ACS OMEGA

Open Access



http://pubs.acs.org/journal/acsodf

Article

# Computational Approach to Elucidating Insulin−Protamine Binding Interactions and Dynamics in Insulin NPH Formulations

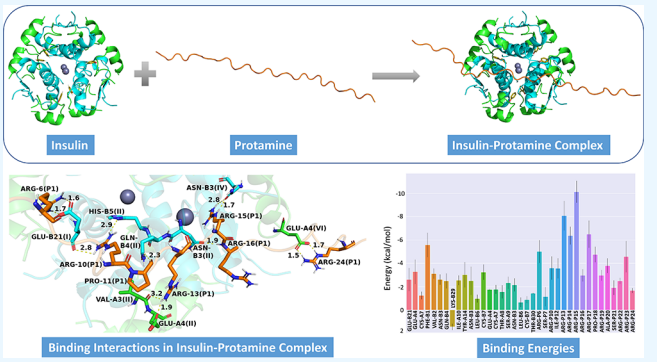
Ketan Kumar Rohilla and Manoj Kumar Pandey*

Cite This: *ACS Omega* 2024, 9, 4857−4869

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

**ABSTRACT:** Insulin NPH is an intermediate-acting insulin. Its protracted action profile is due to the formation of microcrystalline suspensions when insulin is complexed with a basic peptide protamine, zinc ion, and phenolic ligands. Despite advancements in analytical techniques, the binding epitope and binding mode of the protamine in the insulin−protamine complex are still unknown. In this study, we used bioinformatics tools such as molecular docking and molecular dynamics (MD) simulations to compute the binding sites and energetics of the insulin−protamine complex. We have taken four naturally occurring protamine peptides that are independently docked with the insulin R6 hexamer and subjected them to 200 ns MD simulations to observe the dynamics of the complexes and estimate the binding energies. The arginine-rich protamine peptides were found to bind on the surface of the insulin hexamer through hydrogen bonding, hydrophobic, and electrostatic interactions well supported by the calculated negative binding energies. The overall structure of the insulin hexamer was retained upon binding, highlighting its dynamic stability in the complex. Furthermore, the residues at the termini of the protamine peptides in the complex were seen to be highly dynamic, which stabilize toward the end of the simulation.




## 1. INTRODUCTION

Insulin is one of the smallest proteins found in the human body,[1−3] and it is secreted by pancreatic beta cells in the islets of Langerhans.[3,4] Insulin is necessary for the regulation of glucose metabolism in the body.[5,6] However, due to various reasons, either insufficient insulin is produced in the body or insulin produced is inefficient, resulting in failure of glucose metabolism, a condition termed as diabetes mellitus (DM).[7,8] Thus, the required insulin is supplemented through external sources[2,9] to maintain the bolus and basal levels of insulin in the blood plasma (pre- and postmeals). Different insulin analogues with varying action profiles have been developed to maintain constantly changing insulin requirements in the blood plasma. These are classified as fast-acting, intermediate-acting, long-acting, ultralong-acting, and inhaled insulin.[10] The time action profile of these insulin analogues is determined by the release of monomeric insulin from its hexameric state, which is governed by metal ions such as zinc and phenolic ligands and alternatively by the addition of highly basic peptides such as protamine. Insulin−protamine complexes have been found to be very effective in altering the time action profile of insulin drugs while having no adverse effects on the body. The inclusion of protamine in the formulation, together with zinc and phenolic ligands, was the first successful precipitation method of protraction of insulin action.[11] Insulin Neutral Protamine Hagedorn (NPH) is administered as a microcrystalline suspension of preformed protein insulin suspension produced in a 5:1 molar ratio of insulin and protamine.[12] Insulin NPH has a 1−2 h onset of action time, 2−8 h peak action time, and a total duration of action of 14−24 h,[13,14] thus reducing the insulin dosage frequency. It has an advantage over other drugs as it can be used as both bolus and basal insulin.

This study focuses on understanding the mechanistic picture of insulin−protamine interactions. The protamine (Salmine) used in the synthesis of insulin NPH is isolated from the sperm of Salmonide fish,[15,16] which comprises of four main peptides, as detected by reversed-phase high-performance liquid chromatography (RP-HPLC).[17,18] These are referred to as P1 (peptide 1), P2 (peptide 2), P3 (peptide 3), and P4 (peptide 4) in this study. Each of these peptides contains 30−32 amino acids (molecular weight ∼4 kDa) and exists as a random coil due to the lack of any secondary structure, as confirmed by one-dimensional (1D) proton solution NMR, circular dichroism (CD) spectroscopy, and Fourier transform infrared (FTIR) studies.[19,20] Despite advancements in technology, the binding sites and the binding mode of the

Received: October 26, 2023
Revised: December 28, 2023
Accepted: December 28, 2023
Published: January 18, 2024

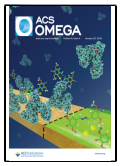


Table 1. The Primary Amino Acid Sequence, as Detected by RP-HPLC,[17,18] of Four Peptides of Protamine Is Written in a Single-Letter Amino Acid Code (P, Proline; S, Serine; R, Arginine; I, Isoleucine; A, Alanine; V, Valine; G, Glycine)[a]

| | |
|---|---|
| Peptide 1 | PRRRRRSSSRPIRRRRRPRASRRRRRGGRRRR |
| Peptide 2 | PRRRRSSRRPVRRRRRPRVSRRRRRGGRRRR |
| Peptide 3 | PRRRRSSSRPVRRRRRPRVSRRRRRGGRRRR |
| Peptide 4 | PRRRRASRRIRRRRRPRVSRRRRRGGRRRR |

[a]The difference in the amino acid sequences of these four peptides is highlighted.

insulin−protamine complex have not been known until now. A previous study on insulin−protamine interactions using X-ray crystallography was done by Norrman et al.[21] However, they were unable to determine the precise location of the protamine's binding to insulin as the electron density was too low to be considered for reliable modeling. Nonetheless, the study emphasized identifying a prominent electron density around $Asn^{B3}$ (asparagine, the third amino acid of the B-chain of insulin). It is therefore important to understand the binding epitope and conformational changes (if any) that occur in insulin following protamine binding, which may also provide information on the crystallization mechanism. In this study, we used computational techniques to investigate the binding. Computational approaches, particularly molecular dynamics simulation, have proven to be quite useful in highlighting various aspects of insulin's structure and function and have been widely used in the past.[22] To begin, we used the molecular docking technique to map the precise binding sites and mode of binding of protamine to insulin. Further evaluation of the stability and dynamic nature of the insulin−protamine complex was performed by using molecular dynamics simulations on docked complexes. To the best of our knowledge, this is the first molecular docking and molecular dynamics simulation study of protamine with an insulin hexamer. Our study revealed important insights into the insulin−protamine complex and its behavior in the presence of a solvent, which may help the pharmaceutical industry to develop more efficient drugs.

## 2. MATERIALS AND METHODS

**2.1. System Preparation.** In the absence of the PDB structures of the four protamine peptides used in this study, their structure predictions were carried out using an AI-based program, AlphaFold.[23,24] Table 1 shows the amino acid sequence used for the structure predictions.[17] Random coiled structures for all four peptides were obtained, which is also in agreement with previous experimental results.[19,20] As insulin in the insulin NPH formulation exists as a hexamer, therefore, all our computational calculations were carried out on the insulin R6 hexamer structure (PDB ID: 1AIY[25]) downloaded from the RCSB Protein Data Bank.[26,27] These structures were further used for molecular docking and molecular dynamics simulations.

**2.2. Molecular Docking.** AlphaFold-predicted PDB files for all four protamine peptides were well equilibrated for 1 ns under the NVT (thermodynamic state defined with a fixed number of atoms (N), volume (V), and temperature (T)) ensemble and 10 ns under the NPT (thermodynamic state defined with a fixed number of atoms (N), pressure (P), and temperature (T)) ensemble using GROMACS before performing the docking analysis with the insulin R6 hexamer. Simulation details and parameters for equilibration are described in the next section. We chose the R conformation of the insulin hexamer structure as predicted by the X-ray crystallography study by Norrman et al.[21] in the insulin NPH formulations. Considering the insulin-to-protamine sulfate ratio (5:1) in the pharmaceutical formulations,[12] we independently docked all four protamine peptides (ligand) to the insulin R6 hexamer (receptor) using the HADDOCK (High Ambiguity Driven protein−protein Docking) 2.4 web server,[28,29] which is driven by experimental evidence rather than energetics considerations for the protein−peptide docking studies. These are introduced in the form of ambiguous interaction restraints (AIRs). In this regard, we have used ambiguous interaction restraints in the form of active residues (based on electron density mapping) reported previously in the X-ray study of insulin NPH.[21] Alternatively, ab initio docking utilizing center-of-mass restraints may also be performed to identify the potential binding interactions between four protamine peptides and the insulin hexamer. However, in the present study, we chose to carry out molecular docking by selecting active residues to correlate the insulin−protamine binding interactions with experimental evidence. The docking protocol is composed of three discrete stages to refine the conformational arrangement of a given system: First, orientation randomization and rigid body energy minimization are implemented, followed by a selection of the top 200 results for further refinement. Subsequently, simulated annealing refinements are performed, allowing for stepwise conformational rearrangements of both side chains and backbone. Finally, a refinement step is conducted using the TIP3P explicit solvent model, employing steepest descent energy minimization.[29]

Minimum energy docked insulin−protamine structures were selected for all of the further studies. The docked insulin−protamine complexes are referred to as P1-R6, P2-R6, P3-R6, and P4-R6 wherein insulin R6 is docked with protamine peptide 1, peptide 2, peptide 3 and peptide 4, respectively. All four insulin−protamine complexes contain two $Zn^{2+}$ ions and six phenol molecules. Analysis of the docked structures was carried out using Ligplot[30] and PyMOL v. 2.5.4.[31] The binding residues were predicted by analyzing hydrogen-bonded interactions, hydrophobic contacts, and salt bridges.

