Received: 13 November 2023 | Revised: 8 February 2024 | Accepted: 6 March 2024

DOI: 10.1002/prot.26685

RESEARCH ARTICLE

PROTEINS WILEY

Molecular basis for differential recognition of an allosteric inhibitor by receptor tyrosine kinases

Jyoti Verma[1] | Harish Vashisth[1,2,3,4]

[1]Department of Chemical Engineering and Bioengineering, University of New Hampshire, Durham, New Hampshire, USA

[2]Department of Chemistry, University of New Hampshire, Durham, New Hampshire, USA

[3]Integrated Applied Mathematics Program, University of New Hampshire, Durham, New Hampshire, USA

[4]Molecular and Cellular Biotechnology Program, University of New Hampshire, Durham, New Hampshire, USA

Correspondence
Harish Vashisth, Department of Chemical Engineering and Bioengineering, University of New Hampshire, 33 Academic Way, Durham, NH 03824, USA.
Email: harish.vashisth@unh.edu

Funding information
National Institute of Health (NIH), Grant/Award Number: R35GM138217; NSF EPSCoR, Grant/Award Number: OIA-1757371

Abstract

Understanding kinase-inhibitor selectivity continues to be a major objective in kinase drug discovery. We probe the molecular basis of selectivity of an allosteric inhibitor (MSC1609119A-1) of the insulin-like growth factor-I receptor kinase (IGF1RK), which has been shown to be ineffective for the homologous insulin receptor kinase (IRK). Specifically, we investigated the structural and energetic basis of the allosteric binding of this inhibitor to each kinase by combining molecular modeling, molecular dynamics (MD) simulations, and thermodynamic calculations. We predict the inhibitor conformation in the binding pocket of IRK and highlight that the charged residues in the histidine-arginine-aspartic acid (HRD) and aspartic acid-phenylalanine-glycine (DFG) motifs and the nonpolar residues in the binding pocket govern inhibitor interactions in the allosteric pocket of each kinase. We suggest that the conformational changes in the IGF1RK residues M1054 and M1079, movement of the αC-helix, and the conformational stabilization of the DFG motif favor the selectivity of the inhibitor toward IGF1RK. Our thermodynamic calculations reveal that the observed selectivity can be rationalized through differences observed in the electrostatic interaction energy of the inhibitor in each inhibitor/kinase complex and the hydrogen bonding interactions of the inhibitor with the residue V1063 in IGF1RK that are not attained with the corresponding residue V1060 in IRK. Overall, our study provides a rationale for the molecular basis of recognition of this allosteric inhibitor by IGF1RK and IRK, which is potentially useful in developing novel inhibitors with improved affinity and selectivity.

KEYWORDS
allosteric inhibitor, docking, insulin-like growth factor receptor, insulin receptor, molecular dynamics simulation, receptor tyrosine kinase

1 | INTRODUCTION

Protein phosphorylation is a key post-translational modification in signaling proteins.[1,2] Specifically, these modifications are synergistically regulated by protein kinases and phosphatases.[3] Protein kinases are phosphotransferase enzymes that catalyze the transfer of the gamma phosphoryl group from Adenosine triphosphate (ATP) to specific

Abbreviations: Akt, an alias for protein kinase B; CK2, casein kinase II alpha; DFG, aspartic acid-phenylalanine-glycine; HRD, histidine-arginine-aspartic acid; IGF1RK, insulin-like growth factor 1 receptor kinase; IRE1, inositol-requiring enzyme 1; IRK, insulin receptor kinase; JAK, Janus kinase; LIMK2, LIM domain kinase; MAP, mitogen-activated protein; MEK1/2, mitogen-activated extracellular signal-regulated kinase; PAK, P21-activated kinase; PI3K, phosphoinositide 3-kinase; PTP1B, protein tyrosine phosphatase 1B.

residues in protein substrates.[4] These signaling proteins are often connected by sequential protein–protein interactions critical for cellular signaling. Given the essential role of kinases in various cellular events, their function is tightly regulated because their dysregulation is frequently found to be oncogenic and associated with the onset of neurological and immunological disorders, and infectious diseases.[5–9] Therefore, kinases are among the most sought-after therapeutic targets in cancer and other metabolic disorders. However, a major challenge in developing a selective inhibitor for a specific kinase among a group of kinases is the sequence conservation across the human kinome.[10,11] A majority of approved and preclinical small-molecule kinase inhibitors have undesired selectivity profiles.[12,13] Besides, low specificity toward the targeted kinase and lack of selectivity for structurally comparable kinases often leads to off-target toxicity, side effects, and drug resistance. The selectivity barriers are often associated with the ATP competitive kinase inhibition (type I and type II) due to remarkable sequence and structural similarity in the ATP-binding pockets in kinases.[14]

