Received <day> <Month>, <2022>; | Revised <day> <Month>, <year>; | Accepted <day> <Month>, <year>

DOI: xxx/xxxx

RESEARCH ARTICLE

Structural Models of Viral Insulin-like Peptides and their Analogs

Viral Insulin-like Peptides

Biswajit Gorai[1] | Harish Vashisth[1,*]

[1]Department of Chemical Engineering, University of New Hampshire, Durham, NH 03824, USA.

Correspondence
*Harish Vashisth, Department of Chemical Engineering, University of New Hampshire, Durham, NH 03824, USA.
Email: harish.vashisth@unh.edu

Abstract

The insulin receptor (IR), the insulin-like growth factor-1 receptor (IGF1R), and the insulin/IGF1 hybrid receptors (hybR) are homologous transmembrane receptors. The peptide ligands, insulin and IGF1, exhibit significant structural homology and can bind to each receptor *via* site-1 and site-2 residues with distinct affinities. The variants of the *Iridoviridae* virus family show capability in expressing single-chain insulin/IGF1 like proteins, termed viral insulin-like peptides (VILPs), which can stimulate receptors from the insulin family. The sequences of VILPs lacking the central C-domain (dcVILPs) are known, but their structures in unbound and receptor-bound states have not been resolved to date. We report all-atom structural models of three dcVILPs (dcGIV, dcSGIV, and dcLCDV1) and their complexes with the receptors (μIR, μIGF1R, and μhybR), and probed the peptide/receptor interactions in each system using all-atom molecular dynamics (MD) simulations. Based on the non-bonded interaction energies computed between each residue of peptides (insulin and dcVILPs) and the receptors, we provide details on residues establishing significant interactions. The observed site-1 insulin/μIR interactions are consistent with previous experimental studies, and a residue-level comparison of interactions of peptides (insulin and dcVILPs) with the receptors revealed that due to sequence differences, dcVILPs also establish some interactions distinct from those between insulin and IR. We also designed insulin analogs and report enhanced interactions between some analogs and the receptors.

KEYWORDS:
viral peptides, insulin, insulin receptor, insulin analogs, molecular dynamics

INTRODUCTION

Insulin (Ins) and insulin-like growth factor-1 (IGF1) are cognate peptides for the insulin receptor (IR) and the insulin-like growth factor-1 receptor (IGF1R), respectively. The binding of these peptides to their cognate receptors regulates key biological processes implicated in glucose homeostasis and cellular growth[1,2,3,4,5,6,7,8,9,10,11,12]. While Ins is a dual-chain (dc) peptide, consisting of an A-chain (21 residues) and a B-chain (30 residues) linked by three disulfide bonds, IGF1 is a single chain (sc) peptide with 62 residues, among which the residues 1-29, 30-41, and 42-62 correspond to the B-, C- and A-domains, respectively. Ins modulates several physiological processes but is primarily responsible for maintaining the glucose level in blood[7,13,14,15]. Insulin deficiency or insulin resistance leads to type 1 or type 2 diabetes mellitus[16,17,18,19,20,21,22]. The uncontrolled level of glucose in the blood can also lead to anxiety, fatigue, seizures, heart disease, kidney disease, blindness, nerve damage, and neurodegenerative Alzheimer's disease[23]. In contrast, IGF1 plays a critical role in development and growth of tissues and organs[24]. The processes underlying normal growth, development, and differentiation of cells are regulated by IGF1. Alteration in IGF1 signaling may also lead to pathological conditions like obesity, hypertension, cardiovascular diseases, and cancer[25,26].

IR and IGF1R are homodimeric glycoproteins of the receptor tyrosine kinase (RTK) superfamily[27]. These glycosylated receptors are comprised of two extracellular α-subunits and two membrane-spanning β-subunits linked by several disufide bridges[28]. The Ins-IR system is highly homologous to IGF1-IGF1R system, but each peptide-receptor system exhibits distinct biological functions[29]. The binding of Ins/IGF1 to the extracellular subunits of their cognate receptors triggers the autophosphorylation in the tyrosine residues of the cytoplasmic tyrosine kinase domain (TKD) of the β-subunit and initiates downstream signaling pathways[7,30]. Several experimental structures of insulin bound to IR have been determined using X-ray crystallography or cryo-electron microscopy (cryo-EM) methods[31,32,33,34,35,36,37,38]. Ins and IGF1 bind to their receptors *via* two binding surfaces known as "site 1" and "site 2". The site 1 of insulin molecule includes residues GlyA1, IleA2, ValA3, GluA4, GlnA5, TyrA19, AsnA21, GlyB8, SerB9, LeuB11, ValB12, TyrB16, GlyB23, PheB24, PheB25, and TyrB26; and the site 2 includes residues SerA12, LeuA13, GluA17, HisB10, GluB13, and LeuB17[13,39]. Similarly, the site 1 of IGF1 includes residues Ala8, Val11, Phe23, Tyr24, Tyr31, Arg36, Arg37, Val44, and Ala62; and the site 2 includes residues Glu9, Asp12, Phe16, Arg21, Asp53, Leu54, Arg56, Glu58, Met59, Tyr60, Lys65, and Lys68[40]. While Ins and IGF1 favorably bind to their cognate receptors, each ligand can also bind to the alternate receptor with a lesser affinity[41].

Given the structural homology between IR and IGF1R, these receptors can also dimerize to form hybrid IR/IGF1R receptors (hybR)[42]. The hybR is a heterotetramer composed of two heterodimers (α/β-subunits) belonging to each receptor[43]. In several mammalian tissues, this receptor is expressed in different proportions along with IR and IGF1R. Notably, the hybR shares a major fraction of receptors in heart, brain, kidney, and skeletal muscles[43,44]. Several studies have shown that IGF1 binds with a higher affinity to hybR in comparison to IR[45,46,47] and thus hybR closely resembles IGF1R. The hybR present in cells exhibits

effective cross-linking between IR and IGF1R[48]. A higher expression of hybR/IGF1R in different cancer cells suggests the association of these receptors with the disease[11,49,50]. However, the unique signaling characteristics and specific physiological roles of the hybR remains enigmatic. Extensive efforts have been invested toward the design of potent fast acting Ins and IGF1 analogs which can bind efficiently to the members of the insulin receptor family[51,52,53,54,55,56,57,58,59,60,61].

Several organisms and pathogens have developed specialized insulin like-peptides for defense, prey, or deter a competitor[62,63,64,65,66]. Viruses are microscopic pathogens which require the intracellular environment of host to survive and replicate. They have evolved critical molecular mimicry mechanisms to express host-like proteins. In a recent study, Altindis et al.[67] used bioinformatics techniques to identify small peptide homologous to Ins or IGF1 in the *Iridoviridae* virus family, termed viral insulin/IGF1-like peptides (VILPs). The VILPs are highly homologous to IGF1 as they are single-chain (sc) peptides and have shown a higher affinity to IGF1R in comparison to IR[68]. In an another study, a double-chain (dc) analog resembling insulin was created by cleaving the C-region of the scVILP, thereby creating a dcVILP[69]. The dcVILPs are dual-chain peptides that stimulate IR and IGF1R phosphorylation *in vitro* and *in vivo*, but they bind with higher affinities to IGF1R than to IR[69]. However, mechanistic details of their interactions with the receptors remain elusive due to the absence of the experimental structures of dcVILPs or their complexes with the receptors[69]. Relevant to this, structural modeling is a promising tool to gain insights into the atomic details of the interactions of insulin-like peptides with the receptors of the insulin family[40]. We have earlier probed the conformational dynamics and interactions between the Ins, IGF1, and IGF2 with the IR, IGF1R, and IGF2R using molecular modeling and molecular dynamics (MD) simulation techniques[70,71,72,73,74].

In our previous study[75], we successfully modeled cone snail insulin (Con-Ins) peptides and studied their interactions with the truncated human insulin receptor (μIR), composed of the first leucine-rich (L1) domain and the C-terminal region of the α-chain (αCT) of the IR, using all-atom MD simulations. The Con-Ins peptides exist in monomeric forms and have shorter sequences compared to Ins. The key Con-Ins residues responsible for their interaction with IR were substituted to derive two insulin analogs. Notably, one of these analogs bound with μIR with a higher affinity than insulin.

Consistent with our previous work, we utilized molecular modeling techniques to predict the tertiary structures of three dcVILPs, including VILPs derived from Grouper Iridovirus (GIV), Singapore Grouper Iridovirus (SGIV), and Lymphocystis disease virus 1 (LCDV1). Using explicit-solvent and all-atom MD simulations, we probed residues involved in the interaction of dcVILPs and Ins with the receptors: IR, IGF1R, and hybR. The non-bonded binding energy contribution of each residue of dcVILPs and Ins with the receptors were judiciously compared to identify favorable substitutions in Ins that may likely enhance the binding affinity of insulin analogs to the receptors. We report further results on four new analogs, three of which are based on residue substitutions identified in this work, and a fourth hybrid analog based on this work and our previous work on cone-snail insulin-like peptides[75].

MATERIALS AND METHODS

Structural modeling of dcVILPs

We obtained the primary sequences of three dcVILPs: dcGIV, dcSGIV, and dcLCDV1, from the previous work by Chrudinová et al.[69]. The alignment of the sequence of each dcVILP with the Ins sequence shows a lower sequence conservation among them, although all cysteine residues are conserved (Figure 1A). The sequence identity between dcLCDV1 and Ins is the highest (∼47%), followed by dcSGIV (∼35%), and dcGIV (∼33%). We used MODELLERv9.20[76] to generate the tertiary structure of each dcVILP based on the template of the structure of Ins (PDB code: 6pxv). Before constructing the tertiary structures of dcVILPs, we modeled any missing residues in the template using MODELLER. We preserved the disulfide bonds and used multi-chain modeling approach to generate the A- and B-chains of dcVILPs during model building. We generated 200 models of each dcVILP and selected the best model based on the lowest discrete optimized protein energy (DOPE) score[77]. Before conducting MD simulations, we used the online portal PROPKA[78] to assign protonation states to the side-chains of ionizable residues at a pH value of 7.

