THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 291, NO. 53, pp. 27023–27041, December 30, 2016

Extending Halogen-based Medicinal Chemistry to Proteins

IODO-INSULIN AS A CASE STUDY[*] [S]

Received for publication, September 29, 2016, and in revised form, October 31, 2016 Published, JBC Papers in Press, November 14, 2016, DOI 10.1074/jbc.M116.761015

Krystel El Hage[‡1,2], Vijay Pandyarajan[§1,3], Nelson B. Phillips[§4], Brian J. Smith[¶5], John G. Menting[‖], Jonathan Whittaker[§4], Michael C. Lawrence[‖6], Markus Meuwly[‡7], and Michael A. Weiss[§‡‡§§8]**

*From the [‡]Department of Chemistry, University of Basel, Klingelbergstrasse 80 CH-4056 Basel, Switzerland, the Departments of [§]Biochemistry, [‡‡]Medicine, and [§§]Biomedical Engineering, Case Western Reserve University, Cleveland, Ohio 44106, the [¶]La Trobe Institute for Molecular Science, La Trobe University, Melbourne, Victoria 3086, Australia, the [‖]The Walter and Eliza Hall Institute of Medical Research, 1G Royal Parade, Parkville, Victoria 3052, Australia, and the **Department of Medical Biology, University of Melbourne, Parkville, Victoria 3010, Australia*

Edited by Norma Allewell

Insulin, a protein critical for metabolic homeostasis, provides a classical model for protein design with application to human health. Recent efforts to improve its pharmaceutical formulation demonstrated that iodination of a conserved tyrosine (Tyr^{B26}) enhances key properties of a rapid-acting clinical analog. Moreover, the broad utility of halogens in medicinal chemistry has motivated the use of hybrid quantum- and molecular-mechanical methods to study proteins. Here, we (i) undertook quantitative atomistic simulations of 3-[iodo-Tyr^{B26}]insulin to predict its structural features, and (ii) tested these predictions by X-ray crystallography. Using an electrostatic model of the modified aromatic ring based on quantum chemistry, the calculations suggested that the analog, as a dimer and hexamer, exhibits subtle differences in aromatic-aromatic interactions at the dimer interface. Aromatic rings (Tyr^{B16}, Phe^{B24}, Phe^{B25}, 3-I-Tyr^{B26}, and their symmetry-related mates) at this interface adjust to enable packing of the hydrophobic iodine atoms within the core of each monomer. Strikingly, these features were observed in the crystal structure of a 3-[iodo-Tyr^{B26}]insulin analog (determined as an R_6 zinc hexamer). Given that residues B24–B30 detach from the core on receptor binding, the environment of 3-I-Tyr^{B26} in a receptor complex must differ from that in the free hormone. Based on the recent structure of a "micro-receptor" complex, we predict that 3-I-Tyr^{B26} engages the receptor via directional halogen bonding and halogen-directed hydrogen bonding as follows: favorable electrostatic interactions exploiting, respectively, the halogen's electron-deficient σ-hole and electronegative equatorial band. Inspired by quantum chemistry and molecular dynamics, such "halogen engineering" promises to extend principles of medicinal chemistry to proteins.

Insulin, a small protein critical to metabolic homeostasis (1), provides a model for studies of protein folding and design (2) with long-standing application to human therapeutics (3). The hormone contains two chains, A and B (Fig. 1*A*), linked by two disulfide bridges (cystines A7–B7 and A20–B19); the A chain is further stabilized by cystine A6–A11. In pancreatic β-cells, insulin is stored within the secretory granules as zinc-coordinated hexamers. This study has exploited insulin semi-synthesis (4) (simplified through the use of norleucine (Nle)[9] at position B29; *arrow* in Fig. 1*A* (5)) to investigate a site-specific modification of an aromatic ring by a single halogen atom (6). The modification, 3-iodo-Tyr at position B26 (3-I-Tyr^{B26}), is associated with enhanced binding to the insulin receptor (IR)

* This work was supported in part by National Institutes of Health Grants R01 DK04949 and DK079233 from NIDDK (to M. A. W.), Swiss National Science Foundation Grant 200021-117810, National Competence Center for Research Molecular and Ultrafast Science and Technology (CCR MUST) (to M. M.), Australian National Health and Medical Research Council Project Grants 1005896 and 1058233, and the Hazel and Pip Appel Fund (to M. C. L.). M. A. W. has equity in Thermalin Diabetes, LLC (Cleveland, OH), where he serves as Chief Scientific Officer; he has also been a consultant to Merck Research Laboratories and DEKA Research and Development Corp. N. B. P. and J. W. are consultants to Thermalin Diabetes, LLC. Part of M. C. L.'s research is funded by Sanofi (Germany). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

This work was stimulated by the Nobel Celebration Symposium in honor of Prof. Martin Karplus, held in San Francisco (Oct. 1, 2014), and is dedicated by the corresponding authors to his mentorship.

[S] This article contains supplemental Figs. S1–S6 and Table S1.

The atomic coordinates and structure factors (code 5EMS) have been deposited in the Protein Data Bank (http://wwpdb.org/).

[1] Both authors contributed equally to this work.

[2] Supported by the NCCR MUST through the Swiss National Science Foundation.

[3] Supported by a predoctoral fellowship of the Case Western Reserve University Medical Scientist Training Program supported by National Institutes of Heath Institutional Grant T32 GM007250 and National Institutes of Health Fellowship F30 DK094685-04.

[4] Supported in part by the American Diabetes Association.

[5] Supported by computational resources at the Victorian Life Sciences Computation Initiative.

[6] Recipient of National Health and Medical Research Council Independent Research Institutes Infrastructure Support Scheme Grant 361646 and Victorian State Government Operational Infrastructure Support Grant (to the Walter and Eliza Hall Institute).

[7] Supported by the NCCR MUST and the University of Basel. To whom correspondence may be addressed. Tel.: 41-61-267-38-21; E-mail: m.meuwly@unibas.ch.

[8] To whom correspondence may be addressed: Depts. of Biochemistry, Medicine, and Biomedical Engineering, Case Western Reserve University, Cleveland, OH 44016. Tel.: 216-368-5991; E-mail: michael.weiss@case.edu.

[9] The abbreviations used are: Nle, norleucine; 3-I-Tyr, 3-iodotyrosine; αCT, α-chain C-terminal segment; CR, Cys-rich domain; IR, insulin receptor; IR-A and IR-B, A and B isoforms of the IR; L1, first Leu-rich repeat domain; L2, second Leu-rich repeat domain; PME, particle mesh Ewald; QM, quantum mechanics; MM, molecular mechanics; MD, molecular dynamics; MTP, electrostatic multipole; PC, point charge; PDB, Protein Data Bank; r.m.s.d., root-mean-square differences; ESP, electrostatic potential.

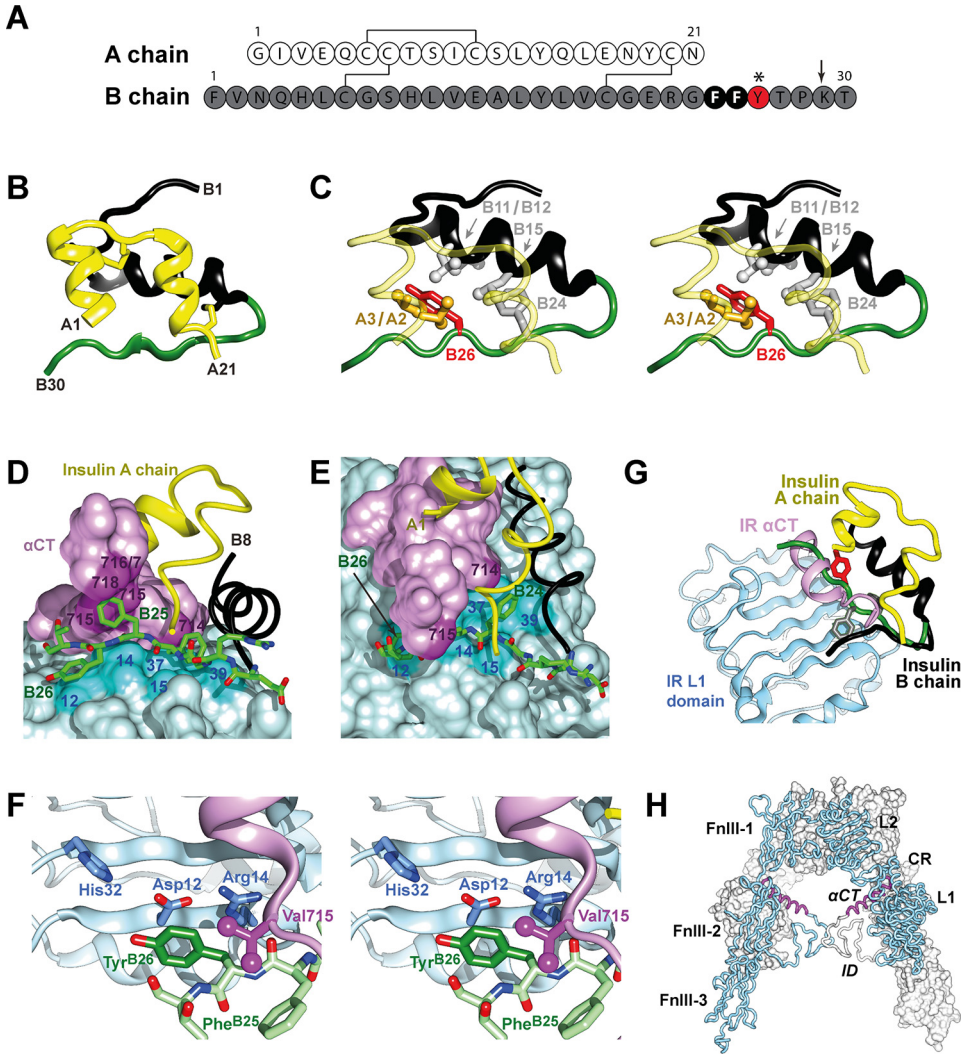

FIGURE 1. **Insulin sequence and structure.** *A,* sequence of WT insulin and modification sites. A and B chains are shown in *white* and *gray*. Conserved aromatic residues Phe^{B24} and Phe^{B25} are highlighted as *black circles*. This study focused on substitutions of Tyr^{B26} (*red circle*); additional substitutions were made at position B29 (Nle; *arrow*) to facilitate semi-synthesis. *B, ribbon* model of insulin monomer (T state extracted from T_6 zinc hexamer) (2). The A chain is shown in *yellow* and the B chain in *black* (B1–B19) and *green* (B20-B30). *C,* environment of Tyr^{B26} (stereo). The side chain of Tyr^{B26} is highlighted in *red*; Ile^{A2}, Val^{A3}, Val^{B12}, Leu^{B15}, and Phe^{B24} are as labeled. *D, stick* representation of residues B20–B27 (carbon atoms (*green*), nitrogen atoms (*blue*) and oxygen atoms (*red*)) packed between αCT and the L1-β_2 sheet. B chain residues B8–B19 are shown as a *black ribbon* and the A chain as a *yellow ribbon*; residues A1–A3 are concealed behind the surface of αCT. Key contact surfaces of αCT with B24–B26 are highlighted in *magenta*, and L1 with B24–B26 are highlighted in *cyan*; L1 and αCT surfaces not in interaction with B24–B26 are shown in *lighter shades*. *E,* orthogonal view to *D*, showing interaction of the side chain of Phe^{B24} with the nonpolar surface of the L1-β_2 sheet. Tyr^{B26} is hidden below the surface of αCT. Engagement of conserved residues A1–A3 against the nonpolar surface of αCT is shown at top. *F,* environment of Tyr^{B26} within site 1 complex (stereo). Neighboring side chains in L1 and αCT are as labeled. Coordinates were obtained from PDB code 4OGA (39). *G,* model of wild-type insulin in its receptor-free conformation overlaid onto the structure of the insulin-bound μIR (71). The L1 domain and part of CR domain are shown in *powder blue*; αCT is shown in *purple*. Residues Phe^{B24} and Tyr^{B26} are as in *A*. The B chain of μIR-bound insulin is shown in *dark gray* (B6–B19); the *brown tube* indicates classical location within the overlay of residues B20–B30 of insulin in its receptor-free conformation, highlighting steric clash of B26–B30 with αCT. In the μIR co-crystal structure, insertion of the insulin B20–B27 segment between L1 and αCT peptide is associated with a small rotation of the B20–B23 β-turn and changes in main-chain dihedral angles flanking Phe^{B24} (39). *H,* inverted V-shaped assembly of IR ectodomain homodimer. One monomer is in *ribbon* representation (*labeled*), the second in *surface* representation. Domains are labeled as follows: *L1,* first leucine-rich repeat domain; *CR,* cysteine-rich domain; *L2,* second leucine-rich repeat domain; *FnIII-1, -2,* and *-3,* first, second, and third fibronectin type III domains, respectively; *ID,* insert domain; and *αCT,* α-chain C-terminal segment. Coordinates were obtained from PDB code 4ZXB (42).

(7–9). The general class of halo-aromatic modifications defines a key frontier of medicinal chemistry (10) and holds promise in the non-standard engineering of proteins through manipulation of π systems and weakly polar interactions (11, 12). In addition, the σ-hole of larger halogens can anchor interactions with surrounding polar groups and water molecules (13–19). 3-[iodo-Tyr^{B26}]Insulin thus provides a model for studies of engineered proteins at the border of molecular mechanics and quantum chemistry.

Crystal structures of insulin (as zinc-free dimers (20) or zinc-stabilized hexamers (21–23)) provide a foundation for its therapeutic formulation (24) and analysis of structure-activity relationships (2, 25, 26). The structure of a monomer in solution (27–30) resembles a crystallographic protomer (as in zinc-free T_2 dimers or T_6 zinc hexamers) (Fig. 1*B*); this "closed" conformation contains an α-helical globular subdomain and tethered C-terminal B-chain β-strand (residues B24–B28). Tyr^{B26} (*red* in Fig. 1*A*, *asterisk*) provides a key contact between the β-strand

and the α-helical subdomain (Fig. 1*C*). The contribution of this side chain to the stability of the insulin monomer has recently been investigated by molecular dynamics (MD) simulations (31) and mutagenesis (32).

Insulin undergoes a change in conformation to "open" on receptor binding (33–37). A recent structural advance exploited domain-minimized models of the α-subunit of the insulin receptor (IR) (38) containing the primary insulin-binding elements (leucine-rich domain 1 (L1) and the C-terminal segment of the α-subunit (αCT)) (39). A co-crystal structure has been determined at 3.5 Å resolution of a ternary complex involving insulin, an L1-CR fragment, and a synthetic αCT peptide (residues 704–719 of receptor isoform A (IR-A)) (39).[10] In this structure (designated the micro-receptor (μIR) complex), the C-terminal segment of the insulin B chain is detached from the hormone's α-helical core; such detachment enables its insertion between L1 and αCT (Fig. 1*D*). The inserted segment includes a conserved triplet of aromatic residues (Phe^{B24}, Phe^{B25}, and Tyr^{B26}) that lie at the μIR interface (39). Whereas Phe^{B24} packs within a classical nonpolar pocket, Tyr^{B26} lies at one edge (Fig. 1*D* and 90° rotated view in Fig. 1*E*). An expanded view of the Tyr^{B26} environment in the μIR (Fig. 1*F*) highlights contacts to conserved residues within L1 (Asp-12, Arg-14, and His-32) and αCT (Val-715). These contacts require repositioning of the C-terminal segment of the insulin B chain from its unbound conformation (*green* in Fig. 1*G*) in the μIR complex (*black*), thereby avoiding a clash between B25–B30 and αCT (*purple*). The L1 and αCT elements of the ectodomain belong to different α-subunits within the multidomain $(\alpha\beta)_2$ IR dimer (Fig. 1*H*). Binding of insulin in *trans* to these elements may alter the orientation between the αβ-subunits as the first step in signal propagation (40–42).[11]

Insulin's dimer interface is remarkable for aromatic-aromatic interactions across eight aromatic side chains (Tyr^{B16}, Phe^{B24}, Phe^{B25}, Tyr^{B26}, and their symmetry-related mates (2, 43)). In this cluster, the side chain of Tyr^{B26} packs against Phe^{B24} and the dimer-related side chains of $Tyr^{B16'}$ and $Phe^{B24'}$ (where ′ indicates that the residue belongs to the alternate monomer within the dimer; Fig. 2*A*), giving rise to complex and asymmetric electrostatic environments (44). Halogen substitutions within these rings would be expected to alter the distribution of π electrons and so modulate such interactions (Fig. 2*B*). In addition, the tyrosine's *para*-OH group would be expected to cause subtle differences in iodine's inductive effects (Fig. 2*C*). Given these features, the present study focused on the effects of iodination of Tyr^{B26}, long known to enhance the affinity of insulin for the IR (7–9) and recently shown to enhance the pharmaceutical properties of a rapid-acting clinical analog, including its stability and resistance to physical degradation (45).[12] Such findings raise salient questions regarding the role of the iodo-aromatic modification on the structure of the free hormone and its potential role at the hormone-receptor interface.

How might iodination of an aromatic residue affect its electrostatic properties and in turn its conformation? What would be the preferred molecular environment of the iodine, and how might its quantum-chemical features be exploited? In particular, how might the asymmetric electronic distribution of the iodo-substituent, in principle capable of halogen bonding (46, 47) and/or halogen-directed hydrogen bonding (18, 48, 49), affect weakly polar interactions within the protein (44)? To what extent might these chemical features underlie the improved biochemical or biophysical properties of such a modified protein? We addressed these questions in three parts. Our study began with molecular dynamics (MD) simulations of 3-[iodo-Tyr^{B26}]insulin with multipolar parameters derived from quantum-mechanical (QM) modeling of the halogenated side chain. We next verified predicted features of such models by determining the crystal structure of a 3-[iodo-Tyr^{B26}]insulin analog, herein described as an R_6 zinc insulin hexamer. The final part of this study sought insight into potential mechanisms by which 3-I-Tyr^{B26} enhances IR binding (7–9). Together, our results highlight the promise of non-standard protein engineering guided by molecular mechanics at the interface of quantum chemistry.

Results

Rigid-body Modeling Distinguished Opposite Edges of the B26 Aromatic Ring

Past ^{1}H NMR studies of insulin as an R_6 zinc hexamer (50, 51) or engineered T_2 dimer (52) demonstrated that all Phe and Tyr side chains exhibit equivalent *meta* resonances (ring positions 3 and 5 as defined in Fig. 2*A*, *left*) and likewise equivalent *ortho* resonances (positions 2 and 6). These observations indicated that the aromatic rings either freely rotate (as do Phe^{B1} and Tyr^{A14}) or undergo rapid 180° "flips" about the C_β–C_γ bond axis (<1 ms on the NMR time scale; Tyr^{A19}, Tyr^{B16}, Phe^{B24}, Phe^{B25}, and Tyr^{B26}). Because within the native state the latter rotations would incur steric clashes, such ^{1}H NMR features reflect the flexibility of the surrounding protein framework (50).

Respective *ortho* and *meta* positions of 3-I-Tyr are in principle not equivalent. Accordingly, which of the two B26 conformations, *i.e.* with the iodine "in" or "out" with respect to the core of a monomer, is preferred? In either orientation, the iodine atom would be inaccessible to solvent within a nonpolar environment. Naive modeling of the insulin dimer suggested that either conformation would encounter marked steric occlusion as follows: at the 3-position (in), an iodine would overlap with the γ-CH_3 of Ile^{A2} and one γ-CH_3 of Val^{A3} (Fig. 3*A*),

[10] The respective mRNAs encoding isoforms A and B of the IR differ by the absence (A) or presence (B) of 36 bases encoded by exon 11 (110).

[11] Because hormone-ectodomain complexes have to date proven refractory to crystallization, it is not known whether or how domains CR, L2, or the three fibronectin-homology domains may contribute to insulin binding or signaling (39). Potential structural differences between isoforms IR-A and IR-B (110) are also not well understood.

[12] Biophysical effects of the iodo-Tyr^{B26} modification were recently studied in the context of insulin *lispro* (containing substitutions $Pro^{B28} \rightarrow Lys$ and $Lys^{B29} \rightarrow Pro$; the active component of Humalog® (Lilly) (45). Although the structure of the modified analog was not determined, iodo-Tyr^{B26} was found to mitigate the loss of stability associated with the paired B28–B29 substitutions. In the context of the present parent analog ([Nle^{B29}]insulin), iodo-Tyr^{B26} does not alter stability, presumably because the native Pro at B28 is retained, *i.e.* the parent's native-like dimer interface is not in need of repair. The iodo-Tyr^{B26} modification enhances receptor binding to a similar extent in both templates.