Nevertheless, a common strategy to achieve inhibitor specificity is to target the inactive conformation in kinases, as more structural variations across kinases are observed in the inactive state.[15–17] One of the examples is the discovery of imatinib, a potent and specific inhibitor of the Abl kinase which fails to inhibit homologous tyrosine kinases like c-Src.[18] The molecular determinants responsible for imatinib's specificity toward Abl have been extensively investigated using molecular simulations and kinetics-based studies.[19–21] Moreover, various structure-based strategies using computational, biophysical, and crystallographic techniques have been used to rationally develop selective kinase inhibitors reaching clinical trials.[12] CX4945, a casein kinase II alpha (CK2) inhibitor, has entered phase I clinical trials, demonstrating the importance of the shape, electrostatics, and the flexibility of the ATP-binding pocket in developing selective inhibitors.[22,23] The discovery of CX4945 also highlighted the importance of water-mediated interactions in the ATP-binding site of kinases.[24] The identification of phosphoinositide 3-kinase (PI3K) isoform-specific inhibitor is another striking example of obtaining selectivity among protein isoforms sharing a high degree of sequence homology in the ATP-binding sites. The design of inhibitors with thienopyrimidine scaffold having selectivity for PI3Kα over PI3Kβ suggests that selectivity can be rationalized by critical backbone interactions and differences in the electrostatic potential among the protein isoforms.[25] Furthermore, other strategies such as the fragment-based approach and the use of metal-based chemical probes in the design of small-molecule kinase inhibitors have been explored to achieve selectivity.[26–29]

Moreover, allosteric inhibition has been proposed as a promising approach to overcome the current limitations of orthosteric inhibitors and achieve selectivity.[30,31] As a consequence of conformational variability among kinases, selectivity can be achieved by targeting non-ATP-binding sites referred to as allosteric sites.[32,33] The allosteric sites are described as type III and type IV pockets, where type III inhibitors bind adjacent to the ATP-binding site and type IV inhibitors bind away from the ATP pocket.[34] The Food and Drug Administration (FDA) approval of trametinib in 2013, the first small-molecule allosteric inhibitor targeting mitogen-activated extracellular signal-regulated kinase (MEK1/2), served as a crucial turning point in the discovery of allosteric inhibitors.[35] Allosteric kinase inhibition strategy evolved rapidly with the progression of more than 10 MEK and protein kinase B (Akt) allosteric inhibitors reaching clinical trials and exploration of allosteric pockets and inhibitors in several kinases such as P21-activated kinase (PAK), inositol-requiring enzyme 1 (IRE1), LIM domain kinase (LIMK2), protein tyrosine phosphatase 1B (PTP1B), Janus kinase (JAK), and many others.[36] The discoveries of allosteric inhibitors have further illuminated the strategies and challenges in allosteric modulator design for kinases.

One of the key propositions to develop allosteric kinase inhibitors is the variability of residue composition and conformations of allosteric pockets in achieving higher inhibitor selectivity. However, designing an inhibitor with selectivity is still challenging for protein isoforms and homologs sharing a high degree of sequence and structural homology. Despite this, efforts have been made to develop isoform-specific and selective inhibitors. The selectivity obtained among the PI3K isoforms is one such example.[25] Further, structural studies of the MAP kinase p38 have revealed that a single residue difference appears to be sufficient for compounds to distinguish among various p38 isoforms.[37] A recent study elucidated a structure-based approach for designing selective covalent allosteric inhibitors for the Akt isoform.[38] Another example of inhibitor selectivity within homologous proteins is presented in a study reporting a novel class of allosteric insulin-like growth factor-I receptor (IGF-1R) inhibitors.[39] In this study, the structure–activity relationship (SAR) studies were used to discover an indole-butyl-amine derivative, MSC1609119A-1, as an allosteric inhibitor of the insulin-like growth factor receptor 1 kinase (IGF1RK). In addition, the cellular assays highlighted the selectivity of the inhibitor toward IGF1RK (IC_{50}: 0.4 μM) over the homologous insulin receptor kinase (IRK) (IC_{50}: 6.9 μM).

IGF1RK and IRK are homologous protein members of the receptor tyrosine kinase (RTK) superfamily involved in mediating conserved signaling pathways.[40] Although these receptors are coupled to similar intracellular signaling networks, their biological functions are distinct. IR plays a key role in regulating metabolic functions, particularly insulin metabolism, whereas IGF-1R is mainly involved in growth regulation functions.[41,42] They have a high degree of homology, reflecting their similar ancestry.[43] Both receptors are expressed at the cellular surface in the $\alpha_2\beta_2$ configuration. The α subunit contains the ligand binding domain on the extracellular region and the β subunit includes a large cytoplasmic domain with tyrosine kinase activity.[44,45] The kinase domains of these proteins share 85% sequence similarity and a high structural homology (Figure 1A,B). IGF1R kinase and IR kinase, like any other kinase domain, have an αC-helix, a glycine-rich loop (GC loop), and two other longer loops, the catalytic loop (C-loop) and the activation loop (A-loop), which are essential for the catalytic activity (Figure 1B). The tertiary structures of the kinase domains of IGF1R and IR are similar in their inactivated, nonphosphorylated, and fully activated phosphorylated forms.[46] Consequently, it is challenging to selectively target the kinase domain of one of the receptors. Therefore, it is crucial to investigate the mechanism by which the allosteric inhibitor, MSC1609119A-1, selectively inhibits IGF1RK over IRK.