Structural modeling of complexes of dcVILPs with different receptors

We first obtained the structures of the full-length IR ectodomain (PDB code: 6pxv), IGF1R ectodomain (PDB code: 6vwg), and hybR (PDB code: 4xss). From these structures, we have considered a truncated version of each receptor (μ-receptor), specifically, μIR composed of the L1 domain (residues 1-155) and the αCT peptide (residues 704-718), μIGF1R composed of the L1 domain (residues 1-155) and the αCT peptide (residues 691-705), and μhybR composed of the L1 domain (residues 1-155) and the αCT peptide (residues 691-705). Each of these receptor constructs retains the primary binding sites of the insulin family of peptides. In Figure S1, we show the sequence alignment of the constituents of each truncated receptor. We first modeled any missing residues in the initial structures of receptors. We superimposed ligands on the Ins/IGF1 of μ-receptors to obtain a complexed structure of each ligand with each receptor, using the 'align' command in PyMOL[79]. Overall, we modeled 12 ligand-receptor complexes: four peptide ligands (Ins, dcGIV, dcSGIV, and dcLCDV1) docked to three different receptors (μIR, μIGF1R, and μhybR), referred hereafter as IR, IGF1R, and hybR, respectively. Additionally, we constructed five structural models of the complexes of insulin analogs with different receptors (analog-1/μIR, analog-2/μIGF1R, analog-3/μhybR, analog-4/μIR, and truncated-analog-4/μIR; see Table S1).

Simulation setup

We conducted all-atom MD simulations of all modeled ligands as well as ligand/receptor complexes using the GROMACSv2020.4[80,81] software combined with the GROMOS 54A8 force-field[82]. Using the SPC water model[83], we solvated each

ligand or the corresponding ligand-receptor complex in a dodecahedron simulation domain with a minimum buffer of 12 Å along each direction. To neutralize and achieve a physiological salt concentration of 140 mM, we added requisite number of sodium and chloride ions. We used the steepest-descent algorithm[84] to minimize each system. We equilibrated each system for 500 ps at 300 K using Berendsen thermostat with a coupling time of 0.1 ps. We also equilibrated each system at 1 atm pressure using Berendsen barostat for 10 ns. During initial equilibration steps, we applied harmonic restraints (k = 1000 kJ/mol/nm^2) on the heavy atoms of proteins. Finally, we removed all restraints and performed a 500 ns long production simulation of each system in the NPT ensemble. We switched to the V-rescale thermostat with a coupling time of 1.0 ps during the long time-scale MD simulations. We used a nonbonded cut-off of 14 Å with periodic boundary conditions, and calculated the long-range electrostatic interactions using the Particle Mesh Ewald method[85]. We applied all-bond constraints using the LINCS algorithm[86] to use a time-step of 2 fs during simulations. We saved all simulation trajectories every 20 ps and carried out the analyses of MD trajectories with GROMACS and VMD[87]. Each set of simulation was performed in triplicate for each ligand and the ligand-receptor complex, resulting in a total simulation time of ∼32 μs. The details of all MD simulations are summarized in Table S1.

RESULTS

Structural models and all-atom MD simulations of unbound structures of dcVILPs

We first generated the tertiary structures of three dcVILPs (dcGIV, dcSGIV, and dcLCDV1) using multi-chain homology modeling approach[88]. The all-atom root mean squared deviation (RMSD) between the Ins template and each modeled dcVILP structure is less than 1 Å, highlighting the structural consistency with the template used. The initial structural models preserved three helices and three disulfide bonds present in the Ins template.

To investigate the conformational flexibilities of the modeled dcVILPs, we conducted three independent all-atom MD simulations (each 500 ns long) in explicit solvent for each peptide. Using GROMOS clustering algorithm (with an RMSD cut-off of 2 Å), we extracted representative conformers from the last 400 ns of each trajectory for Ins and each dcVILP. The superimposition of each representative conformer to its starting structure (Figure 1B-E) exhibits all-atom RMSD values lower than 3 Å. As expected, the RMSD between the Ins structure and the representative conformer of Ins obtained from MD simulations is the lowest (1.4 Å). Among dcVILPs, the MD-equilibrated modeled structures are also stable with dcLCDV1 (RMSD ∼2.2 Å) and dcGIV (RMSD ∼2.7 Å) showing the lowest and the highest RMSD values, respectively. These data reveal that dcLCDV1 with the highest percentage of sequence identity (∼47%) with the Ins sequence exhibits the lowest RMSD, and dcGIV with the lowest percentage of sequence identity (∼33%) exhibits the highest RMSD.

All-atom MD simulations of ligand/receptor complexes

In Figure 2, we show the representative all-atom structures of all dcVILP/receptor complexes, where the C_α atoms of those dcVILP residues that correspond to site 1 and site 2 residues (smaller and larger spheres, respectively) of insulin are also depicted. For each of the four peptide ligands (Ins, dcGIV, dcSGIV, dcLCDV1), we conducted three independent MD simulations in bound states with three different receptors (IR, IGF1R, and hybR), thereby resulting in 36 simulations of each 500 ns length (totaling 18 μs for all peptide/receptor complexes). In Figure S2, we report for each complex the distributions of the center of mass (COM) distance (first row) and the buried surface area (BSA) (second row) between each peptide and the receptor. The average initial COM separation and the BSA between the peptide/receptor pairs are 2.31 $\pm$ 0.03 nm and 21.21 $\pm$ 0.53 nm^2, respectively. The calculated COM distance from MD simulations (varies between $\sim$2.10 nm to 2.35 nm) is only marginally reduced and the BSA (varies between $\sim$19 nm^2 to 27 nm^2) increased notably from the initial values indicating that the ligand/receptor complexes are stable and each ligand maintained contact with each receptor throughout each simulation. The RMSD distributions of each peptide ligand in the unbound state and in the bound state (Figure S2, third row) reveal that the bound peptides exhibit lower RMSDs in comparison to the unbound peptides. These data also show that the mean-values in distributions for bound peptides are mostly observed between 2.5 nm and 3.5 nm, and that for unbound peptides near 5 nm. The peaks in the RMSD distributions for unbound and bound peptides are distinct, except for dcLCDV1 where a partial overlap of peaks for the hybR complex and the unbound form is observed.

Based on the data from all MD trajectories of each peptide, we further calculated the average change in the root mean squared fluctuation (ΔRMSF) of each residue of each bound peptide relative to the unbound state (Figure S3). The ΔRMSF values for the Ins/receptor complexes are minimal, however the per residue fluctuations of dcVILPs notably reduced after binding to the receptors, as indicated by mostly negative values of ΔRMSF. The ΔRMSF data suggest that the conformational fluctuations of each residue of dcGIV and dcSGIV significantly decreased on binding to the receptors. We observed that most residues in the modeled dcVILPs exhibit negative ΔRMSF values after binding to the receptors, except a few residues in the A-chain and the N-terminus of the B-chain of dcLCDV1 (Figure S3D). These observations further suggest that the reductions in the fluctuations of residues are likely due to favorable interactions of specific residues with the receptors.

Peptide/Receptor interactions

The calculated average non-bonded interaction energy (NBIE, cumulative van der Waals and electrostatic interaction energies) between all residues of each peptide-receptor complex from three independent simulations are summarized in Table S2. A

negative NBIE value implies favorable interactions between each peptide and the receptor. Ins binds to IR with the lowest (–938 kJ/mol) NBIE, and at higher NBIE values with IGF1R (–771 kJ/mol) and hybR (–763 kJ/mol), highlighting a stronger interaction with the cognate receptor IR, and weaker interactions with non-cognate receptors, IGF1R and hybR.

Among dcVILPs, dcGIV (–804 kJ/mol) and dcSGIV (–773 kJ/mol) show moderate interaction energies for IR, but dcLCDV1 exhibits the weakest interaction (an NBIE value of –680 kJ/mol) for IR. With IGF1R, the NBIE values suggest that dcSGIV (–931 kJ/mol) shows the most favorable interaction followed by dcGIV (–835 kJ/mol), Ins (–771 kJ/mol), and dcLCDV1 (–631 kJ/mol). The NBIE values of peptides for hybR follow a trend similar to that of IGF1R: dcSGIV (–1053 kJ/mol) < dcGIV (–925 kJ/mol) < Ins (–763 kJ/mol) < dcLCDV1 (–712 kJ/mol). Thus, Ins exhibits the lowest NBIE and thereby strongest interaction with IR, and dcSGIV with both IGF1R and hybR, and dcLCDV1 exhibits the weakest interaction with each receptor.

Based on our simulation data, we further compared known site 1 insulin/IR interactions (Table S3),[39,13] and predicted the site 1 interactions of dcVILPs with three micro-receptors (Table S4). The observed insulin/IR interacting residue-pairs from our work are in good agreement with those previously reported (Table S3)[39]. Due to conformational dynamics of the residue sidechains, which is captured in MD simulations, we also report a few new insulin/IR interactions (marked by # in Table S3). We further calculated the average interaction energy of each residue of insulin with μIR (Figure S4) and confirm that the site 1 residues, and not the site 2 residues, of insulin significantly interact with μIR. Similarly, we have probed the inter-residue interactions between each modeled dcVILP with different micro-receptors (Table S4 and Table S5). These data show that the inter-residue ligand/receptor interactions for dcGIV and dcSGIV are comparably similar to each other than to dcLCDV1. While our data suggest some similarities with insulin for the site 1 interactions of dcVILPs, the differences are primarily due to their unique sequences (Figure 1A).

We further calculated the change in the NBIE, i.e., $\Delta E = NBIE_{dcVILP\_residue} - NBIE_{Ins\_residue}$, where $NBIE_{dcVILP\_residue}$ and $NBIE_{Ins\_residue}$ are the NBIE contribution of each residue of dcVILPs and Ins to the same receptor, respectively (Figure 3). The ΔE value for each residue is useful in identifying those residues in dcVILPs that have stronger interactions with the receptors in comparison to insulin. A negative ΔE value signifies that a specific residue of insulin exhibits lesser or no interactions with the 'site 1' of the receptor in comparison to the equivalent dcVILP residue, or vice-versa that the dcVILP residue exhibits stronger interactions. Overall, dcVILPs have weaker interactions (higher interaction energies) with IR than Ins (Table S2), although a few residues in dcVILPs exhibit stronger interactions (lower ΔE values) than the Ins residues (Figure 3, first row). Specifically, the residues AspA17 (–40.4 kJ/mol), AspB13 (–45.1 kJ/mol), and ArgB30 (–30.7 kJ/mol) of dcGIV complexed with IR exhibit lower ΔE values which reveals that the A17, B13, and B30 residues of Ins do not establish significant interactions with IR, but that of dcGIV do. Similarly, we observed that the key interacting residues of dcSGIV to IR are congruent to those noted for dcGIV. Moreover, GluB17 of dcSGIV also shows a stronger interaction (ΔE = –34.9 kJ/mol) with IR than Ins does. However, only a single residue of dcLCDV1, SerB8, exhibits a significantly stronger interaction (ΔE = –39.4 kJ/mol) with IR in comparison

to Ins. It is noteworthy that the residues AspA17, AspB13, GluB17, and ArgB30 are not conserved between insulin and two dcVILPs (dcGIV and dcSGIV), and similarly SerB8 is not conserved among insulin and dcLCDV1 (Figure 1A).