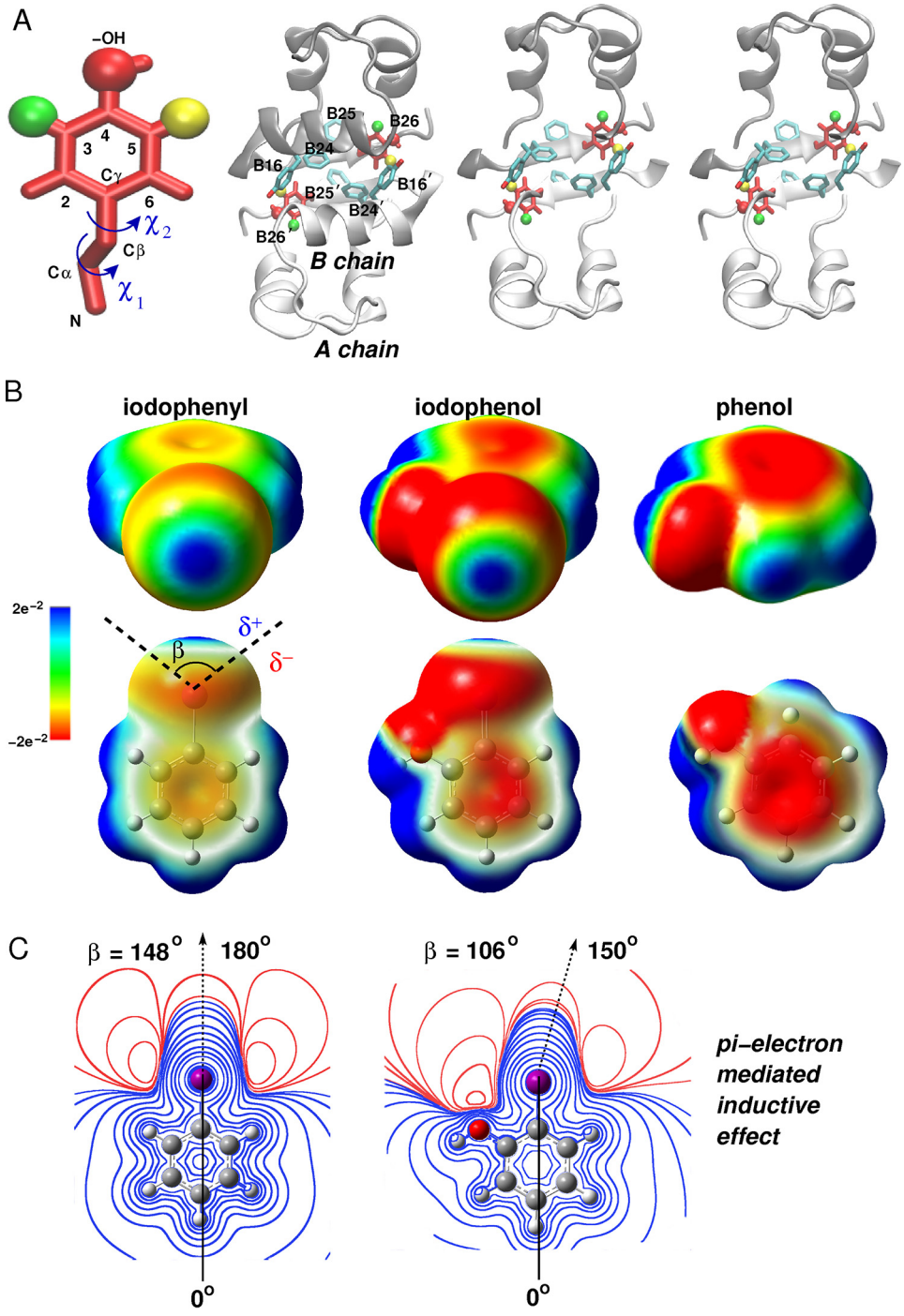

FIGURE 2. **Quantum chemistry of iodo-aromatic system.** *A,* Tyr side chain with iodine in its 3- or 5-ring position (*green* and *yellow*, in and out, respectively) with carbon-atom positions labeled. Rotation angles for rigid-body modeling involve rotations around the C_α–C_β bond (χ_1) and C_β–C_γ bond (χ_2). The insulin dimer with B26 side chains (*red*) and side chains of Tyr^{B16}, Phe^{B24}, Phe^{B25}, and their dimer-related mates (*blue*). The *right-hand side* provides a cross-eyed stereo view of the modified dimer with the B chain helix removed for clarity. *B,* electrostatic potential (*ESP*) surface maps of phenol, iodophenol, and iodophenyl at the 0.001 $e\ bohr^{-3}$ isodensity. The color scale of the surface potential ranges from $-2.12\ e^{-2}$ (*red*) through 0 (*green*) to $2.12\ e^{-2}$ (*blue*). In the *upper row* the iodine (facing the viewer) exhibits the effect of the electron-donating –OH on the σ-hole. The *lower row* shows effects of iodine on the π-system of the phenol ring. In the *1st row* the surface is opaque, and in the *2nd row* the surface is transparent. Angle β represents the σ-hole size as delimited by *black dashed lines*. δ^+ and δ^- represent respective regions of positive and negative charge around the iodine. *C,* ESP contours of iodophenyl (*left*) and 2-iodophenol (*right*), at different isovalues, calculated in the plane of the aromatic ring. The halogen boundary represents a region of an electron isodensity of $10^{-3}\ e\ bohr^{-3}$ (111). Isocontours in the *left* and *right panels* are at the same heights but in uneven separations. The σ-hole size, defined by an angle β (*B*), was calculated from the angular profile of the ESP on the intersection line of the 2D grid and halogen boundary where the ESP changes its sign (55); positive ESP and negative ESP regions are shown in *blue* and *red*, respectively. The *black dashed arrow* indicates directionality of the C–I bond.

whereas at the 5-position (out) it would encounter the side chain of dimer-related Tyr^{B16} and carbonyl oxygen of Gly^{B20} (Fig. 3*B*). The seeming steric incompatibility of either 3-I-Tyr^{B26} or 5-I-Tyr^{B26} in naive models (akin to seeming steric barriers to ring rotation) stood in contrast to its observed stabilization of an insulin analog (5), suggesting that structural

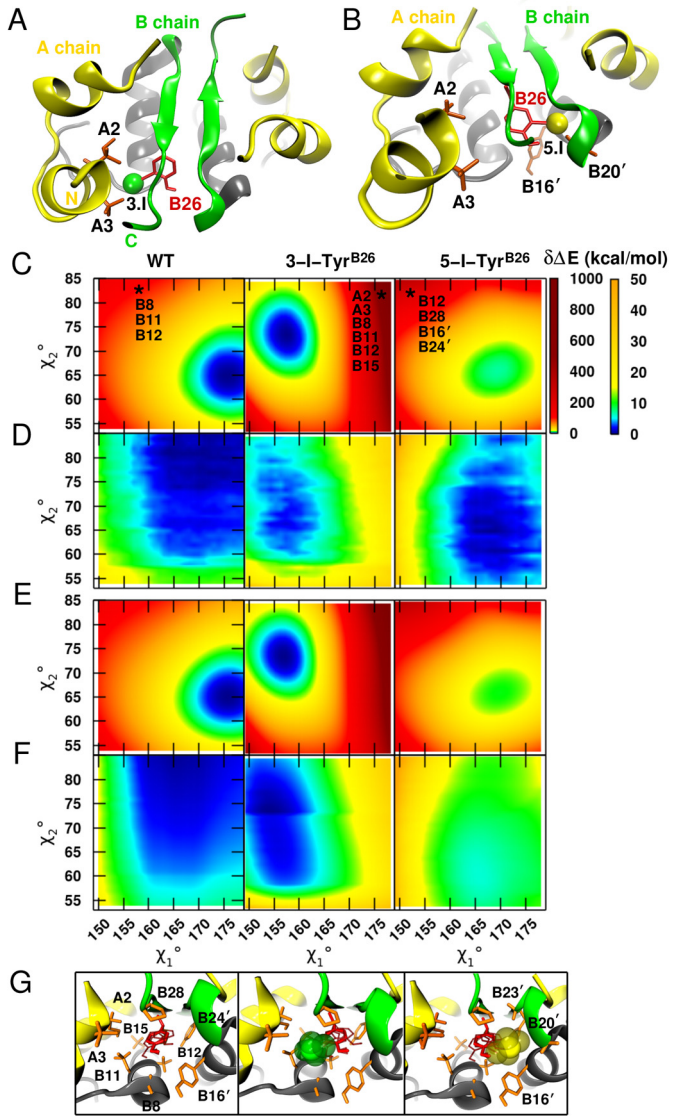

FIGURE 3. **Rigid-body modeling of 3- and 5-[iodo-Tyr^{B26}]insulin analogs.** *A,* naive model of 3-[iodo-Tyr^{B26}]insulin highlights overlap of the iodine with the side chains of Ile^{A2} and Val^{A3}. *B,* analogous model of 5-[iodo-Tyr^{B26}]insulin exhibits clash with Tyr$^{B16'}$ and the backbone oxygen atom of Gly$^{B20'}$. This view is slightly tilted relative to that in *A* to better illustrate the unfavorable contacts. *C,* conformational rotational maps for a PC representation of the B26 side chain around dihedral angles χ_1 and χ_2 for WT (*left*), 3-I-Tyr^{B26} (*middle*), and 5-I-Tyr^{B26} (*right*). The two energy scales highlight large (0–1000 kcal/mol) and finer energy differences (0–50 kcal/mol) relative to the global minimum. The zero of energy for WT and 3-I-Tyr^{B26} is at 0 kcal/mol, whereas the minimum for 5-I-Tyr^{B26} is relative to that of 3-I-Tyr^{B26}. The *stars* indicate one of the (χ_1,χ_2) conformations leading to severe steric clashes with neighboring residues. The geometries are at ⟨157°, 84°⟩, ⟨176°, 83°⟩, and ⟨151°, 84°⟩ for WT, 3-I, and 5-I, respectively, and the clashing residues are listed. *D* is as in *C* but for a relaxed scan over the (χ_1,χ_2) grid. The *panels* show only energies covering χ_1 and χ_2 intervals common to the three dimers as follows: [149°, 180°] for χ_1 and [53°, 85°] for χ_2. *E* and *F* contain respective rigid and relaxed conformational rotational maps for an MTP representation of the B26 side chain. *G,* corresponding structures in the neighborhoods of residue B26 (*transparent red*) highlighting residues affected by dihedral rotations of B26 (*orange*): *left* to *right,* WT insulin, 3-[iodo-Tyr^{B26}]insulin model, and 5-[iodo-Tyr^{B26}]insulin model.

accommodation at one or the other sites is feasible and without free-energy penalty.

To evaluate the two competing modes of iodo-Tyr packing, we first explored potential *local* accommodation by considering B26 (χ_1,χ_2) energy maps in the neighborhood of the native structure. Insight was obtained through the use of increasingly realistic models as follows: first, a conventional point charge (PC (53)) model, and second, an improved representation based on electrostatic multipoles (MTP (54)) as parameterized by *ab initio* quantum-chemical calculations (see "Experimental Procedures"). We then explored potential interplay of *local* and *non-local* mechanisms using MD simulations based on MTPs for the modified side chain. The latter captured essential aspects of quantum chemistry without incurring the computational burden of explicit mixed quantum mechanical/molecular mechanics (QM/MM) simulations. These results are described in turn.

Residue-specific Energy Maps Suggest Local Accommodation Is Possible

For a given representation (PC or MTP) of the B26 side chain, its conformation-dependent empirical energy was first assessed within a rigid protein environment ("unrelaxed maps"). Subsequent energy minimization at each B26 (χ_1,χ_2) setting allowed local structural reorganization of neighboring side chains, yielding the corresponding "relaxed maps." Comparison of unrelaxed (Fig. 3, *C* and *E*) and relaxed (Fig. 3, *D* and *F*) maps suggested that local steric clashes could be mitigated. Comparison of respective PC- and MTP-based maps informed the extent to which such mitigation is influenced by electrostatic features of the iodo-aromatic system.

Representation of Residue B26 by a PC Model—A first step toward distinguishing between "in" and "out" isomers was provided by rigid-body calculations assuming either 3-I-Tyr^{B26} (*green* position in Fig. 2*A*) or 5-[iodo-Tyr^{B26}]insulin (*yellow* position in Fig. 2*A*); in the context of WT insulin, the 3-I and 5-I isomers, respectively, correspond to "iodo-in" and "iodo-out" conformations. To this end, unrelaxed (χ_1,χ_2) energy maps in the neighborhood of the WT structure (±15 in each dihedral angle in steps of 1°) were computed using a conventional PC representation for three B26 side chains (unmodified Tyr^{B26}, 3-I-Tyr^{B26}, and 5-I-Tyr^{B26}; *left* to *right* in Fig. 3*C*). This representation neglected known anisotropic electrostatic features around halogens (48) and their perturbation by the neighboring *para*-OH group (Fig. 2*C*) (18, 55). The shared energy scale of the 3-I and 5-I maps cannot be directly compared with that of WT.

Despite the PC approximation, comparison of qualitative features was informative. (*a*) Consistent with naive modeling, violations (*red regions* in Fig. 3*C*) are more severe for the iodinated residues than for the unmodified residue. (*b*) WT- and 5-I-Tyr^{B26} maps are similar in overall topography; in each case a hydrogen atom at the 3-position points toward the crowded core of the same monomer. (*c*) Surprisingly, in this particular "frozen" protein environment, 3-I-Tyr^{B26} appears to be energetically favored over 5-I-Tyr^{B26}. (*d*) Despite its lower minimum, the 3-I basin is considerably narrower, especially with respect to χ_1 rotation, than is the 5-I basin (or that of the unmodified residue; Fig. 3*C*). The "walls" of this narrow basin were enforced largely by steric clashes with neighboring side chains (Ile^{A2}, Val^{A3}, Ile^{B11}, Val^{B12}, Leu^{B15}, and Pro^{B28}; Fig. 3*G*).

To quantify the effects of relaxing the protein environment on the B26 energy maps, corresponding (χ_1,χ_2) scans were

obtained in which the positions of surrounding side chains were energy-minimized (Fig. 3*D*). As in the unrelaxed calculations, the relaxed 5-I-Tyr^{B26} energy map was similar to that of the unmodified residue. Although the severe clashes found for 3-I-Tyr^{B26} were largely mitigated due to local adjustments, the relaxed PC-based maps predicted that in an insulin dimer the outward positioning of the iodine atom (5-I-Tyr^{B26}) would be preferred by 1.3 kcal/mol relative to the 3-I-Tyr^{B26} isomer. Relaxation of neighboring side chains in the PC model thus reversed the predicted orientation of the iodinated ring from in to out. The 5-I basin also remained wider than that of 3-I.

Although the PC-predicted outward conformation of the iodine atom seemed intuitive given the narrow confines of the hydrophobic core and inferred flexibility of the dimer interface (56–59), we next sought to test whether this prediction remained valid using a more rigorous representation of the halogenated side chain.

Representation of Residue B26 by an MTP Model—Given the limitations of a PC-based force field, we employed an MTP representation of iodo-Tyr parameterized as follows. The electrostatic potential of *ortho*-iodophenol was mapped as a model compound (Fig. 2*B*). The electron-donating property of the *para*-hydroxyl group in iodophenol led to a partial attenuation of iodine's σ-hole together with an increase in its negative equatorial potential region (δ^-; Fig. 2*B*) relative to phenol. The electron-attracting property of the iodine induces non-local decrease of the planar δ^- π-system above and below the plane of the ring.

The unrelaxed MTP energy maps (Fig. 3*E*) validated aspects of the PC calculations (Fig. 3*C*). The WT and 5-I maps were similar to one another, and the 3-I minimum was again found to be more favorable than that of 5-I-Tyr^{B26} but with a narrower basin. In striking contrast to the PC-based calculations, however, the relaxed MTP energy map (Fig. 3*F*) predicted the 3-I minimum remained lower than that of 5-I, in this case by 5 kcal/mol. Hence, the more accurate electrostatic model (MTP) led to a preferred position of the iodine in seeming contradiction to our initial intuition.

B26 Interaction Energy Analysis—To understand the physical origins of the MTP prediction, interaction energies between B26 and the neighboring dimer-related residues were analyzed for the most energetically favorable 3-[iodo-Tyr^{B26}]insulin and 5-[iodo-Tyr^{B26}]insulin dimer structures. Both PC and MTP electrostatic models were employed (supplemental Table S1). The results suggested that, relative to the 5-I orientation, the 3-I conformation provides both (i) favorable hydrophobic interactions with the side chains of Ile^{A2}, Val^{A3}, Leu^{B11}, Val^{B12}, Leu^{B15}, and Pro^{B28}; and (ii) favorable aromatic-aromatic interactions distant from the halogen (Phe$^{B16'}$ and Phe$^{B24'}$). Attractive, weakly polar interactions between 3-I-Tyr^{B26} and Gly^{B8} are also possible. In contrast, 5-I-Tyr^{B26} engages in favorable interactions only in the immediate neighborhood of the iodine, and only with the ends of a solvent-exposed and flexible β-turn (Gly$^{B20'}$ and Gly$^{B23'}$). These key differences rationalize why an inward conformation is preferred despite our initial expectation.

Multipole-based MD Simulations Predicted Structural Features of an Iodinated Insulin Analog

Side-chain conformations were further probed by analyzing 20 ns of four MTP-based MD simulations as follows: the WT and 3-[iodo-Tyr^{B26}]insulin dimers, each in R_2 or T_2 states (supplemental Figs. S1 and S2, respectively). Initial coordinates were obtained from PDB entries 1DPH (T_2) and 1ZNJ (as extracted from a WT R_6 zinc insulin hexamer). We first analyzed respective (χ_1,χ_2) occupancies, $P(\chi_1,\chi_2)$, of residues Tyr^{A19}, Tyr^{B16}, Phe^{B24}, Phe^{B25}, and Tyr^{B26}. Of these, A19 provided a probe of the α-helical core within component protomers, whereas B16 and B24–B26 provided probes of the asymmetric dimer interface. The resulting dihedral angle distributions $P(\chi_1,\chi_2)$ demonstrated that these side chains each adopted conformations consistent with experiments (see symbols in supplemental Figs. S1 and S2) with the exception of Phe^{B25} (whose variable conformation and asymmetry across the dimer interface have previously been noted (2)). $P(\chi_1,\chi_2)$ plots were similar whether the MTP-based MD simulations used T_2 or R_2 starting structures, except for Phe^{B25} (which agreed better with crystal structures when starting from the R_2 dimer).

3-I-Tyr^{B26} Trajectories—Each subunit exhibited similar features with optimal positioning of the iodine atom requiring a change in B26 side-chain conformation ($\Delta\chi_1 = 10(\pm3°)$ and $\Delta\chi_2 = 5(\pm8°)$; supplemental Figs. S1 and S2). Despite this change, the aromatic face of 3-I-Tyr^{B26} adjoined the side chain of Val^{B12} as in WT insulin (supplemental Fig. S3*A*). MTP-based MD simulations established that the 3-I-Tyr^{B26} side chain could pack within a hydrophobic cavity formed by the side chains of Ile^{A2}, Val^{A3}, Leu^{B11}, and Val^{B12} (and the main chain of Gly^{B8}) within the same protomer (supplemental Fig. S3*B*). The predicted environment of 3-I-Tyr^{B26} and its range of conformational excursions were similar in R_2- and T_2-based MD simulations.

The modified aromatic ring of 3-I-Tyr^{B26} represents a perturbation within an anti-parallel dimer-related β-sheet. The subtle changes in the side-chain dihedral angles of 3-I-Tyr^{B26} and 3-I-Tyr$^{B26'}$ observed in the crystal structure of 3-[iodo-Tyr^{B26},Nle^{B29}]insulin (see below; relative to Tyr^{B26} and Tyr$^{B26'}$ in WT R_6 hexamers) do not affect this sheet; its four dimer-related hydrogen bonds exhibit essentially native lengths and angles. Packing at the dimerization interface for the WT dimer (supplemental Fig. S4*A*) and comparison with the predicted 3-[iodo-Tyr^{B26}]insulin dimer show that the iodine within each protomeric core within the R_6 hexamer would (*a*) provide an overall nonpolar environment and (*b*) enable formation of a novel and favorable electrostatic interaction (supplemental Fig. S4*B*) between the *para*-OH of Tyr^{A19} and the equatorial belt surrounding the halogen (Fig. 2*B*). No halogen bonds were observed to the σ-hole of the iodine (supplemental Fig. S4*B*).

5-I-Tyr^{B26} Trajectories—MTP simulations of the alternative 5-I-Tyr^{B26} ring conformation suggested that it could achieve more compact packing at the dimer interface than was observed for WT or 3-I-Tyr^{B26} simulations (supplemental Fig. S3*C*). The simulations further verified the implications of the MTP energy maps that unfavorable steric and electrostatic interactions across the dimer interface render unfavorable 5-I-posi-

TABLE 1
Average r.m.s.d. across dimer interface

Average r.m.s.d. across the WT-T_2, 3I-Tyr^{B26} T_2, and R_2 dimer interface from 20 ns of MD simulation and from the three dimers (D1, D2, and D3) taken from the 3-I-Tyr^{B26} R_6 crystal structure, with respect to WT (PDB code 1DPH). All structures are aligned with respect to the backbone atoms of residues B24–B28 and B24′–B28′. Values reported are r.m.s.d. of the side chains of residues B24–B26 and B24′–B26′.