Such instances reporting inhibitor selectivity might be attributed to the conformational selection mechanism, binding affinity differences, or

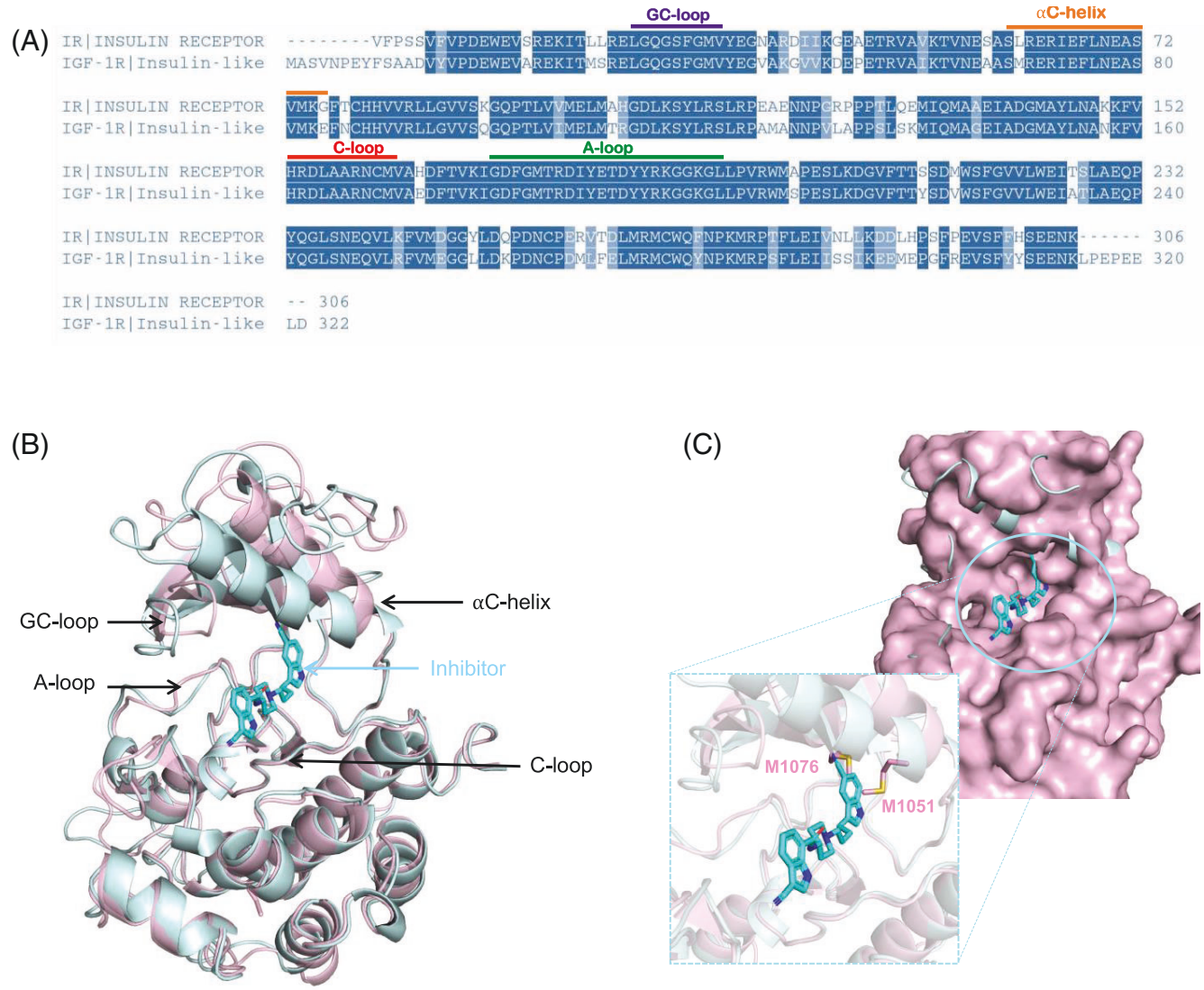

FIGURE 1 Structure and sequence comparison of the insulin-like growth factor-I receptor kinase (IGF1RK) and insulin receptor kinase (IRK) domains. (A) Sequence alignment of IRK and IGF1RK kinase domains. The conserved amino acid residues are highlighted in blue. Identical and similar amino acid residues are shown in dark and light shades, respectively. (B) A structural superimposition of the inhibitor-bound IGF1RK structure (cyan) with the apo IRK structure (pink). (C) The steric hindrance to the inhibitor due to the IRK residues M1051 and M1076 is highlighted via superimposed structures of IGF1RK bound to the inhibitor (cyan) and the apo IRK structure (pink).