With IGF1R (Figure 3, second row), the residues AspA17 (–38.3 kJ/mol), SerB28 (–51.7 kJ/mol), and ArgB30 (–83.4 kJ/mol) of dcGIV, and the residues AspA17 (–41.0 kJ/mol) and ArgB30 (–83.4 kJ/mol) of dcSGIV, exhibit stronger interactions (negative ΔE values). We point out that the dcGIV residue SerB28 is ProB28 in insulin. For dcLCDV1, SerB8 is the only residue showing a significantly stronger interaction (negative ΔE value) with IGF1R. For hybR (Figure 3, third row), three residues AspB13 (–28.9 kJ/mol), GluB17 (–32.8 kJ/mol), and ArgB30 (–34.6 kJ/mol) of dcGIV, and two residues AspA17 (–38.0 kJ/mol) and AspB13 (–77.8 kJ/mol) residues of dcSGIV, exhibit stronger interactions (negative ΔE values) than Ins. The dcLCDV1 residues SerB8 (–36.4 kJ/mol) and ArgB30 (–48.2 kJ/mol) also show significant interactions with hybR in comparison to Ins. These results indicate that the residues in dcVILPs which are distinct from Ins residues may account for differences in their binding affinities to the receptors. These residues are potential candidates as reasonable substitutions in Ins and insulin-like peptides to design new peptide variants targeting the receptors.

In Figure 4, we show unique residue-level interactions observed in the dcVILP/receptor complexes, and also provide a list of contacts in Table S4. The dcGIV residues AspA17 and AspB13 establish stable salt-bridging (SB) interactions with Arg717 of the αCT peptide, and anion-π interactions with Tyr67, respectively, of the IR. The C-terminal residues of the B-chain of dcGIV form hydrogen-bonding (HB) interactions with the residues Gly10, Asp12, and Glu22 in IR. We observed that the acidic residues AspA17, AspB13, and GluB17 of dcSGIV also form SB interactions with Arg717 of the αCT peptide, HB with Phe39 of the L1 domain, and HB with Lys40, respectively. The dcSGIV's basic residue ArgB30 interacts with the residues Glu6 and Asn25 of IR. Only a SerB8 residue of dcLCDV1 forms HB with the Arg65 residue of the L1 domain of IR.

In dcVILP/IGF1R complexes, the A-chain residue AspA17 of dcGIV exhibits a SB interaction with the Arg704 residue of the αCT peptide of IGF1R. The B-chain residues SerB28 and ArgB30 of dcGIV form HB interactions with the Asp8 and Arg18 of the L1 domain of IGF1R, respectively. The ArgB30 of dcGIV also interacts with Gln15 and Glu20 of the L1 domain of IGF1R. The dcSGIV residue AspA17 forms a SB interaction with Arg704 of IGF1R, and ArgB30 interacts with the residues Gly6, Glu26, and Tyr28 of the L1 domain. In dcLCDV1, SerB8 establishes an HB interaction with His697 of the αCT peptide of IGF1R.

The dcVILPs also show significant binding interactions with hybR. Among the crucial residues of dcGIV, AspB13 and GluB17 form HB interactions with the residues Tyr67 and Lys60 of the L1 domain of hybR, respectively. The terminal residue ArgB30 of dcGIV exhibits a SB interaction with the residue Asp12 of the L1 domain of hybR. Only two residues of dcSGIV, belonging to each chain, show significant relative NBIE to hybR (third row in Figure 3B). The A-chain residue AspA17 of dcSGIV establishes a HB interaction with the residue Arg704 of the αCT peptide of hybR, and the B-chain residue Asp13 forms

HB interactions with Arg65 and Tyr67 of the L1 domain of hybR. In dcLCDV1, SerB8 forms favorable interactions with the residues Arg65 and Glu97 of the L1 domain of hybR.

Binding energies of insulin analogs

To rationalize the affinity of analogs derived by substituting equivalent residues of Ins by the key residues of dcVILPs (Figure 5A) underlying stronger interactions with the receptors, we generated their structural models in complex with the respective receptors and performed three independent all-atom MD simulations (Table S1). The analog 1 was derived from the interactions of dcVILPs with the IR, constituted by substituting GluA17, GlyB8, GluB13, LeuB17, and ThrB30 of Ins by Asp, Ser, Asp, Glu, and Arg, respectively. The analog 2 was derived from the interactions of dcVILPs with the IGF1R, constituted by substituting GluA17, GlyB8, ProB28, and ThrB30 of Ins by Asp, Ser, Ser, and Arg, respectively. Similarly, based on the interactions of dcVILPs with hybR, analog 3 was obtained by substituting GluA17, GlyB8, GluB13, LeuB17, and ThrB30 of Ins by Asp, Ser, Asp, Glu, and Arg, respectively. Notably, analog 1 and analog 3 are derived from the interactions of dcVILPs with IR and hybR, however, both have identical residue substitutions in Ins. We have termed them as distinct analogs to distinguish the fact that they are derived from different receptors.

Each analog in complex with their corresponding receptors are subjected to triplicate all-atom MD simulations, each 500 ns long. The average NBIE is then estimated from three independent simulations of each analog with their receptors. The average NBIE of analog 1 (–935 kJ/mol) for IR is comparable to the NBIE of Ins-IR (–938 kJ/mol) complex. However, the average NBIE of analog 2 (–891 kJ/mol) and analog 3 (–910 kJ/mol) are significantly lower in comparison to the NBIE of Ins-IGF1R (–771 kJ/mol) and Ins-hybR (–763 kJ/mol) complexes, respectively. We also estimated the difference of the NBIE (ΔE) of each analog residue from the NBIE of respective residues of Ins complexed with the same receptor (Figure 5B-D). Among the substituted residues, we observed AspA17 as a favorable substitution for GluA17 of Ins which enhances the binding interaction of each analog to their corresponding receptors. We also noted that substituting the residue ThrB30 by ArgB30 enhances the interaction of analog 2 to IGF1R.

In our previous study [75], we probed the interactions of the insulin-like peptides from different variants of cone snails with μIR using MD simulations. We identified key residues interacting with the μIR and reported a shorter insulin analog, lacking the last eight residues of the B-chain of Ins, that binds to the receptor with an enhanced affinity. We discovered that mutating residues LeuB15 and ArgB22 of Ins to Tyr and Asn, respectively, increased the binding interactions of analogs derived from the cone snail peptides with the μIR. Combining observations from our previous work and this work, we created the structural model of 'analog 4' by substituting residues GluA17, LeuB15, and ArgB22 of Ins by Asp, Tyr, and Asn, respectively (Figure S5A). We performed three independent MD simulations of the hybrid analog 4 in complex with IR and obtained the average NBIE value of –984 kJ/mol. Notably, the average NBIE of analog 4 is lower than the average NBIE of analogs derived from the cone snail

peptides[75] and dcVILPs, as well as of Ins for the IR. The difference of the NBIE (ΔE) of each residue of analog 4 from the NBIE of equivalent residue of Ins (Figure S5B) verifies significant contribution of the residues AspA17 and TyrB15 substitutions in Ins for enhancing the binding interaction with the IR. We further performed atomistic simulations of a truncated form of hybrid analog 4 (Figure S6A) in complex with the IR (three independent runs of each 500 ns long) and calculated the total NBIE between the truncated analog and the receptor (Figure S6B). The average NBIE calculated between Ins and IR in the insulin–IR complex is –938 kJ/mol, and between the truncated analog 4 and IR in the complex is –922 kJ/mol. The NBIE are comparable and support the fact that the last three residues of the B-chain of insulin do not contribute significantly for interactions with the site 1 of IR.

DISCUSSION

Certain Iridovirus peptides, termed viral insulin-like peptides (VILPs), have emerged as promising peptides for designing novel insulin/IGF1 mimics targeting receptors[67]. The VILPs are single chain proteins which possess agonist or antagonist properties with high potency toward IR and IGF1R[68]. The double chain variants of VILPs (dcVILPs), which are lacking the C-domain, can stimulate both human receptors, IR and IGF1R[69]. The dcVILPs also have the ability to modulate the glucose level in specific tissues. In particular, dcGIV enhances the Akt phosphorylation and glucose transporter type 4 gene expression in white adipose tissue[69].

In current study, we have performed structural modeling and all-atom MD simulations of three dcVILPs (dcGIV, dcSGIV, and dcLCDV1) in complex with three different micro-receptors (μIR, μIGF1R, and μhybR) to characterize the key residues of dcVILPs significant for their interactions with the receptors. The primary sequence of dcLCDV1 closely resembles to Ins than to the sequences of the peptides dcGIV and dcSGIV. Since the experimental structures of dcVILPs or their complexes with the receptors are not available, we have modeled their all-atom structures in this work and studied their structural stabilities using extensive all-atom MD simulations. Based on the Ins-IR complex (PDB code: 6pxv), we used the Ins structure as a template to predict the structures of dcVILPs and their complexes with the receptors. We observed that the unbound and bound conformers of dcVILPs are stable during MD simulations with subtle fluctuations in the loop regions. The RMSD distributions (Figure S2, third row) showed that the bound structures of dcVILPs resemble more native like conformers than the unbound conformers. The ΔRMSF per residue, which indicates the average change in the residue fluctuation of the peptide ligands in the bound state relative to that from the unbound state, demonstrates negligible variation for Ins than dcVILPs. The higher ΔRMSF per residue of dcVILPs is likely due to their conformational rearrangements in their bound states with the receptors. The negative ΔRMSF per residue value of dcGIV and dcSGIV implies a dampening of residual fluctuations in the bound states. However, the positive

ΔRMSF per residue values of a few A-chain residues and a few N-terminal residues of the B-chain of dcLCDV1 in all three complexes demonstrate the region to be more flexible than in the unbound form.