Comparison	r.m.s.d. (Å)
WT_{1DPH}/WT (20 ns MD)	0.3
WT_{1DPH}/T_2 3-I-Tyr^{B26} (20 ns MD)	0.9
WT_{1DPH}/R_2 3-I-Tyr^{B26} (20 ns MD)	1.0
WT_{1DPH}/D1 of R_6 3-I-Tyr^{B26} X-ray	0.7
WT_{1DPH}/D2 of R_6 3-I-Tyr^{B26} X-ray	0.7
WT_{1DPH}/D3 of $R_6$3-I-Tyr^{B26} X-ray	0.7

TABLE 2
X-ray data processing and refinement statistics

Wavelength (Å)	1.5478
Resolution range (Å)	40.85–2.30 (2.40–2.30)[a]
Space group	$P2_1$
a (Å), b (Å), c (Å), β (°)	46.43, 61.63, 58.58, 111.38
Redundancy	4.76 (2.55)
Completeness (%)	95.6 (80.4)
R_{merge}	0.054 (0.227)
$\langle I/\sigma(I)\rangle$	18.2 (3.9)
$CC_{1/2}$[b]	0.999 (0.934)
Refinement	
Resolution range (Å)	40.85–2.30
No. of reflections	13,255
R_{work}/R_{free}[c]	0.163/0.228
No. of protein atoms	2305
No. of non-protein atoms	121
$\langle B_{iso}\rangle$ protein atoms (Å^2)	41.6
$\langle B_{iso}\rangle$ non-protein atoms (Å^2)	34.0
σ_{bonds} (Å)/σ_{angles} (°)	0.008/1.12
Ramachandran plot	
Favored (%)	100
Outliers (%)	0

[a] Numbers in parentheses refer to the outer resolution shell.
[b] Pearson correlation coefficient between merged intensities of two random halves of the diffraction data set (112).
[c] Free set contained 10% of total observed reflections.

tioning of the iodine atom. The clash between the 5-iodine and the dimer-related side chain of Tyr^{B16} could not be relieved by displacement of B16 because it itself is constrained by Val$^{B12'}$ and Phe^{B25} (supplemental Fig. S3*C*). Although a clash between the 5-iodine and the dimer-related carbonyl oxygen of Gly^{B20} could readily be relieved by a change in B20 main-chain dihedral angles (supplemental Fig. S5), the resulting B20 conformations (with negative ϕ angles in the Ramachandran plane) are less favorable (60). Previous studies of synthetic insulin analogs have shown that reinforcement of the native positive ϕ angle by D-Ala^{B23} stabilizes insulin, whereas perturbation of the B23 conformation by L-Ala or L-Val (the latter associated with human diabetes) is associated with structural frustration and misfolding (60, 61).

The robustness of the MTP-MD simulations was probed through average r.m.s.d. on pairwise comparisons of the conformations sampled in the course of 20-ns trajectories. The values given in Table 1 pertain to side chains at the dimer interface (residues B24, B25, and B26) following alignment based on the main-chain atoms of residues B24–B28 and B24′–B28′. Baseline r.m.s.d. values in the WT T_2 dimer (PDB code 1DPH) were calculated following a control MD simulation (*i.e.* in the absence of a modified B26). For residues in the WT anti-parallel β-sheet (residues B24–B28 and their dimer-related mates), the average main-chain r.m.s.d. was 0.1 Å, and the average side-chain r.m.s.d. was 0.3 Å relative to starting structure (Table 1). Corresponding values for the modified T_2 dimer were 0.3 and 0.9 Å, also relative to the WT crystal structure; similar values were obtained on comparison of the modified R_2 dimer. The larger r.m.s.d. values in the modified dimers reflected consistent conformational adjustments required to accommodate core packing of the iodine atom in the 3-I conformation.

Crystal Structure Verifies Essential Features of MTP Modeling

3-[iodo-Tyr^{B26},Nle^{B29}]Insulin exhibited an affinity (K_d) for the lectin-purified receptor (isoform IR-B) of 21 ± 5 pM under conditions in which WT insulin exhibits an affinity of 62(±8) pM. The analog's stability (ΔG_u = 3.4(±0.1) kcal/mol at 4 °C) was indistinguishable from that of [Nle^{B29}]insulin as probed by guanidine denaturation (32).

The analog was crystallized under conditions that ordinarily facilitate crystallization of WT insulin as a phenol-stabilized R_6 hexamer (23). A monoclinic lattice was observed in which one R_6 hexamer defined the asymmetric unit (for refinement statistics, see Table 2). In this crystal form, each protomer in the hexamer is crystallographically independent (and so may in principle exhibit subtle structural differences). A ribbon model (Fig. 4*A*) highlights the positions of the iodine atoms (*green spheres*) relative to the six R state-specific B1-B19 α-helices (*green*) and A chains (*black*). We first describe the structure and then compare it to the predicted models.

Crystal Structure of 3-[iodo-Tyr^{B26},Nle^{B29}]Insulin Resembles the WT Hormone

No significant differences were observed between the modified hexamer and the corresponding WT R_6 hexamer with respect to secondary structure, chain orientation, mode of assembly, or structures of the Zn^{2+}- and phenol-binding sites. The six independent R state protomers exhibited essentially identical conformations (average pairwise main-chain r.m.s.d. of 0.42 Å and average side-chain r.m.s.d. of 1.43 Å). Tetrahedral coordination of the two axial zinc ions (overlying *red spheres* at center in Fig. 4*A*) by the side chains of His^{B10} (three per R_3 trimer; *light gray* side chains) is essentially identical to that in WT R_6 hexamers (62); the fourth coordination sites contain a presumed chloride anion. In each protomer, the B26 iodine atom is positioned within the α-helical core in accordance with the 3-I-conformational isomer. No excess electron density was observed at ring position 5, providing evidence of a single predominant conformation.

Superposition of a representative analog protomer and WT protomer (*dark* and *light gray ribbons,* respectively, in Fig. 4*B*) yielded the following average pairwise differences between a representative protomer of 3-[iodo-Tyr^{B26},Nle^{B29}]Insulin and the WT R state: main-chain r.m.s.d. = 0.55(±0.08) Å and side-chain r.m.s.d. = 1.94(±0.23) Å. These values are similar to those observed among a collection of independent WT R state protomers[13] (main-chain r.m.s.d. 0.68(±0.26) Å; side-chain r.m.s.d. 1.14(±0.34) Å). Within the crystal structure of 3-[iodo-

[13] WT R state coordinates were obtained from PDB entries 1TRZ, 1TYL, 1TYM, 1RWE, 1DPH, and 1ZNJ.

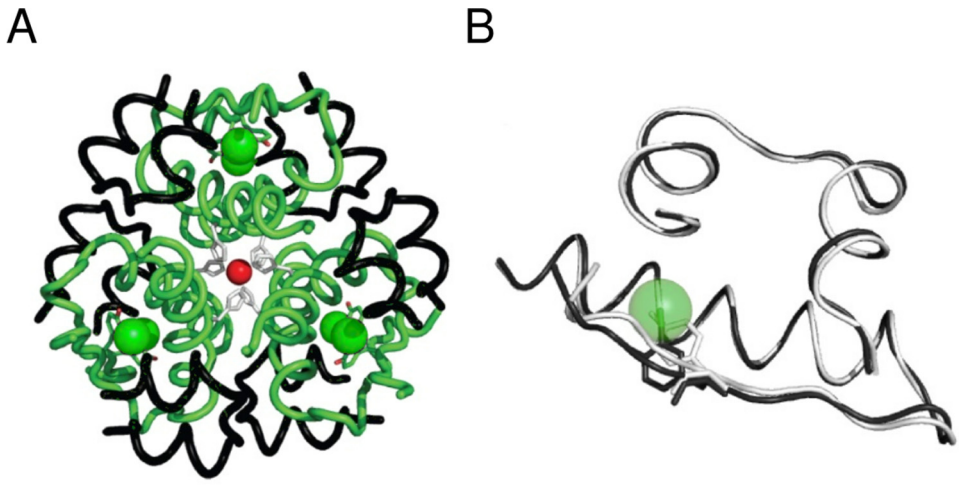

FIGURE 4. **Crystal structure of 3-[iodo-Tyr^{B26},Nle^{B29}]Insulin.** *A,* R_6 hexamer with A and B chains (*black* and *green ribbons*, respectively). Iodine atoms (*green spheres*) and the two axial zinc ions (*red spheres*) are aligned at *center*, each coordinated by 3-fold-related His^{B10} side chains (*light gray*). *B,* superposition of WT protomer (*light gray*) and 3-I-Tyr^{B26} analog (*dark gray*). Side chains of Tyr^{B26} and 3-I-Tyr^{B26} are shown as *sticks*. For clarity, the iodine atom is shown as a transparent *sphere*; Nle^{B29} is not shown.

Tyr^{B26},Nle^{B29}]Insulin, no polypeptide-like ($2F_{obs} - F_{calc}$) electron density (continuous at $>1\sigma$) was observed C-terminal to B28 in any of the modified B chains.

The similarity of the variant and WT structures suggests that the essential features required for dimer formation are not altered by the asymmetric distribution of partial charges in the aromatic ring of 3-I-Tyr^{B26} and its associated pattern of aromatic-aromatic interactions (supplemental Fig. S4*C*). The six independent side chains of 3-I-Tyr^{B26} nonetheless exhibit consistent differences in conformation relative to WT Tyr^{B26} (supplemental Fig. S4*D*). The modified side chain (*dark gray* in supplemental Fig. S4*D*) is rotated by ~12° about the C_α–C_β bond with respect to its WT counterpart (*light gray* in supplemental Fig. S4 *D*). As predicted by the MTP-based calculations, this rotation positions the iodo-group within a non-polar pocket formed by the side chains of residues Ile^{A2}, Val^{A3}, Leu^{B11}, and Val^{B12}, residues conserved among vertebrate insulins and essential for biological activity (2). In the WT structure, the pocket is occupied by the phenolic hydroxyl group of Tyr^{B26}, although its packing within the pocket is less intimate than that of the iodine atom of 3-I-Tyr^{B26}. The consequent displacement of the *para*-hydroxyl group of 3-I-Tyr^{B26} from the pocket results in its greater solvent exposure. Side-chain dihedral angles of three aromatic side chains at or near the dimer interface (B16, B24, and B26) are given in Table 3 in relation to a reference WT R_6 structure.

As expected, the 3-[iodo-Tyr^{B26},Nle^{B29}]Insulin hexamer contains six bound phenol molecules, located at an interface between dimers as in the WT R_6 hexamer (23). The phenol-binding sites are essentially identical. The electron density ($2F_{obs} - F_{calc}$) associated with one such phenol is shown in relation to a superposition of variant and WT structures (*dark* and *light gray* in Fig. 5, *A* and *B*). In each case, a characteristic pair of hydrogen bonds from the phenolic –OH group engages the main-chain carbonyl oxygen (acceptor) and amide group (donor) of Cys^{A6} and Cys^{A11}, respectively. A corresponding depiction of B26 side chain environments highlights the asymmetry in density between the 3- and 5-ring positions (Fig. 5, *C* and *D*).

TABLE 3

Side-chain dihedral angles of aromatic side chains near dimer interface

Phe^{B25} has been excluded due to poor side-chain density found in the crystal structure.

	3-I-Tyr^{B26}		WT insulin[a]	
Residue	χ_1 (°)	χ_2 (°)	χ_1 (°)	χ_2 (°)
Tyr^{B16}	172.4	78.8	174.7	77.6
	176.9	83.1	179.6	80.8
	175.4	82.9	172.9	84.3
	174.5	81.1	177.6	84.6
	177.0	79.2	175.9	75.5
	178.0	79.6	175.3	66.2
Phe^{B24}	63.2	−87.6	62.5	89.1
	69.9	88.4	69.3	83.2
	60.0	−85.3	59.3	83.4
	63.0	−86.9	55.2	−85.0
	60.1	−89.0	56.7	−87.9
	61.0	−87.5	68.8	87.5
Tyr^{B26}	167.8	74.1	−173.9	82.7
	166.6	81.0	168.9	−90.5
	161.4	72.6	−176.7	69.8
	167.7	75.3	176.5	74.7
	168.1	75.7	−173.9	80.1
	165.3	78.0	175.4	72.4

[a] Molecular coordinates were obtained from PDB code 1ZNJ.

Accommodation of the modified side chain does not alter the canonical hydrogen-bonding pattern of the dimer-related antiparallel β-strands (B24–B26 and B26′–B24′ segments) (Fig. 6*A*). Although differences in side-chain conformation (relative to WT) were observed at B25, its side-chain density was poor, suggesting dynamic disorder (Fig. 6*B*). We speculate that these differences reflect slight alterations in backbone geometry; variation in the crystallization milieu cannot be excluded.

Comparison of Observed and Predicted Conformations of 3-I-Tyr^{B26}

The average r.m.s.d. across the dimer interface was calculated for the three 3-I-Tyr^{B26} insulin R_2 dimers with respect to WT R_2 dimers (PDB code 1ZNJ; Table 1). The deviation in the 3-I-Tyr^{B26} R_2 dimers (0.7 and 0.9 Å) arises from the packing of iodine in the hydrophobic pocket. The average r.m.s.d. across the dimer interface was also calculated with respect to the crys-

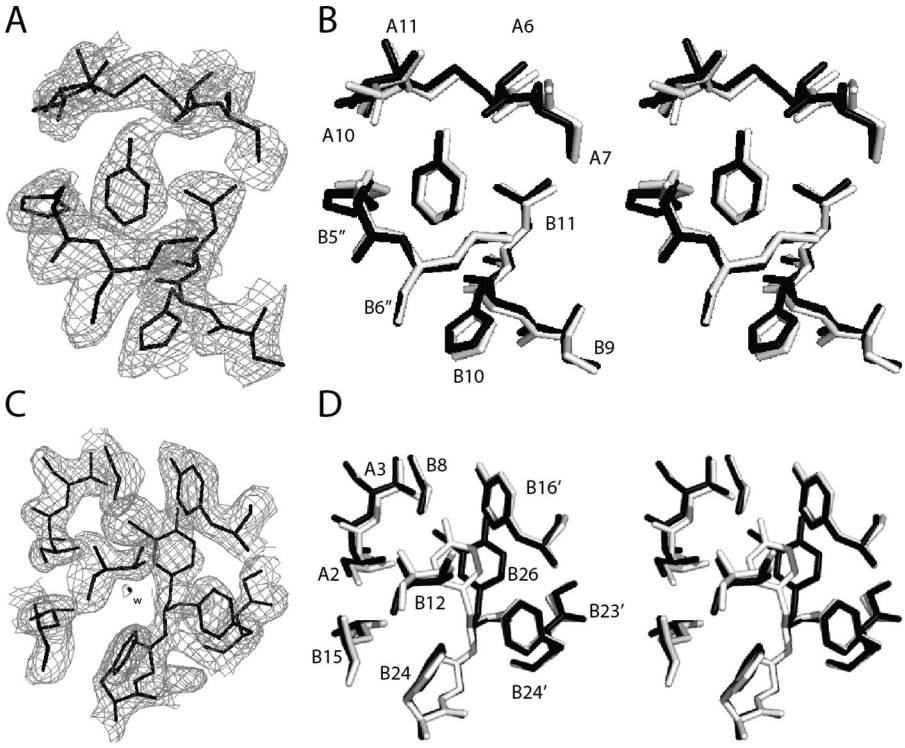

FIGURE 5. **Crystallographic features of R_6 3-[iodo-Tyr^{B26}]insulin hexamer.** *A,* ($2F_{obs} - F_{calc}$) difference σ_A-weighted electron density contoured at the 1σ level of a representative bound phenol molecule. Its *para*-OH group participates in hydrogen bonding with the carbonyl oxygen of Cys^{A6} and amide proton of Cys^{A11} (cystine A6–A11). An edge-to-face interaction occurs with the imidazole ring of His^{B5} from another dimer. *B,* stereo view as in *A* aligning the structure of the analog (*dark gray*) with that of WT insulin as an R_6 hexamer (*light gray*). *C,* electron density of 3-I-Tyr^{B26} and surrounding residues. *D,* stereo view of residues seen in *C* (*stick* representation) superposed as in *B*. WT coordinates for *B* and *D* were obtained from PDB code 1ZNJ.

tal structure of the WT T_2 dimer (PDB code 1DPH). Because in the T state Gly^{B8} lies within a β-turn (with positive ϕ angle) whereas the R state has B8 in an α-helix (negative φ angle), the T_2-related r.m.s.d. was calculated only for backbone atoms of the β-sheets at the dimer interface (β/β').

The χ_1 and χ_2 dihedral angles of selected side chains in the three dimers in the crystallographic hexamer were compared with the predicted dimer dihedral angle distribution from WT- and 3-[iodo-Tyr^{B26}]insulin T_2 dimers (*black stars* for dimer 1, *black squares* for dimer 2, and *black circles* for dimer 3; see supplemental Fig. S2). In the crystallographic hexamer, the aromatic residues locally re-organize such that B24 and B24′ always have opposite χ_2 angles (this is observed for all the dimers in the crystal); in contrast, the B25 side chains are disordered. Simulations, 20 ns in length, based on the R_2 structure (PDB code 1ZNJ) established that the R_2 dimer samples all experimentally observed states. As a control to test whether such consistency would be affected by the TR transition, corresponding trajectories based on a crystallographic T_2 dimer (PDB code 1DPH) were undertaken (supplemental Fig. S2). Although sampling of B25 side-chain conformations (and to some extent that of Phe^{B24}) differed in the T_2 trajectory (supplemental Fig. S2), the results were in good overall agreement with the R_2-based simulations (supplemental Fig. S1).

Molecular Modeling of 3-[iodo-Tyr^{B26}]Insulin at the μIR Interface

How might iodo-Tyr^{B26} enhance the binding of insulin to the receptor? The environment (and conformation) of iodo-Tyr^{B26} in the variant hormone-IR complex is likely to differ from its internal environment in the modified zinc hexamer, given that in the co-crystal structure of the WT μIR complex the B23–B27 segment is displaced from its location in the free hormone (Fig. 1*C*) (39). Such displacement permits the aromatic rings of Phe^{B24}, Phe^{B25}, and Tyr^{B26} to contact the receptor (Fig. 7). The open receptor-bound conformation of the hormone is thus predicted to expose the side chain of 3-I-Tyr^{B26} and in particular enable its engagement at the L1-αCT surface.

We hypothesized that the enhanced affinity of this and related iodo-Tyr^{B26} analogs (7, 63, 64) might be due to a novel interaction between the halogen atom and the IR. MD simulation of 3-I-Tyr^{B26} and 5-I-Tyr^{B26} at the μIR interface (using PDB code 4OGA as starting structure) supported the plausibility of this model (Fig. 7). An iodine at either the 3- or 5-positions (Fig. 7) of Tyr^{B26} could readily be accommodated within the receptor-binding cleft (*i.e.* with the iodo group directed either toward or away from the μIR interface). Whereas the 5-I-Tyr^{B26} conformation did not appear to offer new favorable interactions, our modeling revealed that a 3-iodo-substituent could participate in three novel contacts (Fig. 8*A*) as follows: (i) a halogen bond between its δ^+ region and the backbone oxygen atom of Val-712, and (ii and iii) favorable electrostatic interactions between its δ^- equatorial belt (Fig. 2*B*) and one hydrogen from the ϵ-NH_2 of Gln-34 and one hydrogen from the ϵ-NH_2 of Arg-14. Details are as follows.

The predicted halogen bond at the variant μIR interface represents a "gain of function" by a nonstandard side chain (65). In

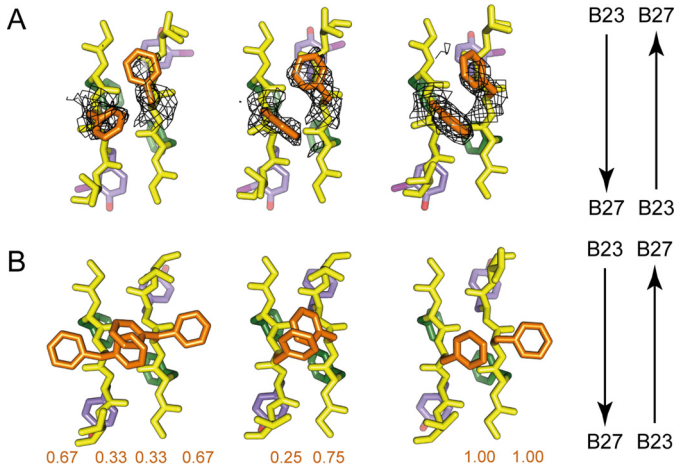

FIGURE 6. **Side-chain arrangements within dimer interfaces.** Residues B23–B26 are shown within respective crystal structures of the 3-[iodo-Tyr^{B26},Nle^{B29}]insulin hexamer (*A*) and WT insulin R_6 zinc hexamer (PDB code 1ZNJ) (*B*). The three subpanels within *A* and *B* correspond to the respective three copies of the dimer interface within the crystallographic asymmetric units of the two structures. Within each subpanel, the side-chain carbon atoms of Phe^{B24} and its non-crystallographic symmetry equivalents are shown in *green*, of Phe^{B25} and its non-crystallographic symmetry equivalents in *orange*, and of 3-I-Tyr^{B26} or Tyr^{B26} and their non-crystallographic symmetry equivalents in *light purple*, whereas all backbone atoms are in *yellow*, as are the side-chain atoms of Thr^{B27} and its non-crystallographic symmetry equivalents. The *arrows* on the *right* assist in identifying the direction of the respective polypeptides within each subpanel. Chains within each subpanel correspond (from *left* to *right*) to chains B, D, F, H, J, and L (respectively) within each structure. Overlaid on the three subpanels in *A* is σ_A-weighted ($2F_{obs} - F_{calc}$) difference electron density contoured at the 0.75 σ level and masked to within 2.5 Å of the side-chain atoms of Phe^{B25} and its symmetry-related equivalents. The values displayed under the respective chains within the subpanels of *B* correspond to the side-chain occupancies of the Phe^{B25} and its respective non-crystallographic symmetry equivalents within PDB code 1ZNJ. The side chain of Nle^{B29} is not shown.