both. Hence, a molecular-level understanding of the inhibitor interactions in the binding pocket is crucial for rationalizing the selectivity differences observed in experiments.[39] Since the sequences and structures of these kinases are significantly conserved, it is challenging to determine the molecular mechanisms responsible for the binding specificity of MSC1609119A-1. However, the cocrystallized structure of the allosteric inhibitor, MSC1609119A-1, with the IGF1RK domain (PDB ID 3LW0) informed on the structural basis for inhibitor binding and interactions.[39] For example, the conformational change in the sidechains of M1051 and M1076 in IRK, resulting in the distinct shape of the allosteric pockets (Figure 1C), provided support for the experimentally observed lack of selectivity of the inhibitor against IRK (IC_{50}: 6.9 μM).[39] However, due to the lack of structural data on the complex of this allosteric inhibitor with IRK, the location of the inhibitor binding pocket, the binding mode of the inhibitor, and its interactions with IRK remain unresolved.

Therefore, based on the experimental results and structural guidance from inhibitor-bound conformation of IGF1RK,[39] we sought to resolve the following questions: (i) the binding mode of the allosteric inhibitor in the IRK pocket, (ii) conformational dynamics of the inhibitor in the binding pocket of each kinase domain, and (iii) the inhibitor binding free energy and residue interactions with each kinase. Specifically, we judiciously employed molecular modeling, docking, all-atom molecular dynamics (MD) simulations, and free-energy calculations to elucidate the factors underlying inhibitor selectivity for IGF1RK over IRK.

2 | METHODS

2.1 | Structural modeling of IGF1RK

The coordinates from the experimentally determined structure[39] of the allosteric inhibitor (MSC1609119A-1) with the IGF1RK served as a starting point for our modeling and simulations of IGF1RK and IRK. Specifically, the crystal structure of IGF1RK in complex with the allosteric inhibitor was obtained from the Protein Data Bank (PDB ID 3LW0).[39,47] Using the Chimera interface of Modeller, we modeled the missing residues A1097 through A1105 and G1169 through G1171 in the IGF1RK structure.[48,49] The atomic coordinates for the allosteric inhibitor and water molecules were retained in the modeled structure. Finally, the modeled IGF1RK structure in complex with the allosteric inhibitor was used for subsequent MD simulations. Furthermore, sequence and structural level comparisons of the IGF1R and IR kinase domains were performed. The amino-acid sequences of the IGF1RK/IRK domains were retrieved in the FASTA format for pairwise sequence alignment using the EMBL-EBI EMBOSS Needle tool.[50] Next, the structural superimposition of the IGF1RK and IRK domains was performed using the PyMOL program.[51] The C_α atoms of the residues were used as the reference for structural alignment.

2.2 | Molecular docking of the allosteric inhibitor with the IRK domain

To predict an initial binding conformation of the allosteric inhibitor in the binding pocket of the IRK domain, molecular docking of the inhibitor with the apo IRK crystal structure (PDB ID 1IRK) was performed. The structural superimposition of apo IRK and inhibitor-bound IGF1RK (PDB ID 3LW0) structures indicated that all residues that form the allosteric pocket of the inhibitor in IGF1RK are conserved in IRK. As a result, we hypothesized the existence of a comparable allosteric binding region in the IRK structure. Furthermore, the allosteric inhibitor coordinates were extracted from the IGF1RK-bound structure (PDB ID 3LW0). The structural optimization of the allosteric inhibitor is required prior to docking. The LigPrep module of the Schrödinger software package was used to parameterize and minimize the ligand structure.[52,53] We optimized the inhibitor structure using the OPLS3e force-field without altering its ionization state. For the stereochemical information, we selected the "retain specified chiralities" option to keep the chirality information from the input crystal structure PDB file. Finally, the inhibitor structure was minimized and the output was generated in the maestro format (.mae/.maegz). The next step was accomplished using the protein preparation wizard of the Schrödinger software suite.[52] In this step, the protein structure (IRK) was first pre-processed with the removal of water molecules, the addition of missing hydrogen atoms, and any missing atoms in residues reported in the crystal structure (PDB ID 1IRK). The next stage of protein preparation was to optimize the hydrogen bonding network. In the H-bond assignment step, the PROPKA program was used to produce optimal residue ionization/tautomer states to maximize the H-bond network. Finally, a restrained energy minimization was performed with the OPLS3e force-field to remove any steric clashes allowing the displacements of the heavy atoms up to 0.3 Å. Subsequently, a grid was generated on the receptor to designate the binding site for docking of the ligand. In the receptor grid generation panel, a grid box of 20 $Å^3$ was generated around the predicted IRK binding site residues, L1030, F1044, E1047, M1051, V1060, M1076, H1130, D1132, D1150, Y1162, L1170, L1171, P1172, and F1186. These residues were selected based on the conserved allosteric pocket amino acids in IGF1RK and IRK. Finally, the Glide module of the Schrödinger suite (v2022.4) was used to perform molecular docking.[54,55] Docking was performed in the extra precision (XP) mode to predict the ligand binding affinity in terms of the GlideScore, and an energetically favorable binding conformation was determined.[56] Eventually, the coordinates from the docked complex of the allosteric inhibitor with the IRK were used for subsequent MD simulations.