**2.3. Molecular Dynamics Simulations and Analysis.** Molecular dynamics simulations were performed on the four protamine peptides (only equilibration), undocked insulin (insulin R6), and docked insulin−protamine complexes (P1-R6, P2-R6, P3-R6, and P4-R6) to observe the stability and dynamics and calculate the binding energy. The starting structure for undocked insulin was taken from the Protein Data Bank (PDB ID: 1AIY) and the docked structures were


https://doi.org/10.1021/acsomega.3c08445
ACS Omega 2024, 9, 4857−4869

obtained from the HADDOCK 2.4 web server as discussed above. The CHARMM-GUI (V3.7) input generator[32−34] was used to generate inputs for GROMACS[35,36] using the CHARMM-GUI online interface's solution builder module. MD simulations with the CHARMM36m all-atom force field were carried out using GROMACS 2020.6. Notably, the CHARMM36m force field in recent times has been extensively used for investigating molecular dynamics of insulin and other proteins.[37−40] MD simulation was carried out in three stages (minimization, equilibration, and production), with a force constant of 1000 kJ/mol $nm^2$ employed during the minimization and equilibration stages to restrict all heavy atoms and maintain the original protein folding. The TIP3P explicit water model was used to solvate the box of the dimension $(74\ Å)^3$ for R6 and the box of the dimensions $(113\ Å)^3$, $(103\ Å)^3$, $(109\ Å)^3$, and $(113\ Å)^3$ for the P1-R6, P2-R6, P3-R6 and P4-R6 complexes, respectively. The type of the box was rectangular, and the crystal type was cubic with $\alpha = 90°$, $\beta = 90°$, and $\gamma = 90°$. The system was neutralized by adding KCl, whose concentration was 0.15 M. (38, 30), (122, 122), (91, 91), (110, 110), and (122, 122) $K^+$ and $Cl^-$ ions were required to completely neutralize the system for R6, P1-R6, P2-R6, P3-R6, and P4-R6, respectively. For ion placement, the Monte Carlo approach was applied. The energy of the system was minimized by using the steepest descent minimization algorithm. To manage long-range electrostatic interactions, the particle mesh Ewald (PME) approach was used. The simulations were carried out using a Nose−Hoover thermostat with 2 fs time step and an isochoric isothermal (NVT) ensemble at 303.15 K. A pressure coupling system based on a Parrinello−Rahman barostat was used. After the system was adequately equilibrated for 1 ns under the NVT ensemble and 10 ns under the NPT ensemble, a 200 ns MD run was performed. The coordinates were saved every 1 ps. VMD was used to analyze the trajectory and estimate salt bridge interactions.[41] PyMOL was used to analyze the binding sites again to check for any difference/change in the binding sites between protamine and insulin. Properties like potential energy, root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration ($R_g$), H-bonding, and solvent-accessible surface area (SASA) were calculated for a 200 ns MD run, and graphs were plotted using xmgrace software (version 5.1.25).

**2.4. Binding Energy Calculation Using MM/GBSA.** For the binding energy calculations, we have used the MM/GBSA (molecular mechanics/generalized Born surface area) methodology,[42−44] which was implemented in gmx_MMPBSA based on MMPBSA.py.[45,46] Binding energy at fixed temperature was calculated for all four docked insulin−protamine complexes as follows:

$$\Delta G_{bind} = \Delta H - T\Delta S \quad (1)$$

here, $\Delta H$ is the enthalpy change, and $T\Delta S$ is the conformational entropy change when the ligand binds to the complex. $\Delta H$ is computed from the following equation,

$$\Delta H = \Delta E_{MM} + \Delta G_{solv} \quad (2)$$

where $\Delta E_{MM}$ is the change of MM (molecular mechanics) energy corresponding to the gas phase, and $\Delta G_{solv}$ is the energy change of solvation, which can be computed from eqs 3 and 4 below, respectively,

$$\Delta E_{MM} = \Delta E_{bonded} + \Delta E_{non\text{-}bonded} = (\Delta E_{bonding} + \Delta E_{angle} + \Delta E_{dihedral}) + (\Delta E_{elec} + \Delta E_{vdW}) \quad (3)$$

$$\Delta G_{solv} = \Delta G_{polar} + \Delta G_{non\text{-}polar} = \Delta G_{GB} + \Delta G_{SASA} \quad (4)$$

where $\Delta E_{MM}$ has the contributions from $\Delta E_{bonded}$ and $\Delta E_{non\text{-}bonded}$, which further includes bonding ($\Delta E_{bonding}$), angle ($\Delta E_{angle}$), and dihedral ($\Delta E_{dihedral}$) energies (for the bonded part) and electrostatic ($\Delta E_{elec}$) and van der Waals ($\Delta E_{vdW}$) energies (for the nonbonded part), respectively. $\Delta G_{solv}$ includes electrostatic (polar) contribution and non-electrostatic (nonpolar) contribution. The polar part is estimated using the GB model here, and the nonpolar part is calculated using solvent-accessible surface area (SASA). The interaction entropy method[47] was employed to compute the conformational entropy component. Binding energy analysis was performed on the last 20000 frames, i.e., last 20 ns of the simulation trajectory using the input file generated by gmx_MMPBSA (version: v1.5.6+21.gc3e03f6 based on MMPBSA.py v.16.0). "PBRadii" was set to 4 for building the amber topology from the GROMACS files. "iGB" (GB model used) was 8, and the internal dielectric constant was 4. All of the other parameters were set to default in the input file.

We have also performed per-residue decomposition analysis to find the binding energy contribution of the active sites for the receptor and ligand complex. Decomposition analysis in MM-GBSA is a method used to estimate the energy contribution of individual residues or groups of residues to the overall binding affinity of a protein−ligand complex. This analysis involves calculating the energy contribution of each residue by summing its interactions with all other residues in the system. The per-residue decomposition method is commonly used to identify key residues that contribute significantly to the binding energy of a complex. This information is essential for experimental studies relying on site-directed mutagenesis or drug design to improve the binding affinity and selectivity of a protein−ligand complex. For this purpose, the residues up to a distance of 6 Å from the active sites of the receptor and ligand have been chosen, and the binding energy has been calculated for those residues that come in that region. "idecomp" was set to 1 for the decomposition analysis.

## 3. RESULTS AND DISCUSSION

**3.1. Binding Site Analysis in the Insulin−Protamine Complex after Docking.** The four protamine peptides were docked independently to determine the precise binding locations and to understand whether any of the peptides bind to insulin in NPH formulations preferentially over the others. We observed the following in the docked insulin−protamine complexes: (a) insulin's hexameric state is retained upon binding with protamine peptides, (b) protamine peptides bind exclusively on the surface of the insulin R6 hexamer, (c) all four protamine peptides bind to insulin in nearly similar orientation, (d) B-chain residues of insulin are mostly involved in binding with the protamine residues, which is consistent with a previous docking study of protamine peptide with an insulin monomer,[48] and (e) upon binding, protamine peptides retained their random coil structure. It is well known that insulin coexists as a monomer and a dimer, which self-



https://doi.org/10.1021/acsomega.3c08445
ACS Omega 2024, 9, 4857−4869

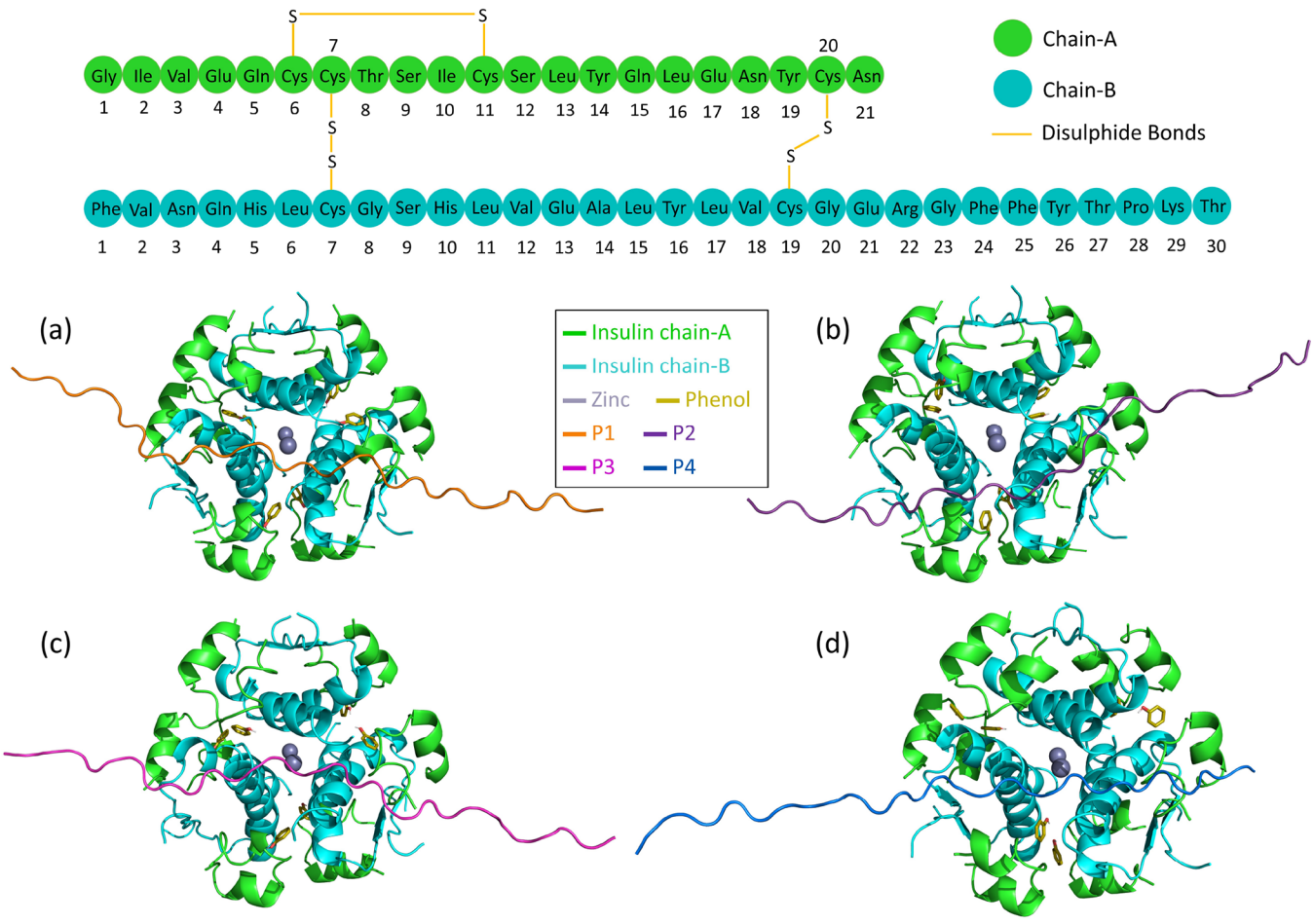



**Figure 1.** Amino acid sequence of an insulin monomer (top panel). Docked hexameric insulin−protamine complexes (a) P1-R6, (b) P2-R6, (c) P3-R6, and (d) P4-R6 (bottom panel). The coloring scheme is shown in the inset.

