Structural resolution of a tandem hormone-binding element in the insulin receptor and its implications for design of peptide agonists

Brian J. Smith[a,1], Kun Huang[b,1], Geoffrey Kong[a], Shu Jin Chan[c], Satoe Nakagawa[c], John G. Menting[a], Shi-Quan Hu[b,d], Jonathan Whittaker[b], Donald F. Steiner[c], Panayotis G. Katsoyannis[d], Colin W. Ward[a], Michael A. Weiss[b,e,2], and Michael C. Lawrence[a,2]

[a]The Walter and Eliza Hall Institute of Medical Research, 1G Royal Parade, Parkville, Victoria 3052, Australia; [b]Departments of Biochemistry and [e]Medicine, Case Western Reserve University, Cleveland, OH 44106; [c]Department of Medicine, University of Chicago, Chicago, IL 60637; and [d]Department of Pharmacology and Biological Chemistry, Mt. Sinai School of Medicine of New York University, New York, NY 10029

Contributed by Donald F. Steiner, February 18, 2010 (sent for review December 24, 2009)

The C-terminal segment of the human insulin receptor α-chain (designated αCT) is critical to insulin binding as has been previously demonstrated by alanine scanning mutagenesis and photo-cross-linking. To date no information regarding the structure of this segment within the receptor has been available. We employ here the technique of thermal-factor sharpening to enhance the interpretability of the electron-density maps associated with the earlier crystal structure of the human insulin receptor ectodomain. The αCT segment is now resolved as being engaged with the central β-sheet of the first leucine-rich repeat (L1) domain of the receptor. The segment is α-helical in conformation and extends 11 residues N-terminal of the classical αCT segment boundary originally defined by peptide mapping. This tandem structural element (αCT-L1) thus defines the intact primary insulin-binding surface of the *apo*-receptor. The structure, together with isothermal titration calorimetry data of mutant αCT peptides binding to an insulin minireceptor, leads to the conclusion that putative "insulin-mimetic" peptides in the literature act at least in part as mimics of the αCT segment as well as of insulin. Photo-cross-linking by novel bifunctional insulin derivatives demonstrates that the interaction of insulin with the αCT segment and the L1 domain occurs in *trans*, i.e., these components of the primary binding site are contributed by alternate α-chains within the insulin receptor homodimer. The tandem structural element defines a new target for the design of insulin agonists for the treatment of diabetes mellitus.

diabetes | IGF-1 receptor | insulin binding | photo-cross-linking | protein crystallography

Binding of insulin to the insulin receptor initiates a signaling cascade in target tissues as the first step in the regulation of metabolic homeostasis. However, a molecular description of how insulin binds and activates its receptor remains elusive. Whereas determination of the structure of insulin (Fig. 1*A*) represented an early triumph of protein crystallography (2), the structure of the much larger receptor ectodomain homodimer (in *apo* form) has only recently been crystallographically analyzed (3). The latter structure and its implications for the nature of the hormone-binding sites have been extensively reviewed (4–7). Briefly, the insulin receptor is a disulfide-linked dimer wherein each proreceptor monomer is proteolytically cleaved into an N-terminal α-chain and a C-terminal β-chain (Fig. 1*B*). A single disulfide bond links the α- and β-chains within each monomer. The extracellular portion of the insulin receptor includes both α-chains as well as 194 residues (Ser724-Lys917) of each β-chain. Each receptor monomer consists of several structural domains (Fig. 1*B*), including a leucine-rich repeat domain L1 (residues 1–157), a cysteine-rich region (CR, residues 158–310), a second leucine-rich repeat domain L2 (residues 311–470), and three fibronectin type-III domains: FnIII-1 (residues 471–595), FnIII-2 (residues 596–808), and FnIII-3 (residues 809–906). FnIII-2 contains a ∼120-residue insert domain (ID, residues 638–756) that contains the α/β cleavage site at residues 720–723. The segments of the ID that lie within the α- and β-chains are termed IDα and IDβ, respectively. C-terminal to the FnIII-3 domain lies a single transmembrane helix, followed by a ∼40-residue intracellular juxtamembrane region (JM), a tyrosine kinase (TK) catalytic domain, and a ∼100-residue C-terminal tail. The closely related Type 1 insulin-like growth factor receptor (IGF-1R) shares a similar domain organization. Recent x-ray crystal studies (3) have shown that the extracellular domains of each monomer assemble as an inverted "V," with the L1-CR-L2 domains forming one leg and the three FnIII domains the other. In the receptor dimer, the second monomer is related to the first by a twofold rotation about the axis of the inverted "V," resulting in the L1-CR-L2 leg of one monomer being packed against the three FnIII domains of the other (Fig. 1*C*).

Current models of the insulin/insulin receptor complex propose that hormone binding is mediated by two adjoining structural elements termed Site 1 and Site 2 (4–7). Extensive studies of mutant receptors and chemical cross-linking have established that Site 1 is comprised of elements from two distinct regions: (i) the central of the three β-sheets that make up the L1 domain, and (ii) the last 16 residues of the α-chain (the so-called αCT segment, labeled by an asterisk in Fig. 1*B*), a critical component in ligand binding (8, 9). A possible location for Site 2 was identified when the structure of the insulin receptor ectodomain dimer was determined and is likely formed by loops at the junction of the FnIII-1 and FnIII-2 domains from the opposite monomer to that which contributes the L1 domain (Fig. 1*C*). Such models of hormone binding are consistent with the "cross-linking," *trans* binding mechanism of receptor activation (7, 10, 11) and signal transduction.