2.3 | MD simulation setup

We conducted all-atom MD simulations of each kinase domain (IRK and IGF1RK) bound to the allosteric inhibitor using GROMACS (v2020.4) combined with the AMBER force-field.[57] The protein structures were parameterized using the AMBER99SB force-field and the coordinates were saved as a gromacs-compatible topology file.[58] The parameters for the small-molecule allosteric inhibitor were based on the GAFF force-field and obtained using the Antechamber module.[59,60] The coordinates of both the protein and the ligand were combined, and the ligand topology information was included in the system topology file. A cubic simulation domain centered around the protein was defined with a minimum distance of 10 Å between the protein surface and the box edges. Each system was solvated with the TIP3P water molecules and Na^+ and Cl^- ions were subsequently added to neutralize the net charge on the system.[61] Each system was initially subjected to energy minimization for 50 000 steps using the steepest descent algorithm.[62] The minimized systems were then equilibrated in the NPT ensemble for 10 ns using the Berendsen barostat at a pressure of 1 atm and a velocity rescale thermostat at 300 K.[63,64] Each system was further equilibrated in the NVT ensemble for 50 ns at 300 K using the velocity rescale thermostat.[64] The heavy atoms of both the protein and the ligand were positionally restrained for the initial minimization and equilibration steps. The bond lengths were constrained using the LINCS algorithm and a time step of 2 fs was used in all MD simulations.[65] The periodic boundary conditions, a cutoff radius of 12 Å for all nonbonded short-ranged interactions, and the Particle Mesh Ewald method were used for calculating the electrostatic interactions.[66] Finally, restraints were removed, and three independent production MD simulations (each 500 ns long) were carried out for each kinase/inhibitor complex. The coordinates from these simulation trajectories were saved every 20 ps.

2.4 | Binding free-energy calculations

To compute the binding free energy of the allosteric inhibitor for each kinase domain, we applied the Molecular Mechanics Generalized-Born

Surface Area (MMGBSA) approach to estimate the enthalpy change and the solvation free energy.[67,68] The MMGBSA method evaluates the free-energy differences between the bound and unbound states of the protein and ligand.[69] The binding free energy of a ligand molecule can be estimated as follows:

$$\Delta G_{\text{bind}} = G_{\text{complex}} - G_{\text{protein}} - G_{\text{ligand}}, \quad (1)$$

where each term on the right side of Equation (1) can be given by

$$G = E_{\text{MM}} + G_{\text{solv}} - \text{TS}; \quad (2)$$

where $E_{\text{MM}} + G_{\text{solv}}$ gives the enthalpy of binding and—TS corresponds to the entropic term. The ΔG_{bind} is the total free energy when the entropic term is eliminated, which is often enough to compare the relative binding free energies of related ligands:[70–72]

$$\Delta G_{\text{bind}} = \Delta E_{\text{MM}} + \Delta G_{\text{solv}}, \quad (3)$$

where ΔE_{MM} and ΔG_{solv} are decomposed into different terms:

$$\Delta E_{\text{MM}} = \Delta E_{\text{bonded}} + \Delta E_{\text{nonbonded}}, \quad (4)$$

$$\Delta E_{\text{MM}} = \Delta E_{\text{internal}} + \Delta E_{\text{vdw}} + \Delta E_{\text{ele}}, \quad (5)$$

and

$$\Delta G_{\text{solv}} = \Delta G_{\text{polar(GB)}} + \Delta G_{\text{nonpolar}}, \quad (6)$$

where ΔE_{MM} corresponds to the energy changes in the gas phase, including the ΔE_{bonded}, also known as the internal energy, and $\Delta E_{\text{nonbonded}}$ corresponding to the van der Waals and the electrostatic contributions. The solvation energy has both polar and nonpolar components. The polar component was estimated using the GB model with the internal and external dielectric constant values 1.0 and 78.5, respectively. The nonpolar solvation-free energy was modeled as a single term linearly proportional to the solvent-accessible surface area, with the surface tension of the solvent as 0.0072 and the surface tension offset value as 0.0. The calculations were performed using the gmx_MMPBSA tool which integrates the AMBER tools python-based script (MMPBSA.py) for performing free-energy calculations using MD trajectories.[71,73] Each energy term was estimated as an average over 100 frames in each 500 ns long MD simulation trajectory of each complex. Further, the average contributions of residues to the binding free energy were also calculated with per-residue decomposition.