We further calculated the average NBIE between each dcVILP and different receptors based on our MD simulations (Table S2). Consistent with the reported experimental K_d values[69], we observed a similar trend of the average NBIE between the peptide ligands and the receptors IR and IGF1R. Insulin is known to bind with a higher affinity to IR than to IGF1R, and we observed a stronger interaction (lower interaction energy) with IR (–938 kJ/mol) than with IGF1R (–771 kJ/mol). For IR, we also observed that dcGIV and dcSGIV had weaker binding energies than insulin, but for IGF1R, dcGIV and dcSGIV had stronger binding energies than insulin. In comparison to other dcVILPs, dcLCDV1 is known to have the lowest affinity for IGF1R[69]. Consistent with this, we also observed that the dcLCDV1/IGF1R interaction is the weakest (–631 kJ/mol) in comparison to the interactions between other two dcVILPs and IGF1R (Table S2). However, we note that the single-chain form of LCDV1 (scLCDV1) showed the highest binding affinity (among VILPs) for IGF1R[67], and it is identified as a potent antagonist of IGF1R[68]. Intriguingly, the activity of the LCDV1 peptide to the receptor is substantially reduced after losing its C-domain in the dc-form[69]. These observations emphasize the significance of the C-domain of LCDV1 for its binding affinity toward the receptors.

The heteromeric hybrid InsR/IGF1R receptors are widely distributed in mammalian tissues, particularly in brain, kidney, skeletal muscle, and heart where they exist in large proportions[43,44]. The structure of hybR (PDB code: 4xss) is composed of the L1 domain of IR and the αCT peptide of IGF1R[42]. For hybR, we observed that dcSGIV exhibits the strongest interactions among all peptide ligands, followed by dcGIV, Ins, and dcLCDV1. We note that dcVILPs, dcSGIV and dcGIV, bind to IGF1R and hybR with higher affinities, whereas, Ins shows a higher binding affinity for IR. Our results suggest that dcVILPs resemble closely with IGF1 and are consistent with the reported experimental studies which demonstrate the preference of hybR for IGF1 than Ins[42,89].

It is important to note that the primary sequences of Ins and dcVILPs differ significantly (Figure 1A). The multiple sequence alignment demonstrates that the dcVILP residues at the equivalent positions belonging to site 1 of Ins also vary notably. Specifically, only 5 site-1 residues of insulin (GlyA1, TyrA19, AsnA21, LeuB11, and GlyB23) and corresponding residues in dcVILPs at equivalent positions are fully conserved. We list residue-level contacts between insulin and IR in Table S3, and those between dcVILPs and three receptors IR, IGF1R, and hybR in Table S4. In Table S5, we also report the interaction energies of residues belonging to site 1 of Ins with different receptors and the equivalent residues of dcVILPs with the receptors. The interaction energy of each insulin residue with μIR is shown in Figure S4. For each dcVILP/receptor complex, the difference in the average interaction energy of a given dcVILP residue relative to the corresponding residue in insulin is shown in Figure 3.

These data confirm that insulin/μIR contacts observed in our simulations are consistent with those reported previously (Table S3)[39] and the site-1 residues of insulin have stronger interactions with μIR than the site 2 residues (Figure S4). Moreover, the receptor interaction energies of 5 site-1 residues of insulin (GlyA1, TyrA19, AsnA21, LeuB11, and GlyB23), that are

conserved in dcVILPs (Figure 1A), are also comparable for each receptor (Figure 3 and Table S5). The differences in interactions of dcVILPs and insulin with different receptors are also observed, especially at those positions where the residue between a dcVILP and insulin is not conserved. For example, dcGIV showed interactions stronger than insulin (a negative ΔE value in Figure 3) with different receptors for the residues AspA17 (GluA17 in insulin), AspB13 (GluB13 in insulin), GluB17 (LeuB17 in insulin), SerB28 (ProB28 in insulin), and ArgB30 (ThrB30 in insulin). For dcSGIV as well, stronger interactions than insulin are observed with different receptors for the residues AspA17, AspB13, GluB17, and ArgB30, which are conserved among dcSGIV and dcGIV (Figure 1A). While it has been suggested that SerB28 of dcGIV could explain its lower affinity than dcSGIV for the IR-A isoform[69], we do not see any significant differences in the interaction energies of SerB28 (dcGIV) or ProB28 (dcSGIV) with μIR (Figure 3). However, for μIGF1R, dcGIV (but not dcSGIV) showed a stronger interaction (a negative ΔE value in Figure 3) than insulin at SerB28. For dcLCDV1, only one non-conserved residue SerB8 (GlyB8 in insulin) showed interactions stronger than insulin for each of the three receptors.

It was pointed out earlier[69] that several residues that are involved in insulin binding to the site 2 of IR are replaced in two dcVILPs (dcGIV and dcSGIV) by residues that are identical to IGF-1. For example, these are dcGIV/dcSGIV residues GluB10 (Glu9 in IGF-1 and HisB10 in insulin), AspB13 (Asp12 in IGF-1 and GluB13 in insulin), and AspB21 (Asp20 in IGF-1 and GluB21 in insulin). It was further noted that two of these three residues of IGF-1 (Glu9 and Asp12) are involved in site 1 binding to IGF1R, and also that the dcGIV/dcSGIV peptides may preferentially bind to the site 1 of IGF1R than to site 2 of IR[69]. We also observed that three conserved residues among dcGIV and dcSGIV (AspA17, AspB13, and GluB17), that correspond to the site 2 residues in insulin (GluA17, GluB13, and LeuB17), show stronger site-1 interactions with μIGF1R and/or μIR in our structural models (Figure 3). Specifically, we find that the residue AspA17 (GluA17 in insulin) has stable salt-bridging interactions with Arg717 of αCT in IR and with homologous residue Arg704 of αCT in IGF1R. We analyzed several crystal structures of insulin/IR complexes and found that the residue GluA17 of insulin (AspA17 in dcGIV/dcSGIV) is located in proximity of Arg717 of αCT in IR (Figure S7), thereby suggesting the possibility of this salt-bridging interaction.

Based on these results, we designed several insulin analogs by substituting those residues in insulin where dcVILPs exhibited stronger binding to the receptors than the equivalent Ins residues. Based on the interactions of dcVILPs with IR, IGF1R, and hybR, we designed three insulin analogs and computed their average binding energies with μIR, μIGF1R, and μhybR from three independent MD simulations (Figure 5). The analog 1 binds to IR with a binding energy comparable to that of Ins for IR, but the analogs 2 and 3 demonstrate enhanced binding energies for IGF1R and hybR, respectively, than that of Ins for both of these receptors. We observed that mutating the insulin residue GluA17 to AspA17 (which forms a salt-bridging interaction with αCT; Figure S7) contributed to enhanced binding energy of these analogs.

In our previous work[75], we studied the interactions between IR and insulin-like peptides from venoms of fish hunting cone snails[66,90,91,92,93] and reported two insulin analogs with enhanced binding energies. We derived six key substitutions from the

cone snail insulin and reported that the mutations LeuB15Tyr and ArgB22Asn in insulin are promising. Notably, each key substitution identified from the cone snail insulins is localized in the B-chain of Ins. Moreover, based on the current study, we identified a favorable substitution in the A-chain (GluA17Asp) in insulin. Taken together, we proposed a hybrid analog that has features of both key interacting residues of cone snail insulin-like peptides and dcVILPs. Importantly, the hybrid analog exhibited a higher binding energy for μIR in comparison to analogs derived from cone snail peptides and dcVILPs, and also higher than Ins (Figure S5). We did not observe a significant change in the IR binding energy (Figure S6) of a truncated form of this hybrid analog lacking the B28-B30 sequence, thereby suggesting that these residues are not contributing significantly to the interaction of this analog with IR. Briefly, we note that in our previous study[75] and this work, only the site-1 containing domains of the receptors (L1/αCT) are included to study the binding interactions of peptides, but including the fibronectin domains in future studies may offer further enhanced understanding of the site 2 interactions of these peptides.

CONCLUSION

We report all-atom structural models of three dcVILPs (dcGIV, dcSGIV, and dcLCDV1) and their complexes with the micro-receptors (μIR, μIGF1R, and μhybR). We further conducted all-atom MD simulations based on these structural models to understand the conformational flexibility and the stability of each dcVILP in unbound and receptor-bound states. We observed that the ligand/receptor complexes are stable, and each dcVILP engages in interactions similar to insulin for conserved residues as well as in some unique interactions for non-conserved residues. The overall trends in interaction energies of dcVILPs with IR and IGF1R are consistent with experimentally measured affinities for these receptors. Specifically, dcGIV and dcSGIV have stronger interactions with IGF1R, and dcLCDV1 has the weakest (among all dcVILPs) interaction with each receptor. Our results also show that some residues in dcVILPs (e.g., AspA17, AspB13, and GluB17) that correspond to site 2 residues in insulin show stronger site 1 interactions due to unique residue contacts (e.g., AspA17-Arg717/Arg704 salt-bridging interactions in dcGIV/dcSGIV and αCT). We further designed four insulin analogs and reported enhanced interactions between some analogs and the receptors. We suggest that the findings from our study could offer future possibilities for developing new therapies for treatment of diseases related to proteins of the insulin receptor family.

NOTES

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We acknowledge financial support from the National Institutes of Health (NIH) through grant R35GM138217 (H.V). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. We thank computational support through the following resources: Premise, a central shared high-performance computing cluster at the University of New

Hampshire supported by the Research Computing Center; and BioMade, a heterogeneous CPU/GPU cluster supported by the NSF EPSCoR award (OIA-1757371).