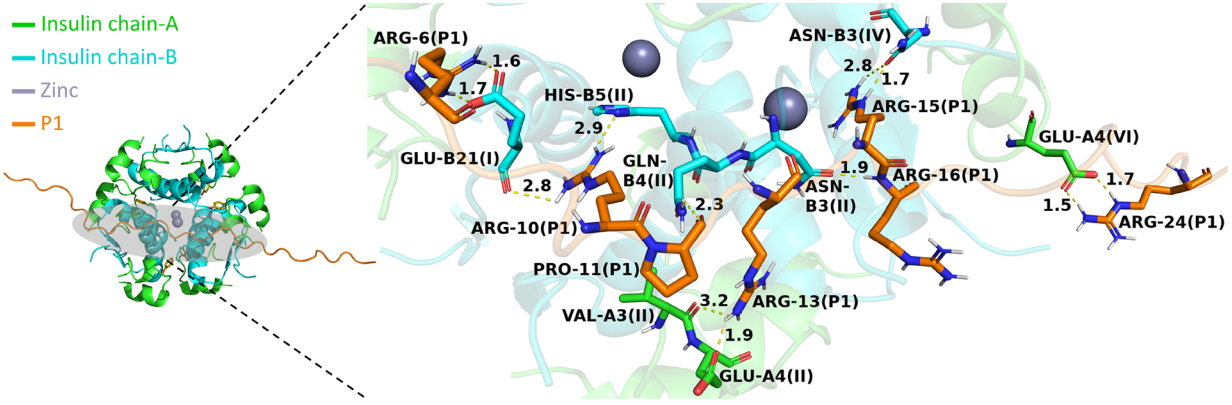



**Figure 2.** Residues involved in the H-bonding interactions between insulin and protamine in the docked insulin−protamine complex P1-R6. The residues are written in a three-letter amino acid code along with their chain and residue number. The coloring scheme is shown in the inset. Six insulin monomers in its hexameric form are named I, II, III, IV, V, and VI for the A- and B-chains. The monomeric unit is represented in parentheses. The H-bonds are shown with yellow dashed lines, and the distance between the donor and acceptor atoms is indicated in the Å unit.

assembles to form an inactive hexameric state, which is key for insulin's storage and prolonged release into the bloodstream. The residues that contribute primarily to forming the dimeric interfaces of insulin are the B-chain's C-terminal aromatic residues ($Phe^{B24}$, $Phe^{B25}$, and $Tyr^{B26}$). Residues $Ala^{A14}$, $Leu^{A17}$, $Leu^{B13}$, $Tyr^{B14}$, and $Glu^{B7}$ further strengthen the hexameric state through the formation of the trimer of dimers.[22] In addition, two bivalent zinc metal ions ($Zn^{2+}$) coordinated with six histidine residues [($Zn^{2+}$---N(His)] to help in the self-assembly of the insulin monomers into the hexameric state, which is important for the storage and their extended time of action. Interestingly, the binding of protamines to insulin does not disturb the above interactions in all four docked insulin−protamine complexes. Our docking results also show that $Zn^{2+}$---N (His) distances in all four insulin−protamine



https://doi.org/10.1021/acsomega.3c08445
ACS Omega 2024, 9, 4857−4869

complexes are constant. Hence, protamine binding has no adverse effect on the self-associated oligomeric state of insulin.

We believe that the retention of the inner core region of insulin's hexameric state is one of the factors for protamine peptides to bind only on the surface of insulin and not the inner core region of insulin (Figure 1a–d). As seen in Figure 1a–d, all four protamine peptides bind to insulin in a similar orientation and surface, which is also expected given their nearly similar amino acid sequence and structure. Along with $Zn^{2+}$ and phenolic ligands, the binding of protamine to insulin provides added stability to the insulin–protamine complex by holding together two dimers with additional binding contacts (Figure 1), which are discussed in detail below. Primarily, our study focused on three types of interactions that are present between insulin and protamine in the docked complexes: (a) H-bonding, (b) hydrophobic, and (c) salt bridge interactions. The hydrogen bonds between insulin and protamine were calculated using PyMOL. The maximum and minimum cutoff distances for the hydrogen bond between the donor and acceptor atoms were set to 3.5 and 1.5 Å, respectively. Mainly arginine residues of protamine peptides and $Glu^{A4}$, $Cys^{A7}$, $Ser^{A9}$, $Ile^{A10}$, $Val^{B2}$, $Asn^{B3}$, $Gln^{B4}$, $His^{B5}$, $Glu^{B21}$, and $Pro^{B28}$ residues of insulin are involved in H-bonding. For demonstration, the H-bonding interactions in the P1-R6 complex are shown in Figure 2, while these interactions in P2-R6, P3-R6, and P4-R6 complexes can be found in Figure S1a–c of the Supporting Information. Table S1 lists the specific insulin and protamine residues that form the H-bonds. Although the location of the binding of the four protamines on the surface of insulin is nearly similar (Figure 1), the H-bonding partners are not exactly identical in all four docked complexes resulting from minor variations in the amino acid sequences of the protamine peptides (Figure 2 and Table S1). The hydrophobic interactions for all four docked insulin–protamine complexes are shown in Figure S2 of the Supporting Information. Both N- and C-terminal residues of the B-chain of insulin are involved in hydrophobic interactions with protamine peptides. The residues of insulin's B-chain involved in these interactions are $Phe^{B1}$, $Val^{B2}$, $Asn^{B3}$, $Gln^{B4}$, $Leu^{B6}$, $Pro^{B28}$, and $Thr^{B30}$. The N-terminal $Phe^{B1}$ is the most common amino acid that is involved in hydrophobic interaction with the protamine peptides. Additionally, three A-chain amino acid residues $Cys^{A7}$, $Thr^{A8}$, and $Ser^{A9}$ were also found to interact hydrophobically with the protamine peptides. On protamine peptides, proline, valine, and alanine residues are mainly involved in the hydrophobic interactions with insulin. Notably, these residues of protamine do not participate in the H-bonding interactions. $Glu^{A4}$, $Glu^{A17}$, and $Glu^{B21}$ of insulin were found to be involved in the salt bridge interactions with arginine amino acids of protamine peptides, providing the insulin–protamine complex extra stability, which is discussed in detail in Section 3.4.2.

The docking study reveals that the side chain of arginine residues of protamine peptides, which contain the guanidinium group, mostly participates in the binding interactions with insulin. However, no preferential site could be located on the four protamine peptides for binding with insulin. This could be due to high sequence similarity, random coil structure, and high arginine content (~70%) in the four protamine peptides. Interestingly, the protamine peptides did not undergo any significant structural change after binding to insulin, suggesting no specific folding of protamine peptides upon binding with insulin. Our results corroborate well with the X-ray study performed by Norrman et. al.[21] In their study, the electron density of the guanidinium group of arginine side chains was predicted in the vicinity of $Asn^{B3}$; however, the electron density was too weak for reliable molecular modeling. In this study, the side chain of arginine residues of protamine peptides is found to be located in between $Asn^{B3}$ residues of the neighboring monomers as shown for P1-R6 in Figure 3. The

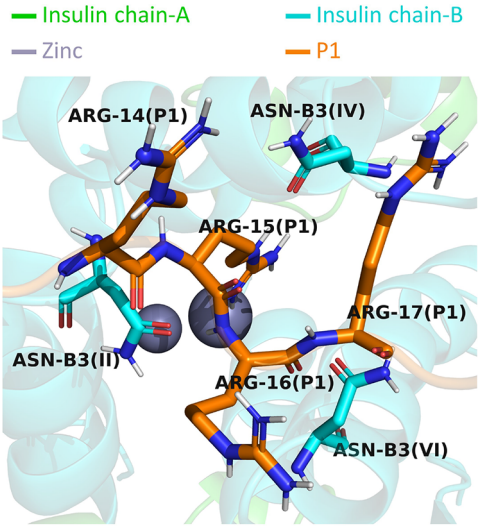



**Figure 3.** Presence of the side chain of arginine residues of protamine peptide in between the $Asn^{B3}$ residues of neighboring monomers in the docked P1-R6 complex.

$Asn^{B3}$ residues of adjacent monomers of insulin are in the proximity of the binding surface of the insulin–protamine complex. This observation is common to all remaining three insulin–protamine complexes, P2-R6, P3-R6, and P4-R6 (refer to Figure S3a–c of the Supporting Information). The side chain of arginine residues of the protamine peptides provides insulin–protamine complex additional binding interactions, which may be essential for the prolonged time action profile of insulin NPH drugs.