A major omission from these models, however, is the structure and location of the αCT segment. This segment was not visualized in the initial structure of the insulin receptor ectodomain (3). However, we have previously speculated that the αCT segment might be associated with a tube-like segment of electron density observed lying on the central β-sheet of the L1 domain within the

Author contributions: D.F.S., P.K., C.W.W., M.A.W., and M.C.L. designed research; B.J.S., K.H., G.K., S.J.C., S.N., J.G.M., S.-Q.H., J.W., and M.C.L. performed research; S.-Q.H. and P.K. contributed new reagents/analytic tools; B.J.S., K.H., G.K., S.J.C., J.G.M., J.W., M.A.W., and M.C.L. analyzed data; and B.J.S., D.F.S., C.W.W., M.A.W., and M.C.L. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

[1]B.J.S. and K.H. contributed equally to this work.

[2]To whom correspondence may be addressed. E-mail: lawrence@wehi.edu.au or michael.weiss@case.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/1001813107/DCSupplemental.

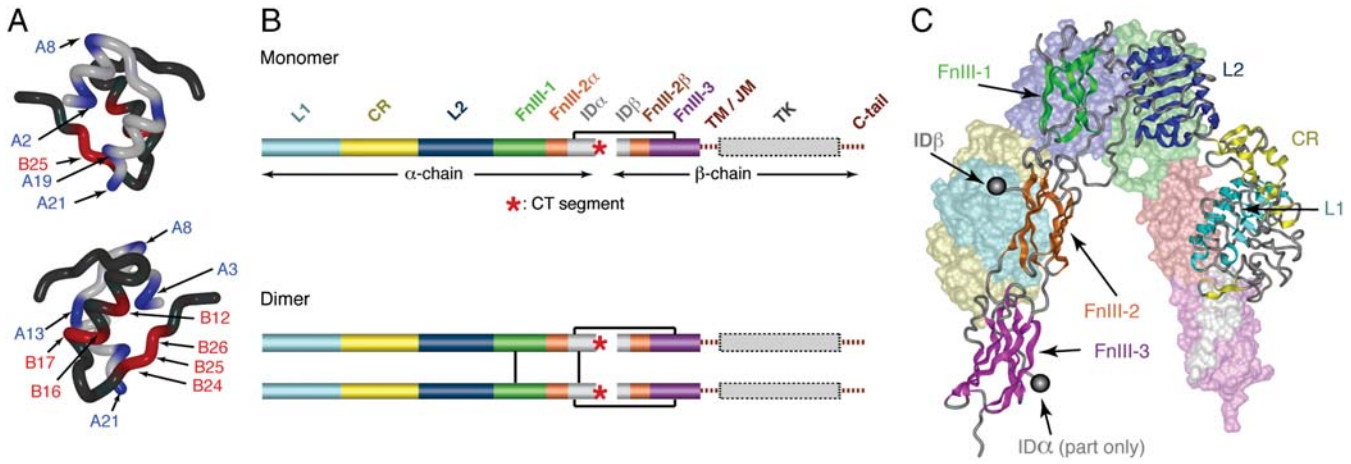

Fig. 1. Structure of insulin and insulin receptor. (*A*) Alternate views of insulin monomer. A- and B-chains are shown in *Light Gray* and *Dark Gray*, respectively, with receptor-binding residues (1) in the A- and B-chains highlighted in *Blue* and *Red*, respectively. (*B*) Domain organization within the insulin receptor monomer and dimer. L1, L2: first and second leucine-rich-repeat domains, CR: cysteine-rich region, FnIII-1, -2, -3: first, second and third Type III fibronectin domains, IDα and IDβ: α-chain and β-chain components of the insert domain ID, TM/JM: trans- and juxta-membrane regions, TK: tyrosine kinase domain. Connecting *Black Line* segments denote disulphide bonds. Intracellular components are shown in *Dashed* representation, ectodomain components are shown in *Solid* representation. A *Red Asterisk* marks the location of the α-chain C-terminal segment αCT within each monomer. (*C*) Crystal structure of the insulin receptor ectodomain. One receptor monomer is in secondary structure representation, the alternate receptor monomer in surface representation; domain color scheme is as in *B*. *Gray Spheres* within the foreground insulin receptor monomer depict the observed C terminus (residue 655) of the α-chain and the observed N terminus (residue 755) of the β-chain; intervening IDα residues 656–719 and IDβ residues 724–754 are not resolved.

crystallographic maps (3). Recent availability of new crystallographic techniques and software has now enabled us to obtain improved electron-density maps of this region of the insulin receptor and construct a crystallographic model of the αCT segment and its interactions with the adjacent L1 domain. We have also recently proposed (12) that there is a close structural relationship between the αCT segment and the so-called Site 1 insulin-mimetic peptides developed by Schäffer and coworkers (13). Here our refined structure, together with extensions of earlier isothermal titration calorimetry (ITC) studies, provides additional support that this is indeed the case. Whether the αCT segment binds to the L1 domain in *cis* or *trans* is then addressed by the use of bifunctional photo-cross-linking insulin derivatives (14).

