349 and R_{work}/R_{free} = 0.289/0.335, respectively (Supplementary Table 1). Ramachandran statistics for both complexes are: favoured region 89%, acceptable region 7% and disallowed region 4%.

Structure solution and refinement of complex D

Molecular replacement with PHASER located sequentially two copies each of a module consisting of Fab 83-14 in complex with the FnIII-1 domain, an L1–CR module and an L2 domain, using search models derived from PDB entry 3LOH[12]. The assembly was subject to rigid-body and TLS refinement using autoBUSTER, followed by individual restrained isotropic B-factor and coordinate refinement, subject both to LSSRs to PDB entries 2HR7-A (L1-CR-L2), 3LOH-E (FnIII-1), 1PLG-H (83-14 heavy chain) and 2VXU-M (83-14 light chain) and to non-crystallographic symmetry (NCS) restraints, yielding $R^{xpct}_{work}/R^{xpct}_{free}$ = 0.308/0.299. Difference electron density analysis revealed: (1) two helix-like features of electron density, each connected to a respective insulin receptor residue Thr 593 and lying on the L1–β_2 surface of an adjacent monomer, thus identifiable as the αCT segment of IR593.αCT; and (2) two clusters of three helix-like features, each cluster lying adjacent to a respective one of the above αCT segments and putatively identifiable as insulin. Insulin was then docked into each of the above putative insulin–αCT clusters of features by means of an exhaustive six-dimensional ESSENS search. The highest scoring solution in each instance overlaid insulin accurately onto the respective three-helix cluster (Supplementary Fig. 10d) and was well discriminated from lower scoring solutions; the pose of insulin and αCT with respect to L1–β_2 was effectively identical at the two NCS related sites. The two insulins were then included in the model followed by further refinement. COOT was then used to position poly-alanine α-helices into the respective density features corresponding to the αCT segment (sharpened by the application of a negative B-factor), followed by mutation of the sequence to that of αCT residues 704–715. Assignment of the αCT register presented no more than a single residue prima facie ambiguity (depending on the way in which residues at the FnIII-1–αCT peptide junction were built). One alternative was strongly favoured, aligning independently with that determined for complexes A, B and C above. Residues B5–B7 of each insulin could be built into electron density following a path similar to that seen in certain high-affinity insulin analogues[18]. Further crystallographic refinement was then undertaken with autoBUSTER, yielding final $R^{xpct}_{work}/R^{xpct}_{free}$ = 0.244/0.260 (Supplementary Table 1). Ramachandran statistics are favoured region 88%, acceptable region 7% and disallowed region 5%.

Holoreceptor photo-crosslinking

Production and photo-crosslinking of derivatized holoreceptors were performed as described previously[25].

Photo-crosslinking of Pap-derivatized insulins to IR468

Photo-activatable insulin derivatives Pap^{A3}-Ins, Pap^{A8}-DKP-Ins, Pap^{B16}-DKP-Ins, Pap^{B24}-DKP-Ins, [Pap-^{B25}]-DP-Ins were synthesized as described[28,51] and iodinated using the chloramine-T oxidation method. αCT(703–719)-Myc peptide was synthesized in-house. ^{125}I-photo-activatable insulin analogues were incubated overnight at 4 °C with IR468 and 10^{-5}M αCT(703–719)-Myc in 100 mM HEPES (pH7.5), 100 mM NaCl, 10 mM $MgCl_2$, 0.05% bovine serum albumin, 0.025% Triton X-100. One aliquot was exposed to short-wavelength ultraviolet light (Minera-light UVG-54) for 30s at 1 cm and a second aliquot treated with 0.2 mM disuccinimidyl suberate (DSS; Pierce). Samples were diluted with Laemmli sample buffer, resolved on 10% Tricine SDS-PAGE and auto-radiographed.