2.5 | Conformational analyses

All MD trajectories were analyzed using the Gromacs trajectory analysis tools from the Gromacs package (v2020.4) and the Visual Molecular Dynamics (VMD) program.[74,75]

2.5.1 | Root-mean-squared deviation/fluctuation (RMSD/RMSF)

To determine the conformational changes, we calculated the RMSD of the protein backbone heavy atoms, and the RMSF was evaluated for the C_α atom of each residue from its mean position. The IGF1RK/inhibitor cocrystallized complex and the IRK/inhibitor docked complex were used as the initial reference structures to compute the RMSD values.

2.5.2 | Principal component analysis

The principal component analysis (PCA)[76] was performed to understand the conformational transitions and the principal motions of the kinase/inhibitor complexes. A covariance matrix of the eigenvectors representing the atomic fluctuations in the protein backbone was constructed and diagonalized yielding a set of eigenvectors with their corresponding eigenvalues. The principal components (PCs) are the eigenvectors with the highest eigenvalues representing the critical motions of each complex. Therefore, the essential subspace from the protein–ligand complex dynamics was defined by the projection of the first two PCs, PC1 and PC2. Further, the lowest energy conformational ensembles of both complexes were evaluated through the free-energy surface (FES) projected along each principal component.

2.5.3 | Protein–inhibitor interactions

Based on MD simulations, the protein–inhibitor interactions were determined using LigPlot$^+$.[77] The hydrogen bonds and their occupancies between the inhibitor and the protein residues from the MD trajectories were calculated using VMD. A cutoff distance of 3.5 Å was used between the donor and acceptor atoms (N, O, and F). The PyMOL software was used to visualize and analyze the complexes.[51]

3 | RESULTS

3.1 | Sequence and structural comparison of IRK and IGF1RK

In the present study, we focus on characterizing the binding conformation of the allosteric inhibitor in the IRK domain and deciphering the rationale for differences in the selectivity of the inhibitor among IRK and the homologous kinase IGF1RK. We first performed a sequence and structure-level comparison of both kinases. The global alignment of the IGF1RK/IRK sequences revealed 85.1% similarity and 74.8% identity (Figure 1A). The A-loop, C-loop, and the αC-helix residues are highly conserved. The structural superimposition of the IGF1RK (cocrystallized with the allosteric inhibitor) and IRK domains showed a significant difference in the position of the GC loop and the αC-helix resulting in an altered shape of the allosteric pocket (Figure 1B).[39] In addition, all amino acid residues that interact with the