Results

Crystallographic model. The original 3.8 Å resolution crystal structure of the insulin receptor ectodomain homodimer (Protein Data Bank accession code 2DTG) was derived from x-ray diffraction data collected from a crystal containing a complex of the insulin receptor ectodomain homodimer with two pairs of F_{ab} fragments 83-7 and 83-14 (3). We have now reintegrated the reflection intensities from the diffraction images associated with this structure (*SI Appendix*, Table S1) and rerefined the 2DTG crystallographic model against the reintegrated data, in both cases using alternative software packages to those employed originally. We then calculated $(2mF_o\text{-}DF_c)$ and $(mF_o\text{-}DF_c)$ σ_A-weighted difference electron-density maps (15), which were then artificially sharpened by multiplying the constituent structure-factor amplitudes by $\exp(-B_{sharp}\sin^2\theta/\lambda^2)$, where B_{sharp} is the negative of the thermal factor derived from a Wilson plot of the data, θ the reflecting angle, and λ the x-ray wavelength (16). This technique has recently proved particularly effective in the rerefinement of the three crystal structures of the full-length ATPase p97 (17). The tube-like segment of density originally observed on the surface of the L1 domain (3) was resolvable in these maps as being an α-helix (Fig. 2*A*) and could be assigned unambiguously to insulin receptor residues 693–710 (*SI Appendix*). A model for this segment was then built and incorporated into the F_{ab}-complexed insulin receptor atomic model and the structure further refined. The final crystallographic refinement statistics are excellent given the resolution of the data, *viz.* $R_{work} = 0.225/R_{free} = 0.285$ (*SI Appendix*, Table S1). The low R-factors are likely attributed (16) to the fact that the higher-resolution fragment structures of the L1-CR-L2 domains (2.3 Å resolution) and of prototypical F_{ab}s underpin approximately ∼75% of the original model (3). A more detailed discussion of the data processing and map sharpening strategies applied here is presented in *SI Appendix*.

The interaction of the 693–710 segment with the surface of the central β-sheet of the L1 domain has the following salient features (Fig. 2*B*): (i) The side chains of residues Phe701 and Phe705 are packed adjacent to each other in a hydrophobic pocket formed by the side chains of the residues Leu62, Phe64, Phe88, Phe89, Tyr91, Val94, Phe96, and Arg118 of the L1 domain; (ii) the side chain of Tyr708 is packed approximately parallel to the surface of the central β-sheet of the L1 domain, in proximity to the side chains of residues Arg14, Gln34, Leu36, and Phe88 of the L1 domain; and (iii) the side chains of the residue pair Glu698/Arg702 lie in close proximity to each other and interact with the side chains of the L1 domain residue-pair Arg118/Glu120, respectively, the four side chains forming a charge-compensating cluster. There is also a hydrophobic interaction between the side chain of Leu709 and the side chains of Leu37 and Phe64. On the surface of the helix distal to the interface with the L1 domain, the side chains of the residue-pair Lys703/Asp707 are in proximity to each other and likely charge compensate. The validity of the model is indirectly supported by observations that (i) the three aromatic residues (Phe701, Phe705, and Tyr708) are conserved in type (as Tyr688, Phe692, and Phe695, respectively) in IGF-1R, as might be expected from the fact that the αCT peptide of IGF-1R can substitute for that of insulin receptor in conferring hormone-binding capability on domain-minimized receptors (18); (ii) the subsegment of the helix that is in closest proximity to the L1 domain surface (*viz.* residues Glu698 to Leu709) corresponds almost precisely to a segment of the α-chain C terminus previously predicted to have helical structure (12); and (iii) the shape complementarity of the modeled αCT segment/L1 domain interface is high (0.72), commensurate with a well-packed interaction (19). The total buried molecular surface area is 938 $Å^2$; details of the degree of solvent accessibility of individual residues within the interface are presented in *SI Appendix*, Table S2.

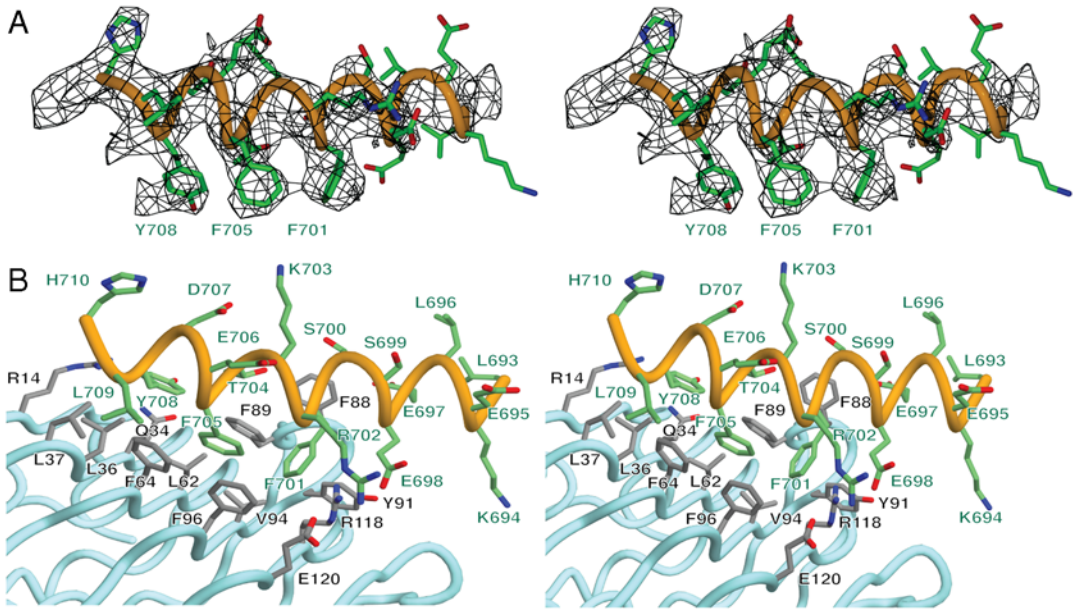

Fig. 2. The L1-domain/αCT segment tandem element. (*A*) Stereo view of the thermal factor sharpened (mF_o-DF_c) electron density overlaid with final atomic model of insulin receptor residues 693–710 (map contoured at 1.8σ). Residues within the aromatic motif used for sequence register assignment (*SI Appendix*) are labeled. (*B*) Stereo view of the interaction between insulin receptor residues 693–710 (*Yellow* backbone, *Green* carbons, and residue labels) and the central β-sheet of L1 (*Cyan* backbone, *Gray* carbons, and *Black* residue labels). The view of the αCT segment in (*A*) is rotated ~45° around the horizontal axis with respect to that in (*B*) in order to show clearly the density associated with the aromatic motif.