Isothermal titration calorimetry (ITC)

ITC analysis of insulin binding to IR310.T and its Fab 83-7 complex followed protocols identical to those described previously[52]. The affinity of zinc-free human insulin for IR310.T in the presence of a 10× molar ratio of αCT(704–719) was determined to be K_d =30 nM (Supplementary Fig. 8e) and the affinity of zinc-free human insulin for IR310.T–Fab 83-7 in the presence of a 10× molar ratio of αCT(704–719) to be K_d =48 nM (Supplementary Fig. 8f). ITC analysis of mutant αCT peptides binding to IR485 and of insulin binding to IR485 in complex with mutant αCT peptide were also performed as described previously[52], with samples prepared in TBSA (25 mM Tris-HCl, 137 mM sodium chloride, 2.7 mM potassium chloride, pH 8.0 + 0.02% NaN_3).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This Letter is dedicated to our co-author, the late Guy Dodson, in recognition of his lifetime contribution to the study of the structure of insulin. This work was supported by Australian National Health and Medical Research Council (NHMRC) Project grants 516729, 575539 and 1005896 and the Hazel and Pip Appel Fund (to M.C.L.), NHMRC Independent Research Institutes Infrastructure Support Scheme Grant 361646 and Victorian State Government Operational Infrastructure Support Grant (to the Walter and Eliza Hall Institute of Medical Research), NIH grant no. DK40949(to M. A. W. and J.W.)and American Diabetes Association grantno.1-11INI-31 (to J.W.), Grant Agency of the Czech Republic grant P207/11/P430 (to L.Z.), Research Project of the Academy of Sciences of the Czech Republic RVO:61388963 (to the Institute of Organic Chemistry and Biochemistry), NIH grants DK13914 and DK20595 to D.F.S.),a BBSRC PhD studentship (to C.J.W.)and the UoYResearch Priming Fund(to the York Structural Biology Laboratory). Part of this research was undertaken on the MX2 beamline at the Australian Synchrotron (AS), Victoria, Australia.We thank the DLS for access to beamline I24 and the Australian International Synchrotron Access Program for travel funds. We thank P. Colman and J. Gulbis, our colleagues at CSIRO and the AS beam line staff for their support; J. Turkenburg for assistance in collecting data at DLS; K. Huang for assistance with midi-receptor photo-crosslinking; Q.-X. Hua and Y. Yang for discussion of NMR studies of insulin; S.-Q. Hu, S. H. Nakagawa, N. F. Phillips and S. Wang for assistance with insulin analogue synthesis; P. G. Katsoyannis for advice about the synthesis of photo-reactive insulin analogues and for providing an initial set of Pap analogues; K. Siddle for supplying the 83-7 and 83-14 hybridomas; L. Lu and the fermentation group CSIRO Materials Science and Engineering for large-scale cell culture.

References

1. Taniguchi CM, Emanuelli B, Kahn CR. Critical nodes in signalling pathways: insights into insulin action. Nature Rev. Mol. Cell Biol. 2006; 7:85–96. [PubMed: 16493415]