**3.2. Probing Stability of Insulin–Protamine Complexes Using MD Simulations.** A combinatorial approach of molecular docking and MD simulations for a deeper understanding of binding events at the molecular level is often employed in drug research and development.[49,50] In this study, we performed MD simulations for a time duration of 200 ns to determine whether or not the docked insulin–protamine complexes remain dynamically stable over time. For a stable insulin–protamine complex to form, the binding residues must also be stable over the simulation duration, and the binding energy should be energetically favorable for these residues. Figure 4 shows snapshots of the insulin–protamine complexes after a 200 ns MD simulation run. The insulin–protamine binding residues, as observed from molecular docking (Figure 2 and Figure S1a–c), were retained after the MD run. Moreover, we identified additional insulin–protamine interactions that further stabilize the docked complexes. At the beginning of the simulation, the protamine terminal residues did not interact with insulin and were free to move due to its random coil structure. However, after around 60–70 ns, the termini of protamine peptides were found to be interacting with insulin and therefore suggesting added stability to the hexameric assembly of insulin.


https://doi.org/10.1021/acsomega.3c08445
ACS Omega 2024, 9, 4857–4869

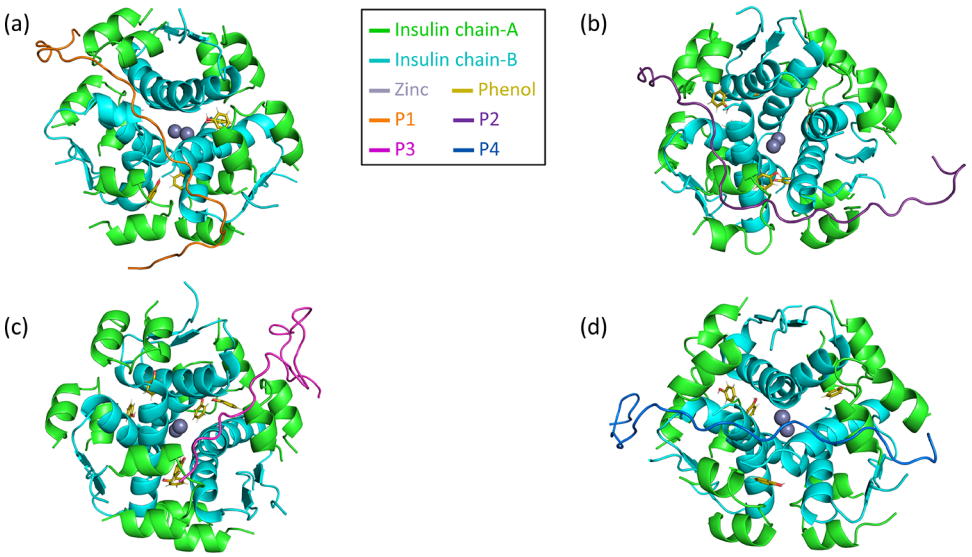



**Figure 4.** Snapshots of the insulin−protamine complex taken after a 200 ns MD run. (a) P1-R6, (b) P2-R6, (c) P3-R6, and (d) P4-R6.

To further substantiate the above findings, potential energy calculations were performed on all four docked complexes (bound state) and compared with the insulin R6 hexamer in the presence of a solvent (Figure 5). The potential energy of

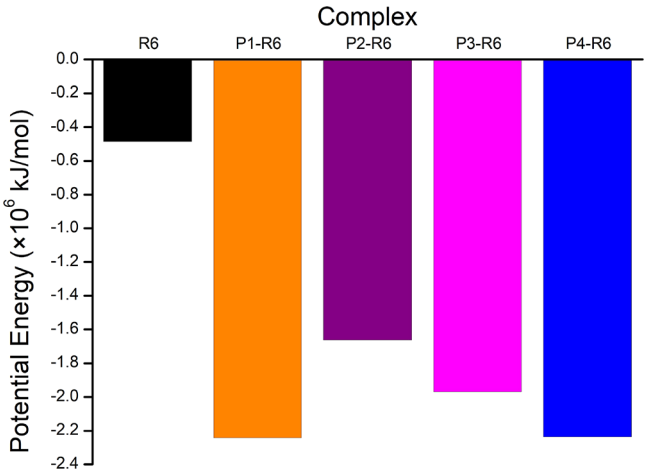



**Figure 5.** Potential energy plots for undocked insulin R6 and docked insulin−protamine complexes P1-R6, P2-R6, P3-R6, and P4-R6.

the system is determined by the summation of all constituent energy terms and can be mathematically expressed as the sum of energy contributions resulting from bonding and non-bonding (Lennard-Jones/van der Waals and Coulomb/electrostatic potentials) interactions. The calculation of potential energy is performed on the most stable structure following the energy minimization step once the system has attained an equilibrium state and the potential energy has stabilized. Potential energy values were observed to be lower in the docked insulin−protamine complexes than in the undocked insulin. This again suggests that the insulin−protamine complex gains stability with the addition of protamine. Moreover, all four docked complexes showed comparable potential energies, indicating that each of these peptides behaved similarly during the course of the MD simulations.

To analyze the stability of insulin in the docked complexes, the root mean square deviation (RMSD) values of insulin were calculated. RMSD is a measure of the deviation of the atomic positions of a molecule or a complex of molecules from a reference structure (the structure before the MD run) over the course of the simulation. The RMSD values less than 4 Å (Figure 6a) for all docked complexes are found to be consistent with the earlier reported literature.[22] Therefore, the docked complexes remain relatively stable and do not significantly deviate from the starting structure across the simulation run. A comparison of the RMSD values of the four docked complexes is shown in Figure 6a. As expected, the RMSD values follow a nearly similar trend due to the high structural similarity of the protamines. Out of the four complexes, P1-R6 (orange) shows a slightly lower value of RMSD after around 60 ns, which is also consistent with the results based on the potential energy calculations. Furthermore, RMSD values for the final 20 ns of the simulation run were found to have a constant trend and remained below 4 Å, highlighting the stability of all four insulin−protamine complexes toward the end of simulations (Figure S4 of the Supporting Information).

As the RMSD analysis only informs on the overall structural stability of the complex, we further evaluated residue-specific fluctuations in the insulin−protamine complexes. For this purpose, the root mean square fluctuation (RMSF) values corresponding to each residue of insulin in the docked complexes were computed. The RMSF values indicate how much a particular atom or group of atoms moves or fluctuates relative to its average position during the MD simulation run. The RMSF value of insulin's residues in the docked complexes is shown in Figure 6b. An increased RMSF value is observed for the residues located at the N- and C-termini and the aromatic residues of insulin. This could potentially be due to the binding of protamine at the center of hexamer and possibly a slight opening of the termini, making them more flexible. Moreover, aromatic residues like phenylalanine, tyrosine, etc., can undergo ring flipping (refer to Figure S5 of the Supporting Information) and therefore show higher RMSF values. The important region to concentrate in insulin is the one with the minimum fluctuations to locate the residues that are


https://doi.org/10.1021/acsomega.3c08445
ACS Omega 2024, 9, 4857−4869

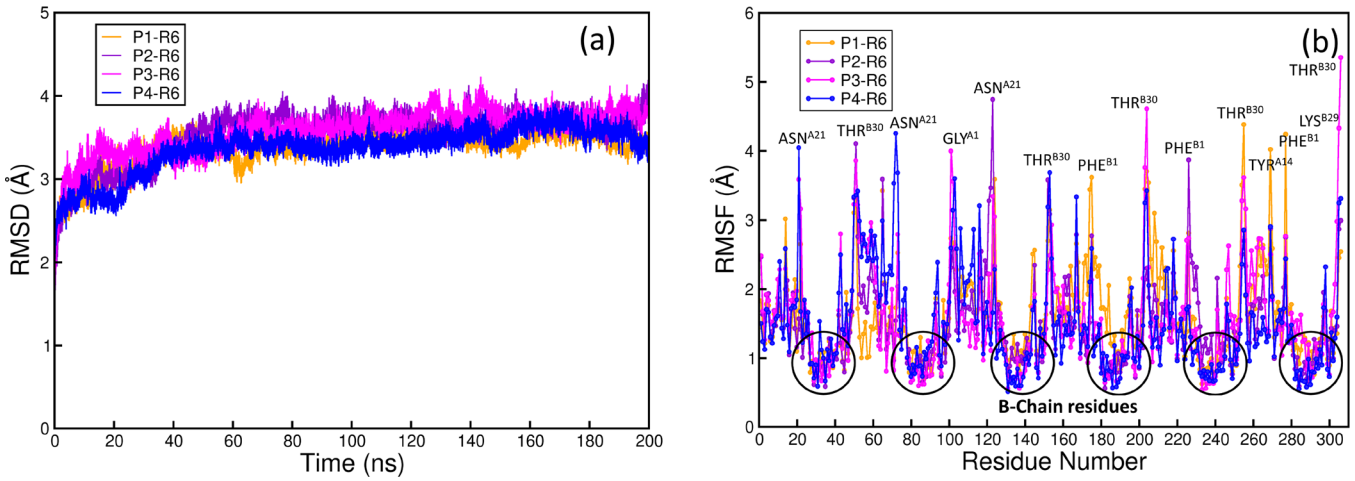



**Figure 6.** Analysis of the trajectory of docked complexes during a 200 ns MD run. (a) RMSD vs simulation time plot. (b) RMSF vs residue number plot for insulin in the docked insulin−protamine complexes.