Binding of Mutant αCT Peptides to an Insulin Minireceptor. Insulin-mimetic peptides have been discovered by phage display technology and classified as Site 1, 2, or 3 on the basis of competition of binding to insulin receptor (13). The affinity-matured Site 1 peptides are characterized by a FYXWF motif (13); selected Site 1 and 2 peptides have been covalently tethered to yield agonists with up to picomolar affinity for insulin receptor (20). We have proposed (12) a sequence relationship between the αCT segment and the prototypic Site 1 mimetic peptide that places the αCT segment residues Phe701 and Phe705 in respective alignment with the two flanking phenylalanine residues in the FYXWF motif (*SI Appendix*, Fig. S1). This relationship was used to explain the competitive binding of the αCT peptide and the prototypic Site 1 mimetic peptide to the insulin minireceptor IR485, a construct that consists of the receptor L1-CR-L2 domains only (12). If this relationship is correct, then mutation of Arg702 and/or Thr704 within the αCT segment to either tyrosine or tryptophan should lead to significantly higher affinity of the segment for the insulin minireceptor. It should also then be possible to model these substitutions directly onto the structure reported here.

These hypotheses have now been tested. Table 1 presents ITC-derived dissociation constants for a set of mutant αCT peptides titrated against the insulin minireceptor IR485. Progressive inclusion of aromatic residues at positions 702 and 704 is seen to result in an up to 100-fold increase in affinity, supporting the view that there is a structural relationship between the Site-1 mimetic peptides and the native αCT segment. Modeling of the interaction of the mutant αCT peptides with the L1 domain (*Materials and Methods*) reveals how the aromatic side chains of αCT residues Trp702 and Tyr704 are likely to interact with the surface of the L1 domain and enhance the affinity of the interaction. At position 702 the aromatic side chain is docked into a pocket formed by the Phe96 and the alkyl portion of the side chain of Lys121; the hydroxyl group of a variant tyrosine residue can in addition form a hydrogen bond with the Lys121 ε-amino group. At position 704 the aromatic side chain of the variants is docked against the side chains of L1 domain residues Phe88 and Phe89.

Table 1. Dissociation constants of mutant αCT peptides bound to insulin minireceptor IR485

αCT peptide*	K_d peptide (nM)
Wild-type	1340 ± 230†
T704Y	740 ± 210
R702W	309 ± 36
R702Y	249 ± 3
T704W	201 ± 16
R702Y/T704W	18 ± 3

*αCT peptides span residues 698–719 and have a biotin moiety attached at the N terminus.
†Error estimates are the standard error of the mean.

Photo-Cross-Linking of Insulin to the αCT Segment and the L1 Domain. It is not apparent in the above structure whether the interaction between the C-terminal segment of the receptor α-chain and the L1 domain is *cis* or *trans* within the receptor homodimer, given that linking residues 656–692 are unresolved and that distance constraints do not rule out either option. It has been argued (3, 5) that a *trans* association would place the α-chain C terminus in close proximity to the N terminus of the β-chain of the opposite monomer to that providing the L1 domain. Given that during biosynthesis proteolytic cleavage occurs after assembly of the homodimer (21), such *trans* association would appear to be more natural, because a *cis* association would require relocation of the αCT segment to the opposite L1 domain surface after cleavage (5). A *trans* arrangement is also supported by complementation analysis using coexpression of pairs of differentially tagged insulin midi-receptors carrying single mutations in either the L1 domain or the αCT segment (22).

We demonstrate directly here the *trans* association of these elements within the hormone-binding site by cross-linking experiments with photo-labeled insulin derivatives containing *para*-azido-phenylalanine (Pap) substitutions. We exploited the fact that Pap^{B25} and Pap^{A3} in respective monofunctional insulin derivatives contact αCT (23, 24) , whereas Pap^{B16} contacts the L1 domain in a monofunctional derivative (25). Accordingly, a bifunctional [Pap^{B16}, Pap^{B25}]-insulin derivative was prepared (*SI Appendix*, Fig. S2), with its B-chain N terminus labeled with