2. Cohen P. The twentieth century struggle to decipher insulin signalling. Nature Rev. Mol. Cell Biol. 2006; 7:867–873. [PubMed: 17057754]
3. Talbot K, et al. Demonstrated brain insulin resistance in Alzheimer's disease patients is associated with IGF-1 resistance, IRS-1 dysregulation, and cognitive decline. J. Clin. Invest. 2012; 122:1316–1338. [PubMed: 22476197]
4. Pollak M. The insulin and insulin-like growth factor receptor family in neoplasia: an update. Nature Rev. Cancer. 2012; 12:159–169. [PubMed: 22337149]
5. Lemmon MA, Schlessinger J. Cell signaling by receptor tyrosine kinases. Cell. 2010; 141:1117–1134. [PubMed: 20602996]
6. Adams MJ, et al. Structure of rhombohedral 2 zinc insulin crystals. Nature. 1969; 224:491–495.
7. McKern NM, et al. Structure of the insulin receptor ectodomain reveals a folded-over conformation. Nature. 2006; 443:218–221. [PubMed: 16957736]
8. Seino S, Bell GI. Alternative splicing of human insulin receptor messenger RNA. Biochem. Biophys. Res. Commun. 1989; 159:312–316. [PubMed: 2538124]
9. Schäffer LA. model for insulin binding to the insulin receptor. Eur. J. Biochem. 1994; 221:1127–1132. [PubMed: 8181471]
10. De Meyts P. Insulin and its receptor: structure, function and evolution. Bioessays. 2004; 26:1351–1362. [PubMed: 15551269]
11. Ward CW, Lawrence MC. Ligand-induced activation of the insulin receptor: a multi-step process involving structural changes in both the ligand and the receptor. Bioessays. 2009; 31:422–434. [PubMed: 19274663]
12. Smith BJ, et al. Structural resolution of a tandem hormone-binding element in the insulin receptor and its implications for design of peptide agonists. Proc. Natl Acad. Sci. USA. 2010; 107:6771–6776. [PubMed: 20348418]
13. Ward CW, Lawrence MC. Similar but different: ligand-induced activation of the insulin and epidermal growth factor receptor families. Curr. Opin. Struct. Biol. 2012; 22:360–366. [PubMed: 22521506]
14. Kiselyov VV, Versteyhe S, Gauguin L, De Meyts P. Harmonic oscillator model of the insulin and IGF1 receptors' allosteric binding and activation. Mol. Syst. Biol. 2009; 5:243. [PubMed: 19225456]
15. 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]
16. Lou M, et al. The first three domains of the insulin receptor differ structurally from the insulin-like growth factor 1 receptor inthe regions governing ligand specificity. Proc. Natl Acad. Sci. USA. 2006; 103:12429–12434. [PubMed: 16894147]
17. Garrett TP, et al. Crystal structure of the first three domains of the type-1 insulinlike growth factor receptor. Nature. 1998; 394:395–399. [PubMed: 9690478]
18. Jira´cˇek J, et al. Implications for the active form of human insulin based on the structural convergence of highly active hormone analogues. Proc. Natl Acad. Sci. USA. 2010; 107:1966–1970. [PubMed: 20133841]
19. Karplus PA, Diederichs K. Linking crystallographic model and data quality. Science. 2012; 336:1030–1033. [PubMed: 22628654]
20. Evans P. Resolving some old problems in protein crystallography. Science. 2012; 336:986–987. [PubMed: 22628641]
21. Brünger AT, DeLaBarre B, Davies JM, Weis WI. X-ray structure determination at low resolution. Acta Crystallogr. D. 2009; 65:128–133. [PubMed: 19171967]
22. Smart OS, et al. Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. Acta Crystallogr. D. 2012; 68:368–380. [PubMed: 22505257]
23. Murshudov GN, et al. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. D. 2011; 67:355–367. [PubMed: 21460454]
24. Hua QX, Shoelson SE, Kochoyan M, Weiss MA. Receptor binding redefined by a structural switch in a mutant human insulin. Nature. 1991; 354:238–241. [PubMed: 1961250]