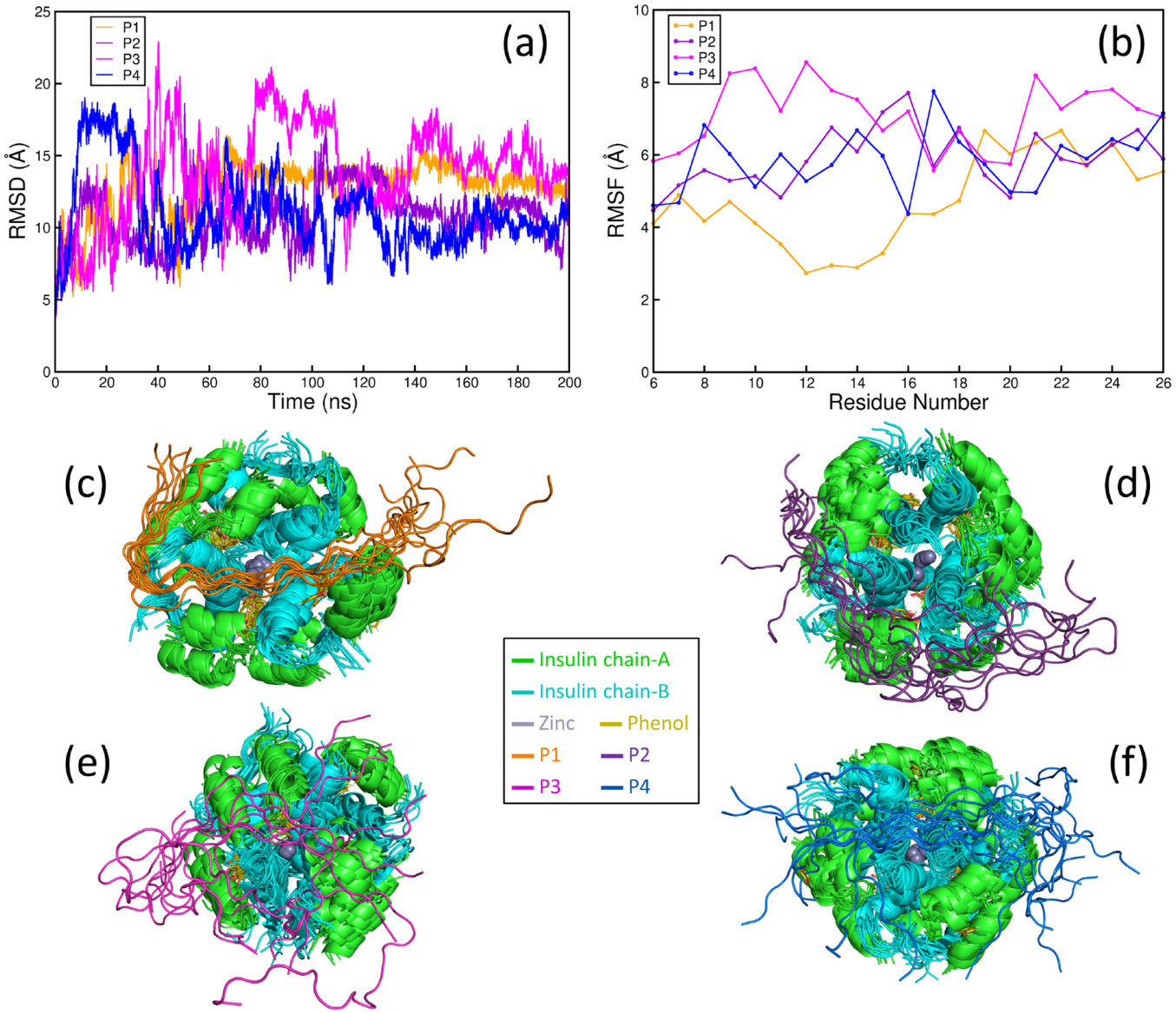



**Figure 7.** Analysis of simulation trajectories of all four protamines during a 200 ns MD run: (a) RMSD vs simulation time plot; (b) RMSF vs residue number plot for protamines in docked complexes. (c−f) Superimposed snapshots of insulin−protamine complexes at an interval of 20 ns.



https://doi.org/10.1021/acsomega.3c08445
ACS Omega 2024, 9, 4857−4869

dynamically stable over the simulation run. As seen by the RMSF versus residue number plot, the B-chain residues have very low RMSF values (less than 1 Å), explaining their stability on protamine binding. While protamine primarily binds to only a few initial N-terminal B-chain residues, the fluctuations in the majority of the B-chain are lower, indicating the enhanced stability of the B-chain due to protamine binding. The RMSF values of A-chain residues show a higher variation in comparison to those of B-chain residues. Consequently, we may conclude that the binding of protamine to insulin occurs primarily via the B-chain of insulin.

We also calculated the radius of gyration and SASA to further substantiate the above findings. The radius of gyration ($R_g$) of a protein structure dictates its compactness, and it is calculated as the root-mean-square distance of all atoms in the molecule or protein from its center of mass. It indicates how well a structure is folded over the MD simulation run. In our study, the radius of gyration ranges from 18.5 to 19.5 Å for the four docked complexes (Figure S6a). As expected, due to the binding of insulin with protamine, $R_g$ values for the four docked complexes are slightly higher than that for insulin. This also confirms that insulin in the docked complexes did not undergo any significant conformational changes during the simulation. The solvent-accessible surface area (SASA) of a biomolecule is the surface area that is accessible to a solvent. Figure S6b shows SASA for the four docked complexes. The higher the value of SASA, the greater is the disruption of the insulin hydration sphere. The plot shows that SASA values for all four docked complexes are almost identical, which eventually decrease as the simulation duration increases. The binding sites on insulin become more buried and inaccessible to solvent molecules, resulting in a decrease in SASA at that site, also indicating the stable binding of protamine to insulin.

**3.3. Dynamic Behavior of Protamine Peptides.** As discussed in the previous section, the insulin structure is quite stable after binding to protamine peptides. Additionally, we also evaluated protamine peptides during the MD simulation run. It was observed that all four protamine peptides retained the random coiled structure after the 200 ns MD run. Subsequently, they are expected to be more dynamic in nature, which is also evident from their exceptionally high RMSD values (5−25 Å) when compared to insulin (below 4 Å) (Figure 7a). The highly dynamic nature of protamine may be essential for the release of insulin from crystals when injected into the bloodstream. At the beginning of the MD simulation, the N- and C-termini of the protamine peptides do not interact with insulin. As a result, their termini are free to interact with the solvent and change position frequently over time, as also seen in the RMSD vs time plot (Figure 7a). Moreover, it is only after some time (60 ns) that the N- and C-termini also start to interact with insulin and tend to stabilize after ∼140 ns duration. Moreover, all four protamine peptides show variable RMSD values and therefore cannot provide information on their stability on binding. In this regard, the RMSF values for protamine peptides were calculated for residues 6 to 26 (excluding highly mobile N- and C-terminal residues) to rightly predict the minor fluctuations on binding (Figure 7b). We observed that P1 showed lower fluctuations than other protamine peptides, which is also seen from the H-bond analysis and binding energy calculations discussed further in this article.

**Overall structure:** The fluctuations in the structure of the insulin−protamine complexes can be seen in Figure 7c−f, where the snapshots of the insulin−protamine complex are captured every 20 ns and superimposed on each other, which shows how the structure is changing as the simulation time increases. The starting and end (200 ns MD run) structures of the insulin−protamine complex are shown in Figures 1 and 4, respectively. During the simulation time, there is a significant difference in the orientation of protamine peptides, but the alignment of insulin remains less significant. The binding interactions of the middle portion (residue numbers 6 to 26) of protamine peptides are more stable as compared to the outer flexible part over most of the time during the simulation run. This is evident from Figure 7c−f, where we can see that as the simulation time increases, insulin is quite stable, but the protamine peptides are highly dynamic, which happens due to the random coiled structure and the free termini of protamine peptides as discussed earlier. Toward the end of the run, the fluctuations in the protamine peptides are reduced drastically and the insulin−protamine complex becomes more stable as compared to the starting structure.

**3.4. Interactions Providing Stability during MD Simulation.** *3.4.1. Hydrogen Bonding Interactions.* Hydrogen bonding in proteins is the most common interaction between amino acids, as they contain electronegative atoms in abundance. The greater the number of H-bonds in a protein−ligand complex, the stronger the interactions and the more stable the complex. For the docked complexes, hydrogen bonding analysis was performed between protamine and insulin over 200 ns MD simulation time using GROMACS. For the calculation of H-bonding, a cutoff distance between the donor and acceptor atoms was considered as 3.5 Å and the angle cutoff between H−D−A (hydrogen−donor−acceptor) was taken as 30°. Figure 8 depicts the total number of hydrogen bonds between protamine and insulin in all four docked complexes during the course of 200 ns simulation. It shows that P1 forms a slightly higher number of H-bonds and, therefore, it suggests that the P1-R6 complex is slightly more stable among all four docked complexes. All other docked complexes (P2-R6, P3-R6, and P4-R6) showed almost the

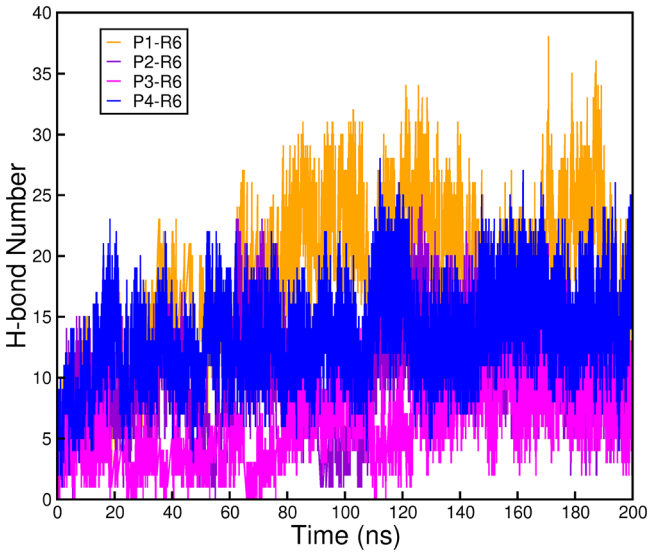



**Figure 8.** Comparison of the change in the total number of hydrogen bonds with the increase in simulation time between protamine and insulin in the docked complexes: (a) P1-R6, (b) P2-R6, (c) P3-R6, and (d) P4-R6.