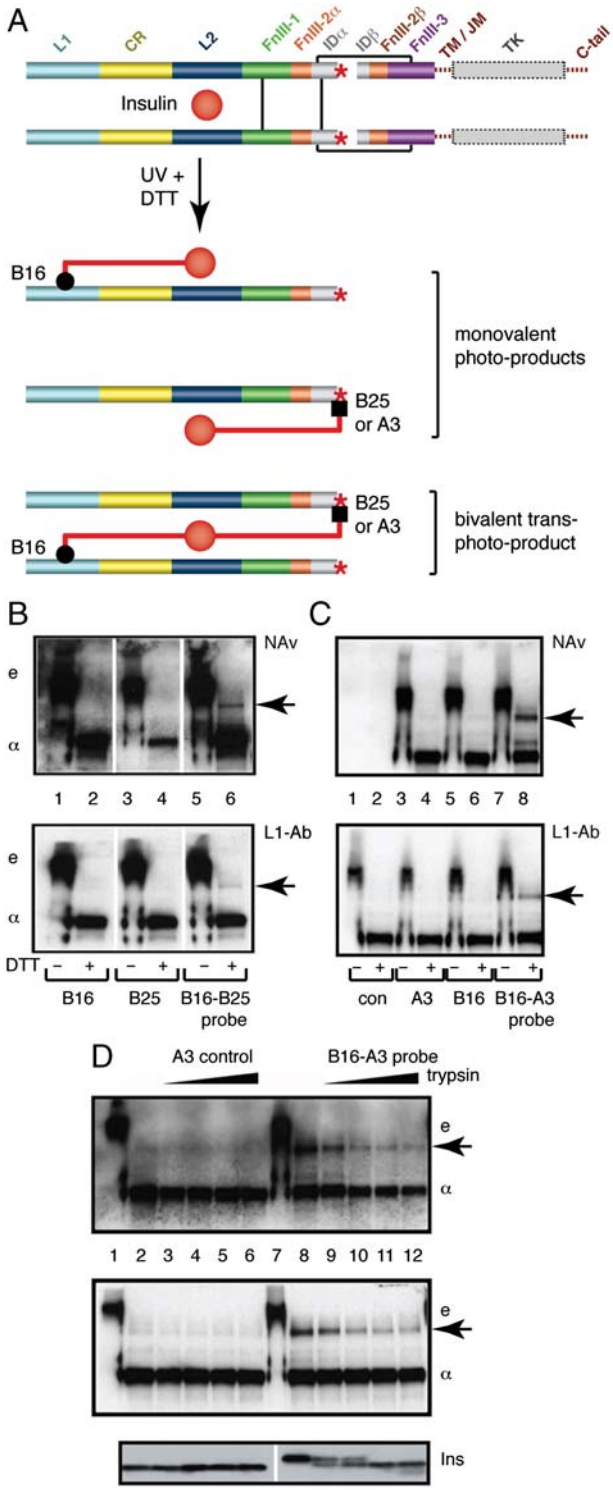

Fig. 3. Photo-cross-linking studies. (*A*) Schematic model of holoreceptor complex showing the production of photo-products by mono- and bifunctional insulins. (*B*) SDS–PAGE analysis of photo-cross-linking by Pap^{B16}, Pap^{B25}, and [Pap^{B16}, Pap^{B25}]-insulin derivatives with or without (+/−) DTT reduction. Arrow at right indicates covalent α-B-α complex. Gel positions of ectodomain and isolated α-subunit are indicated by "e" and "α," respectively. Detection of photo-products is by NeutraAvidin (NAv, *Top*) and by an L1-specific antiserum (L1-Ab, *Lower*). (*C*) Corresponding studies utilizing a [Pap^{B16}, Pap^{A3}]-single-chain insulin (SCI) derivative. Again, arrow at right indicates covalent α-B-α complex. (*D*) Control studies for (*C*) in which interchain linker in SCI is progressively cleaved by trypsin. Coelectrophoresed molecular weight markers (not shown) confirm that the positions of the "e," "α," and arrowed bands correspond in molecular weight across (*B*), (*C*) and (*D*).

biotin to permit detection. The bifunctional derivative should be capable of cross-linking the αCT segment and the receptor L1 domain (Fig. 3*A*). Photo-products of singly- and doubly-derivatized Pap analogs were analyzed by sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS–PAGE) with or without reduction by dithiothreitol (DTT). In the absence of DTT, detection by NeutrAvidin (NAv; Fig. 3*B Upper*) or by an L1-specific antiserum (L1-Ab, Fig. 3*B Lower*) revealed a covalent hormone-ectodomain complex and, upon reduction, an isolated insulin B-chain/insulin receptor α-chain adduct. The bifunctional insulin derivative uniquely gave rise to an additional photo-product of molecular mass consistent with the sum of two insulin receptor α-chains and one insulin B-chain, implying that insulin cross-linked L1 and αCT from separate α-chains, as proposed in Fig. 3*A*. Similar results were obtained (Fig. 3*C*) using a single-chain insulin analog (SCI) with Pap^{A3} and Pap^{B16} substitutions and an N-terminal biotin tag (*SI Appendix*, Fig. S3); an SCI was employed to avoid separation of the A3 and B16 photo-products following reduction of interchain disulfide bridges. Because the peptide connecting the A- and B-chains in the SCI contained a tryptic cleavage site, a control was provided by pretreatment of the photo-products with trypsin prior to SDS–PAGE. This pretreatment led to disproportionate attenuation of the covalent α-SCI-α band relative to a monovalent Pap^{A3}-related complex (Fig. 3*D*).

Discussion

Our results define the structure of a tandem hormone-recognition element in the insulin receptor dimer. The tandem element comprises the amphipathic α-helix (αCT) packed against the predominantly nonpolar surface of the central β-sheet of the L1 domain. The *trans* character of the tandem element is built into

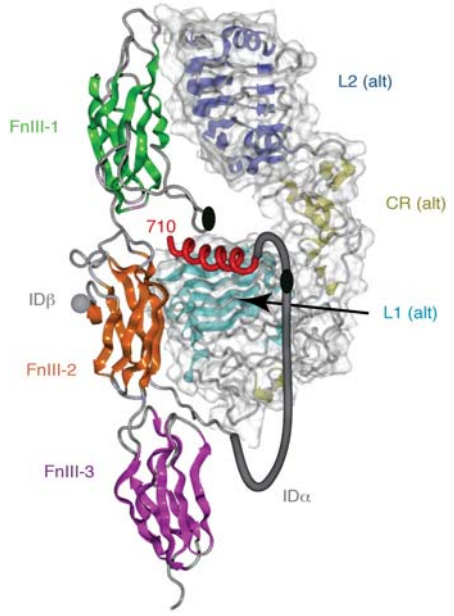

Fig. 4. Schematic diagram showing the *trans* assembly of the tandem hormone-binding element within one leg of the homodimeric insulin receptor ectodomain. The fibronectin module of the insulin receptor monomer that contributes the αCT segment is shown in secondary structure representation while the L1, CR, and L2 domains of the alternate monomer within the same leg of the homodimeric are shown in molecular surface-encased secondary structure representation. All domains are colored as in Fig. 1 *B* and *C*, with the αCT segment shown as a *Red* helix. A *Gray* tube depicts in stylized fashion the connection between the last ordered residue of the α-chain as it emerges from the FnIII-2 domain and the start of the αCT segment. The *Black* ellipses depict the location of the inter-α-chain disulfides at Cys524 and at the (Cys682, Cys683, Cys685) cluster.