25. Whittaker J, et al. α-Helical element at the hormone-binding surface of the insulin receptor functions as a signaling element to activate its tyrosine kinase. Proc. Natl Acad. Sci. USA. 2012; 109:11166–11171. [PubMed: 22736795]
26. Kristensen C, Wiberg FC, Andersen AS. Specificity of insulin and insulin-like growth factor I receptors investigated using chimeric mini-receptors. Role of C-terminal of receptor α subunit. J. Biol. Chem. 1999; 274:37351–37356. [PubMed: 10601304]
27. Xu B, et al. Diabetes-associated mutations ininsulin: consecutive residuesin the B chain contact distinct domains of the insulin receptor. Biochemistry. 2004; 43:8356–8372. [PubMed: 15222748]
28. Xu B, et al. Decoding the cryptic active conformation of a protein by synthetic photo-scanning. Insulin inserts a detachable arm between receptor domains. J. Biol. Chem. 2009; 284:14597–14608. [PubMed: 19321435]
29. Xu B, et al. Diabetes-associated mutations in insulin identify invariant receptor contacts. Diabetes. 2004; 53:1599–1602. [PubMed: 15161767]
30. Weiss MS. Global indicators of X-ray data quality. J. Appl. Cryst. 2001; 34:130–135.
31. Hoogenboom HR, et al. Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. Nucleic Acids Res. 1991; 19:4133–4137. [PubMed: 1908075]
32. Stanley P. Chinese hamster ovary cell mutants with multiple glycosylation defects for production of glycoproteins with minimal carbohydrate heterogeneity. Mol. Cell. Biol. 1989; 9:377–383. [PubMed: 2710109]
33. Soos MA, et al. Monoclonal antibodies reacting with multiple epitopes on the human insulin receptor. Biochem. J. 1986; 235:199–208. [PubMed: 2427071]
34. Evan GI, Lewis GK, Ramsay G, Bishop JM. Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product. Mol. Cell. Biol. 1985; 5:3610–3616. [PubMed: 3915782]
35. Denley A, et al. Structural determinants for high-affinity binding of insulin-like growth factor II to insulin receptor (IR)-A, the exon 11 minus isoform of the IR. Mol. Endocrinol. 2004; 18:2502–2512. [PubMed: 15205474]
36. Markussen J, Halstrom J, Wiberg FC, Scha¨ffer L. Immobilized insulin for high capacity affinity chromatography of insulin receptors. J. Biol. Chem. 1991; 266:18814–18818. [PubMed: 1918001]
37. Žáková L, et al. Insulin analogues with modificationsat position B26. Divergence of binding affinity and biological activity. Biochemistry. 2008; 47:5858–5868. [PubMed: 18452310]
38. McPhillips TM, et al. Blu-ice and the distributed control system: software for data acquisition and instrument control at macromolecular crystallography beamlines. J. Synchrotron Radiat. 2002; 9:401–406. [PubMed: 12409628]
39. Kabsch W. Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr. D. 2010; 66:133–144. [PubMed: 20124693]
40. Collaborative Computing Project, No The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D. 1994; 50:760–763. [PubMed: 15299374]
41. McCoy AJ. Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallogr. D. 2007; 63:32–41. [PubMed: 17164524]
42. Bricogne G, et al. BUSTER version 2.10 (Global Phasing Ltd, 2011).
43. Blanc E, et al. Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. Acta Crystallogr. D. 2004; 60:2210–2221. [PubMed: 15572774]
44. Kleywegt GJ, Jones TA. Template convolution to enhance or detect structural features in macromolecular electron-density maps. Acta Crystallogr. D. 1997; 53:179–185. [PubMed: 15299952]
45. Pettersen EF, et al. UCSF Chimera—a visualization system for exploratory research and analysis. J. Comp. Chem. 2004; 25:1605–1612. [PubMed: 15264254]
46. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr. D. 2004; 60:2126–2132. [PubMed: 15572765]
47. Jones TA, Thirup S. Using known substructures in protein model building and crystallography. EMBO J. 1986; 5:819–822. [PubMed: 3709525]

48. Lu¨thy R, Bowie JU, Eisenberg D. Assessment of protein models with three-dimensional profiles. Nature. 1992; 356:83–85. [PubMed: 1538787]

49. Sparrow LG, et al. N-linked glycans of the human insulin receptor and their distribution over the crystal structure. Proteins Struct. Funct. Bioinform. 2008; 71:426–439.

50. Vagin A, Teplyakov A. Molecular replacement with MOLREP. Acta Crystallogr. D. 2010; 66:22–25. [PubMed: 20057045]

51. Hua QX, et al. Enhancing the activity of a protein by stereospecific unfolding. The conformational life cycle of insulin and its evolutionary origins. J. Biol. Chem. 2009; 248:14586–14596. [PubMed: 19321436]

52. Menting JG, Ward CW, Margetts MB, Lawrence MC. A thermodynamic study of ligand binding to the first three domains of the human insulin receptor: relationship between the receptor α-chain C-terminal peptide and the site1insulin mimetic peptides. Biochemistry. 2009; 48:5492–5500. [PubMed: 19459609]