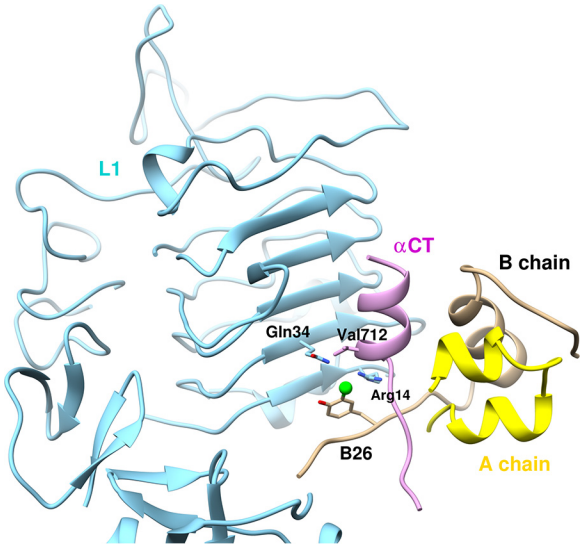

FIGURE 7. **Homology model of μIR/insulin interface.** Docked structure of 3-[iodo-Tyr^{B26}]insulin bound to L1 (Arg-14 and Gln-34) and αCT (Val-712) of the μIR. 5-[iodo-Tyr^{B26}]insulin has the iodine away from the interface (through rotation around the C_β–C_γ axis (χ_2 180°); see Fig. 2*A*) and is expected to interact less favorably with the μIR.

the predicted halogen bond, the average distance between the iodine and the carbonyl oxygen of Val-712 is 3.6 Å, which is less than the sum of their van der Waals radii (4.1 Å) (Fig. 8*B*, *upper panel*). However, the estimated σ-hole size of iodine bound to a phenol, and so *ortho* to its hydroxyl group, is smaller due to the latter's electron-donating properties (see Fig. 2*B*). The σ-hole size is represented by angle β (Fig. 2*C*), typically 148° for iodine bound to phenyl; the directionality of the halogen bond is along the C–I bond axis (55). For 2-iodophenol, this angle is smaller (~106°), and the halogen-bond direction shifted from the C–I bond by ~30° (Fig. 2*C*). Thus, the σ-hole bond angle θ is within the δ^+ region that ranges from 127° $\langle\theta_{C\text{-}I\cdot O}\rangle$ to 233° (Fig. 2*C*). The σ-hole bond angle distribution $\theta_{C\text{-}I\cdot O}$, of the iodine atom (I) with the backbone O of Val-162 from 1-ns MD, ranges from ~127 to 170° and peaks at 145°, whereas the $\theta_{C\text{-}I\cdot H}$ distributions are within the δ^- region (Fig. 8*C*). The I···O distance of 3.6 Å and the C-I···O angle of ~145° favor formation of a strong halogen bond in the complex as suggested by previous quantum-chemical calculations (66) and MD studies of other systems (55, 67, 68).

In essence, our modeling suggested that 3-I-Tyr^{B26} leads to an increased number of local interactions at the μIR interface with retention of native contacts. All-atom simulations thus rationalized the increased affinity of such insulin analogs for the intact receptor. As a further control to test whether such increased affinity is electrostatic or nonpolar (van der Waals) in nature, we undertook additional simulations with an atom-centered PC force field and with a simplified force field in which the iodine only engaged in van der Waals interactions. In contrast to the MTP electrostatic model (*upper panel* in Fig. 8*B*), the PC model exhibited only one of the above three contacts: that to Gln-34 (*lower panel*). Moreover, a "neutral" iodine ($q = 0$) led to detachment of Tyr^{B26} from the μIR surface after 150 ps (Fig. 9*A*). These control simulations suggest that the increased receptor-binding affinity of 3-[iodo-Tyr^{B26}]insulin is driven by electrostatics at the level of quantum chemistry rather than a consequence of the hydrophobicity of iodine (Fig. 9*B*). The potential role of water molecules at the modified interface is discussed below.

Discussion

This study has focused on position B26, broadly conserved as Tyr among vertebrate insulins and as Phe among insulin-like growth factors (69). Although non-polar and charged side chains at B26 are compatible with high affinity for the IR, such insulin analogs are unstable and prone to fibrillation (32). Of the natural amino acids at B26, Tyr thus appears to offer the best combination of activity and stability. How might the expanded chemical space of unnatural mutagenesis (70) be exploited to enhance the biophysical properties of an active insulin molecule? We approached this question in three parts. We first simulated the structure of [iodo-Tyr^{B26}]insulin (as a dimer) to predict how the iodine atoms could be accommodated. This simulation, critically dependent on the quantum-chemical properties of an iodo-aromatic group, suggested that the iodine might enhance (rather than perturb) native packing interactions within the core of the free hormone. We next verified the predicted conformational preference for the 3-I-Tyr^{B26} state through crystallographic studies. Finally, we investigated possible mechanisms by which 3-I-Tyr^{B26} enhances IR binding (7–9). Such enhancement posed a seeming paradox as a modification that "closes" the *free* conformation of insulin

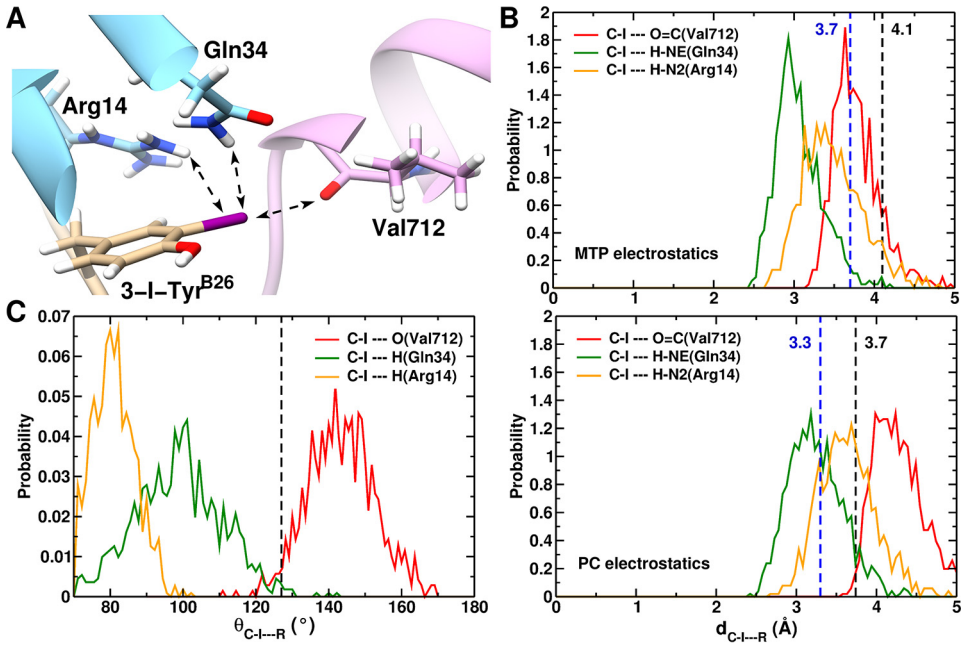

FIGURE 8. **MD-based model of μIR/3-[iodo-Tyr^{B26}]insulin interface.** *A,* structure of 3-[iodo-Tyr^{B26}]insulin bound to the μIR. Only μIR residues interacting with 3-I-Tyr^{B26} are illustrated. Potential hydrogen/halogen bonds with iodine are shown as *dashed arrows*. *B,* probability distribution along the C-I···R distance, where R is (O=C(Val-365)) (*red line*); R is (H-Nϵ(Gln-81)) (*green line*); or R is (N2(Arg-61)) (*dashed orange line*). The *upper panel* is from simulations with MTP electrostatics, whereas the *lower panel* uses point charges. The *black dashed lines* at 4.1 Å (3.7 Å, *lower panel*) represents the C-I···O(Val-365) interaction limit using optimized van der Waals radii for the iodine and oxygen atoms. *Dashed lines* at 3.7 Å (3.3 Å, *lower panel*) indicate the C-I···H (Gln-34 and Arg-14) distance using optimized van der Waals radii for iodine and polar H-atoms. *C,* probability distribution of the halogen/hydrogen bond angular variation $\theta_{C\text{-}R}$ from 1 ns of MD simulation. The *black dashed line* at 127° represents the boundary between the negative ($\delta^- < 127°$) and positive electrostatic region ($127° \langle \delta^+ \rangle 233°$) for I.

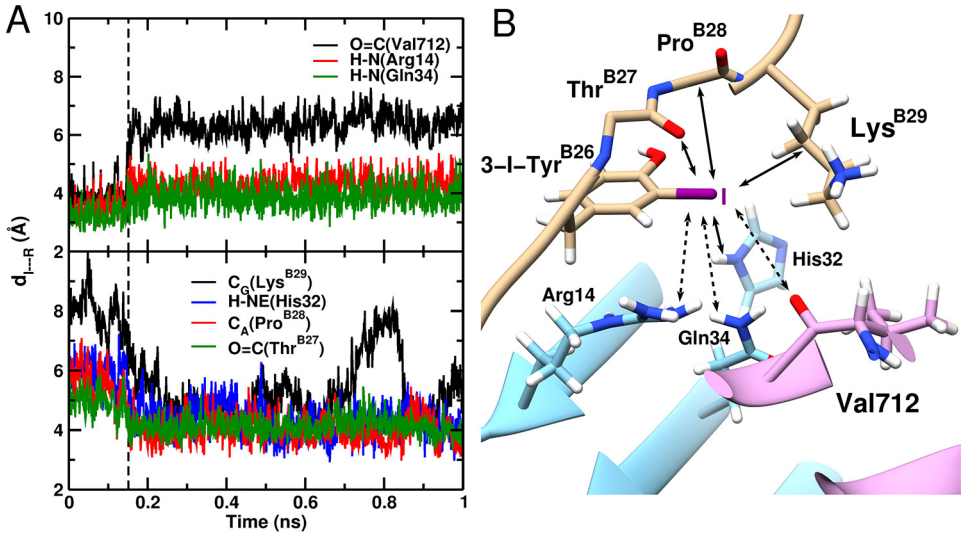

FIGURE 9. **MD-based model of μIR/3-[iodo-Tyr^{B26}]insulin interface assuming neutral iodine.** *A,* time evolution of the I···R distance (distance of Tyr^{B26}-I to the interacting insulin/μIR residues). The *upper* and *lower panels* show the increase of the I···O=C(Val-365), I···H-N(Arg-61), and I···H-N(Gln-81) bond lengths, respectively, and the decrease of the I···C$_\gamma$(Lys^{B29}), I···H-NE(His-79), I···C$_\alpha$(Pro^{B28}), and I···O=C(Thr^{B27}) bond lengths in the course of the MD simulation. The *black dashed line* at 150 ps represents the point when the electrostatically driven interactions dissociate and the van der Waal-driven interactions form. *B,* snapshot structure of 3-[iodo-Tyr^{B26}]insulin bound to the μIR. Only the residues interacting with 3-I-Tyr^{B26} are illustrated. Bond formation/dissociation with the iodine atom are shown as *full* and *dashed line arrows*, respectively.

(45) might have been expected to impair its ability to open on receptor binding (39, 71).

The structure and properties of 3-[iodo-Tyr^{B26}]insulin reflect general physico-chemical principles. Packing of an iodo-aromatic modification within the core of a globular protein in principle reflects its overall hydrophobicity (72) and stereo-electronic properties (18). Indeed, modification of one edge of an aromatic system both introduces a unique local electronic distribution (*i.e.* at the halogen) and alters the overall electron density of the π-electronic cloud, including at the opposite edge. These features are exemplified by crystal structures of thyroid hormone bound to its nuclear receptor (73) or carrier proteins (74). Because thyroid hormone may have evolved from an ancestral iodo-Tyr (as in its route of biosynthesis (75)), we may regard iodinated derivatives of insulin as models for iodo-aromatic chemistries related to this evolutionary innovation.

The present structure and MD simulations highlight that aromatic-rich protein interfaces are not classical ball-and-stick objects.

Modification of Residue B26 Enhances Packing Efficiency

3-I-Tyr^{B26} represents an apparent perturbation within an anti-parallel dimer-related β-sheet. The subtle changes in the side-chain dihedral angles of 3-I-Tyr^{B26} and 3-I-Tyr$^{B26'}$ observed in the crystal structure of 3-[iodo-Tyr^{B26},Nle^{B29}] Insulin (relative to Tyr^{B26} and Tyr$^{B26'}$ in WT R_6 hexamers) do not affect this sheet; its four dimer-related hydrogen bonds exhibit essentially native lengths and angles. Side-chain packing schemes in the WT insulin dimer within the R_6 hexamer (Fig. 4*A*) and its comparison with the 3-[iodo-Tyr^{B26}]insulin dimer demonstrate that the iodine atoms both (*a*) reside within an overall nonpolar environment within each protomeric core, and (*b*) enable formation of a novel and favorable electrostatic interaction (supplemental Fig. S4*B*), *i.e.* between the *para*-OH of Tyr^{A19} and the equatorial belt surrounding the halogen (Fig. 2*B*). The latter contact is analogous to weakly polar interactions between hydrogen-bond donors (such as the carboxamide NH_2 of Asn or Gln) and the planar π cloud of aromatic rings (44). No halogen bonds were observed (supplemental Fig. S4, *B* and *D*).

Despite the larger size of iodine relative to hydrogen, only a slight re-arrangement occurs within the modified hormone core. Indeed, the internal location of the iodine atom in the crystal structure of 3-[iodo-Tyr^{B26},Nle^{B29}]Insulin highlights a potential packing "defect" in WT insulin. Such gaps are in general widespread among crystal structures of globular proteins, as reflected by mean side-chain packing efficiencies of 70–80% (76). Perfect packing efficiency is not attainable given the distinct sizes, shapes, and preferred dihedral angles of amino acids. The existence and ubiquity of such cavities was highlighted in seminal studies of xenon-saturated cores (77). The packing of 3-I-Tyr^{B26} near the side chains of A2, A3, and A19 in the present structure is thus reminiscent of the accommodation of large xenon atoms within the core of myoglobin (77).

3-I-Tyr^{B26} Modulates Aromatic-Aromatic Interactions

The dimer interface of insulin exhibits successive aromatic-aromatic interactions involving eight residues: Tyr^{B16}, Phe^{B24}, Phe^{B25}, Tyr^{B26}, and their dimer-related mates (2). Whereas the two B25 side chains lie at the periphery of this interface (and exhibit alternative or multiple conformations among WT crystal structures (2)), the remaining six aromatic rings conform to favorable pairwise edge-to-face orientations (44). Respective aromatic-aromatic interactions at the dimer interface of WT insulin and 3-[iodo-Tyr^{B26},Nle^{B29}]Insulin are shown in Fig. 10*B*. Inter-ring centroid distances and orientations are defined as illustrated in Fig. 10*A* (44). In a representative WT R_6 structure (determined at a resolution of 2.0 Å with one hexamer in the asymmetric unit such that the three dimer interfaces are crystallographically independent; PDB code 1ZNJ), the dimer interface contains the following structural relationships.

Dimer-related Neighbors—The nearest pairs of rings in the WT R_6 hexamer are across the dimer interface as follows: B16–B26′ and B26–B16′ (5.8/5.7, 5.8/5.5, and 5.7/5.5 Å in respective interfaces BD, FH, and JL as defined in the PDB); B24–B24′ (5.8, 5.7, and 5.9 Å), B24–B26′ and B26–B24′ (6.2/6.2, 5.9/5.8, and 5.8/6.1 Å). Detailed differences pertaining to B16–B26′/B16′–B26 and to B24–B26′/B24′–B26 distances reflect subtle asymmetries at each interface. Structural relationships between B16–B26′ and between B16′–B26 closely conform, in each of these six pairs, to the canonical edge-to-face packing of benzene rings (44) with mean polar angles θ = 136(±4) and ϕ = 49(±4)° and mean inter-plane dihedral angle ψ = 136(±4)° using polar coordinates as defined in Fig. 10*A*. The three B24–B24′ pairs exhibit displaced edge-to-face packing.

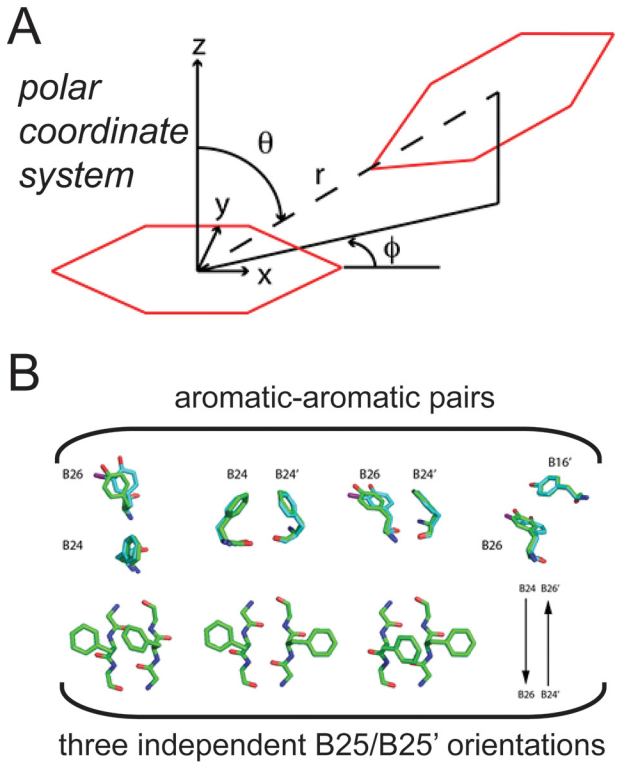

FIGURE 10. **Aromatic-aromatic interactions.** *A*, axes and definition of polar coordinates (r, ϕ, and θ) as originally defined by Burley and Petsko (44). Ψ provides the dihedral angle between the two planes formed by each of the aromatic rings. The two interacting aromatic rings are shown in *red*. *B*, interacting pairs of aromatic rings at the dimer interface of WT insulin and the 3-I-Tyr^{B26} analog. *Upper panel*, Phe^{B24}/Tyr^{B26}, Phe^{B24}/Phe$^{B24'}$, Phe^{B24}/Tyr^{B26}, and Tyr^{B26}/Tyr$^{B16'}$; primed residue numbers indicate the dimer-related residue A representative WT structure (*green*) is overlaid in comparison with the side chains of the 3-[iodo-Tyr^{B26},Nle^{B29}]insulin structure (*cyan*). *Lower panel*, Phe^{B25}/Phe^{B25} interaction pair and its three possible conformations. Images of representative Phe^{B25} side chains from the crystal structure of 3-I-Tyr^{B26}; the side chains of [Nle^{B29}]insulin are not shown due to dynamic disorder. WT coordinates were obtained from PDB code 1ZNJ.

Intra-chain Relationships—Within each WT B chain, the side chains of Phe^{B24} and Tyr^{B26} project from the same side of a β-strand but (due to their respective χ_1 and χ_2 dihedral angles) exhibit edge-to-face interactions rather than π stacking (Fig. 10*B*). Their mean centroid distance is (r) 7.5(±0.1) Å with average θ values of 116(±2) and ϕ = 54(±3)° and with inter-plane dihedral angle 62(±8)°. Corresponding intra-chain centroid distances between Tyr^{B16} and Tyr^{B26} are more distant (r = 13.4 (±0.1) Å), beyond the range of a favorable weakly polar interaction (44).

Alternative Occupancies of Phe^{B25}—Although the side chains of Phe^{B25} in the WT crystal form can adopt two conformations, one mode corresponds to displaced π stacking (B25–B25′) with a centroid distance of 5.6 Å. This mode represents a distinct motif of aromatic-aromatic interaction from a quantum-chemical perspective (44). It is also possible for the two B25/B25′ aromatic rings to each point inward (*i.e.* toward Tyr^{A19} in its own protomer), attenuating their aromatic-aromatic interaction.