the hormone-free structure of the receptor (Fig. 4) and so is unrelated to potential intermonomer cross-linking event movements that effect negative cooperativity (11). The present study thus reveals the three-dimensional structure of Site 1, the primary insulin-binding site.

The structure of the tandem hormone-binding element leads to a reinterpretation of alanine scanning mutagenesis data for the insulin receptor. Of the 10 residues from the 693–710 segment that contribute to the αCT segment/L1 domain interface (*SI Appendix*, Table S2), single alanine mutations at four sites directly involved in the interface (Thr704, Phe705, Tyr708, and Leu709) are known to severely impair binding of insulin to its receptor (26, 27). We suggest that the loss of insulin binding in these instances is indirect, reflecting rather structural perturbation of the interface within the tandem element. However, alanine substitutions at two sites distal to the interface (Glu706 and His710) also impair insulin binding; here we propose that these residues interact directly with insulin. Phe714, also critical for insulin binding, lies downstream of the segment modeled here and is disordered in the ectodomain crystal structure; here we predict that this residue too is involved in direct interaction with insulin. The L1 domain residues Gln34, Leu36, Leu37, Phe64, Phe88, Phe89, and Glu120 have less than 28% of their potential surface area available for possible interaction with insulin (*SI Appendix*, Table S2), yet at these positions alanine substitutions leads to insulin affinities of 0.04% to 40% relative to wild-type. We therefore propose that their primary contribution to insulin binding is through stabilization of the tandem element. In contrast, Asp12, Arg14, Phe39, Asn90, Tyr91, and Lys121 have a more limited engagement in the αCT segment/L1 domain interface, with 80% to 98% of their surfaces remaining accessible (*SI Appendix*, Table S2), and yet alanine substitutions at these sites also impair insulin binding (affinities 0.1% to 33% of wild-type). We speculate that these side chains directly contact insulin and reflect a diversity of charged, polar, aromatic, and aliphatic interactions.

The picomolar receptor-binding affinity achieved by linking a Site 1 mimetic peptide C-terminal to a Site 2 mimetic peptide (as exemplified by the 36-residue peptide termed S519) (20) may reflect an economical structural mechanism of receptor binding and activation. Our results support the hypothesis (12) that Site 1 mimetic peptides (exemplified by the 16 C-terminal residue segment of the S519 peptide, termed S519C16) bind to insulin receptor by αCT mimicry. Site 2 mimetic peptides, however, must bind differently, because a prototypic Site 2 peptide (the N-terminal 20-residue segment of the S519 peptide, termed S519N20) does not bind to the insulin minireceptor IR485 (12). The S519N20 peptide contains a disulfide bridge at its C terminus and, like the S519C16 peptide, is predicted to be helical in structure (7). The full-length S519 peptide may thus form an antiparallel two-helix bundle similar to insulin (7) and also exhibit *trans* binding to the L1 domain surface and the FnIII-1/FnIII-2 loops. This suggests that the insulin-binding cavity within the receptor homodimer (Fig. 1*C*) may provide a target for the design of helical mimetic nonpeptide agonists, perhaps achievable in part by molecules the size of macrolide antibiotics and could be "druggable."

A critical unanswered question is whether, after insulin binds, the receptor αCT segment remains in the conformation and location observed here. A recent model of the complex (28) would imply significant steric overlap between insulin and the L1-bound α-chain segment. Studies of photo-activatable insulin derivatives have shown that azido probes at B16, B24 (as either the L- or D- enantiomer) and B26 cross-link to the receptor L1 domain, whereas probes at A1 (as the D-enantiomer), A2, A3, A4, A8, A14, B25, and B27 cross-link to the C-terminal region of the receptor α-chain (14). These experiments suggest that αCT stabilizes a structural transition in the bound insulin monomer involving displacement of the B-chain C-terminal segment from its classical conformation in insulin dimers and hexamers (14). Such change of conformation of the insulin B-chain is also supported by recent analysis of crystal structures of nonstandard insulin analogs (29). There is a weak sequence relationship between the insulin B-chain and Site 1 mimetic peptides (7), but it does not include the "footprint" residues Phe701, Phe705, and Tyr708. This suggests that if the insulin B-chain does displace the αCT segment, then its mode of L1 domain binding would likely be different from that of αCT. We have shown (*SI Appendix*) that the tighter binding of the variant aromatic-substituted αCT peptides to the receptor is associated with an approximately commensurate loss of insulin affinity. Unfortunately, this does not answer the question of αCT mobility, as the observed loss of affinity can be attributed to either a reduction in a required mobility of the αCT peptide or steric hindrance from the introduced aromatic residues. Final resolution of this issue must await determination of the structure of the hormone-bound receptor.