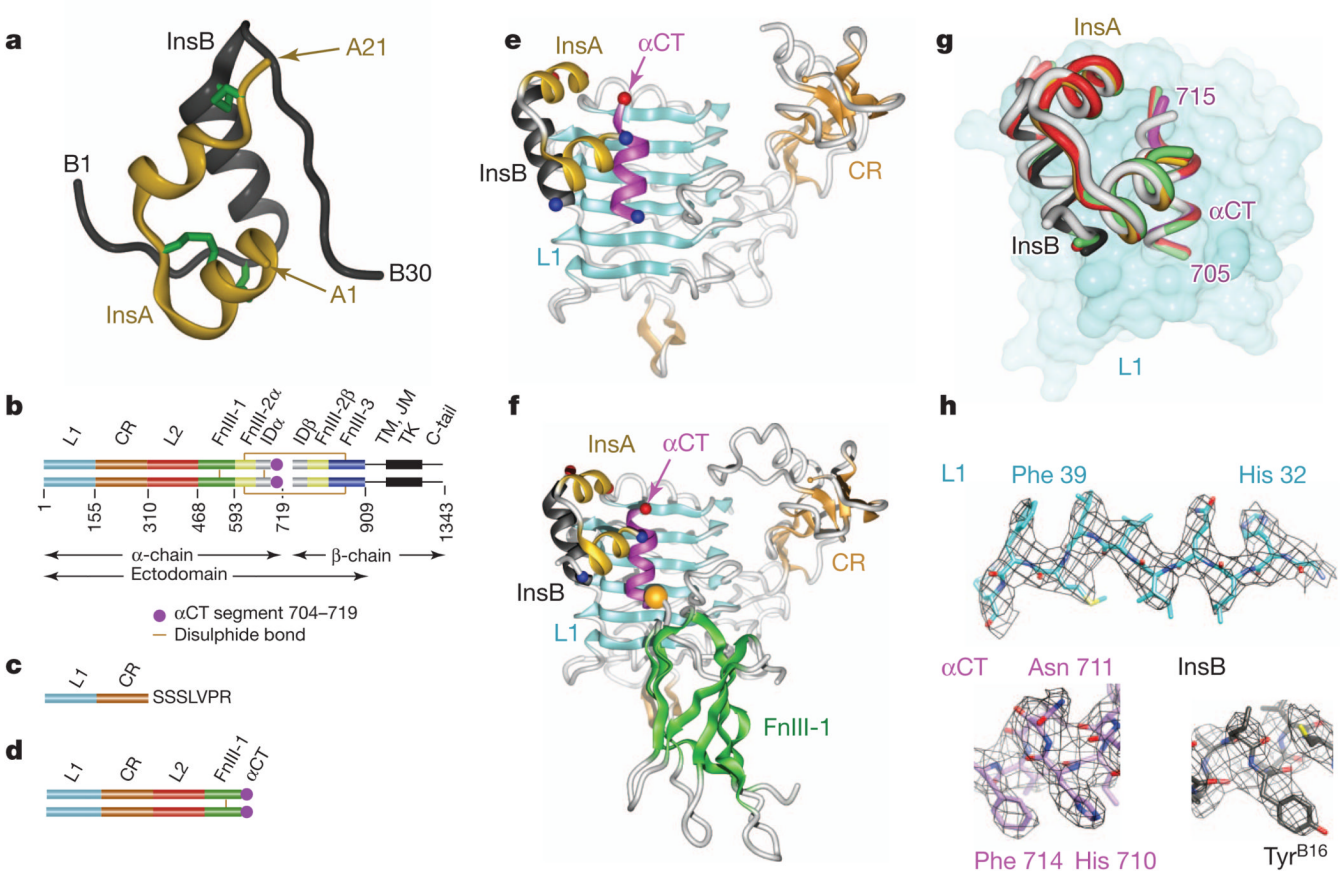

Figure 1. Structure of insulin, insulin receptor and the site 1 complexes

a, Insulin. InsA, A chain; InsB, B chain. **b**, Insulin receptor. CR, Cys-rich domain; FnIII-1, FnIII-2, FnIII-3, first, second and third fibronectin type III domains; ID, insert domain; L1, L2, first and second leucine-rich-repeat domains; TK, tyrosine kinase; TM, JM, transmembrane and juxtamembrane segments. **c, d**, IR310.T and IR593.αCT domain structure, respectively. **e, f**, Insulin-bound site 1 in complexes A and D, respectively. Blue, red spheres, observed chain termini; orange sphere, FnIII-1–αCT junction. **g**, Overlay of insulin-bound site 1 in complexes A (coloured as in **e**), B (red), C (green) and D (white). **h**, Sample $2F_{obs} - F_{calc}$ map volumes (B_{sharp} = −160 Å^2; contours = 1.1–1.5σ) for complex A.