The crystal structure of 3-[iodo-Tyr^{B26},Nle^{B29}]Insulin exhibits only subtle differences relative to the above. The altered χ_1 and χ_2 dihedral angles of 3-I-Tyr^{B26} and 3-I-Tyr$^{B26'}$ (χ_1 = 166 (±2.5) and χ_2 = 76(±3)°) enable packing of the iodine atom in a gap bounded by conserved side chains in both chains (Ile^{A2}, Val^{A3}, Tyr^{A19}, Leu^{B11}, Val^{B12}, Leu^{B7}, and cystine A7–B7; see supplemental Fig. S4*D*). Respective side-chain conformations of residues B26 and B26′ across each of the three (crystallographically independent) dimer interfaces are similar but not identical (Table 3). Repositioning of the B26/B26′ side chains is in turn associated with subtle changes in geometric parameters describing B26–B16′ (and likewise B26′–B16), B26–B24′ (also B26′–B24), and B26–B24 (Fig. 10*B* and Table 3). Of these, the most distinct pairwise orientations (relative to the WT reference structure) are exhibited by B26–B16′ and B24–B26′. These in turn necessitate a small change in the relative B24–B24′ positions.

Of future interest would be comparison of iodo-Tyr^{B26} with iodo-Phe^{B26} as a test of the predicted effect of the former's *para*-OH group on the position and intensity of iodine's σ-hole. Comparison of such modifications with smaller halo-aromatic substitutions would likewise enable local effects of the halogen's electronic distribution to be distinguished from general long range effects on the overall dipole moment of the modified aromatic system.

Application of MTP Modeling to the Variant Hormone-Receptor Complex

Our efforts to apply MM methods to the modified insulin dimer highlighted the importance of accurate representations of the iodo-aromatic ring and its interactions. Indeed, a PC-based model of the iodo-Tyr predicted (incorrectly) an outward orientation of the iodine atom (5-I-Tyr^{B26}), whereas an MTP-based model favored the observed inward orientation (3-I-Tyr^{B26}). These findings are in accordance with previous simulations of small molecules wherein halogenated compounds required an MTP treatment, although electronically less demanding building blocks (such as *N*-methylacetamide) did not (67, 78). Experimental verification of MTP-based predictions in the case of the insulin hexamer encouraged us to simulate the possible function of iodo-Tyr^{B26} at the surface of the μIR complex (39, 71).

In the WT insulin-μIR complex, invariant L1 residues Arg-14 and Gln-34 are of special interest (39, 71). The side chain of Arg-14 contacts the main chain of Phe^{B25} and defines one edge of the crevice (together with Asp-12) in which the main chain and side chain of Tyr^{B26} loosely pack. The side-chain carboxamide of Gln-34 forms a hydrogen bond with the carboxylate of Asp-12, which in turn forms bidentate charge-stabilized hydrogen bonds to an ϵ-NH_2 and δ-NH of Arg-14 below an aromatic ring of Tyr^{B26} (Fig. 11*A*). This canonical Asp-Arg motif appears to be critical as Ala substitution of either residue markedly impairs the binding of insulin (71, 79, 80). In the predicted structure of the 3-[iodo-Tyr^{B26}]insulin/μIR interface, these native-like contacts are retained and extended by favorable electrostatic interactions with the equatorial belt of the halogen (Fig. 11*B*). In particular, one hydrogen in the key Arg^{14}ϵ-NH_2 group hydrogen-bonds with Asp-12, whereas the other hydrogen engages the δ^--equatorial zone of the iodine. Both hydrogens of the second ϵ-NH_2 group of Arg-14 interact with the carbonyl oxygen of Val-713 in αCT; this bifurcated pair of hydrogen bonds has lengths of 1.8 and 2.3 Å with an acute angle (~60°) between the two hydrogen bonds. Our model thus integrates the asymmetric electronic distribution around the iodine atom within a pre-existing charge-stabilized hydrogen-bond network inferred to exist in the WT complex. The conformationally averaged I···H distances to the respective side chains of Gln-34 and Arg-14 are 2.9 and 3.2 Å.

Water Molecules Are Integral to the Predicted Hormone-Receptor Interface

In the WT-μIR complex Val-712 lies at the C terminus of the αCT α-helix, and its carbonyl oxygen participates only in a weak-capping (i,i + 3) hydrogen bond to the main-chain NH of non-helical residue Val-715 (distance 3.6 Å and angle 113°; Fig. 11*A*). In the predicted structure of the [iodo-Tyr^{B26}]insulin complex, the carbonyl oxygen of Val-712 forms bifurcating interactions, the native hydrogen bond to Val-715, and the novel halogen bond to the σ-hole of the iodine. The angle between the hydrogen bond and halogen bond is ~75°. This interface is thus remarkable for the large number of stabilizing electrostatic interactions (Fig. 11*B*). Although the μIR co-crystal structure was of insufficient resolution to define the bound structure in detail, the B26-related crevice is exposed to solvent, and our MD simulations predicted formation of a bound water network anchored by the *para*-OH of Tyr^{B26} and the carbonyl oxygens of Asn-711, Val-712, and Phe-714 (Fig. 11C). In the [iodo-Tyr^{B26}]insulin complex, this network is retained and reinforced by interactions between the iodine and hydrogen atoms of three water molecules (Fig. 11*D*); one of these water molecules bridges the neighboring iodine and *para*-OH of the modified Tyr^{B26} (Wat3 in Fig. 11, *E* and *F*).

Together, the δ^--equatorial zone of the iodine atom is thus predicted to engage five hydrogen atoms, three from water molecules and two from L1 side chains Arg-14 and Gln-34 (as above; Fig. 11, *E* and *F*). Given the atomic radius of iodine (2.4 Å), its circumference may be estimated as ~30 Å, sufficient to accommodate these five contacts. One of the above three water molecules (Wat1) also contacts the water molecule (Wat2) anchored by the carbonyl oxygen of Val-712 (Fig. 11*F*) and in turn a hydrogen-bonded network involving Val-712 (C=O)···H(W2)···O(W1)···I(Tyr^{B26}). A striking prediction of this model is thus that a water-αCT network bridges the δ^--equatorial zone of the iodine atom with its δ^+ σ-hole.

Might the predicted iodine-anchored network of interfacial water molecules be observable in future co-crystal structures? To address this issue, we computed thermal *B*-factors from the

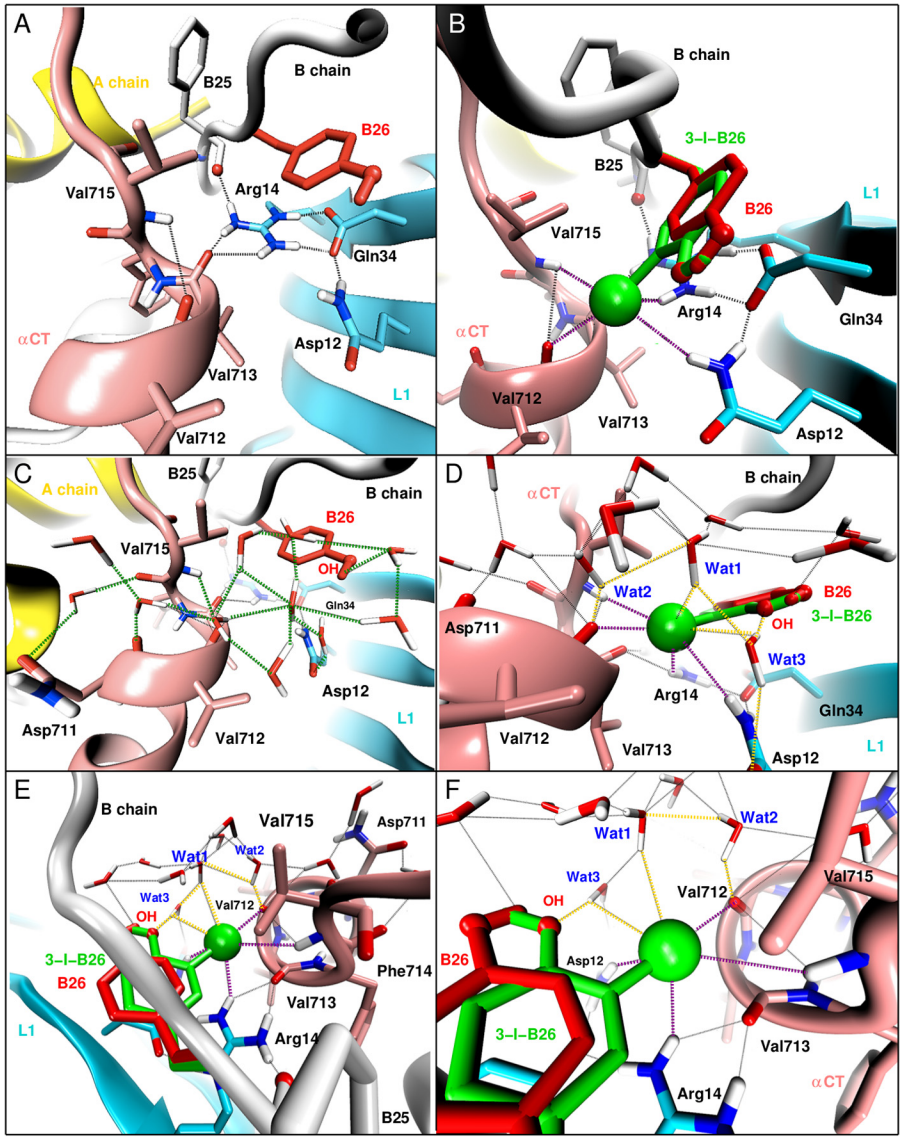

FIGURE 11. **Predicted water network at the μIR/insulin interface.** *A* and *B*, local interaction at the interface around Tyr^{B26} in WT insulin (*A* and *C*) and 3-I-Tyr^{B26} insulin (*B* and *D–F*). The interactions involve Asp-12, Arg-14, and Gln-34 of L1; Val-712, Val-713, and Val-715 of αCT; and with Phe^{B25} and Tyr^{B26} of insulin. Naive WT interactions are shown as *black dashed lines*, and the newly introduced hydrogen-halogen interactions through the 3-I-Tyr^{B26} mutation are shown as *dashed purple lines*. *C–F*, predicted water network at the μIR/insulin interface around Tyr^{B26}. *C*, formation of a water network anchored by the *para*-OH of Tyr^{B26} and the carbonyl oxygens of Asn-711, Val-712, and Phe-714 (highlighted by *green dashed lines*) in WT. *D–F*, reinforcement of the pre-existing water network by interactions (*yellow dashed lines*) between the iodine and hydrogen atoms of three water molecules labeled *Wat1*, *Wat2*, and *Wat3*; Wat3 bridges the iodine and *para*-OH of the modified Tyr^{B26} and the carboxyl oxygen of Asp-12 of the L1 domain (*D*); and Wat1 and Wat2 bridge the iodine atom to the carbonyl oxygen of Val-712 (*E* and *F*).

fluctuation around average positions. We first focused on the three putative water molecules (Wat1, Wat2, and Wat3) involved in hydrogen bonds with (or bridging between) the iodine atom, the *para*-OH of Tyr^{B26}, and the carbonyl oxygen of Val-712. Each of these H_2O oxygen atoms was predicted to exhibit a *B*-factor ~40 $Å^2$, significantly lower than that of other water molecules at this interface (>100 $Å^2$) and similar to the *B*-factors of the C_γ atoms of Val-712, Arg-14, and Gln-34 (*i.e.* ~30 $Å^2$). Moreover, the number of water molecules within 4 Å of the iodine atom at the 3-I-Tyr^{B26}/μIR/water interface ranging from 1 to 5 with an average of 3 during 1 ns of equilibrium MD simulation (supplemental Fig. S6). A similar analysis of the WT insulin/μIR/water interface also found strongly interacting water molecules, albeit with increased *B*-factors (~50 $Å^2$). We therefore anticipate that the three predicted iodine-anchored structural water molecules should be observable by crystallography in a sufficiently well ordered crystal.

Concluding Remarks

The promise of non-standard insulin analogs to enhance the treatment of diabetes mellitus represents an important frontier of molecular pharmacology (81). Indeed, substitution of Tyr^{B26} by 3-I-Tyr in the rapid-acting analog insulin *lispro* enhances its stability and resistance to fibrillation while maintaining its biological activity (45). The present crystal structure has demonstrated how the modified side chain pivots to enable burial of the iodine atom in the hydrophobic core. Despite such apparent optimization of the free state (45), 3-I-Tyr^{B26} also enhances

receptor binding (7–9). Our MD simulations suggest that enhancement is mediated by quantum chemistry: direct electrostatic effects of the iodine exploiting its δ^+ σ-hole and δ^- equatorial belt (46). This mechanism envisions that 3-I-Tyr^{B26} switches from an iodo-in conformation (free) to an iodo-out conformation (IR bound). Testing this proposal will require higher resolution structures of WT- and iodo-modified hormone-receptor complexes.

Experimental Procedures

Preparation of Insulin Analogs—Analogs were made by trypsin-catalyzed semi-synthesis using insulin fragment *des*-octapeptide(B23–B30)-insulin and modified octapeptides (4). The fragment was generated via cleavage of human insulin with trypsin and purified by reverse-phase HPLC; octapeptides were synthesized by solid-phase synthesis (82). Formation of a peptide bond between Arg^{B22} and the synthetic octapeptide was mediated by trypsin (in a mixed solvent system containing 1,4-butanediol and dimethylacetamide) (83). Insulin analogs were purified by preparative reverse-phase C4 HPLC (Higgins Analytical Inc., Proto 300 C4 10 μM, 250 × 20 mm), and their purity assessed by analytical rp-C4 HPLC (Higgins C4 5 μM, 250 × 4.6 mm). Molecular masses were verified using an Applied Biosystems 4700 proteomics analyzer (matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry).

Circular Dichroism—Far-ultraviolet CD spectra were obtained on an AVIV spectropolarimeter equipped with an automated syringe-driven titration unit. The proteins were made 50 μM in 10 mM potassium phosphate (pH 7.4) and 50 mM KCl. Spectra were obtained from 190–250 nm as described (84). Thermodynamic stabilities were probed by guanidine hydrochloride-induced denaturation monitored by CD at 222 nm. Data were fit by non-linear least squares to a two-state model (85) as described (86).

Receptor Binding Assays—Affinities for IR-B were measured by a competitive-displacement scintillation proximity assay. This assay employed solubilized receptor with C-terminal streptavidin-binding protein tags purified by sequential wheat germ agglutinin and StrepTactin-affinity chromatography from detergent lysates of polyclonal stably transfected 293PEAK cell lines expressing each receptor. The details of this assay were recently described (32). To obtain analog dissociation constants, competitive binding data were analyzed by non-linear regression (87).

X-ray Crystallography—Crystals of HPLC-purified 3-[iodo-Tyr^{B26},Nle^{B29}]insulin were obtained via hanging-drop vapor diffusion at 25 °C. 1-μl drops containing the protein at 10 mg/ml in 0.02 N HCl were mixed with a 1-μl drop of reservoir buffer containing 0.1 M sodium citrate, 0.08% zinc acetate, and 2% phenol. Drops were suspended over 1 ml of reservoir buffer. A single crystal was transferred to a solution containing 30% glycerol in the mother liquor for flash freezing. Diffraction data were obtained using an in-house X-ray source consisting of a Rigaku rotating-anode X-ray generator (MicroMax^{TM} 007HF) with VariMax confocal optics, a Saturn 944+ CCD X-ray detector, and an X-Stream 2000 cryogenic crystal cooling system (located at Case Western Reserve University). Data analysis employed XDS (88). The structure was determined by molecular replacement using PDB code 1ZNJ as a search model, followed by iterative refinement and model building using PHENIX (89) and COOT (90), respectively. The refinement strategy included both TLS refinement (translation, libration, and screw rotation) and torsional non-crystallographic symmetry (NCS) restraints between related chains. Coordinates were deposited in the Protein Data Bank (code 5EMS).

QM-parameterized MM Calculations—MD simulations employed CHARMM version c40a1 (91) with the "all-atom" force field CHARMM22 (53) using the correction map (CMAP) potential for the backbone dihedral (ϕ,ψ) to correct for the α-helical bias. To account for electronic anisotropy of *ortho*-iodophenol (employed as a model compound for purposes of MM parametrization), an electrostatic multipole model (MTP) (54) was obtained for the phenolic ring of Tyr^{B26} and iodophenolic ring of I-Tyr^{B26}. The MTP implementation used in this work is that of Bereau *et al.* (54), which uses up to quadrupolar moments on each interaction site (87). Atomic multipoles are assigned to all heavy atoms (but not the hydrogens). The parametrization protocol followed a recently developed strategy that includes optimization of multipole moments to best represent the electrostatic potential and van der Waals parameters to correctly describe experimental solution phase data, including the hydration free energy of iodophenol (92). MTPs were derived from *ab initio* calculations at the MP2/aug-cc-pVDZ (93) level of theory using GAUSSIAN09 (94). The iodine atom was treated by the aug-cc-pVDZ-PP basis set with an effective core potential (95). For the rest of the system, a point-charge (PC) model was used; water molecules were treated with the TIP3P model (96).

Starting coordinates of the insulin dimers were first taken from the T_2 zinc-free structure of WT insulin (PDB code 1DPH, resolution 1.9 Å (97)) and extracted from the R_6 structure of the WT zinc hexamer (PDB 1ZNJ, resolution 2.0 Å), respectively, and subsequently extended to the present crystal structure of a 3-I-Tyr^{B26} insulin analog. For corresponding simulations of the variant μIR-insulin complex, starting coordinates were obtained from the lowest energy initial model (see below). The systems were first minimized by steepest descent for 5×10^4 steps. The proteins were solvated in TIP3P water molecules (96) equilibrated at 300 K and 1 atm (within a 52.77Å cubic box for the dimers and a 93 Å cubic box for the μIR-insulin with sodium and chloride ions added to neutralize the system). The box was heated from 0 to 300 K for 30 ps, equilibrated for 500 ps, and then subjected to 10 ns of production MD with periodic boundary conditions. The particle mesh Ewald (PME) method (98) was used for PC-PC interactions, with grid spacing of 1 Å, a relative tolerance of 10^{-6}, an interpolation order of 4 for long-range electrostatics, and a cutoff of 14 Å together with a 12-Å switching threshold for L-J interactions. Bonds involving hydrogen were constrained by SHAKE (99).

Because of aromatic ring rotation about the C_β–C_γ bond axis, the mono-iodo derivative of [Tyr^{B26}]insulin may in principle form either 3-I-Tyr^{B26} or 5-I-Tyr^{B26} conformational isomers (with the iodine atom at ring positions 3 or 5; *i.e. ortho* to the phenolic hydroxyl group and *meta* to C_γ). Hence, depending on the B26 χ_2 angle, isomerization of 3-I-Tyr^{B26} to 5-I-Tyr^{B26}, and vice versa, is possible during the dynamics. Because the relative stabilities of these conformational isomers were not known *a priori*, two independent MD dimer simulations of 20

ns each were carried out starting from initial 3-I or 5-I B26 conformations, respectively. These simulations began from a T state crystallographic protomer (PDB code 1DPH (97)) because the conformation of an insulin monomer in solution resembles the T state (28, 29).

Rigid-body Calculations—Naive replacement of B26's H atoms (atomic radius 1.20 Å and CH bond length 1.08 Å) by iodine (atomic radius 2.40 Å and CI bond length 2.08 Å) in positions 3 and 5 was carried out, starting from the coordinates of T_2 zinc-free structure of WT insulin (PDB code 1DPH, resolution 1.9 Å (97)). First, the three systems (WT, 3-I, and 5-I) were energy-minimized by steepest descent for 5×10^4 steps using an optimized GBSW (generalized Born with a smoothed switching function (100, 101)) implicit solvent force field (78). In a next step, energies from two scans around χ_1–χ_2 dihedral angles (±20° in each dihedral angle in steps of 1°) were computed, using a PC and an MTP representation for the Tyr^{B26} side chain as follows: (i) a rigid scan, where single point energies are calculated upon χ_1–χ_2 rotations, and (ii) a relaxed scan, where the protein side chains were energy-minimized for 100 steps of steepest descent after each χ_1–χ_2 rotations.

Quantum-Chemical Calculations—The molecular electrostatic potential (incorporated electron density) from the same electronic structure calculations was employed for the MTP parameter fitting as mapped at the 10^{-3} *e* atomic units^{-3} isodensity surface using Gaussview5 (102). All *ab initio* calculations were carried out with Gaussian09 (94), and optimized structures were used. The size of the iodine σ-hole size was measured as the angular profile of the ESP intersection line of the grid and the halogen boundary (defined as a surface of electron isodensity of 10^{-3} *e* atomic units^{-3}) (55). Such *ab initio* calculations were undertaken solely for the purpose of parameterizing an MM model of the modified insulin. Explicit QM/MM simulations were not performed.