The interaction between the αCT segment and L1 domain observed here is likely present in IGF-1R, given its close relationship to the insulin receptor (4, 28). Of the 22 L1 domain residues of insulin receptor involved in the interaction with the αCT segment, 16 are conserved in IGF-1R, with at least three of the remaining six residues being conservatively substituted (*SI Appendix*, Table S2). Of the 10 residues of the αCT segment involved in the interaction with the insulin receptor L1 domain, five are conserved in IGF-1R (*SI Appendix*, Table S2), with three of the remaining five residues being conservatively substituted. The remaining two residues (Lys694 and Ser700) have only limited involvement in the interaction. These observations rationalize the affinities of the insulin-mimetic peptides for both receptors (20) and permit construction of comparative structural models for IGF-1R and as well as for the hybrid (insulin receptor/IGF-1R) receptors. Together, our results have the potential to lead to new avenues of drug design for the treatment of diabetes and cancer.

Materials and Methods

Detailed Materials and Methods are available in *SI Appendix*.

Crystallography. The original structure determination employed a single crystal of the insulin receptor ectodomain complexed with four F_{ab}s (3). Diffraction images from this study were reprocessed using XDS and XSCALE (30) to provide revised data for refinement. Reflections were assigned the same free-R flag as they had in the original dataset associated with PDB entry 2DTG (3). The original model, following minor in-house improvement, was then subjected to rigidbody crystallographic refinement against the revised diffraction amplitudes, followed by atomic coordinate, atomic displacement parameter, and TLS refinement within PHENIX (31). Difference electron maps were sharpened by application of a *B*-factor [equal to the negative (−153.6 Å^2) of that determined by a Wilson plot], to constituent structure-factor amplitudes (16). The tube-like segment of electron density originally observed on the surface of the central β-sheet of the L1 domain (3) now exhibited a clear helical conformation and was present as the largest feature in the sharpened (mF_o-DF_c) difference electron-density map. Sufficient side-chain electron density was visible to allow sequence assignment of the helix to residues 693–710 of the insulin receptor α-chain. Crystallographic refinement of the structure inclusive of the 693–710 segment then followed (atomic-coordinate, atomic displacement parameter, and TLS refinement within PHENIX). Coordinates and structure factors have been deposited with Protein Data Bank (accession code 3LOH).

Isothermal Titration Calorimetry. Biotinylated αCT peptides at >75% purity spanning residues 698–719 of the insulin receptor were obtained from Genscript Inc. The insulin minireceptor IR485 construct was produced and purified as previously described (28), omitting the final ion-exchange chromatography step. The ITC cell contained insulin minireceptor IR485 prepared at 10 μM concentration in Tris-buffered saline plus azide (TBSA; 24.8 mM Tris-HCl (pH 8.0), 137 mM NaCl, 2.7 mM KCl, and 0.02% sodium azide), and the syringe contained the peptide prepared at 60 μM concentration in TBSA. Injection protocols and data processing were as described previously (12).

Molecular Modeling. The docking of the single- and double-mutant αCT peptides to the surface of the L1 domain was investigated by molecular dynamics simulation using the GROMACS v4.0 suite (32) with the OPLS-aa force field (33).

Synthesis of Photo-Reactive Insulin Analogs. Photo-activatable two-chain and single-chain insulin derivatives were prepared by solid-phase peptide synthesis using the photo-stable precursor *para*-amino-Phe (Pmp). Analogs containing B- and A-chain photo-probes were labeled by 6-biotinylamidocaproyl-Phe^{B1} or 6-biotinylamidocaproyl-D-Lys^{A1}, respectively. Two-chain derivatives at positions B16, B25, A3, or A14 were prepared as described (24, 25). A bifunctional derivative containing Pmp^{B16} and Pmp^{B25} was prepared similarly. Chain combination was not significantly impeded by the Pmp substitutions. Single-chain derivatives, employing a 57-residue framework with connecting peptide GGGPRR, were prepared by native-ligation chemical synthesis and refolded as described (34). Pmp derivatives were converted to corresponding *para*-azido-Phe (Pap) derivatives as described and repurified in the dark by reverse-phase HPLC (8). Predicted molecular masses of Pmp and Pap derivatives were in each case verified by mass spectrometry.

Receptor Expression and Purification. The insulin receptor (B isoform) was expressed in stably transfected dihydrofolate-deficient Chinese Hamster Ovary cell line P3-A, extracted, and WGA-purified as described (23).

Photo-Cross-Linking. Protocols for photo-cross-linking were identical to those employed previously (25). After cross-linking and reduction (100 mM DTT) of the complexes, the solutions were run on a 10–20% gradient SDS–PAGE. The separated proteins were blotted onto a nitrocellulose membrane and probed with NeutrAvidin (Pierce, IL) and a polyclonal antibody (catalog number SC-710, Santa Cruz Biotech., CA), which recognizes the insulin receptor α-chain residues 1–20.

ACKNOWLEDGMENTS. This work was supported in part by the Australian National Health and Medical Research Council (NHMRC) (Project Grants 516729 and 575539), the NHMRC Independent Research Institutes Infrastructure Support Scheme (Grant 361646), a Victorian State Government Operational Infrastructure Support Grant (M.C.L., B.J.S., G.K., J.G.M., C.W.W.), National Institutes of Health grants to the University of Chicago Diabetes Research and Training Center (DK20595), P.G.K. (DK56673), D.F.S. (DK13914), and M.A.W. (DK40949).