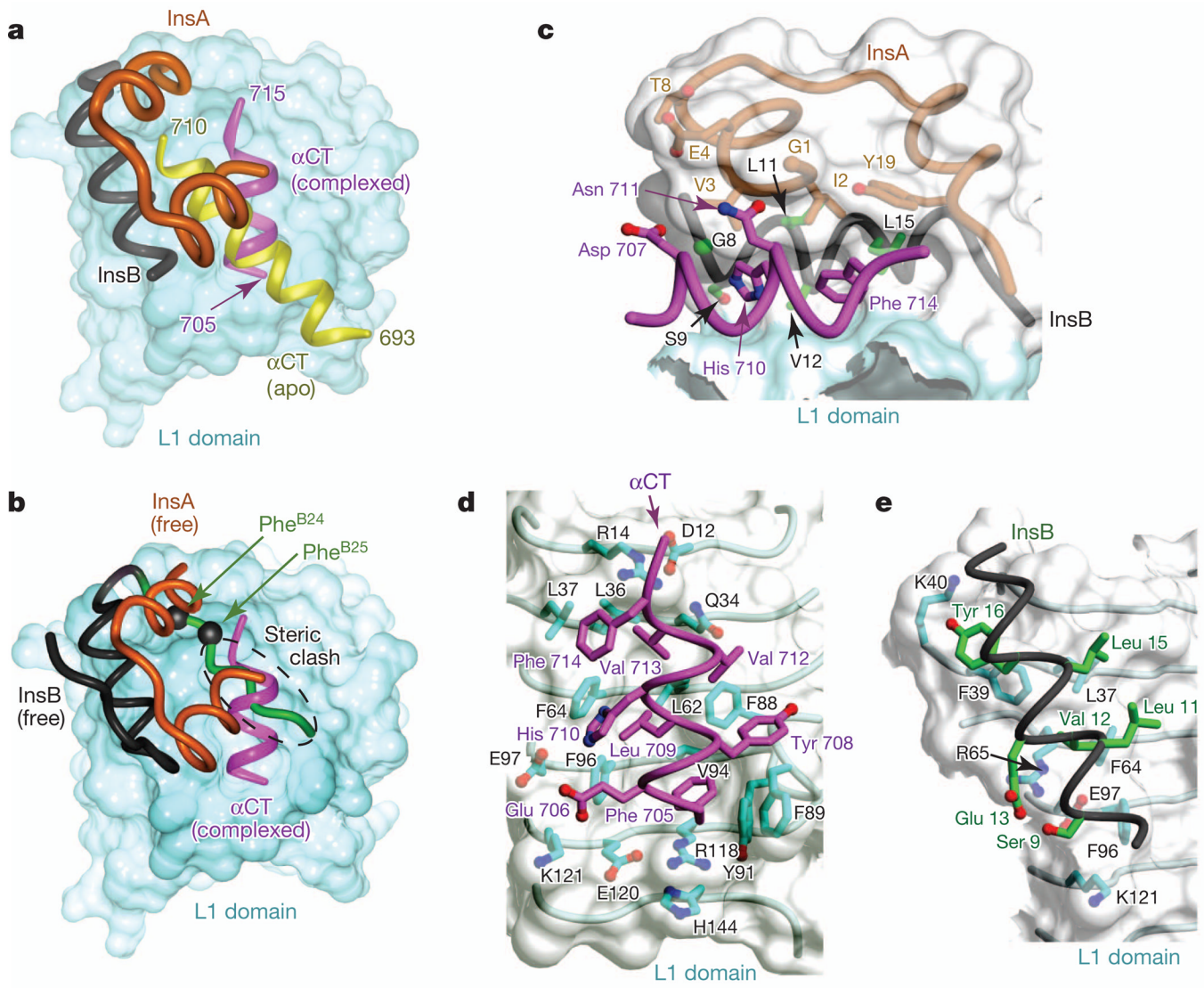

Figure 2. The insulin–site 1 interaction

a, Altered disposition of αCT with respect to that in *apo*-insulin receptor[12]. **b**, Superposition (via the A- and B-chain helices) of receptor-free insulin onto the insulin–site 1 complex, indicating steric clash of the B-chain C-terminal segment (green) with αCT. See also Supplementary Fig. 3a. **c**, Interaction (at the bulk side-chain level of detail) between αCT(704–719) and insulin. View direction is parallel to L1–β2 surface (cyan); white shading shows insulin surface. **d, e**, Interaction (at the bulk side-chain level of detail) between L1 and αCT and between L1 and insulin B-chain helix, respectively. Panels based on complex A.