Modeling of the Variant Hormone-Receptor Complex—To explore how iodo-Tyr^{B26} might pack within the insulin-IR complex, the potential environment of the modified ring was considered in the context of the WT insulin-μIR complex (PDB code 4OGA (39)). Several subsets of residues disordered in the crystal structure (IR residues Cys-159–Asn-168 and Lys-265–Gln-276 and insulin residues B28–B30) were included, as were *N*-linked *N*-acetylglucosamine modifications at sites Asn-16, Asn-25, Asn-111, Asn-215, and Asn-255 (103). An initial set of 50 models was created, and the structure with lowest empirical energy was selected for MD simulations. A model of the variant μIR complex was constructed in two stages. Preliminary MD studies of the variant hormone-μIR complex were first performed using a coarse model of the modified side chain within the GROMACS package (version 4.6.1 (104)) with OPLS-all atom force field (96, 105) as described below. The lowest energy model emerging from this simulation then provided a starting point for QM-parameterized MD as outlined above. Initial MD simulations of the variant μIR complex exploited an approximate model of 3-I-Tyr in which a virtual site was placed near the iodine atom to mimic the σ-hole. This site's position was determined by minimizing the error of the fit of the atom-centered charges to the molecular electrostatic potential for 2-iodo-4-methylphenol, calculated at the HF/6–311G(d,p) level; the optimal position of the virtual site was 1.5 Å from the iodine, co-linear with the C–I bond. Respective partial charges on the virtual site, iodine and carbon attached to iodine, were 0.115*e*, −0.322*e*, and 0.207*e*; charges on all other atoms were adopted from the OPLS-aa parameters for Tyr. Proteins were solvated in a cubic box of TIP4P water molecules (96); the box extended 10 Å beyond any protein atom. Ionizable residues and protein termini were set in their charged states. Sodium and chloride ions were added to neutralize the system at a final ionic strength of 0.10 M. Protein and solvent (including ions) were coupled separately to a thermal bath at 300 K employing velocity rescaling (106) with coupling time 1.0 ps. Pressure was maintained at 1 bar using a Berendsen barostat (107) with coupling constant 5.0 ps and compressibility 4.5×10^{-5} bar. The time step was 2 fs. Simulations were performed with a single non-bonded cutoff of 10 Å and neighbor-list update frequency of 10 steps (20 fs). The PME method modeled long-range electrostatics (108); the grid width was 1.2 Å with fourth-order spline interpolation. Bond lengths were constrained using LINCS (109). The MD protocol consisted of an initial minimization of water molecules, followed by 100 ps of MD with the protein restrained to permit equilibration of the solvent. Calculations were continued for 200 ns from the geometries obtained after initial positionally restrained MD at a temperature of 300 K.

Database Deposition—Atomic coordinates of the crystal of 3-[iodo-Tyr^{B26},Nle^{B29}]insulin have been deposited in the Protein Data Bank (code 5EMS).

Supplemental Information—The supplemental Table S1 provides interaction energies contributing to dimerization. The supplemental Figs. S1 and S2 provide side-chain dihedral-angle distributions (χ_1,χ_2). The supplemental Fig. S3 illustrates predicted packing schemes at the variant dimer interfaces. The supplemental Fig. S4 illustrates predicted and observed structural relationships in 3-[iodo-Tyr^{B26}]insulin. The supplemental Fig. S5 provides details concerning the alternative (and less favorable) 5-[iodo-Tyr^{B26}]insulin dimer interface. The supplemental Fig. S6 depicts water molecules at the variant μIR interface.

Author Contributions—Molecular dynamics simulations were performed by K. E. H., B. J. S., and M. M. The *de novo* quantum simulations and electrostatic multipole parametrization was done by K. E. H. and M. M. Biochemical and biophysical assays were performed by V. P., N. B. P., and J. W. Insulin analogs were prepared by V. P. and N. B. P. Crystallization trials and structure determination were undertaken by V. P. Refinement was performed by V. P., J. G. M., and M. L. C. Molecular modeling of the variant hormone-μIR interface was undertaken by K. E. H., B. J. S., and M. M., with the assistance of J. G. M., M. L. C., and M. A. W. The overall program of research was guided by M. M. and M. A. W. Each of the authors contributed to the manuscript.

Acknowledgments—We thank Drs. Q.-X. Hua, W. Jia, L. Whittaker, S. H. Nakagawa, and Z.-L. Wan for advice regarding experimental procedures. We thank N. Rege for assistance with manuscript preparation and journal editors C. Goodman and M. Spiering for helpful suggestions. We also thank Profs. P. Arvan, T. L. Blundell, F. Ismail-Beigi, M. Karplus, P. G. Katsoyannis, and M. Liu for general discussion.

References

1. Dodson, G., and Steiner, D. (1998) The role of assembly in insulin's biosynthesis. *Curr. Opin. Struct. Biol.* **8,** 189–194
2. Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. M., Hubbard, R. E., Isaacs, N. W., and Reynolds, C. D. (1988) The structure of 2Zn pig insulin crystals at 1.5 Å resolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **319,** 369–456
3. Brange, J., Ribel, U., Hansen, J. F., Dodson, G., Hansen, M. T., Havelund, S., Melberg, S. G., Norris, F., Norris, K., and Snel, L. (1988) Monomeric insulins obtained by protein engineering and their medical implications. *Nature* **333,** 679–682
4. Inouye, K., Watanabe, K., Tochino, Y., Kobayashi, M., and Shigeta, Y. (1981) Semisynthesis and properties of some insulin analogs. *Biopolymers* **20,** 1845–1858
5. Pandyarajan, V., Smith, B. J., Phillips, N. B., Whittaker, L., Cox, G. P., Wickramasinghe, N., Menting, J. G., Wan, Z.-L., Whittaker, J., Ismail-Beigi, F., Lawrence, M. C., and Weiss, M. A. (2014) Aromatic anchor at an invariant hormone-receptor interface function of insulin residue B24 with application to protein design. *J. Biol. Chem.* **289,** 34709–34727
6. Scholfield, M. R., Zanden, C. M., Carter, M., and Ho, P. S. (2013) Halogen bonding (X-bonding): A biological perspective. *Protein Sci.* **22,** 139–152
7. Linde, S., Sonne, O., Hansen, B., and Gliemann, J. (1981) Monoiodoinsulin labelled in tyrosine residue 16 or 26 of the insulin B-chain. Preparation and characterization of some binding properties. *Hoppe Seylers Z Physiol. Chem.* **362,** 573–579
8. Peavy, D. E., Abram, J. D., Frank, B. H., and Duckworth, W. C. (1984) Receptor binding and biological activity of specifically labeled [^{125}I]- and [^{127}I]monoiodoinsulin isomers in isolated rat adipocytes. *Endocrinology* **114,** 1818–1824
9. Burant, C. F., Treutelaar, M. K., Peavy, D. E., Frank, B. H., and Buse, M. G. (1988) Differential binding of monoiodinated insulins to muscle and liver derived receptors and activation of the receptor kinase. *Biochem. Biophys. Res. Commun.* **152,** 1353–1360
10. Swallow, S. (2012) in *Fluorine in Pharmaceutical and Medicinal Chemistry: From Biophysical Aspects to Clinical Applications* (Gouverneur, V., and Muller, K., eds) pp. 34, Imperial College Press, London
11. Kiga, D., Sakamoto, K., Kodama, K., Kigawa, T., Matsuda, T., Yabuki, T., Shirouzu, M., Harada, Y., Nakayama, H., Takio, K., Hasegawa, Y., Endo, Y., Hirao, I., and Yokoyama, S. (2002) An engineered *Escherichia coli* tyrosyl-tRNA synthetase for site-specific incorporation of an unnatural amino acid into proteins in eukaryotic translation and its application in a wheat germ cell-free system. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 9715–9720
12. Chatterjee, A., Xiao, H., and Schultz, P. G. (2012) Evolution of multiple, mutually orthogonal prolyl-tRNA synthetase/tRNA pairs for unnatural amino acid mutagenesis in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **109,** 14841–14846
13. Lommerse, J. P., Stone, A. J., Taylor, R., and Allen, F. H. (1996) The nature and geometry of intermolecular interactions between halogens and oxygen or nitrogen. *J. Am. Chem. Soc.* **118,** 3108–3116
14. Matter, H., Nazaré, M., Güssregen, S., Will, D. W., Schreuder, H., Bauer, A., Urmann, M., Ritter, K., Wagner, M., and Wehner, V. (2009) Evidence for C-Cl/C-Br.π interactions as an important contribution to protein-ligand binding affinity. *Angew. Chem. Int. Ed. Engl.* **48,** 2911–2916
15. Müller, K., Faeh, C., and Diederich, F. (2007) Fluorine in pharmaceuticals: looking beyond intuition. *Science* **317,** 1881–1886
16. Metrangolo, P., Neukirch, H., Pilati, T., and Resnati, G. (2005) Halogen bonding based recognition processes: a world parallel to hydrogen bonding. *Acc. Chem. Res.* **38,** 386–395
17. Metrangolo, P., Meyer, F., Pilati, T., Resnati, G., and Terraneo, G. (2008) Halogen bonding in supramolecular chemistry. *Angew. Chem. Int. Ed. Engl.* **47,** 6114–6127
18. Politzer, P., Murray, J. S., and Clark, T. (2013) Halogen bonding and other σ-hole interactions: a perspective. *Phys. Chem. Chem. Phys.* **15,** 11178–11189
19. El Hage, K., Piquemal, J. P., Hobaika, Z., Maroun, R. G., and Gresh, N. (2015) Could the "Janus-like" properties of the halobenzene CX bond (X=Cl, Br) be leveraged to enhance molecular recognition? *J. Comput. Chem.* **36,** 210–221
20. Badger, J., Harris, M. R., Reynolds, C. D., Evans, A. C., Dodson, E. J., Dodson, G. G., and North, A. C. (1991) Structure of the pig insulin dimer in the cubic crystal. *Acta Crystallogr. B Struct. Sci.* **47,** 127–136
21. Adams, M. J., Blundell, T. L., Dodson, E. J., Dodson, G. G., Vijayan, M., Baker, E. N., Hardine, M. M., Hodgkin, D. C., Rimer, B., and Sheet, S. (1969) Structure of rhombohedral 2 zinc insulin crystals. *Nature* **224,** 491–495
22. Bentley, G., Dodson, E., Dodson, G., Hodgkin, D., and Mercola, D. (1976) Structure of insulin in 4-zinc insulin. *Nature* **261,** 166–168
23. Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Reynolds, C. D., Smith, G. D., Sparks, C., and Swenson, D. (1989) Phenol stabilizes more helix in a new symmetrical zinc insulin hexamer. *Nature* **338,** 594–596
24. Hirsch, I. B. (2005) Insulin analogues. *N. Engl. J. Med.* **352,** 174–183
25. Liang, D. C., Chang, W. R., and Wan, Z. L. (1994) A proposed interaction model of the insulin molecule with its receptor. *Biophys. Chem.* **50,** 63–71
26. De Meyts, P., and Whittaker, J. (2002) Structural biology of insulin and IGF1 receptors: implications for drug design. *Nat. Rev. Drug Discov.* **1,** 769–783
27. Hua, Q. X., and Weiss, M. A. (1991) Comparative 2D NMR studies of human insulin and *des*-pentapeptide insulin: sequential resonance assignment and implications for protein dynamics and receptor recognition. *Biochemistry* **30,** 5505–5515
28. Hua, Q. X., Hu, S. Q., Frank, B. H., Jia, W., Chu, Y. C., Wang, S. H., Burke, G. T., Katsoyannis, P. G., and Weiss, M. A. (1996) Mapping the functional surface of insulin by design: structure and function of a novel A-chain analogue. *J. Mol. Biol.* **264,** 390–403
29. Olsen, H. B., Ludvigsen, S., and Kaarsholm, N. C. (1996) Solution structure of an engineered insulin monomer at neutral pH. *Biochemistry* **35,** 8836–8845
30. Hua, Q. X., Jia, W., and Weiss, M. A. (2011) Conformational dynamics of insulin. *Front. Endocrinol.* **2,** 48
31. Papaioannou, A., Kuyucak, S., and Kuncic, Z. (2015) Molecular dynamics simulations of insulin: elucidating the conformational changes that enable its binding. *PLoS One* **10,** e0144058
32. Pandyarajan, V., Phillips, N. B., Rege, N., Lawrence, M. C., Whittaker, J., and Weiss, M. A. (2016) Contribution of Tyr^{B26} to the function and stability of insulin. Structure-activity relationships at a conserved hormone-receptor interface. *J. Biol. Chem.* **291,** 12978–12990
33. Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Bing, X., and Markussen, J. (1991) X-ray analysis of the single chain B29-A1 peptide-linked insulin molecule. A completely inactive analogue. *J. Mol. Biol.* **220,** 425–433
34. Hua, Q. X., Shoelson, S. E., Kochoyan, M., and Weiss, M. A. (1991) Receptor binding redefined by a structural switch in a mutant human insulin. *Nature* **354,** 238–241
35. Mirmira, R. G., and Tager, H. S. (1991) Disposition of the phenylalanine B25 side chain during insulin-receptor and insulin-insulin interactions. *Biochemistry* **30,** 8222–8229
36. Hua, Q. X., Xu, B., Huang, K., Hu, S. Q., Nakagawa, S., Jia, W., Wang, S., Whittaker, J., Katsoyannis, P. G., and Weiss, M. A. (2009) Enhancing the activity of insulin by stereospecific unfolding. Conformational life cycle of insulin and its evolutionary origins. *J. Biol. Chem.* **284,** 14586–14596
37. Xu, B., Huang, K., Chu, Y. C., Hu, S. Q., Nakagawa, S., Wang, S., Wang, R. Y., Whittaker, J., Katsoyannis, P. G., and Weiss, M. A. (2009) Decoding the cryptic active conformation of a protein by synthetic photoscanning: Insulin inserts a detachable arm between receptor domains. *J. Biol. Chem.* **284,** 14597–14608
38. Kristensen, C., Andersen, A. S., Ostergaard, S., Hansen, P. H., and Brandt, J. (2002) Functional reconstitution of insulin receptor binding site from non-binding receptor fragments. *J. Biol. Chem.* **277,** 18340–18345
39. Menting, J. G., Yang, Y., Chan, S. J., Phillips, N. B., Smith, B. J., Whittaker, J., Wickramasinghe, N. P., Whittaker, L. J., Pandyarajan, V., Wan, Z. L., Yadav, S. P., Carroll, J. M., Stokes, N., Roberts, C. T., Jr., Ismail-Beigi, F., *et al.* (2014) A structural hinge in insulin enables its receptor engagement. *Proc. Natl. Acad. Sci. U.S.A.* **111,** E3395–E3404

40. Ward, C. W., Menting, J. G., and Lawrence, M. C. (2013) The insulin receptor changes conformation in unforeseen ways on ligand binding: sharpening the picture of insulin receptor activation. *Bioessays* **35,** 945–954
41. Kavran, J. M., McCabe, J. M., Byrne, P. O., Connacher, M. K., Wang, Z., Ramek, A., Sarabipour, S., Shan, Y., Shaw, D. E., Hristova, K., Cole, P. A., and Leahy, D. J. (2014) How IGF-1 activates its receptor. *eLife* **3,** e03772
42. Croll, T. I., Smith, B. J., Margetts, M. B., Whittaker, J., Weiss, M. A., Ward, C. W., and Lawrence, M. C. (2016) Higher-resolution structure of the human insulin receptor ectodomain: multi-modal inclusion of the insert domain. *Structure* **24,** 469–476
43. Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. C., Mercola, D. A., and Vijayan, M. (1971) Atomic positions in rhombohedral 2-zinc insulin crystals. *Nature* **231,** 506–511
44. Burley, S. K., and Petsko, G. A. (1988) Weakly polar interaction in proteins. *Adv. Protein Chem.* **39,** 125–189
45. Pandyarajan, V., Phillips, N. B., Cox, G. P., Yang, Y., Whittaker, J., Ismail-Beigi, F., and Weiss, M. A. (2014) Biophysical optimization of a therapeutic protein by non-standard mutagenesis. Studies of an iodo-insulin derivative. *J. Biol. Chem.* **289,** 23367–23381
46. Politzer, P., Lane, P., Concha, M. C., Ma, Y., and Murray, J. S. (2007) An overview of halogen bonding. *J. Mol. Model.* **13,** 305–311
47. Poznański, J., Winiewska, M., Czapinska, H., Poznańska, A., and Shugar, D. (2016) Halogen bonds involved in binding of halogenated ligands by protein kinases. *Acta Biochim. Pol.* **63,** 203–214
48. Clark, T., Hennemann, M., Murray, J. S., and Politzer, P. (2007) Halogen bonding: the σ-hole. Proceedings of "Modeling interactions in biomolecules II," Prague, September 5th–9th, 2005. *J. Mol. Model.* **13,** 291–296
49. Politzer, P., Murray, J. S., and Clark, T. (2010) Halogen bonding: an electrostatically-driven highly directional noncovalent interaction. *Phys. Chem. Chem. Phys.* **12,** 7748–7757
50. Jacoby, E., Hua, Q. X., Stern, A. S., Frank, B. H., and Weiss, M. A. (1996) Structure and dynamics of a protein assembly. ^{1}H NMR studies of the 36 kDa R_6 insulin hexamer. *J. Mol. Biol.* **258,** 136–157
51. Chang, X., Jorgensen, A. M., Bardrum, P., and Led, J. J. (1997) Solution structures of the R6 human insulin hexamer. *Biochemistry* **36,** 9409–9422
52. Kristensen, S. M., Jørgensen, A. M., Led, J. J., Balschmidt, P., and Hansen, F. B. (1991) Proton nuclear magnetic resonance study of the B9(Asp) mutant of human insulin. Sequential assignment and secondary structure. *J. Mol. Biol.* **218,** 221–231
53. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T., Mattos, C., *et al.* (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102,** 3586–3616
54. Bereau, T., Kramer, C., and Meuwly, M. (2013) Leveraging symmetries of static atomic multipole electrostatics in molecular dynamics simulations. *J. Chem. Theory Comput.* **9,** 5450–5459
55. Kolář, M., Hostaš, J., and Hobza, P. (2014) The strength and directionality of a halogen bond are co-determined by the magnitude and size of the σ-hole. *Phys. Chem. Chem. Phys.* **16,** 9987–9996
56. Mark, A. E., Berendsen, H. J., and van Gunsteren, W. F. (1991) Conformational flexibility of aqueous monomeric and dimeric insulin: a molecular dynamics study. *Biochemistry* **30,** 10866–10872
57. Dodson, G. G., Dodson, E. J., Turkenburg, J. P., and Bing, X. (1993) Molecular recognition in insulin assembly. *Biochem. Soc. Trans.* **21,** 609–614
58. Falconi, M., Cambria, M. T., Cambria, A., and Desideri, A. (2001) Structure and stability of the insulin dimer investigated by molecular dynamics simulation. *J. Biomol. Struct. Dyn.* **18,** 761–772
59. Zoete, V., Meuwly, M., and Karplus, M. (2004) A comparison of dynamic behavior of monomeric and dimeric insulin shows structural rearrangements in the active monomer. *J. Mol. Biol.* **342,** 913–929
60. Nakagawa, S. H., Hua, Q. X., Hu, S. Q., Jia, W., Wang, S., Katsoyannis, P. G., and Weiss, M. A. (2006) Chiral mutagenesis of insulin. Contribution of the B20-B23 β-turn to activity and stability. *J. Biol. Chem.* **281,** 22386–22396
61. Mulder, F. A., Mittermaier, A., Hon, B., Dahlquist, F. W., and Kay, L. E. (2001) Studying excited states of proteins by NMR spectroscopy. *Nat. Struct. Biol.* **8,** 932–935
62. Smith, G. D., Ciszak, E., Magrum, L. A., Pangborn, W. A., and Blessing, R. H. (2000) R_6 hexameric insulin complexed with *m*-cresol or resorcinol. *Acta Crystallogr. D Biol. Crystallogr.* **56,** 1541–1548
63. Hamlin, J. L., and Arquilla, E. R. (1974) Monoiodoinsulin. Preparation, purification, and characterization of a biologically active derivative substituted predominantly on tyrosine A14. *J. Biol. Chem.* **249,** 21–32
64. Frank, B. H., Peavy, D. E., Hooker, C. S., and Duckworth, W. C. (1983) Receptor binding properties of monoiodotyrosyl insulin isomers purified by high performance liquid chromatography. *Diabetes* **32,** 705–711
65. Auffinger, P., Hays, F. A., Westhof, E., and Ho, P. S. (2004) Halogen bonds in biological molecules. *Proc. Natl. Acad. Sci. U.S.A.* **101,** 16789–16794
66. Riley, K. E., and Hobza, P. (2013) On the importance and origin of aromatic interactions in chemistry and biodisciplines. *Acc. Chem. Res.* **46,** 927–936
67. Jorgensen, W. L., and Schyman, P. (2012) Treatment of halogen bonding in the OPLS-AA force field: application to potent anti-HIV agents. *J. Chem. Theory Comput.* **8,** 3895
68. El Hage, K., Bereau, T., Jakobsen, S., and Meuwly, M. (2016) Impact of quadrupolar electrostatics on atoms adjacent to the sigma-hole in condensed-phase simulations. *J. Chem. Theory Comput.* **12,** 3008–3019
69. Conlon, J. M. (2001) Evolution of the insulin molecule: insights into structure-activity and phylogenetic relationships. *Peptides* **22,** 1183–1193
70. Liu, C. C., and Schultz, P. G. (2010) Adding new chemistries to the genetic code. *Annu. Rev. Biochem.* **79,** 413–444
71. Menting, J. G., Whittaker, J., Margetts, M. B., Whittaker, L. J., Kong, G. K., Smith, B. J., Watson, C. J., Záková, L., Kletvíková, E., Jiráček J, Chan, S. J., Steiner, D. F., Dodson, G. G., Brzozowski, A. M., Weiss, M. A., Ward, C. W., and Lawrence, M. C. (2013) How insulin engages its primary binding site on the insulin receptor. *Nature* **493,** 241–245
72. Küpper, F. C., Feiters, M. C., Olofsson, B., Kaiho, T., Yanagida, S., Zimmermann, M. B., Carpenter, L. J., Luther, G. W., 3rd., Lu, Z., Jonsson, M., and Kloo, L. (2011) Commemorating two centuries of iodine research: an interdisciplinary overview of current research. *Angew. Chem. Int. Ed. Engl.* **50,** 11598–11620
73. Sandler, B., Webb, P., Apriletti, J. W., Huber, B. R., Togashi, M., Cunha Lima, S. T., Juric, S., Nilsson, S., Wagner, R., Fletterick, R. J., and Baxter, J. D. (2004) Thyroxine-thyroid hormone receptor interactions. *J. Biol. Chem.* **279,** 55801–55808
74. Wagner, R. L., Huber, B. R., Shiau, A. K., Kelly, A., Cunha Lima, S. T., Scanlan, T. S., Apriletti, J. W., Baxter, J. D., West, B. L., and Fletterick, R. J. (2001) Hormone selectivity in thyroid hormone receptors. *Mol. Endocrinol.* **15,** 398–410
75. Björn, L. O. (2010) Comment: evolutionary roots of iodine and thyroid hormones in cell-cell signaling. *Integr. Comp. Biol.* **50,** 138–140
76. Richards, F. M. (1974) The interpretation of protein structures: total volume, group volume distributions and packing density. *J. Mol. Biol.* **82,** 1–14
77. Tilton, R. F., Jr., Kuntz, I. D., Jr., and Petsko, G. A. (1984) Cavities in proteins: structure of a metmyoglobin-xenon complex solved to 1.9 A. *Biochemistry* **23,** 2849–2857
78. Hédin, F., El Hage, K., and Meuwly, M. (2016) A toolkit to fit nonbonded parameters from and for condensed phase simulations. *J. Chem. Inf. Model* **56,** 1479–1489
79. Williams, P. F., Mynarcik, D. C., Yu, G. Q., and Whittaker, J. (1995) Mapping of an NH_2-terminal ligand binding site of the insulin receptor by alanine scanning mutagenesis. *J. Biol. Chem.* **270,** 3012–3016
80. Mynarcik, D. C., Williams, P. F., Schaffer, L., Yu, G. Q., and Whittaker, J. (1997) Identification of common ligand binding determinants of the insulin and insulin-like growth factor 1 receptors. Insights into mechanisms of ligand binding. *J. Biol. Chem.* **272,** 18650–18655
81. Pandyarajan, V., and Weiss, M. A. (2012) Design of non-standard insulin analogs for the treatment of diabetes mellitus. *Curr. Diab. Rep.* **12,** 697–704
82. Barany, G., and Merrifield, R. B. (1980) in *The Peptides* (Gross, E., and Meienhofer, J. eds) pp. 273–284, Academic Press, New York
83. Kubiak, T., and Cowburn, D. (1986) Enzymatic semisynthesis of porcine despentapeptide (B26–30) insulin using unprotected desoctapeptide (B23–30) insulin as a substrate. *Int. J. Pept. Protein Res.* **27,** 514–521