References

1. Ward CW, Lawrence MC. Ligand-induced activation of the insulin receptor: a multi-step process involving structural changes in both the ligand and the receptor. *Bioessays* 2009; 31(4): 422–434.
2. Adams TE, Epa V, Garrett T, Ward C. Structure and function of the type 1 insulin-like growth factor receptor. *Cell Mol Life Sci* 2000; 57(7): 1050–1093.
3. Lou M, Garrett TP, McKern NM, et al. The first three domains of the insulin receptor differ structurally from the insulin-like growth factor 1 receptor in the regions governing ligand specificity. *Proc Natl Acad Sci USA* 2006; 103(33): 12429–12434.
4. McKern NM, Lawrence MC, Streltsov VA, et al. Structure of the insulin receptor ectodomain reveals a folded-over conformation. *Nature* 2006; 443(7108): 218–221.
5. Keyhanfar M, Booker GW, Whittaker J, Wallace JC, Forbes BE. Precise mapping of an IGF-I-binding site on the IGF-1R. *Biochem J* 2007; 401(1): 269–277.
6. Ward CW, Lawrence MC. Landmarks in insulin research. *Front Endocrinol* 2011; 2: 76.
7. De Meyts P. Insulin and its receptor: structure, function and evolution. *Bioessays* 2004; 26(12): 1351–1362.
8. Werner H, Sarfstein R, Laron Z. The Role of Nuclear Insulin and IGF1 Receptors in Metabolism and Cancer. *Biomolecules* 2021; 11(4): 531.
9. Cignarelli A, Genchi VA, Perrini S, Natalicchio A, Laviola L, Giorgino F. Insulin and insulin receptors in adipose tissue development. *Int J Mol Sci* 2019; 20(3): 759.
10. Whittaker J, Whittaker LJ, Roberts CT, et al. α-Helical element at the hormone-binding surface of the insulin receptor functions as a signaling element to activate its tyrosine kinase. *Proc Natl Acad Sci USA* 2012; 109(28): 11166–11171.
11. Singh P, Alex JM, Bast F. Insulin receptor (IR) and insulin-like growth factor receptor 1 (IGF-1R) signaling systems: novel treatment strategies for cancer. *Med Oncol* 2014; 31(1): 805.
12. Gorai B, Vashisth H. Progress in simulation studies of insulin structure and function. *Front. Endocrinol.* 2022; 13: 908724.
13. Lawrence MC. Understanding insulin and its receptor from their three-dimensional structure. *Mol Metab* 2021; 52: 101255.
14. Schäffer L. A model for insulin binding to the insulin receptor. *Eur J Biochem* 1994; 221(3): 1127–1132.
15. Rahman MS, Hossain KS, Das S, et al. Role of insulin in health and disease: an update. *International Journal of Molecular Sciences* 2021; 22(12): 6403.

16. Chen C, Cohrs CM, Stertmann J, Bozsak R, Speier S. Human beta cell mass and function in diabetes: recent advances in knowledge and technologies to understand disease pathogenesis. *Mol Metab* 2017; 6(9): 943–957.

17. Alam U, Asghar O, Azmi S, Malik RA. General aspects of diabetes mellitus. *Handbook of Clinical Neurology* 2014; 126: 211–222.

18. Kaul K, Tarr JM, Ahmad SI, Kohner EM, Chibber R. Introduction to diabetes mellitus. *Diabetes* 2013; 771: 1–11.

19. Tisch R, McDevitt H. Insulin-dependent diabetes mellitus. *Cell* 1996; 85(3): 291–297.

20. Lamster IB, Lalla E, Borgnakke WS, Taylor GW. The relationship between oral health and diabetes mellitus. *J Am Dent Assoc* 2008; 139: 19S–24S.

21. Bellamy L, Casas JP, Hingorani AD, Williams D. Type 2 diabetes mellitus after gestational diabetes: a systematic review and meta-analysis. *Lancet* 2009; 373(9677): 1773–1779.

22. Bach JF. Insulin-dependent diabetes mellitus as an autoimmune disease. *Endocr Rev* 1994; 15(4): 516–542.

23. Berlanga-Acosta J, Guillén-Nieto G, Rodríguez-Rodríguez N, et al. Insulin resistance at the crossroad of Alzheimer disease pathology: A review. *Frontiers Endocrinol* 2020; 11: 560375.

24. LeRoith D, Holly JM, Forbes BE. Insulin-like growth factors: Ligands, binding proteins, and receptors. *Mol Metab* 2021; 52: 101245.

25. Shanmugalingam T, Bosco C, Ridley AJ, Van Hemelrijck M. Is there a role for IGF-1 in the development of second primary cancers?. *Cancer Med* 2016; 5(11): 3353–3367.

26. Hua H, Kong Q, Yin J, Zhang J, Jiang Y. Insulin-like growth factor receptor signaling in tumorigenesis and drug resistance: a challenge for cancer therapy. *J Hematol Oncol* 2020; 13: 1–17.

27. Lemmon MA, Schlessinger J. Cell signaling by receptor tyrosine kinases. *Cell* 2010; 141(7): 1117–1134.

28. Cabail MZ, Li S, Lemmon E, Bowen ME, Hubbard SR, Miller WT. The insulin and IGF1 receptor kinase domains are functional dimers in the activated state. *Nat Commun* 2015; 6(1): 1–8.

29. Lawrence MC, McKern NM, Ward CW. Insulin receptor structure and its implications for the IGF-1 receptor. *Curr Opin Struct Biol* 2007; 17(6): 699–705.

30. White M, Kahn CR. Insulin action at a molecular level–100 Years of progress. *Mol Metab* 2021; 52: 101304.

413 31. Menting JG, Whittaker J, Margetts MB, et al. How insulin engages its primary binding site on the insulin receptor. *Nature* 414 2013; 493(7431): 241–245.

415 32. Menting JG, Yang Y, Chan SJ, et al. Protective hinge in insulin opens to enable its receptor engagement. *Proc Natl Acad* 416 *Sci USA* 2014; 111(33): E3395–E3404.

417 33. Scapin G, Dandey VP, Zhang Z, et al. Structure of the insulin receptor–insulin complex by single-particle cryo-EM analysis. 418 *Nature* 2018; 556(7699): 122–125.

419 34. Weis F, Menting JG, Margetts MB, et al. The signalling conformation of the insulin receptor ectodomain. *Nat Commun* 420 2018; 9(1): 1–10.

421 35. Uchikawa E, Choi E, Shang G, Yu H, Bai Xc. Activation mechanism of the insulin receptor revealed by cryo-EM structure 422 of the fully liganded receptor–ligand complex. *Elife* 2019; 8: e48630.

423 36. Gutmann T, Schäfer IB, Poojari C, et al. Cryo-EM structure of the complete and ligand-saturated insulin receptor 424 ectodomain. *J Cell Biol* 2020; 219(1).

425 37. Nielsen J, Brandt J, Boesen T, et al. Structural investigations of full-length insulin receptor dynamics and signalling. *J Mol* 426 *Biol* 2022; 434(5): 167458.

427 38. Li J, Park J, Mayer JP, et al. Synergistic activation of the insulin receptor via two distinct sites. *Nat Struct Mol Biol* 2022; 428 29(4): 357–368.

429 39. De Meyts P. Insulin/receptor binding: the last piece of the puzzle? What recent progress on the structure of the 430 insulin/receptor complex tells us (or not) about negative cooperativity and activation. *Bioessays* 2015; 37(4): 389–397.

431 40. Vashisth H. Theoretical and computational studies of peptides and receptors of the insulin family. *Membranes* 2015; 5(1): 432 48–83.

433 41. De Meyts P, Sajid W, Palsgaard J, et al. Insulin and IGF-I receptor structure and binding mechanism. In: Springer. 2007 434 (pp. 1–32).

435 42. Menting JG, Lawrence CF, Kong GKW, Margetts MB, Ward CW, Lawrence MC. Structural congruency of ligand binding 436 to the insulin and insulin/type 1 insulin-like growth factor hybrid receptors. *Structure* 2015; 23(7): 1271–1282.

437 43. Belfiore A, Frasca F, Pandini G, Sciacca L, Vigneri R. Insulin receptor isoforms and insulin receptor/insulin-like growth 438 factor receptor hybrids in physiology and disease. *Endocr Rev* 2009; 30(6): 586–623.

44. Bailyes EM, Navé BT, Soos MA, Orr SR, Hayward AC, Siddle K. Insulin receptor/IGF-I receptor hybrids are widely distributed in mammalian tissues: quantification of individual receptor species by selective immunoprecipitation and immunoblotting. *Biochem J* 1997; 327(1): 209–215.

45. Kasuya J, Paz IB, Maddux BA, Goldfine ID, Hefta SA, Fujita-Yamaguchi Y. Characterization of human placental insulin-like growth factor-I/insulin hybrid receptors by protein microsequencing and purification. *Biochemistry* 1993; 32(49): 13531–13536.

46. Slaaby R, Schaffer L, Lautrup-Larsen I, et al. Hybrid receptors formed by insulin receptor (IR) and insulin-like growth factor I receptor (IGF-IR) have low insulin and high IGF-1 affinity irrespective of the IR splice variant. *J Biol Chem* 2006; 281(36): 25869–25874.

47. Slaaby R. Specific insulin/IGF1 hybrid receptor activation assay reveals IGF1 as a more potent ligand than insulin. *Sci Rep* 2015; 5(1): 1–5.

48. Pierre-Eugene C, Pagesy P, Nguyen TT, et al. Effect of insulin analogues on insulin/IGF1 hybrid receptors: increased activation by glargine but not by its metabolites M1 and M2. *PloS one* 2012; 7(7): e41992.

49. Chen J, Nagle AM, Wang YF, Boone DN, Lee AV. Controlled dimerization of insulin-like growth factor-1 and insulin receptors reveals shared and distinct activities of holo and hybrid receptors. *J Biol Chem* 2018; 293(10): 3700–3709.

50. Crudden C, Girnita A, Girnita L. Targeting the IGF-1R: the tale of the tortoise and the hare. *Front Endocrinol* 2015; 6: 64.

51. Pandyarajan V, Weiss MA. Design of non-standard insulin analogs for the treatment of diabetes mellitus. *Curr Diab Rep* 2012; 12(6): 697–704.

52. Mathieu C, Gillard P, Benhalima K. Insulin analogues in type 1 diabetes mellitus: getting better all the time. *Nat Rev Endocrinol* 2017; 13(7): 385–399.

53. Moore MC, Smith MS, Sinha VP, et al. Novel PEGylated basal insulin LY2605541 has a preferential hepatic effect on glucose metabolism. *Diabetes* 2014; 63(2): 494–504.

54. Weiss MA, others . Design of ultra-stable insulin analogues for the developing world. *J Health Spec* 2013; 1: 59–70.