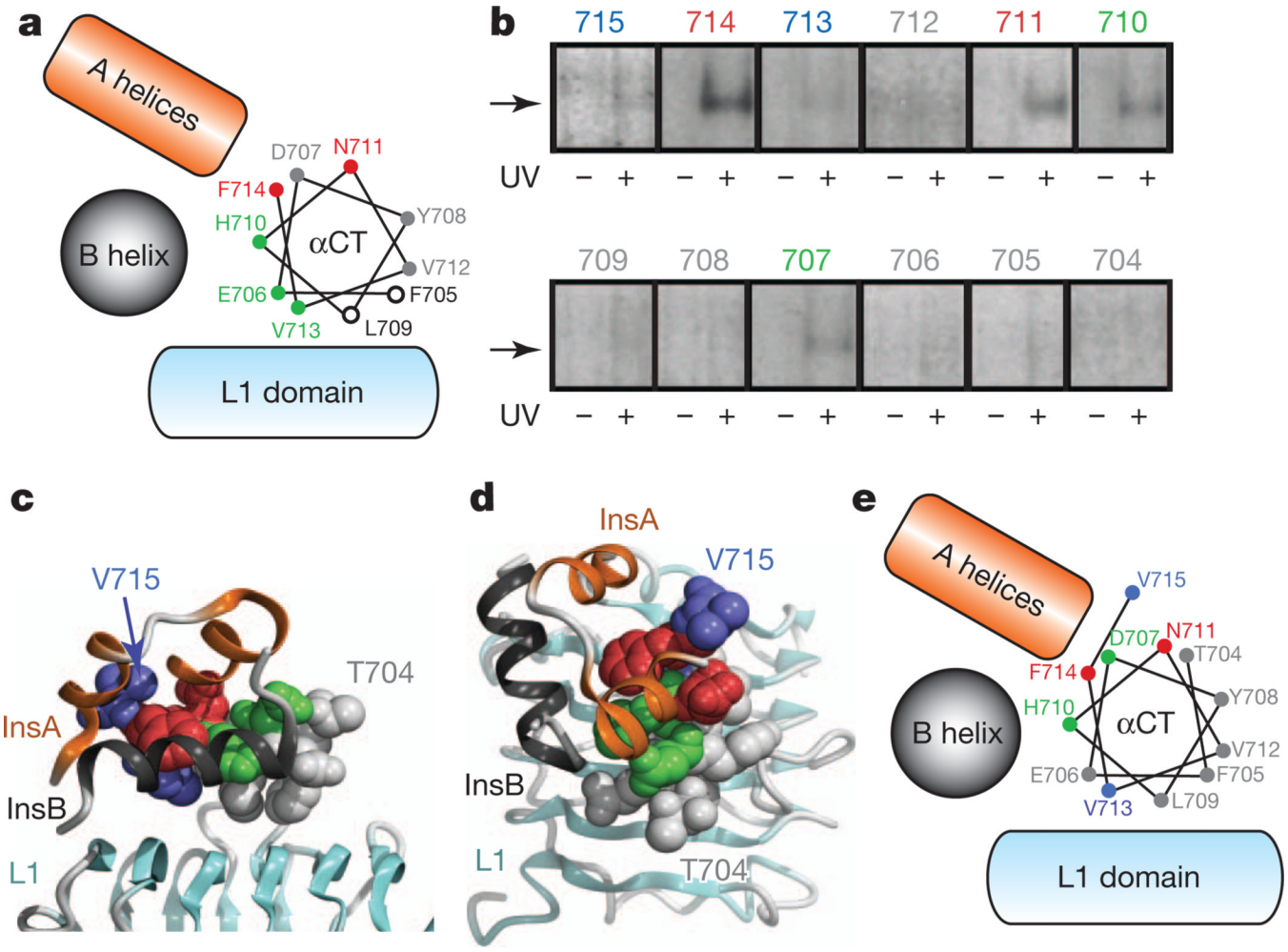

Figure 3. Insulin interactions in L1–CR–L2 mini-insulin receptor and *holo* insulin receptor

a, Helical wheel representation of ITC-derived insulin affinities for IR485 in the presence of Ala-substituted αCT(704–719) (red, > 100× reduction upon Ala substitution; green, > 10× reduction; grey, < 10× reduction; open circle, not determined). **b**, Reducing gel autoradiograms obtained from *holo*-insulin receptor after photo-crosslinking of αCT helix to bound ^{125}I-[Tyr^{A14}]-insulin. Arrowed band indicates crosslinked insulin receptor α-chain–insulin A chain. Colours indicate crosslinking efficiency (red, strong; green, medium; blue, weak; grey, none). **c, d**, Qualitative crosslinking efficiency from **b** mapped onto αCT segment within the site 1 complexes. **e**, Helical wheel representation of crosslinking data presented in **b–d**

Table 1

X-ray diffraction data collection statistics

	Complex A, data set 1	Complex A, data set 2	Complex B	Complex C	Complex D
	Human insulin + IR310.T–83-7+ αCT(704–719)	**Human insulin + IR310.T–83-7+αCT(704–719)**	**[D-Pro^{B26}]-DTI-NH$_2$ + IR310.T–83-7+αCT(704–719)**	**[D-Pro^{B26}]-DTI-NH$_2$ + IR310.T–83-7 +αCT(697–719)**	**Bovine insulin + IR593.αCT/83-14**