https://doi.org/10.1021/acsomega.3c08445
ACS Omega 2024, 9, 4857−4869

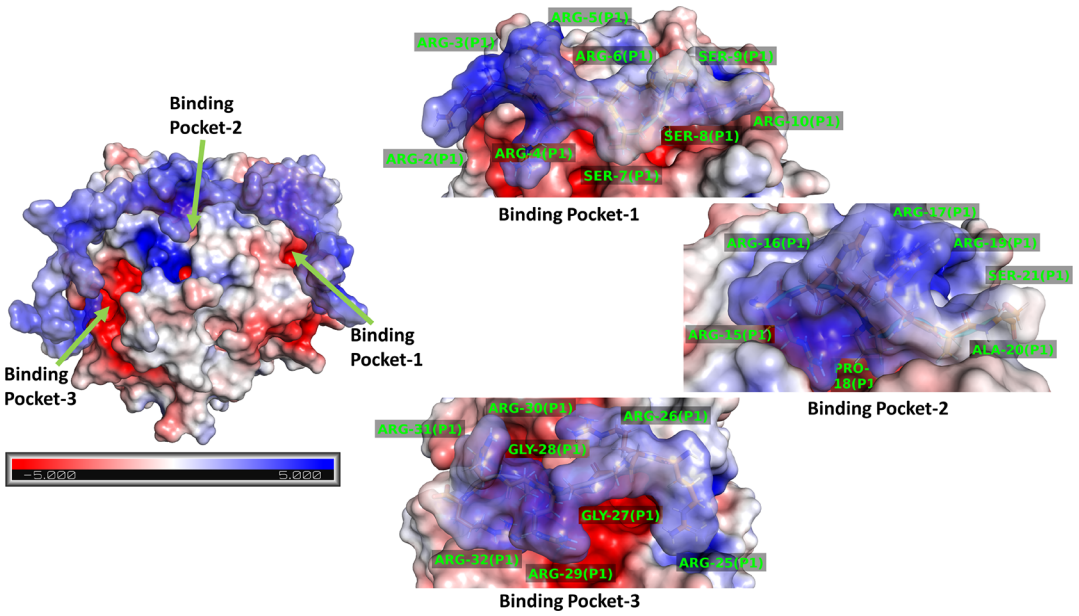



**Figure 9.** Electrostatic potential energy surface representation of the P1-R6 complex. The insulin−protamine complex is rendered as an electrostatic surface where the red color indicates the region with negative potential and the blue color represents the region with positive potential. Stick representation of peptide 1 is indicated in orange color. Three binding pockets with negative electrostatic potential and their interaction with protamine are shown as expanded regions.

same number of H-bonds, which fluctuates over the simulation run. The fact that the P1-R6 complex is the most stable in comparison to others on the basis of H-bonding analysis is also consistent with the lowest RMSD value (Figure 6) that was shown in the previous section.

*3.4.2. Electrostatic Interactions.* Along with H-bonding and hydrophobic interactions, salt bridge interactions between a positively charged amino acid and a negatively charged amino acid play a major role in providing stability to a protein−ligand complex. These are essential for the protein's conformational stability and specificity. In the current study, salt bridge interactions were identified using VMD software, having a cutoff distance of less than 3.2 Å between any of the nitrogen atoms of a basic amino acid and the oxygen atoms of an acidic amino acid. Salt bridge interactions were calculated for all of the docked complexes over a 200 ns timescale. It was observed that the acidic amino acid residues (glutamic acid) $Glu^{A4}$, $Glu^{A17}$, and $Glu^{B21}$ of insulin form salt bridges with basic amino acid (arginine) residues of protamine peptides. We have taken snapshots of the insulin−protamine complex every 20 ns to see whether salt bridge interactions were intact over the complete simulation time duration. Table S2 of the Supporting Information shows salt bridge interactions between insulin and protamine at every 20 ns in the P1-R6 complex over 200 ns simulation time. Highlighted residues in Table S2 are the salt bridge interactions retained over the simulation run. While $Glu^{A4}$(II)-$Arg^{P13}$ and $Glu^{A4}$(II)-$Arg^{P10}$ interactions are retained over the whole 200 ns MD run, a new salt bridge interaction $Glu^{A17}$(III)-$Arg^{P31}$ is formed after 80 ns, which is retained up to 200 ns simulation time. These electrostatic interactions account for the stability of the insulin−protamine complex. Salt bridge interactions in the other docked complexes follow a similar trend (refer to Tables S3−S5 of the Supporting Information).

Above inference can also be supported by mapping the electrostatic potential of the docked complexes after a 200 ns MD run. For this purpose, we have used the APBS (adaptive Poisson−Boltzmann solver) electrostatics plugin in PyMOL. The docked complex P1-R6 after 200 ns is rendered as a surface that is colored according to the electrostatic potential (Figure 9). The lower and upper ranges of the potential energy surface are −5 and +5 kT/e, respectively. We have identified three negatively charged surface pockets (red color), binding pocket-1, binding pocket-2, and binding pocket-3, on the surface of insulin. As protamine contains a large number of arginine amino acid residues, the overall surface charge of protamine is expected to be positive (blue color), which will interact with negatively charged pockets on the insulin hexamer. We can see that the residues of protamine that are present in these binding pockets are mainly arginine residues essential for the stability of the insulin−protamine complex through these electrostatic interactions.

**3.5. Binding Energy Analysis.** Finally, we performed binding energy analysis on the last 20 ns of the trajectory. The binding energies between insulin and protamine in the insulin−protamine complexes, calculated using the MM/GBSA method, are shown in Table 2. The $\Delta G_{bind}$ values of all docked complexes are negative, which means that binding of all four protamine peptides to insulin is energetically favorable and P1 is most strongly bound to insulin as reflected by the higher negative value of $\Delta G_{bind}$ for P1-R6 (−114.79 kcal/mol). This outcome is consistent with the observed lower RMSD value and the highest number of H-bonds observed for the P1-R6 complex. The $\Delta G_{bind}$ values for P2-R6, P3-R6, and P4-R6 complexes are −39.47, −33.49, and −42.47 kcal/mol, respectively. These values are in the acceptable range for a receptor−ligand complex. The most favorable binding of P1 to R6 can be explained based on the different polarities of protamine peptides. In the RP-HPLC profile of protamine peptides,[17,18] it was observed that P1 is the most polar peptide of protamine as it elutes first. Hence, the protamine peptide that binds most efficiently to insulin should be the most polar. This can also be seen through the increased electrostatic


https://doi.org/10.1021/acsomega.3c08445
ACS Omega 2024, 9, 4857−4869

Table 2. Binding Energy Calculation Results (±Standard Deviation) Obtained from MM/GBSA for All Four Docked Complexes in the Unit of kcal/mol[a]

| complex | $\Delta E_{vdW}$ | $\Delta E_{elec}$ | $\Delta E_{GB}$ | $\Delta E_{SURF}$ | $\Delta G_{gas}$ | $\Delta G_{solv}$ | $-T\Delta S$ | $\Delta G_{bind}$ |
|---|---|---|---|---|---|---|---|---|
| P1-R6 | −125.22 ± 3.99 | −1606.69 ± 6.12 | 1536.61 ± 5.49 | −21.13 ± 0.49 | −1731.91 ± 9.09 | 1515.47 ± 5.51 | 101.65 ± 0.05 | −114.79 ± 10.45 |
| P2-R6 | −94.38 ± 3.17 | −1523.18 ± 4.52 | 1460.75 ± 3.67 | −15.48 ± 0.39 | −1617.56 ± 7.72 | 1445.28 ± 3.69 | 132.81 ± 0.05 | −39.47 ± 8.91 |
| P3-R6 | −66.29 ± 1.33 | −1287.82 ± 13.38 | 1232.54 ± 6.52 | −10.93 ± 0.04 | −1354.11 ± 14.48 | 1221.61 ± 6.52 | 99.01 ± 0.05 | −33.49 ± 11.98 |
| P4-R6 | −80.05 ± 0.40 | −1489.96 ± 7.57 | 1425.80 ± 4.71 | −15.24 ± 0.07 | −1570.02 ± 9.09 | 1410.56 ± 4.71 | 116.98 ± 0.05 | −42.47 ± 11.35 |

[a]In the calculations, insulin acts as a receptor and protamine acts as a ligand.

energy $\Delta E_{elec}$ for the docked P1-R6 complex. $\Delta E_{elec}$ is a significant contributor to the binding energy. The fact that $\Delta E_{vdW}$ values are substantially lower than $\Delta E_{elec}$ values indicates that hydrophobic interactions are not as significant as electrostatic interactions, which is consistent with the experimental study done on an equimolar mixture of insulin and protamine.[48] Therein, it was shown that the binding interactions between insulin and protamine are hydrophilic in nature. This is also expected as protamine contains arginine amino acids, which mainly interact through H-bonding and electrostatic interactions. It is important to note that the presence of serine residues in protamine peptides along with arginine residues is also a key factor in forming the binding interface between insulin and protamine. Serine residues are involved in the formation of H-bonds with insulin residues through the side-chain −OH group.

Figure 10 shows the per-residue decomposition analysis for the four docked complexes. The residues predicted to be involved in the binding surface formed after docking showed negative values in the per-residue decomposition analysis. The majority of these residues are of the B-chain of insulin. Lysine residues of insulin show a positive value of energy in the decomposition analysis, which could be due to repulsion from positively charged arginine residues. The P1-R6 complex has the greatest number of residues that favorably bind to the protamine peptide. This is also evident from the overall binding energy analysis of the P1-R6 complex, which is the most negative among all four docked insulin−protamine complexes. We found that the residues that are predicted to take part in the binding show negative binding energy and are actively taking part in the interactions. So, our results are consistent with the docking results, and the complex is quite stable under solvent and force-field conditions.