84. Sreerama, N., and Woody, R. W. (1993) A self-consistent method for the analysis of protein secondary structure from circular dichroism. *Anal. Biochem.* **209,** 32–44
85. Sosnick, T. R., Fang, X., and Shelton, V. M. (2000) Application of circular dichroism to study RNA folding transitions. *Methods Enzymol.* **317,** 393–409
86. Pace, C. N., and Shaw, K. L. (2000) Linear extrapolation method of analyzing solvent denaturation curves. *Proteins* **4,** 1–7
87. Wang, Z. X. (1995) An exact mathematical expression for describing competitive binding of two different ligands to a protein molecule. *FEBS Lett.* **360,** 111–114
88. Kabsch, W. (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 133–144
89. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 213–221
90. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132
91. Brooks, B. R., Brooks, C. L., 3rd., Mackerell, A. D., Jr., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., *et al.* (2009) CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **30,** 1545–1614
92. Kramer, C., Bereau, T., Spinn, A., Liedl, K. R., Gedeck, P., and Meuwly, M. (2013) Deriving static atomic multipoles from the electrostatic potential. *J. Chem. Inf. Model* **53,** 3410–3417
93. Feller, D. (1996) The role of databases in support of computational chemistry calculations. *J. Comput. Chem.* **17,** 1571–1586
94. Frisch, M., Trucks, G., Schlegel, H. B., Scuseria, G., Robb, M., Cheeseman, J., Scalmani, G., Barone, V., Mennucci, B., and Petersson, G. (2009) Gaussian 09, revision A. 02, Gaussian Inc., Wallingford, CT
95. Peterson, K. A., Shepler, B. C., Figgen, D., and Stoll, H. (2006) On the spectroscopic and thermochemical properties of ClO, BrO, IO, and their anions. *J. Phys. Chem. A* **110,** 13877–13883
96. Jorgensen, W. L., and Tirado-Rives, J. (1988) The OPLS potential functions for proteins. Energy minimizations for crystals of cyclic peptides of crambin. *J. Am. Chem. Soc.* **110,** 1657–1666
97. Gursky, O., Badger, J., Li, Y., and Caspar, D. L. (1992) Conformational changes in cubic insulin crystals in the pH range 7–11. *Biophys. J.* **63,** 1210–1220
98. Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98,** 10089–10092
99. Van Gunsteren, W., and Berendsen, H. (1977) Algorithms for macromolecular dynamics and constraint dynamics. *Mol. Physics* **34,** 1311–1327
100. Im, W., Lee, M. S., and Brooks, C. L., 3rd. (2003) Generalized born model with a simple smoothing function. *J. Comput. Chem.* **24,** 1691–1702
101. Im, W., Feig, M., and Brooks, C. L., 3rd (2003) An implicit membrane generalized born theory for the study of structure, stability, and interactions of membrane proteins. *Biophys. J.* **85,** 2900–2918
102. Dennington, R., Keith, T., and Millam, J. (2009) GaussView, Version 5.0, Semichem, Inc., Shawnee Mission, KS
103. Elleman, T. C., Frenkel, M. J., Hoyne, P. A., McKern, N. M., Cosgrove, L., Hewish, D. R., Jachno, K. M., Bentley, J. D., Sankovich, S. E., and Ward, C. W. (2000) Mutational analysis of the *N*-linked glycosylation sites of the human insulin receptor. *Biochem. J.* **347,** 771–779
104. Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **4,** 435–447
105. Damm, W., Frontera, A., Tirado-Rives, J., and Jorgensen, W. L. (1997) OPLS all-atom force field for carbohydrates. *J. Comput. Chem.* **18,** 1955–1970
106. Bussi, G., Donadio, D., and Parrinello, M. (2007) Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126,** 014101
107. Berendsen, H. J., Postma, J. P., van Gunsteren, W. F., DiNola, A., and Haak, J. (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81,** 3684
108. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys.* **103,** 8577–8593
109. Hess, B. (2008) P-LINCS: A parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.* **4,** 116–122
110. Moller, D. E., Yokota, A., Caro, J. F., and Flier, J. S. (1989) Tissue-specific expression of two alternatively spliced insulin receptor mRNAs in man. *Mol. Endocrinol.* **3,** 1263–1269
111. Bader, R. F., Carroll, M. T., Cheeseman, J. R., and Chang, C. (1987) Properties of atoms in molecules: atomic volumes. *J. Am. Chem. Soc.* **109,** 7968–7979
112. Karplus, P. A., and Diederichs, K. (2012) Linking crystallographic model and data quality. *Science* **336,** 1030–1033

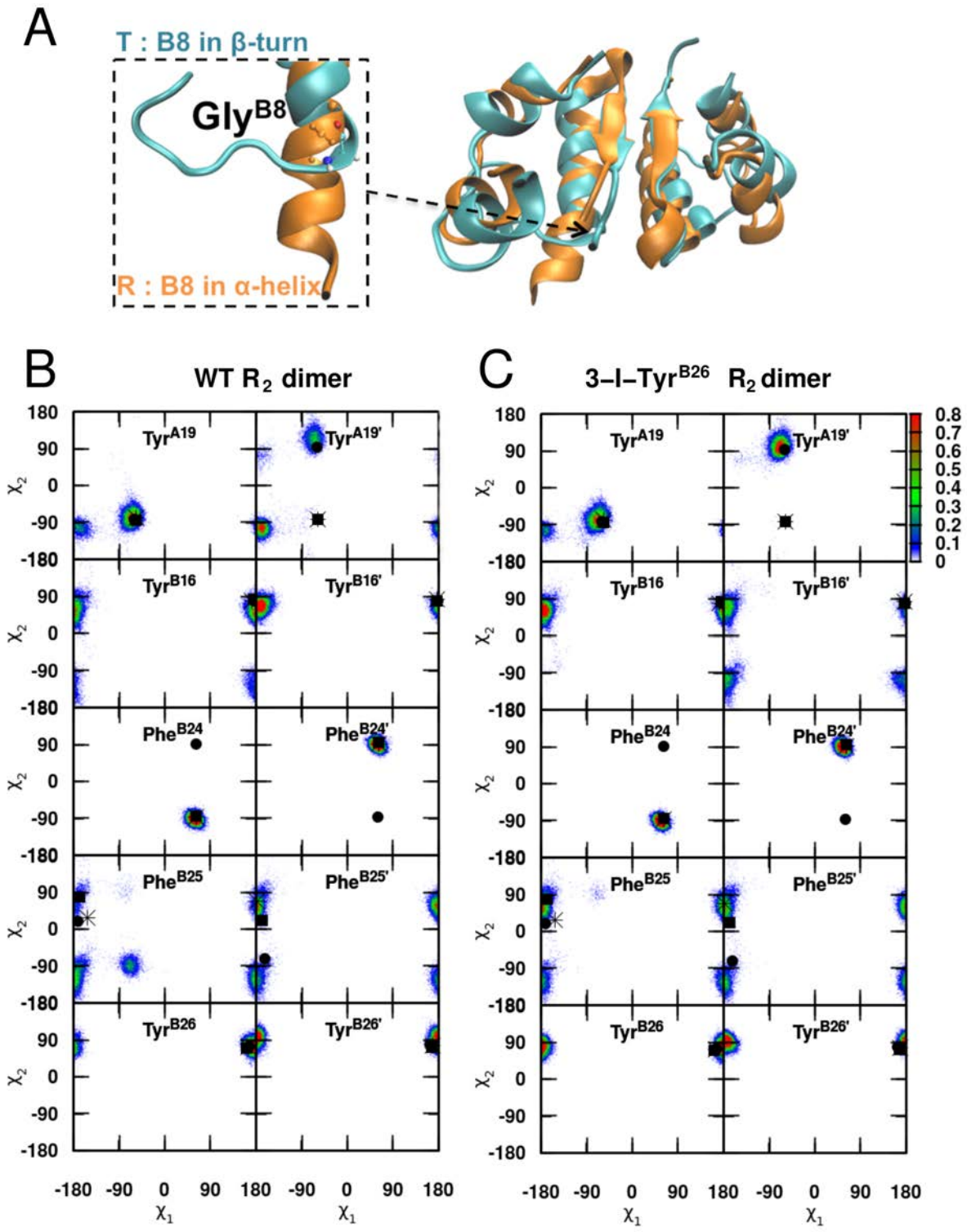

FIGURE S1. **Dihedral probability distribution functions for the R_2 dimers.** (*A*) Structure of the R- (α-helix B1-B8; *gold*) and T (disordered strand B1-B8; *cyan*) -states of insulin dimer. The *inset* indicates the placement of the invariant Gly^{B8} as a function of T/R state. Only Gly^{B8} is illustrated. (*B* and *C*) Probability distribution functions $P(\chi_1,\chi_2)$ for Tyr^{A19}, Tyr^{B16}, Phe^{B24}, Phe^{B25}, Tyr^{B26} and their dimer related mates (indicated by *primes*) for (*B*) the WT R_2 insulin dimer, and (*C*) the 3-iodo-Tyr^{B26} R_2 insulin dimer. The distributions were built from 20 ns equilibrium MD simulations. The starting dimer structure was taken from a dimer of WT R_6 zinc

insulin crystal structure (PDB code 1ZNJ). Symbols (*box, star, circle*) indicate positions in the X-ray crystal structure of the 3-I-Tyr^{B26}-Nle^{B29}–insulin hexamer.

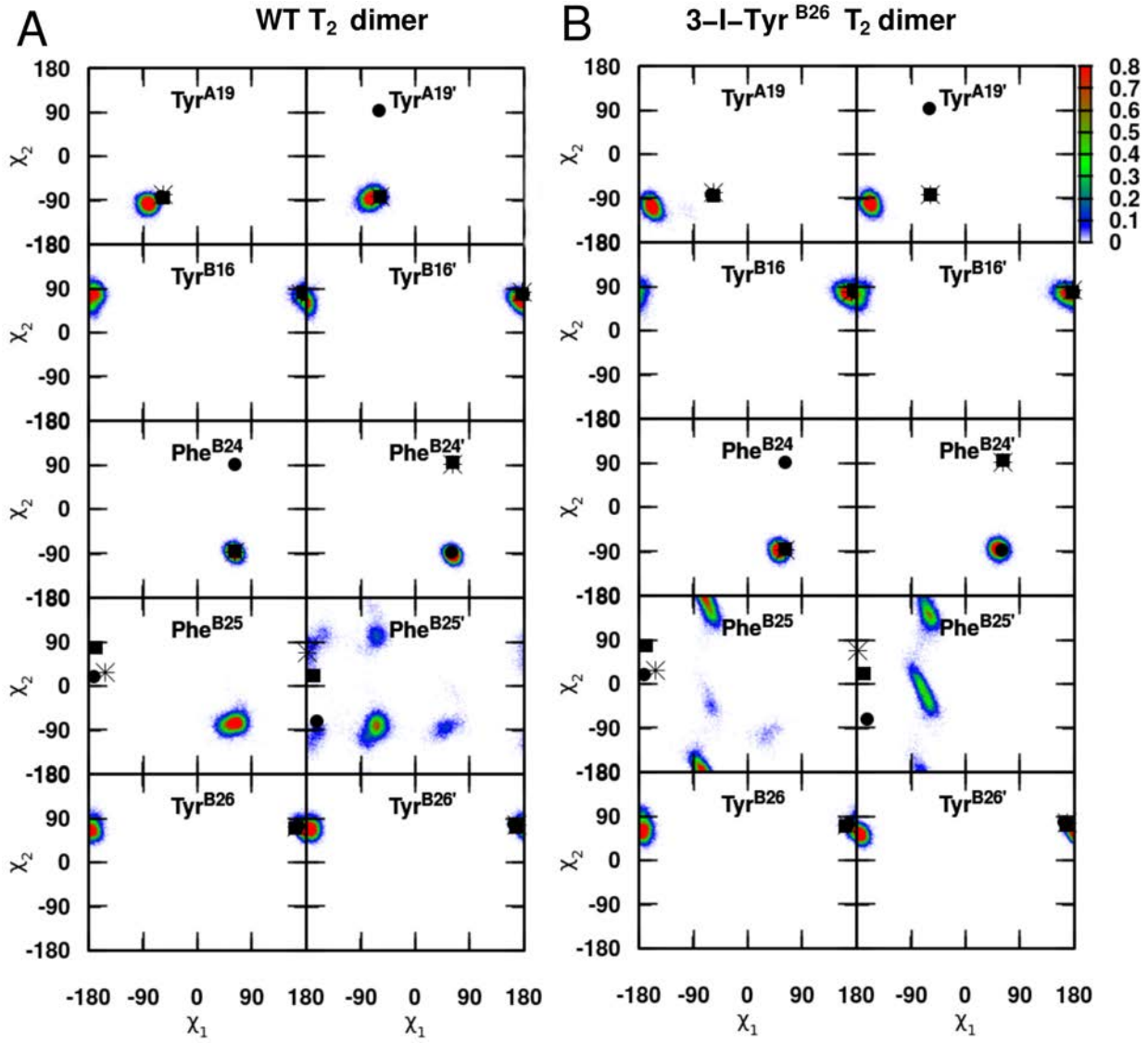

FIGURE S2. **Dihedral probability distribution functions for the T_2 dimers.** Probability distribution functions $P(\chi_1,\chi_2)$ for Tyr^{A19}, Tyr^{B16}, Phe^{B24}, Phe^{B25}, Tyr^{B26} and their dimer-related mates (indicated by *primes*) for (*A*) the WT T_2 dimer, and (*B*) the 3-iodo-Tyr^{B26} T_2 dimer. The distributions were built from 20 ns of equilibrium MD simulations. The starting dimer structure was taken from the T_2 zinc-free dimer WT-insulin structure (PDB code 1DPH). Symbols (*box, star,* and *circle*) indicate positions in the X-ray crystal structure of the 3-I-Tyr^{B26}-Nle^{B29}-insulin hexamer. The side chain of Phe^{B25} is disordered; see main text.

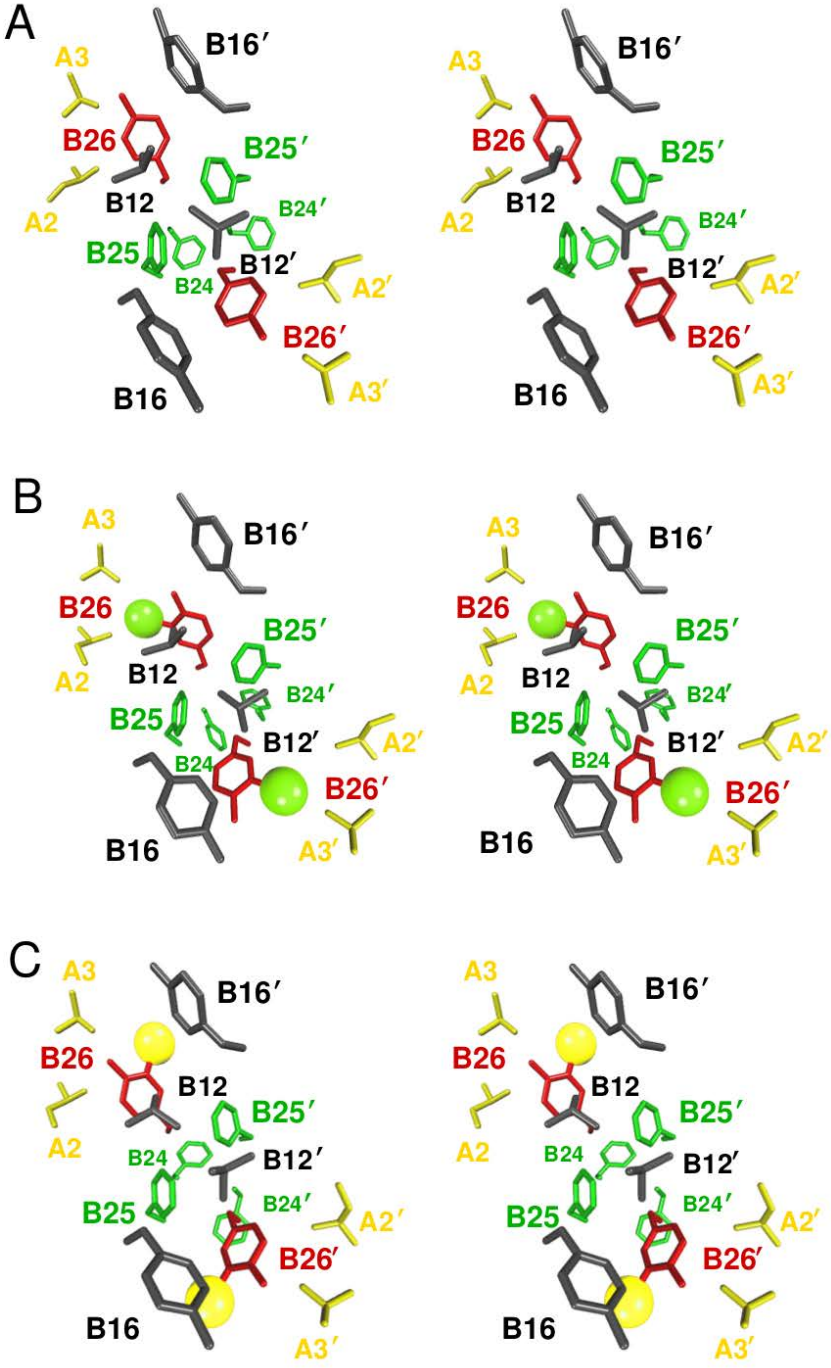

FIGURE S3. **Packing prediction at the dimer interface.** Stereo view of the predicted conformational maps of B-chain aromatic residues from 20 ns MD simulation: (*A*) Packing in the WT-, (*B*) 3-I-Tyr^{B26}- and (*C*) 5-I-Tyr^{B26}-insulin dimers. Residues Ile^{A2}, Val^{A3}, Val^{B12}, Tyr^{B16}, Phe^{B24}, Phe^{B25}, and Tyr^{B26} are shown explicitly (in *licorice*) together with their dimer-related mates. Note the displacement of Val^{B12} in (*C*) due to 5-I-Tyr^{B26}, compared to WT- and 3-I-Tyr^{B26} dimers (see main text). The starting dimer structures were taken from the T_2 zinc-free dimer WT-insulin structure (PDB code 1DPH). The iodine atoms are shown as *green* and *yellow* spheres for 3-I-Tyr^{B26}- and (*C*) 5-I-Tyr^{B26} -insulin dimers, respectively. For an overall view, see also Figure S4.