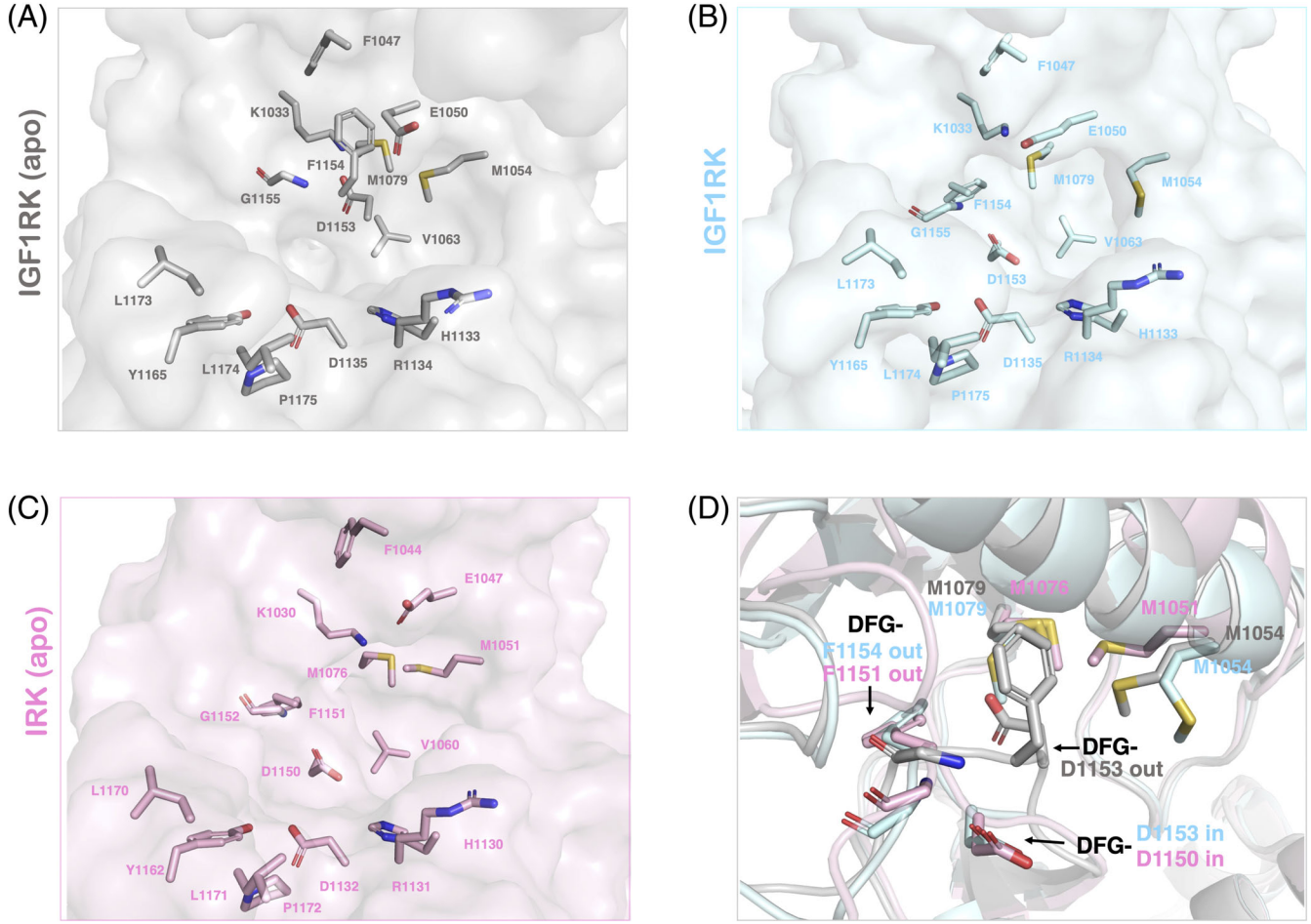

FIGURE 2 Residues forming the allosteric binding pocket in the insulin-like growth factor-I receptor kinase (IGF1RK) and insulin receptor kinase (IRK) structures. (A) Residues in the apo IGF1RK structure (PDB ID 1P4O), (B) Residues in the inhibitor-bound IGF1RK structure (inhibitor removed for comparison) (PDB ID 3LW0), and (C) Residues in the apo insulin receptor kinase (IRK) structure (PDB ID 1IRK). (D) A comparison of the residue sidechains for the DFG motif and the IGF1RK/IRK residue pairs M1054/M1051 and M1079/M1076.

allosteric inhibitor in the IGF1RK pocket are conserved in the IRK domain (Figure 2). The allosteric binding pocket residues and the conformations of their sidechains for apo IGF1RK (PDB ID 1P40), IGF1RK cocrystallized with the allosteric inhibitor (PDB ID 3LW0), and apo IRK (PDB ID 1IRK) are shown in Figure 2A–C, respectively.

Within the apo- and inhibitor-bound forms of the IGF1RK domain, significantly altered conformations were observed for the aspartic acid-phenylalanine-glycine (DFG) motif and the sidechains of the residues M1054 and M1079 (Figure 2D). In this work, we discuss the DFG motif conformations (in and out) with respect to the allosteric pocket. The apo IGF1RK structure has its DFG motif with "D-out" conformation (with respect to the allosteric site), but the inhibitor-bound IGF1RK revealed a "D-in" conformation (Figure 2D and Figure S2A). Additionally, a conformational difference is observed for the phenylalanine ring of the residue F1154 in both structures (Figure 2D and Figure S2A). Hence, this alteration in the DFG motif is responsible for the striking conformational difference observed for the A-loop in the apo and inhibitor-bound IGF1RK structures. Furthermore, in the apo IGF1RK structure, the sidechains of the residues M1054 and M1079 are oriented such that they partially block the allosteric pocket, as observed for the corresponding residues M1051 and M1076 in the apo IRK structure (PDB ID 1IRK) (Figure S2B). In contrast, the conformations of the residues M1054 and M1079 in the inhibitor-bound IGF1RK are considerably different (Figure S2A,C). The DFG motif conformation in the IRK, on the other hand, is similar to that reported in the IGF1RK inhibitor-bound structure (Figure 2D and Figure S2C). Thus, it is significant to note these similarities and differences regarding the conformations of the DFG motif and the residues M1051 and M1076 in apo IRK in comparison to both apo and inhibitor-bound IGF1RK. These structural analyses prompted us to probe the binding conformation for the allosteric inhibitor within the IRK pocket, as well as to explore conformational mechanisms and interactions responsible for the inhibitor selectivity.