## 4. CONCLUSIONS

Insulin−protamine interactions are critical to insulin's precipitation and, henceforth, its delayed release. At present, these interactions are largely understood based on the physical properties of the sample, although no clear descriptions are available. This is largely due to the limitations of available analytical tools; for example, previous studies on the insulin−protamine complex could only propose a potential binding pocket near $Asn^{B3}$ residue based on the electron density but did not infer much on the binding epitope, which could potentially play an important role in its stability. Another HDX-MS study performed on insulin analogues discussed fewer deuterium incorporation for the two intermediate-acting groups (insulin NPH and NPL (Neutral Protamine Lispro)) than human insulin, and other fast-acting analogues such as insulin Lispro and insulin Aspart demonstrated that the slower HDX reactivity of insulin analogues at neutral pH was associated with their oligomeric stability in the presence of protamine peptides.[51] Nonetheless, this study only informed on global stability and did not yield information on site-specific binding. This work describes a combination of molecular docking and molecular dynamics simulations that can be utilized to locate the binding sites and calculate the binding energies of the insulin−protamine complex. We have carried out a comparative study of the binding events of the four basic protamine (arginine-rich) peptides with nearly similar primary amino acid sequences in complexation with the insulin hexamer. Results from molecular docking reveal that the four


https://doi.org/10.1021/acsomega.3c08445
ACS Omega 2024, 9, 4857−4869

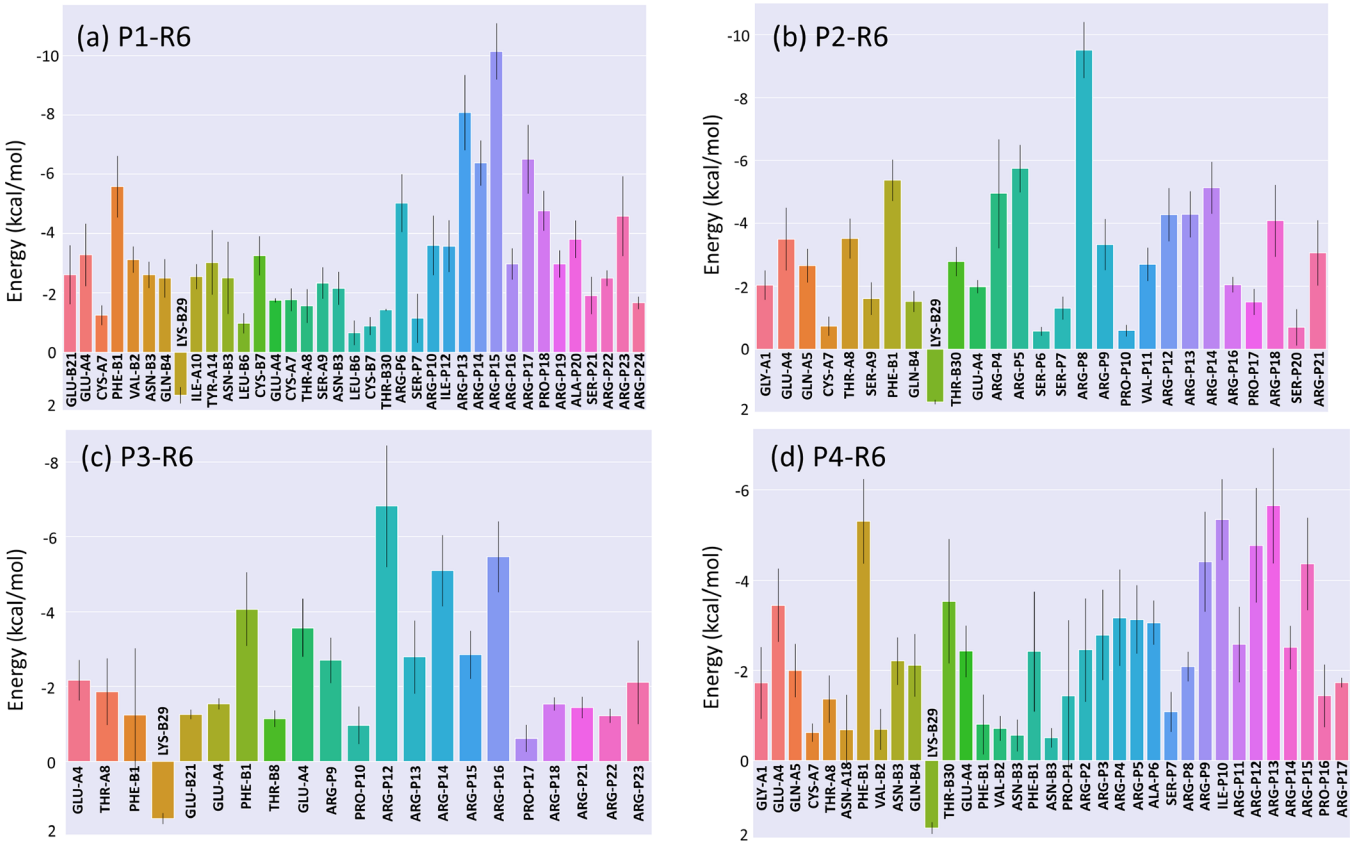



**Figure 10.** Residue decomposition plot (MM/GBSA) for (a) P1-R6, (b) P2-R6, (c) P3-R6, and (d) P4-R6 complexes representing the binding energy contribution of the active residues. The residues actively taking part in the interactions have negative values of energy in the decomposition plot.

protamine peptides show nearly similar propensity toward binding with the R6 insulin hexamer. These results are important, as they suggest that minor variations in the stoichiometry of four peptides may not impact the insulin−protamine complex. An earlier experimental and docking study by Aggarwal et. al.[48] suggested increased flexibility of the complex upon binding, which is contrary to our results. Our simulations clearly indicated lower potential energy throughout the simulation run when compared with human insulin in the absence of protamine. The stability of the insulin−protamine complex is explained through H-bonding, hydrophobic, and electrostatic (glutamic acid residues of insulin and arginine residues of protamine peptides) interactions. Additionally, the calculated overall binding energy for the P1-R6 complex is found to be more negative and therefore most likely forms a more stable complex in comparison to the remaining three insulin−protamine complexes. We speculate a higher polarity of P1 in comparison to the other three peptides as revealed by its largest negative electrostatic energy in the bound state with the insulin hexamer. Our study suggests the binding of protamine on the initial half of the B-chain, which may also lead to reduced levels of deamination in the insulin−protamine complex as reported previously.[48] Throughout the MD run, insulin's hexameric state in the insulin−protamine complex is retained, accounting for its dynamic stability. On the other hand, the terminal residues of the protamine peptides acquire dynamic stability in the insulin−protamine complex toward the end of the MD run. We believe that results emerging from this study may be a step forward to aid the pharmaceutical industry in improving existing formulations through site-directed mutagenesis and modifying the pharmacokinetics of insulin NPH formulations. Furthermore, information on binding sites might help in the design of drugs that can alter the absorption time profile inside the body.

## ■ ASSOCIATED CONTENT

### Data Availability Statement

In this study, structure predictions of protamine peptides were carried out using Alphafold (https://colab.research.google.com/github/deepmind/alphafold/blob/main/notebooks/AlphaFold.ipynb). The crystal structure of the insulin R6 hexamer was downloaded from the RCSB Protein Data Bank (https://www.rcsb.org/). Molecular docking studies were performed using the HADDOCK 2.4 web server (https://wenmr.science.uu.nl/haddock2.4/). MD simulations were carried out using GROMACS 2020.6 (https://www.gromacs.org/) with the topology and mdp files generated from the CHARMM-GUI online interface (https://www.charmm-gui.org/). PyMOL (http://www.pymol.org/2/) and VMD (https://www.ks.uiuc.edu/Research/vmd/) were used for visualization and analysis of the data. MM/GBSA binding energy calculations were performed using the gmx_MMPBSA module (https://valdes-tresanco-ms.github.io/gmx_MMPBSA/). Data analyzed in this study will be made available upon request.

### Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.3c08445.



https://doi.org/10.1021/acsomega.3c08445
ACS Omega 2024, 9, 4857−4869

Figures showing the H-bonding interactions in docked P2-R6, P3-R6, and P4-R6 complexes (Figure S1), hydrophobic interactions in all four docked complexes (Figure S2), arginine side chains between $Asn^{B3}$ residues in docked P2-R6, P3-R6, and P4-R6 complexes (Figure S3), last 20 ns RMSD plot (Figure S4), ring flipping in Tyr amino acid (Figure S5), and $R_g$ and SASA plots for the 200 ns simulation trajectory (Figure S6); tables listing the H-bonding partners between insulin and protamine peptides (Table S1) and salt bridge interactions in insulin−protamine complexes over the 200 ns trajectory (Tables S2−S5) (PDF)

## AUTHOR INFORMATION


### Corresponding Author

**Manoj Kumar Pandey** − *Department of Chemistry, Indian Institute of Technology Ropar, Rupnagar, Punjab 140001, India;* orcid.org/0000-0002-4015-8764; Email: mkpandey@iitrpr.ac.in

### Author

**Ketan Kumar Rohilla** − *Department of Chemistry, Indian Institute of Technology Ropar, Rupnagar, Punjab 140001, India*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.3c08445


### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

M.K.P. would like to acknowledge the Science and Engineering Research Board (SERB)−Department of Science and Technology (DST), Government of India (SERB DST Grant ECR/2017/000713) and IIT Ropar for the financial support. K.K.R. acknowledges the financial support from the University Grants Commission (UGC), Government of India.

## REFERENCES


(1) Rosenfeld, L. Insulin: Discovery and Controversy. *Clin. Chem.* **2002**, *48* (12), 2270−2288.

(2) Vecchio, I.; Tornali, C.; Bragazzi, N. L.; Martini, M. The Discovery of Insulin: An Important Milestone in the History of Medicine. *Front. Endocrinol.* **2018**, *9*, 613 DOI: 10.3389/fendo.2018.00613.

(3) Michael, D. F.; Steiner, L. H. *Insulin Biosynthesis, Secretion, Structure, and Structure-Activity Relationships*; MDText.com, Inc. 2015.

(4) Dodson, G.; Steiner, D. The Role of Assembly in Insulin's Biosynthesis. *Curr. Opin. Struct. Biol.* **1998**, *8* (2), 189−194.

(5) Petersen, M. C.; Shulman, G. I. Mechanisms of Insulin Action and Insulin Resistance. *Physiol. Rev.* **2018**, *98* (4), 2133−2223.

(6) Rahman, M. S.; Hossain, K. S.; Das, S.; Kundu, S.; Adegoke, E. O.; Rahman, M. A.; Hannan, M. A.; Uddin, M. J.; Pang, M.-G. Role of Insulin in Health and Disease: An Update. *Int. J. Mol. Sci.* **2021**, *22* (12), 6403.