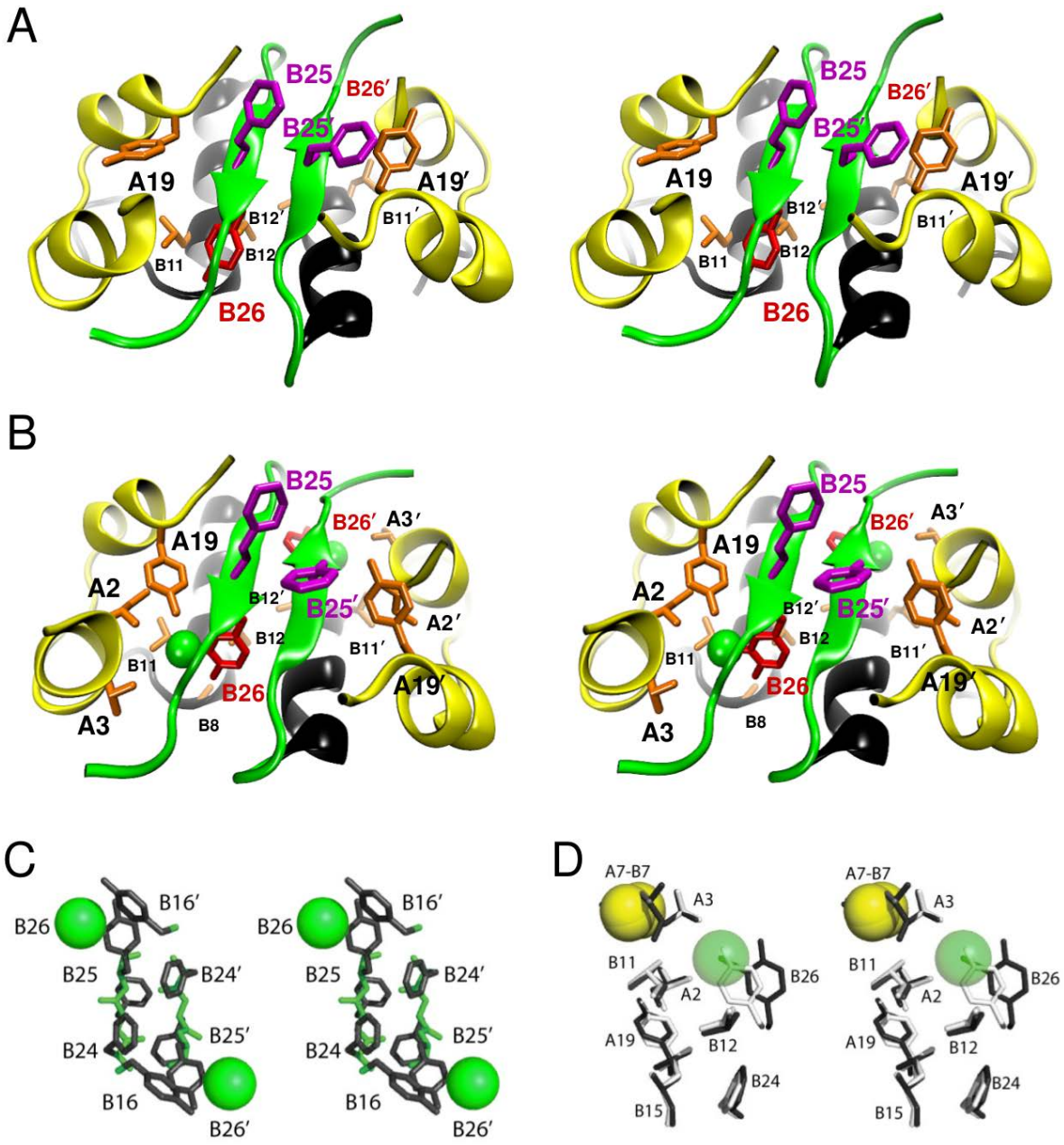

FIGURE S4. **Packing at the dimer interface of 3-I-Tyr^{B26} insulin analog: predicted (A,B) vs. crystal structure (C, D) .** (*A*) WT dimer: Residues Val^{B12}, Tyr^{A19}, Phe^{B25}, and Tyr^{B26} are shown (in *licorice*) together with dimer-related mates. (*B*) 3-I-Tyr^{B26} dimer: Local interactions with residue Ile^{A2}, Val^{A3}, Gly^{B8}, Leu^{B11}, Val^{B12}, and Tyr^{A19} are explicitly shown. Iodine atoms are shown as *green* spheres. For a simplified view, see also Figure S3. (*C*) Stereo view of aromatic-rich dimer interface. The side chains of Tyr^{B16}, Phe^{B24}, Phe^{B25} and 3-I-Tyr^{B26} (*dark gray* sticks) are shown in relation to their dimer-related partners and a portion of the anti-parallel β-sheet (*green*; main chain of residues B24-B26 and B24'-B26'). (*D*) Expanded view of corresponding WT and variant B26 side-chain environments in relation to an inter-chain crevice containing Ile^{A2}, Val^{A3} and Val^{B12}. Neighboring side chains are as labeled; the sulfur atoms of cystine A7-B7 are shown as *yellow* spheres (van der Waals radii). WT coordinates for panels *C* and *D* were obtained from PDB entry 1ZNJ.

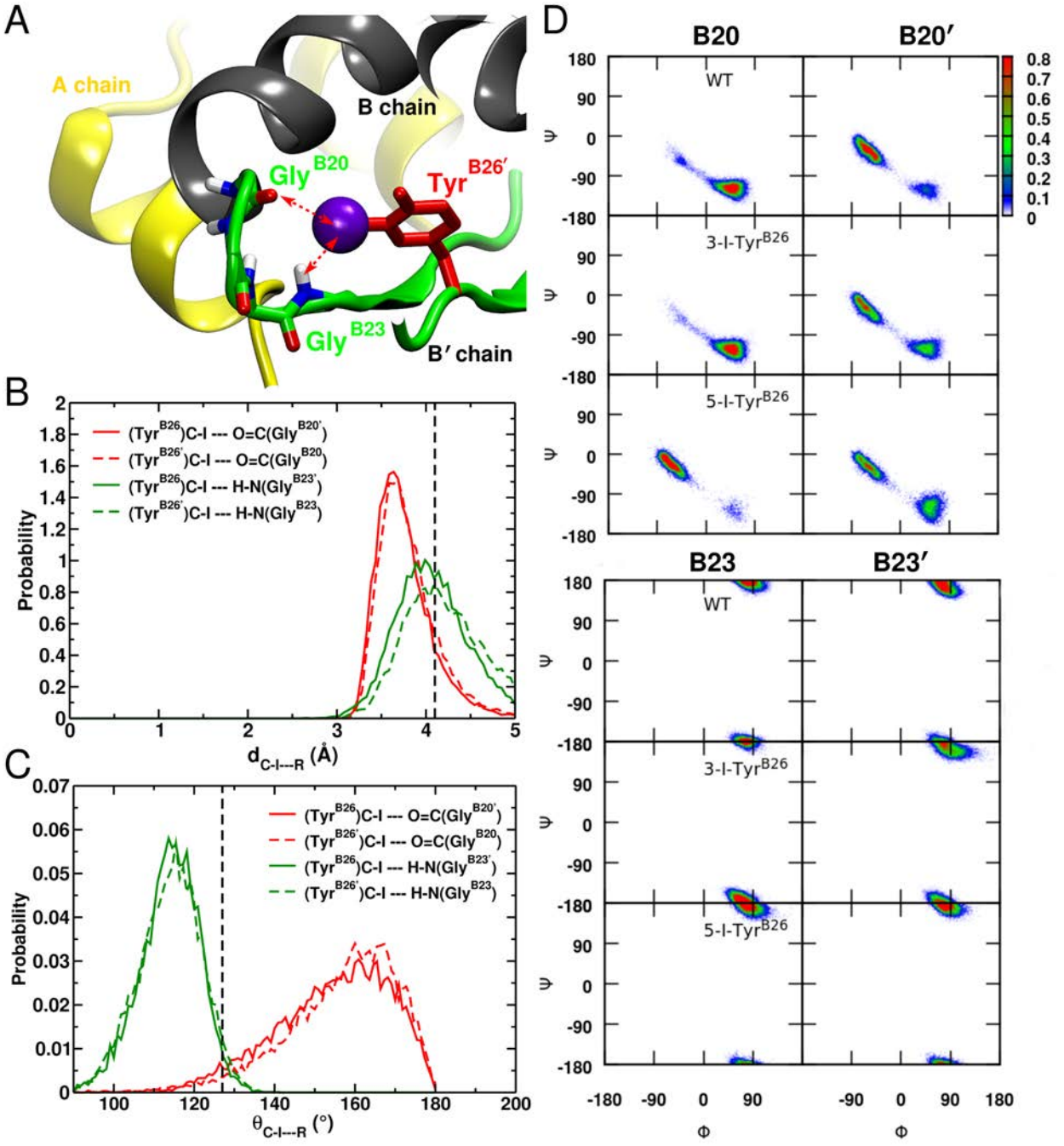

FIGURE S5. **5-I-Tyr^{B26}-insulin dimerization interface.** (*A*) Structure of 5-I-Tyr$^{B26'}$-insulin monomer packing against its related partner illustrates local interaction network between iodine on Tyr$^{B26'}$ and the backbone oxygen of Gly^{B20} and the backbone NH of Gly^{B23}. The iodine atom is shown as a *purple* sphere. Only residues interacting with 5-I-Tyr$^{B26'}$ are illustrated. Potential hydrogen/halogen bonds with I are shown as *dashed red lines*. (*B*) Distance probability distribution of I(Tyr^{B26})—O(Gly$^{B20'}$) (*solid red*), I(Tyr$^{B26'}$)—O(Gly^{B20}) (*dashed red*), I(Tyr^{B26})—HN(Gly$^{B23'}$) (*solid green*), and I(Tyr$^{B26'}$)—HN(Gly^{B23}) (*dashed green*) from 20 ns of MD simulation. The *black dashed line* at 4.1 Å represents the I—O distance interaction limit. (*C*) Angle probability distribution of C-I(Tyr^{B26})—O(Gly$^{B'20'}$) (*solid red*), C-I(Tyr$^{B26'}$)—O(Gly^{B20}) (*dashed red*), C-I(Tyr^{B26})—HN(Gly$^{B23'}$) (*solid green*), and C-I(Tyr$^{B26'}$)—HN(Gly^{B23}) (*dashed green*) from 20 ns MD simulation. The *black dashed line* at 127° represents the angular limit for I between negative electrostatic region ($\delta^- < 127°$) and positive electrostatic region ($127° < \delta^+ <$

233°). (*D*) Predicted backbone dihedral angle distributions (ϕ, ψ) of Gly^{B20} (*upper* panels) and Gly^{B23} (*lower* panels) in the B20-B23 β-turn in WT, 3-I-Tyr^{B26} and 5-I-Tyr^{B26} insulin dimer from 20 ns of MD simulation, and compared to their dimer-related partners (indicated by primes). The starting dimer structure was taken from the T_2 Zinc-free dimer WT-insulin structure (PDB code 1DPH). 5-I-Tyr^{B26} dimer exhibit increased interaction energy along the dimerization interface, compared to WT and 3-I-Tyr^{B26}, but note that the way iodine interacts with the backbone O(Gly^{B20}) leads to its accommodation in a region in the Ramachandran plot that is in principle permitted for glycine but which is empirically unfavorable in the context of the native conformation of insulin (Nakagawa, S. Hua, Q.-X., Jia, W., Wang, S., Katsoyannis, P.G. and Weiss, M.A. (2006) Chiral Mutagenesis of Insulin. CONTRIBUTION OF THE B20-B23 β-TURN TO ACTIVITY AND STABILITY. *J. Biol. Chem*. **281**, 22386-96).

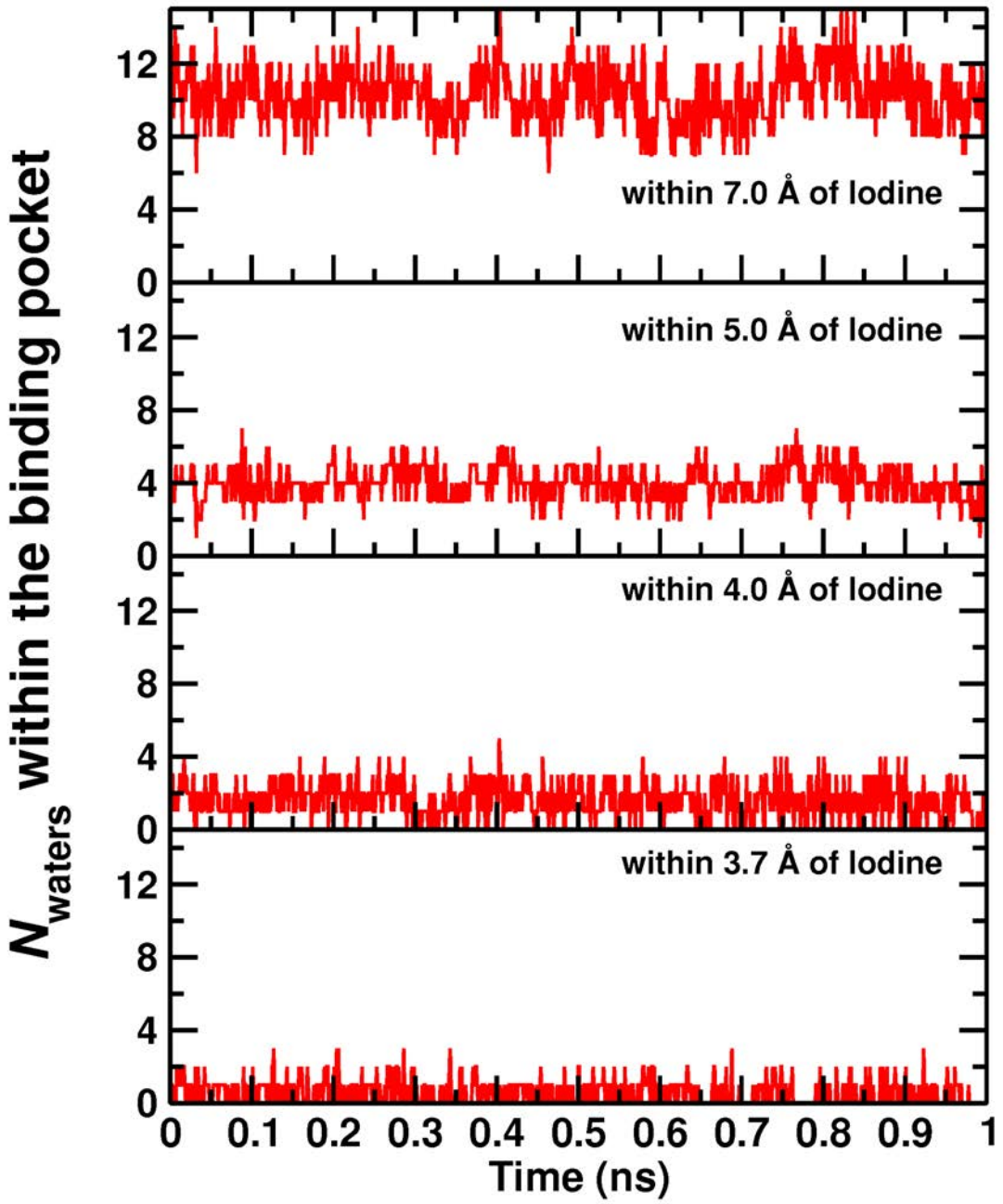

FIGURE S6. **Population of water molecules at the 3-I-Tyr^{B26}/μIR interface.** The number *N* of water molecules present in a 3.7, 4.0, 5.0 and 7.0 Å spheres centered around the iodine atom within the 3-I-Tyr^{B26}/μIR binding pocket during the 1 ns MD simulations.

Table S1. Iodo-Tyr^{B26} –induced interaction energies contributing to the dimerization. Interaction energies E_{total} (sum of van der Waals, E_{vdW}, and electrostatic, E_{elec}, terms) between Tyr^{B26} and neighboring residues calculated for both PC (A) and MTP (B) electrostatics.

(A)

E_{inter}	3-I-Tyr^{B26}			5-I-Tyr^{B26}		
(kcal/mol)	E_{vdW}	E_{elec}	E_{total}	E_{vdW}	E_{elec}	E_{total}
Ile^{A2}	-1.25	-0.06	-1.31	-0.87	-0.02	-0.90
Leu^{A3}	-0.89	-0.18	-1.07	-0.46	-0.18	-0.65
Gly^{B8}	-1.52	-0.41	-1.93	-0.98	0.06	-0.92
Leu^{B11}	-0.94	-0.49	-1.42	-1.13	-0.47	-1.60
Val^{B12}	-2.45	-0.36	-2.81	-2.38	-0.37	-2.75
Leu^{B15}	-0.85	0.02	-0.83	-0.91	0.02	-0.89
Pro^{B28}	-3.53	-0.09	-3.63	-3.23	-0.09	-3.32
Tyr$^{B16'}$	-2.34	-0.89	-3.23	-2.60	-0.32	-2.92
Gly$^{B20'}$	-0.74	-1.77	-2.51	-0.33	-0.04	-0.38
Gly$^{B23'}$	-1.06	0.84	-0.22	-1.39	0.31	-1.08
Phe$^{B24'}$	-2.62	-0.68	-3.30	-3.47	-0.64	-4.11
Total	-18.19	-4.06	-22.25	-17.75	-1.75	-19.50

(B)

E_{inter}	3-I-Tyr^{B26}			5-I-Tyr^{B26}		
(kcal/mol)	E_{vdW}	E_{elec}	E_{total}	E_{vdW}	E_{elec}	E_{total}
Ile^{A2}	-0.95	-0.47	-1.41	-0.54	-0.39	-0.94
Leu^{A3}	-0.84	-0.34	-1.17	-0.29	-0.30	-0.59
Gly^{B8}	-1.31	0.57	-0.73	-0.54	-0.01	-0.55
Leu^{B11}	-0.94	-0.63	-1.58	-0.66	-0.46	-1.12
Val^{B12}	-2.39	-0.91	-3.30	-2.04	-0.83	-2.87
Leu^{B15}	-0.78	-0.41	-1.19	-0.74	-0.38	-1.12
Pro^{B28}	-3.10	-0.89	-3.99	-2.24	-0.78	-3.02
Tyr$^{B16'}$	-2.52	-1.39	-3.91	0.28	-1.96	-1.68
Gly$^{B20'}$	-0.60	-0.92	-1.52	-0.74	-1.73	-2.47
Gly$^{B23'}$	-0.96	-0.60	-1.56	-0.73	-2.16	-2.89
Phe$^{B24'}$	-2.95	-0.02	-2.97	-2.15	-0.14	-2.29
Total	-17.32	-6.01	-23.33	-10.39	-9.14	-19.53

Extending Halogen-based Medicinal Chemistry to Proteins: IODO-INSULIN AS A CASE STUDY

Krystel El Hage, Vijay Pandyarajan, Nelson B. Phillips, Brian J. Smith, John G. Menting, Jonathan Whittaker, Michael C. Lawrence, Markus Meuwly and Michael A. Weiss

J. Biol. Chem. 2016, 291:27023-27041.
doi: 10.1074/jbc.M116.761015 originally published online November 14, 2016