55. Guan X, Chaffey PK, Wei X, et al. Chemically precise glycoengineering improves human insulin. *ACS Chem Biol* 2018; 13(1): 73–81.

56. Mao R, Chen Y, Chi Z, Wang Y. Insulin and its single-chain analogue. *Appl Microbiol Biotechnol* 2019; 103(21): 8737–8751.

57. Kruger DF, Novak LM. Role of ultrafast-acting insulin analogues in the management of diabetes. *J Am Assoc Nurse Pract* 2019; 31(9): 537–548.

58. Guan J, Gluckman PD. IGF-1 derived small neuropeptides and analogues: a novel strategy for the development of pharmaceuticals for neurological conditions. *Br J Pharmacol* 2009; 157(6): 881–891.

59. Park SE, Lawson M, Dantzer R, Kelley KW, McCusker RH. Insulin-like growth factor-I peptides act centrally to decrease depression-like behavior of mice treated intraperitoneally with lipopolysaccharide. *J Neuroinflammation* 2011; 8(1): 1–17.

60. Deacon RM, Glass L, Snape M, et al. NNZ-2566, a novel analog of (1–3) IGF-1, as a potential therapeutic agent for fragile X syndrome. *Neuromolecular Med* 2015; 17(1): 71–82.

61. Guan J, Harris P, Brimble M, et al. The role for IGF-1-derived small neuropeptides as a therapeutic target for neurological disorders. *Expert Opin Ther Targets* 2015; 19(6): 785–793.

62. Baig AM, Khaleeq A. First Reports of Effects of Insulin, Human-like Insulin Receptors and Adapter Proteins in Acanthamoeba castellanii. *Sci Rep* 2020; 10(1): 1–14.

63. Vitali V, Horn F, Catania F. Insulin-like signaling within and beyond metazoans. *Biol Chem* 2018; 399(8): 851–857.

64. Schwartz TS, Bronikowski AM. Evolution and function of the insulin and insulin-like signaling network in ectothermic reptiles: some answers and more questions. *Integr Comp Biol* 2016; 56(2): 171–184.

65. Huang Q, Kahn CR, Altindis E. Viral hormones: expanding dimensions in endocrinology. *Endocrinology* 2019; 160(9): 2165–2179.

66. Safavi-Hemami H, Gajewiak J, Karanth S, et al. Specialized insulin is used for chemical warfare by fish-hunting cone snails. *Proc Natl Acad Sci USA* 2015; 112(6): 1743–1748.

67. Altindis E, Cai W, Sakaguchi M, et al. Viral insulin-like peptides activate human insulin and IGF-1 receptor signaling: A paradigm shift for host–microbe interactions. *Proc Natl Acad Sci USA* 2018; 115(10): 2461–2466.

68. Zhang F, Altindis E, Kahn CR, DiMarchi RD, Gelfanov V. A viral insulin-like peptide is a natural competitive antagonist of the human IGF-1 receptor. *Mol Metab* 2021; 53: 101316.

69. Chrudinová M, Moreau F, Noh HL, et al. Characterization of viral insulins reveals white adipose tissue-specific effects in mice. *Mol Metab* 2021; 44: 101121.

70. Vashisth H, Abrams CF. Docking of insulin to a structurally equilibrated insulin receptor ectodomain. *Proteins* 2010; 78(6): 1531–1543.

71. Vashisth H, Abrams CF. All-atom structural models for complexes of insulin-like growth factors IGF1 and IGF2 with their cognate receptor. *J Mol Biol* 2010; 400(3): 645–658.

72. Vashisth H, Abrams CF. All-atom structural models of insulin binding to the insulin receptor in the presence of a tandem hormone-binding element. *Proteins* 2013; 81(6): 1017–1030.

73. Vashisth H. Flexibility in the insulin receptor ectodomain enables docking of insulin in crystallographic conformation observed in a hormone-bound microreceptor. *Membranes* 2014; 4(4): 730–746.

74. Mohammadiarani H, Vashisth H. All-atom structural models of the transmembrane domains of insulin and type 1 insulin-like growth factor receptors. *Front Endocrinol* 2016; 7: 68.

75. Gorai B, Vashisth H. Structures and Interactions of Insulin-like Peptides from Cone Snail Venom. *Proteins* 2022; 90: 680-690.

76. Martí-Renom MA, Stuart AC, Fiser A, Sánchez R, Melo F, Šali A. Comparative protein structure modeling of genes and genomes. *Annu Rev Biophys Biomol Struct* 2000; 29(1): 291–325.

77. Shen My, Sali A. Statistical potential for assessment and prediction of protein structures. *Protein Sci* 2006; 15(11): 2507–2524.

78. Li H, Robertson AD, Jensen JH. Very fast empirical prediction and rationalization of protein pKa values. *Proteins* 2005; 61(4): 704–721.

79. DeLano WL, others . Pymol: An open-source molecular graphics tool. *CCP4 Newsletter on Protein Crystallography* 2002; 40(1): 82–92.

80. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ. GROMACS: fast, flexible, and free. *J Comput Chem* 2005; 26(16): 1701–1718.

81. Abraham MJ, Murtola T, Schulz R, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 2015; 1: 19–25.

82. Reif MM, Winger M, Oostenbrink C. Testing of the GROMOS force-field parameter set 54A8: structural properties of electrolyte solutions, lipid bilayers, and proteins. *J Chem Theory Comput* 2013; 9(2): 1247–1264.

83. Berendsen HJ, Postma JP, Gunsteren vWF, Hermans J. Interaction models for water in relation to protein hydration. In: Springer. 1981 (pp. 331–342).

84. Chester C, Friedman B, Ursell F. An extension of the method of steepest descents. In: . 53. Cambridge University Press. ; 1957: 599–611.

85. Darden T, York D, Pedersen L. Particle mesh Ewald: An N log (N) method for Ewald sums in large systems. *J Chem Phys* 1993; 98(12): 10089–10092.

86. Hess B, Bekker H, Berendsen HJ, Fraaije JG. LINCS: a linear constraint solver for molecular simulations. *J Comput Chem* 1997; 18(12): 1463–1472.

87. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graphics* 1996; 14(1): 33–38.

88. Webb B, Sali A. Comparative protein structure modeling using MODELLER. *Curr Protoc Bioinformatics* 2016; 54(1): 5–6.

89. Benyoucef S, Surinya KH, Hadaschik D, Siddle K. Characterization of insulin/IGF hybrid receptors: contributions of the insulin receptor L2 and Fn1 domains and the alternatively spliced exon 11 sequence to ligand binding and receptor activation. *Biochem J* 2007; 403(3): 603–613.

90. Robinson SD, Li Q, Bandyopadhyay PK, et al. Hormone-like peptides in the venoms of marine cone snails. *Gen Comp Endocrinol* 2017; 244: 11–18.

91. Ahorukomeye P, Disotuar MM, Gajewiak J, et al. Fish-hunting cone snail venoms are a rich source of minimized ligands of the vertebrate insulin receptor. *Elife* 2019; 8: e41574.

92. Xiong X, Menting JG, Disotuar MM, et al. A structurally minimized yet fully active insulin based on cone-snail venom insulin principles. *Nat Struct Mol Biol* 2020; 27(7): 615–624.

93. Xiong X, Blakely A, Kim JH, et al. Symmetric and asymmetric receptor conformation continuum induced by a new insulin. *Nature Chemical Biology* 2022; 18(5): 511–519.

94. De Meyts P, Whittaker J. Structural biology of insulin and IGF1 receptors: implications for drug design. *Nat Rev Drug Discov* 2002; 1(10): 769–783.

FIGURE AND TABLE LEGENDS

FIGURE 1

Sequence and structure alignment. (A) The alignment of the primary sequences of Ins and dcVILPs (dcGIV, dcSGIV, and dcLCDV1) is shown. The residue numbering at the top corresponds to the Ins sequence. The conserved cysteine residues are enclosed within orange boxes. The residues in dcVILPs similar to Ins are shown in blue. The dissimilar residues between dcGIV and dcSGIV are shown in red. The dcSGIV residue identical to Ins and different to dcGIV is shown in red and underlined. Shown are the superimposed structures of the representative (B) Ins, (C) dcGIV, (D) dcSGIV, and (E) dcLCDV1 structure of the largest cluster from each simulation trajectory (darker colors) on the starting structure (lighter colors). The RMSD values along with the standard deviations after the structural alignment of each representative peptide structure (from each of the three MD simulations) to their initial structure is depicted below each superimposed structure. The A-chain and B-chain of the peptides are shown in brown and magenta, respectively, and the disulfide bonds are represented in red sticks.

FIGURE 2

Structural models of the dcVILP/receptor complexes. The dcVILP chains and the receptor domains in each dcVILP/receptor complex are shown in cartoon representations and uniquely labeled/colored: A-chain of each dcVILP (orange), B-chain of each dcVILP (magenta), the L1 domain (cyan), and the αCT peptide (blue). The C_α-atoms of the residues in dcVILPs corresponding to the site-1 and site-2 residues in insulin at equivalent positions are depicted using smaller-size and larger-size spheres, respectively. The spheres representing the C_α-atoms of these residues are depicted in the same color as the A-chain and the B-chain in each dcVILP. The snapshots shown in colored boxes in each column are for the complexes of dcVILPs (dcGIV, orange box; dcSGIV, blue box; and dcLCDV1, magenta box) with μIR (top row), μIGF1R (middle row), and μhybR (bottom row).

FIGURE 3

Per-residue non-bonded binding energy difference. The total non-bonded energy difference (ΔE) of each residue of (A) dcGIV, (B) dcSGIV, and (C) dcLCDV1 with the IR (first row), IGF1R (second row), and hybR (third row) from the equivalent residues of Ins are depicted. A more negative value for a given residue indicates a stronger interaction of the dcVILP residue in comparison to the insulin residue at equivalent position. The dcVILP residue depicting significant negative (ΔE) values are labeled, where the substituted residue at the equivalent position of Ins is in red and the dcVILP residue is in black. The residues belonging to the A- and B-chains of peptides are depicted in light yellow and orange bars, respectively.