(7) Colwell, A. R., Sr. The Banting Memorial Lecture 1968: Fifty Years of Diabetes in Perspective. *Diabetes* **1968**, *17* (10), 599−610.

(8) Marwood, S. F. Diabetes Mellitus–Some Reflections. *J. R. Coll. Gen. Pract.* **1973**, *23* (126), 38−45.

(9) Aschner, P. Insulin Therapy in Type 2 Diabetes. *Am. J. Ther.* **2020**, *27* (1), e79−e90.

(10) Jarosinski, M. A.; Dhayalan, B.; Chen, Y. S.; Chatterjee, D.; Varas, N.; Weiss, M. A. Structural Principles of Insulin Formulation and Analog Design: A Century of Innovation. *Mol. Metab.* **2021**, *52* (101325), No. 101325.

(11) Hagedorn, H. C. Protamine Insulinate. *J. Am. Med. Assoc.* **1936**, *106* (3), 177.

(12) Simkin, R. D.; Cole, S. A.; Ozawa, H.; Magdoff-Fairchild, B.; Eggena, P.; Rudko, A.; Low, B. W. Precipitation and Crystallization of Insulin in the Presence of Lysozyme and Salmine. *Biochim. Biophys. Acta* **1970**, *200* (2), 385−394.

(13) Saleem, F.; Sharma, A. NPH Insulin. *In StatPearls*; StatPearls Publishing, 2023. https://www.ncbi.nlm.nih.gov/books/NBK549860/.

(14) Heise, T.; Mathieu, C. Impact of the Mode of Protraction of Basal Insulin Therapies on Their Pharmacokinetic and Pharmacodynamic Properties and Resulting Clinical Outcomes: HEISE AND MATHIEU. *Diabetes Obes. Metab.* **2017**, *19* (1), 3−12.

(15) Lewis, J. D.; Song, Y.; de Jong, M. E.; Bagha, S. M.; Ausió, J. A Walk Though Vertebrate and Invertebrate Protamines. *Chromosoma* **2003**, *111* (8), 473−482.

(16) Balhorn, R. The Protamine Family of Sperm Nuclear Proteins. *Genome Biol.* **2007**, *8* (9), 227.

(17) Hoffmann, J. A.; Chance, R. E.; Johnson, M. G. Purification and Analysis of the Major Components of Chum Salmon Protamine Contained in Insulin Formulations Using High-Performance Liquid Chromatography. *Protein Expr. Purif.* **1990**, *1* (2), 127−133.

(18) Hvass, A.; Skelbaek-Pedersen, B. Determination of Protamine Peptides in Insulin Drug Products Using Reversed Phase High Performance Liquid Chromatography. *J. Pharm. Biomed. Anal.* **2005**, *37* (3), 551−557.

(19) Awotwe-Otoo, D.; Agarabi, C.; Keire, D.; Lee, S.; Raw, A.; Yu, L.; Habib, M. J.; Khan, M. A.; Shah, R. B. Physicochemical Characterization of Complex Drug Substances: Evaluation of Structural Similarities and Differences of Protamine Sulfate from Various Sources. *AAPS J.* **2012**, *14* (3), 619−626.

(20) Gucinski, A. C.; Boyne, M. T., 2nd; Keire, D. A. Modern Analytics for Naturally Derived Complex Drug Substances: NMR and MS Tests for Protamine Sulfate from Chum Salmon. *Anal. Bioanal. Chem.* **2015**, *407* (3), 749−759.

(21) Norrman, M.; Hubálek, F.; Schluckebier, G. Structural Characterization of Insulin NPH Formulations. *Eur. J. Pharm. Sci.* **2007**, *30* (5), 414−423.

(22) Gorai, B.; Vashisth, H. Progress in Simulation Studies of Insulin Structure and Function. *Front. Endocrinol. (Lausanne)* **2022**, *13*, No. 908724.

(23) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583−589.

(24) Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; Žídek, A.; Green, T.; Tunyasuvunakool, K.; Petersen, S.; Jumper, J.; Clancy, E.; Green, R.; Vora, A.; Lutfi, M.; Figurnov, M.; Cowie, A.; Hobbs, N.; Kohli, P.; Kleywegt, G.; Birney, E.; Hassabis, D.; Velankar, S. AlphaFold Protein Structure Database: Massively Expanding the Structural Coverage of Protein-Sequence Space with High-Accuracy Models. *Nucleic Acids Res.* **2022**, *50* (D1), D439−D444.

(25) Chang, X.; Jorgensen, A. M.; Bardrum, P.; Led, J. J. Solution Structures of the R6 Human Insulin Hexamer. *Biochemistry* **1997**, *36* (31), 9409−9422.

(26) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F., Jr; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. The Protein Data Bank: A Computer-Based Archival File for Macromolecular Structures. *J. Mol. Biol.* **1977**, *112* (3), 535−542.

(27) Berman, H. M.; Battistuz, T.; Bhat, T. N.; Bluhm, W. F.; Bourne, P. E.; Burkhardt, K.; Feng, Z.; Gilliland, G. L.; Iype, L.; Jain, S.; Fagan, P.; Marvin, J.; Padilla, D.; Ravichandran, V.; Schneider, B.; Thanki, N.; Weissig, H.; Westbrook, J. D.; Zardecki, C. The Protein

Data Bank. *Acta Crystallogr. D Biol. Crystallogr.* **2002**, *58* (Pt 61), 899−907.
(28) de Vries, S. J.; van Dijk, M.; Bonvin, A. M. J. J. The HADDOCK Web Server for Data-Driven Biomolecular Docking. *Nat. Protoc.* **2010**, *5* (5), 883−897.
(29) Dominguez, C.; Boelens, R.; Bonvin, A. M. J. J. HADDOCK: A Protein-Protein Docking Approach Based on Biochemical or Biophysical Information. *J. Am. Chem. Soc.* **2003**, *125* (7), 1731−1737.
(30) Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. LIGPLOT: A Program to Generate Schematic Diagrams of Protein-Ligand Interactions. *Protein Eng. Des. Sel.* **1995**, *8* (2), 127−134.
(31) Schrödinger, L. L. C. *The PyMOL molecular graphics system*, version 2.5.4, 2022.
(32) Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Lemkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y.; Jo, S.; Pande, V. S.; Case, D. A.; Brooks, C. L., III; MacKerell, A. D., Jr; Klauda, J. B.; Im, W. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **2016**, *12* (1), 405−413.
(33) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J. Comput. Chem.* **2008**, *29* (11), 1859−1865.
(34) Brooks, B. R.; Brooks, C. L., III; 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.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30* (10), 1545−1614.
(35) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible, and Free. *J. Comput. Chem.* **2005**, *26* (16), 1701−1718.
(36) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1−2*, 19−25.
(37) Papaioannou, A.; Kuyucak, S.; Kuncic, Z. Molecular Dynamics Simulations of Insulin: Elucidating the Conformational Changes That Enable Its Binding. *PLoS One* **2015**, *10* (12), No. e0144058.
(38) Raghunathan, S.; El Hage, K.; Desmond, J. L.; Zhang, L.; Meuwly, M. The Role of Water in the Stability of Wild-Type and Mutant Insulin Dimers. *J. Phys. Chem. B* **2018**, *122* (28), 7038−7048.
(39) Li, Z.-L.; Buck, M. Modified Potential Functions Result in Enhanced Predictions of a Protein Complex by All-Atom Molecular Dynamics Simulations, Confirming a Stepwise Association Process for Native Protein−Protein Interactions. *J. Chem. Theory Comput.* **2019**, *15* (8), 4318−4331.
(40) Banerjee, P.; Bagchi, B. Dynamical Control by Water at a Molecular Level in Protein Dimer Association and Dissociation. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117* (5), 2302−2308.
(41) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33−38.
(42) Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA Methods to Estimate Ligand-Binding Affinities. *Expert Opin. Drug Discovery* **2015**, *10* (5), 449−461.
(43) Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **2011**, *51* (1), 69−82.
(44) Weng, G.; Wang, E.; Chen, F.; Sun, H.; Wang, Z.; Hou, T. Assessing the Performance of MM/PBSA and MM/GBSA Methods. 9. Prediction Reliability of Binding Affinities and Binding Poses for Protein-Peptide Complexes. *Phys. Chem. Chem. Phys.* **2019**, *21* (19), 10135−10145.
(45) Miller, B. R., 3rd; McGee, T. D., Jr; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. MMPBSA.Py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* **2012**, *8* (9), 3314−3321.
(46) Valdés-Tresanco, M. S.; Valdés-Tresanco, M. E.; Valiente, P. A.; Moreno, E. Gmx_MMPBSA: A New Tool to Perform End-State Free Energy Calculations with GROMACS. *J. Chem. Theory Comput.* **2021**, *17* (10), 6281−6291.
(47) Duan, L.; Liu, X.; Zhang, J. Z. H. Interaction Entropy: A New Paradigm for Highly Efficient and Reliable Computation of Protein-Ligand Binding Free Energy. *J. Am. Chem. Soc.* **2016**, *138* (17), 5722−5728.
(48) Aggarwal, S.; Tanwar, N.; Singh, A.; Munde, M. Formation of Protamine and Zn-Insulin Assembly: Exploring Biophysical Consequences. *ACS Omega* **2022**, *7* (45), 41044−41057.
(49) Alonso, H.; Bliznyuk, A. A.; Gready, J. E. Combining Docking and Molecular Dynamic Simulations in Drug Design. *Med. Res. Rev.* **2006**, *26* (5), 531−568.
(50) Santos, L. H. S.; Ferreira, R. S.; Caffarena, E. R. Integrating Molecular Docking and Molecular Dynamics Simulations. *Methods Mol. Biol.* **2019**, *2053*, 13−34.
(51) Nakazawa, S.; Hashii, N.; Harazono, A.; Kawasaki, N. Analysis of Oligomeric Stability of Insulin Analogs Using Hydrogen/Deuterium Exchange Mass Spectrometry. *Anal. Biochem.* **2012**, *420* (1), 61−67.


https://doi.org/10.1021/acsomega.3c08445
*ACS Omega* 2024, 9, 4857−4869