Space group	*P*23	*P*23	*P*23	*P*23	*I*2
Cell dimensions					
a, b, c (Å)	168.15, 168.15, 168.15	168.91, 168.91, 168.91	169.49, 169.49, 169.49	169.23, 169.23, 169.23	118.15, 140.10, 190.02
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 95.04, 90
Completeness (%)	99.8 (99.8) *	98.3 (99.5)	99.7 (100.0)	98.5 (98.1)	87.7 (84.9) †
Resolution (Å)	59.5–4.0 (4.1–4.0)	46.8–3.9 (4.0–3.9)	29.6–4.3 (4.53–4.3)	29.5–4.30 (4.54–4.3)	56.3–4.4 (4.5–4.4)
R_{merge}	0.156 (2.36)	0.089 (1.57)	0.188 (2.69)	0.101 (1.09)	0.206 (1.31)
‡R_{pim}	0.039 (0.57)	0.048 (0.83)	0.067 (0.97)	0.053 (0.59)	0.143 (0.89)
$I/\sigma I$	11.5 (1.6)	9.0 (1.0)	7.8 (1.0)	8.2 (1.4)	2.8 (0.8)
Redundancy	17.3 (17.4)	4.1 (4.2)	8.8 (8.8)	4.4 (4.3)	2.9 (2.8)
§$CC_{1/2}$	0.998 (0.580)	0.998 (0.342)	0.996 (0.330)	0.997 (0.438)	0.980 (0.232)

Each data set was collected from a single crystal.

*Highest-resolution shell statistics are shown in parentheses.

†Data completeness was compromised by radiation damage.

‡R_{pim}, precision indicating merging *R* factor[30].

§$CC_{1/2}$, Pearson correlation coefficient between independently merged halves of the data set[19]. Highest-resolution shell $CC_{1/2}$ values are significant at at least $P = 0.001$.