FIGURE 4

Unique inter-residue interactions between dcVILPs and the receptors. The inter-residue interactions between each peptide ligand (dcGIV, dcSGIV, and dcLCDV1) and the receptors (IR, IGF1R, and hybR) are shown. The side-chains of residues in the peptides are depicted as pink sticks, and of residues in the receptors are depicted as white sticks. Shown are the interactions consistently observed in at least two representative structures from the dominant conformational clusters derived from three independent MD simulations of each dcVILP/receptor complex. See also Table S4 for additional residue-level dcVILP/receptor contacts.

FIGURE 5

Sequence alignment and per residue energy contribution of proposed insulin analogs. (A) Sequence alignment of hIns, analog 1, analog 2, and analog 3 are shown. The substituted residues are shown in red. The total non-bonded energy difference (ΔE) of each residue of (B) analog 1, (C) analog 2, (D) and analog 3 with the IR, IGF1R, and hybR, respectively, from the respective residues of Ins are depicted. The total non-bonded binding energy between the Ins and each analog in complex with the each receptor are depicted in blue and purple text, respectively. The substituted residues exhibiting lower binding energy than the equivalent Ins residues are also labeled.

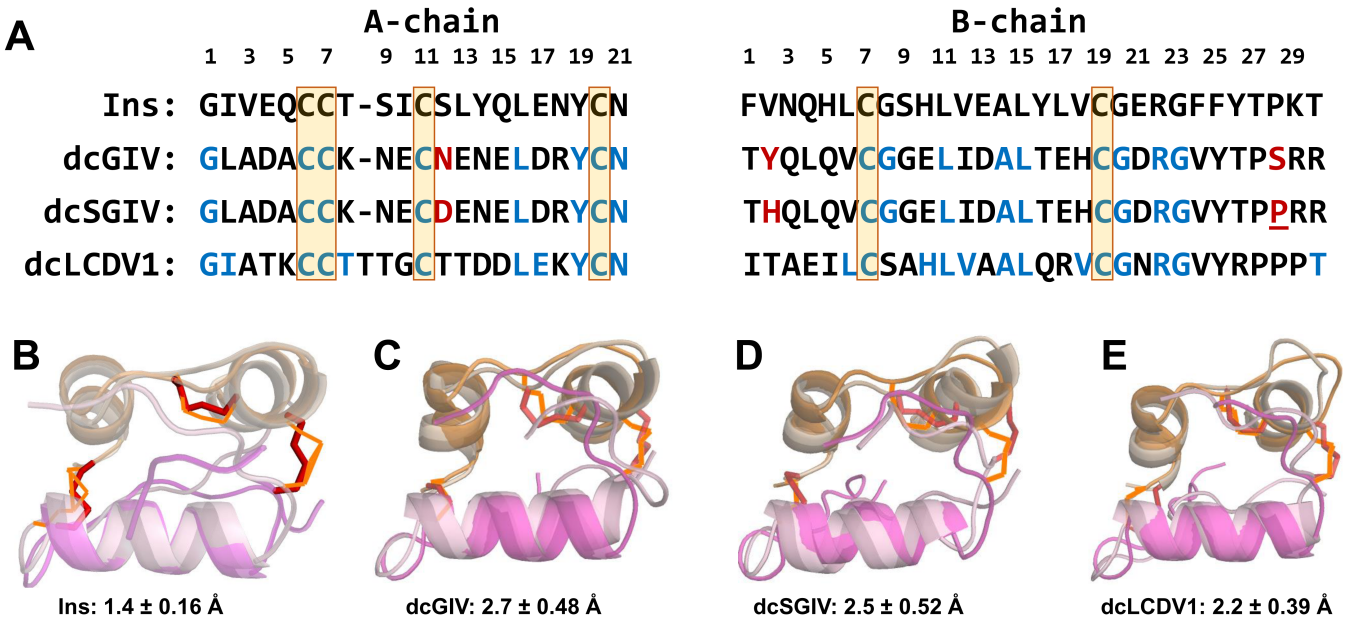

FIGURE 1 Sequence and structure alignment. (A) The alignment of the primary sequences of Ins and dcVILPs (dcGIV, dcSGIV, and dcLCDV1) is shown. The residue numbering at the top corresponds to the Ins sequence. The conserved cysteine residues are enclosed within orange boxes. The residues in dcVILPs similar to Ins are shown in blue. The dissimilar residues between dcGIV and dcSGIV are shown in red. The dcSGIV residue identical to Ins and different to dcGIV is shown in red and underlined. Shown are the superimposed structures of the representative (B) Ins, (C) dcGIV, (D) dcSGIV, and (E) dcLCDV1 structure of the largest cluster from each simulation trajectory (darker colors) on the starting structure (lighter colors). The RMSD values along with the standard deviations after the structural alignment of each representative peptide structure (from each of the three MD simulations) to their initial structure is depicted below each superimposed structure. The A-chain and B-chain of the peptides are shown in brown and magenta, respectively, and the disulfide bonds are represented in red sticks.

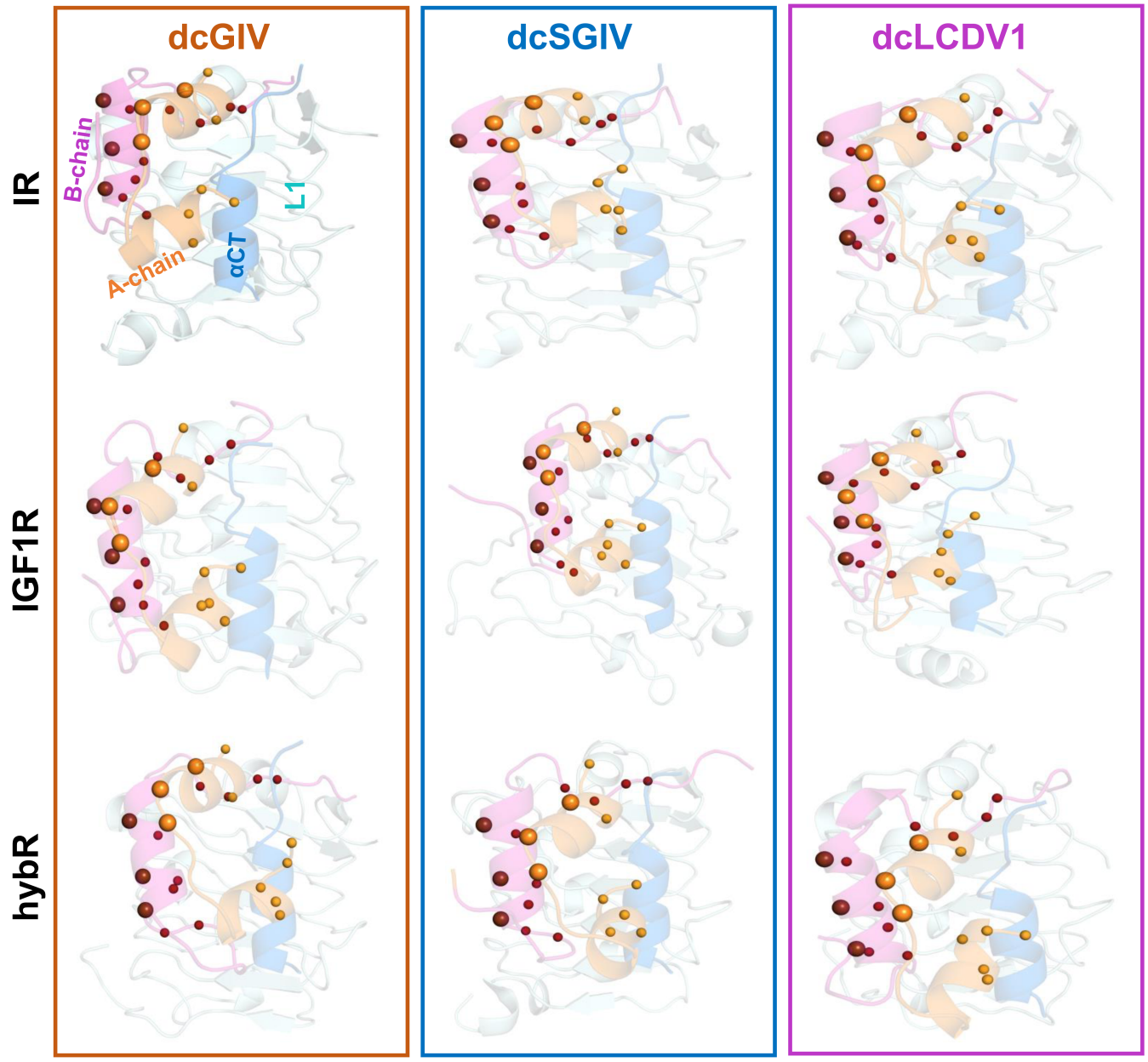

FIGURE 2 Structural models of the dcVILP/receptor complexes. The dcVILP chains and the receptor domains in each dcVILP/receptor complex are shown in cartoon representations and uniquely labeled/colored: A-chain of each dcVILP (orange), B-chain of each dcVILP (magenta), the L1 domain (cyan), and the αCT peptide (blue). The C_α-atoms of the residues in dcVILPs corresponding to the site-1 and site-2 residues in insulin at equivalent positions are depicted using smaller-size and larger-size spheres, respectively. The spheres representing the C_α-atoms of these residues are depicted in the same color as the A-chain and the B-chain in each dcVILP. The snapshots shown in colored boxes in each column are for the complexes of dcVILPs (dcGIV, orange box; dcSGIV, blue box; and dcLCDV1, magenta box) with μIR (top row), μIGF1R (middle row), and μhybR (bottom row).