1. De Meyts P (2004) Insulin and its receptor: Structure, function and evolution. *Bioessays* 26:1351–1362.
2. Adams MJ, et al. (1969) Structure of rhombohedral 2 zinc insulin crystals. *Nature* 224:491–495.
3. McKern NM, et al. (2006) Structure of the insulin receptor ectodomain reveals a folded-over conformation. *Nature* 443:218–221.
4. Lawrence MC, McKern NM, Ward CW (2007) Insulin receptor structure and its implications for the IGF-1 receptor. *Curr Opin Struct Biol* 17:699–705.
5. Ward CW, Lawrence MC, Streltsov VA, Adams TE, McKern NM (2007) The insulin and EGF receptor structures: New insights into ligand-induced receptor activation. *Trends Biochem Sci* 32:129–137.
6. Ward C, et al. (2008) Structural insights into ligand-induced activation of the insulin receptor. *Acta Physiol (Oxf)* 192:3–9.
7. Ward CW, Lawrence MC (2009) Ligand-induced activation of the insulin receptor: A multi-step process involving structural changes in both the ligand and the receptor. *Bioessays* 31:422–434.
8. Kurose T, et al. (1994) Cross-linking of a B25 azidophenylalanine insulin derivative to the carboxyl-terminal region of the α-subunit of the insulin receptor. Identification of a new insulin-binding domain in the insulin receptor. *J Biol Chem* 269:29190–29197.
9. Mynarcik DC, Williams PF, Schäffer L, Yu GQ, Whittaker J (1997) Analog binding properties of insulin receptor mutants. Identification of amino acids interacting with the COOH terminus of the B-chain of the insulin molecule. *J Biol Chem* 272:2077–2081.
10. De Meyts P, Palsgaard J, Sajid W, Theede AM, Aladdin H (2004) Structural biology of insulin and IGF-1 receptors. *Novart Fdn Symp* 262:160–171.
11. Kiselyov VV, Versteyhe S, Gauguin L, De Meyts P (2009) Harmonic oscillator model of the insulin and IGF1 receptors' allosteric binding and activation. *Mol Syst Biol* 5:243.
12. Menting JG, Ward CW, Margetts MB, Lawrence MC (2009) 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 Site 1 insulin mimetic peptides. *Biochemistry* 48:5492–5500.
13. Pillutla RC, et al. (2002) Peptides identify the critical hotspots involved in the biological activation of the insulin receptor. *J Biol Chem* 277:22590–22594.
14. Xu B, et al. (2009) Decoding the cryptic active conformation of a protein by synthetic photo-scanning. Insulin inserts a detachable arm between receptor domains. *J Biol Chem* 284:14597–14608.
15. Read RJ (1986) Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr A* 42:140–149.
16. Brünger AT, DeLaBarre B, Davies JM, Weis WI (2009) X-ray structure determination at low resolution. *Acta Crystallogr D* 65:128–133.
17. Davies JM, Brünger AT, Weis WI (2008) Improved structures of full-length p97, an AAA ATPase: Implications for mechanisms of nucleotide-dependent conformational change. *Structure* 16:715–726.
18. Kristensen C, Wiberg FC, Andersen AS (1999) 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* 274:37351–37356.
19. Lawrence MC, Colman PM (1993) Shape complementarity at protein/protein interfaces. *J Mol Biol* 234:946–950.
20. Schäffer L, et al. (2003) Assembly of high-affinity insulin receptor agonists and antagonists from peptide building blocks. *Proc Natl Acad Sci USA* 100:4435–4439.
21. Bass J, Chiu G, Argon Y, Steiner DF (1998) Folding of insulin receptor monomers is facilitated by the molecular chaperones calnexin and calreticulin and impaired by rapid dimerization. *J Cell Biol* 141:637–646.
22. Chan SJ, Nakagawa S, Steiner DF (2007) Complementation analysis demonstrates that insulin cross-links both α subunits in a truncated insulin receptor dimer. *J Biol Chem* 282:13754–13758.
23. Xu B, et al. (2004) Diabetes-associated mutations in insulin: consecutive residues in the B chain contact distinct domains of the insulin receptor. *Biochemistry* 43:8356–8372.
24. Huang K, et al. (2007) The A-chain of insulin contacts the insert domain of the insulin receptor. Photo-cross-linking and mutagenesis of a diabetes-related crevice. *J Biol Chem* 282:35337–35349.
25. Huang K, et al. (2004) How insulin binds: The B-chain α-helix contacts the L1 β-helix of the insulin receptor. *J Mol Biol* 341:529–550.
26. Mynarcik DC, Yu GQ, Whittaker J (1996) Alanine-scanning mutagenesis of a C-terminal ligand binding domain of the insulin receptor α subunit. *J Biol Chem* 271:2439–2442.
27. Whittaker J, Whittaker L (2005) Characterization of the functional insulin binding epitopes of the full-length insulin receptor. *J Biol Chem* 280:20932–20936.
28. Lou M, et al. (2006) The first three domains of the insulin receptor differ structurally from the insulin-like growth factor 1 receptor in the regions governing ligand specificity. *Proc Natl Acad Sci USA* 103:12429–12434.
29. Jiráček J, et al. (2010) Implications for the active form of human insulin based on the structural convergence of highly active hormone analogues. *Proc Natl Acad Sci USA*, in press.
30. Kabsch W (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J Appl Crystallogr* 26:795–800.
31. Adams PD, et al. (2002) PHENIX: Building new software for automated crystallographic structure determination. *Acta Crystallogr D* 58:1948–1954.
32. Van der Spoel D, Lindahl E, Hess B (2005) Gromacs: Fast, flexible and free. *J Comput Chem* 26:1701–1718.
33. Jorgensen WL, Tirado-Rives J (1988) The OPLS potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J Am Chem Soc* 110:1657–1666.
34. Hua QX, et al. (2008) Design of an active ultrastable single-chain insulin analog. Synthesis, structure, and therapeutic implications. *J Biol Chem* 283:14703–14716.