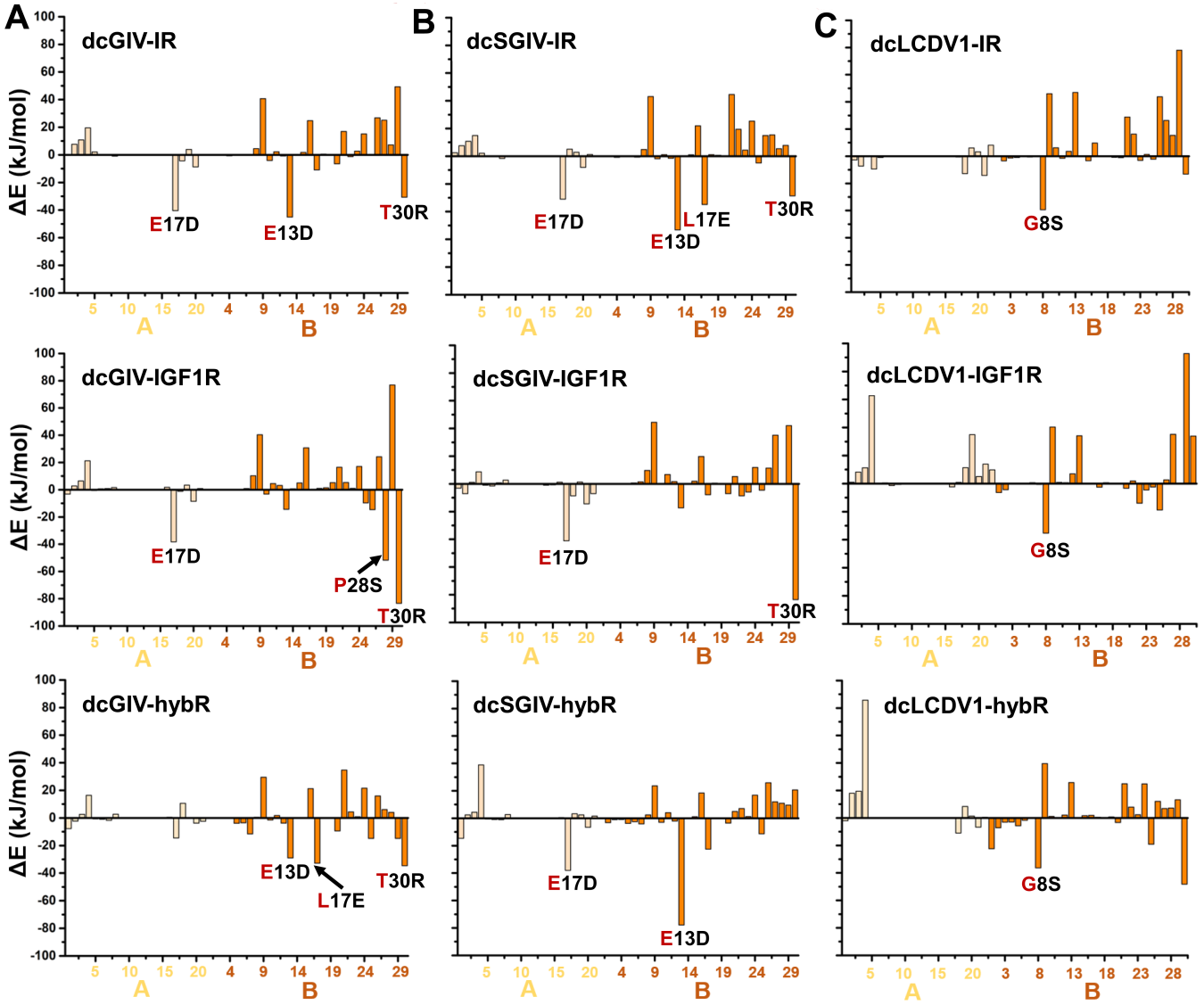

FIGURE 3 Per-residue non-bonded binding energy difference. The total non-bonded energy difference (ΔE) of each residue of (A) dcGIV, (B) dcSGIV, and (C) dcLCDV1 with the IR (first row), IGF1R (second row), and hybR (third row) from the equivalent residues of Ins are depicted. A more negative value for a given residue indicates a stronger interaction of the dcVILP residue in comparison to the insulin residue at equivalent position. The dcVILP residue depicting significant negative (ΔE) values are labeled, where the substituted residue at the equivalent position of Ins is in red and the dcVILP residue is in black. The residues belonging to the A- and B-chains of peptides are depicted in light yellow and orange bars, respectively.

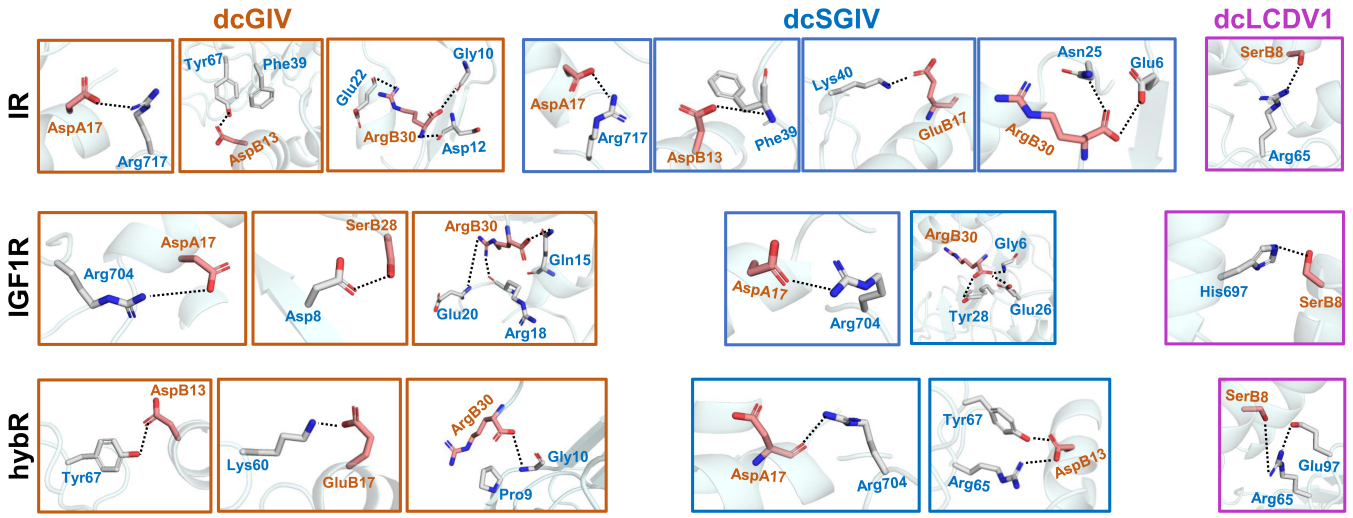

FIGURE 4 Unique inter-residue interactions between dcVILPs and the receptors. The inter-residue interactions between each peptide ligand (dcGIV, dcSGIV, and dcLCDV1) and the receptors (IR, IGF1R, and hybR) are shown. The side-chains of residues in the peptides are depicted as pink sticks, and of residues in the receptors are depicted as white sticks. Shown are the interactions consistently observed in at least two representative structures from the dominant conformational clusters derived from three independent MD simulations of each dcVILP/receptor complex. See also Table S4 for additional residue-level dcVILP/receptor contacts.

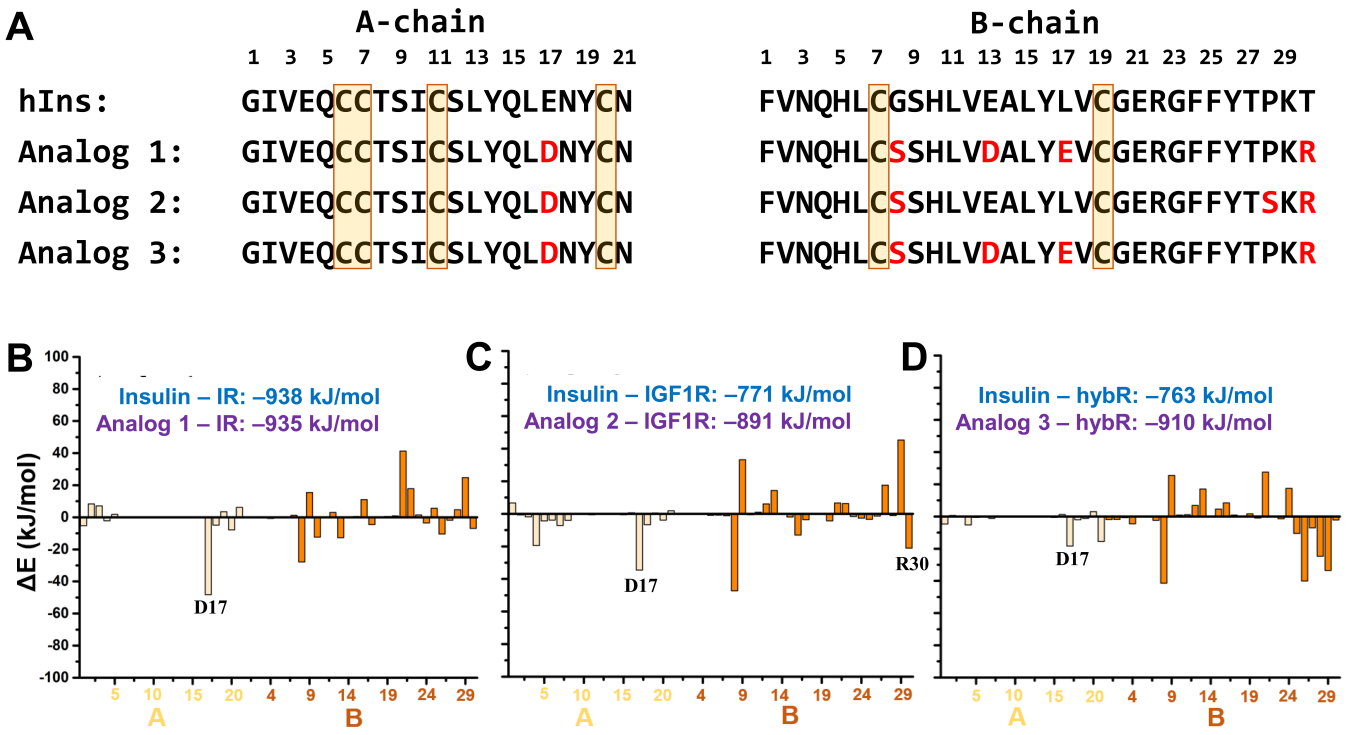

FIGURE 5 Sequence alignment and per residue energy contribution of proposed insulin analogs. (A) Sequence alignment of hIns, analog 1, analog 2, and analog 3 are shown. The substituted residues are shown in red. The total non-bonded energy difference (ΔE) of each residue of (B) analog 1, (C) analog 2, (D) and analog 3 with the IR, IGF1R, and hybR, respectively, from the respective residues of Ins are depicted. The total non-bonded binding energy between the Ins and each analog in complex with the each receptor are depicted in blue and purple text, respectively. The substituted residues exhibiting lower binding energy than the equivalent Ins residues are also labeled.

579 .