3.2 | Conformational dynamics of the inhibitor complexes of IGF1RK/IRK

Using the molecular docking approach, we modeled the interaction between the allosteric inhibitor and the IRK domain (Figure 3). The

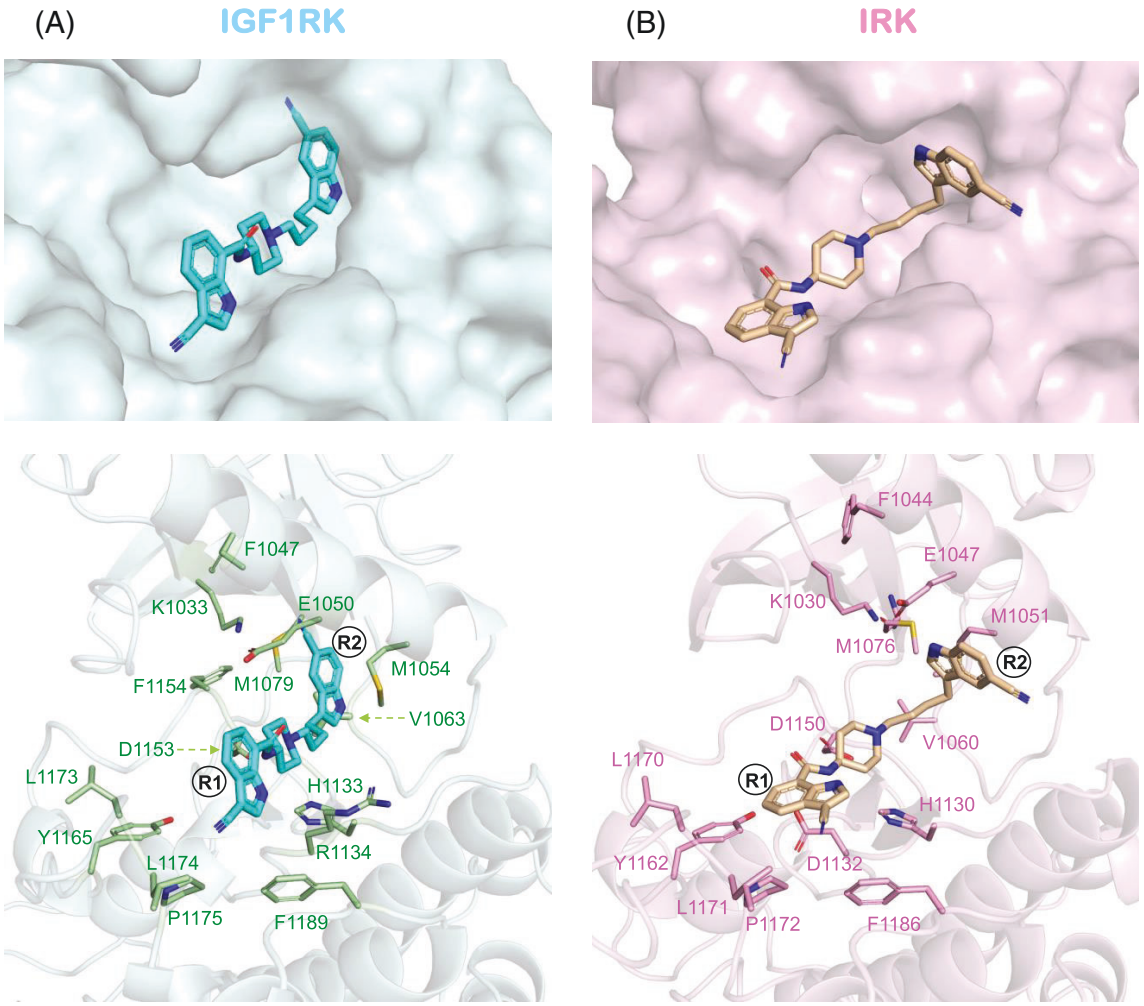

FIGURE 3 Binding conformations of the allosteric inhibitor. (A) The x-ray structure of IGF1RK cocrystallized with the allosteric inhibitor (cyan), and (B) The docked conformation of the allosteric inhibitor (wheat) in the insulin receptor kinase (IRK) pocket. In bottom panels, the binding site residues are labeled and shown in stick representations for each kinase.

initial docked position of the inhibitor in the binding pocket of IRK is distinct from the one observed in the crystal structure of IGF1RK (PDB ID 3LW0), likely due to the altered shape of the binding pocket.[39] The allosteric inhibitor in the IRK docked state has its indole rings (R1 and R2) positioned in a different conformation from the IGF1RK bound state (Figure 3B). The 3-cyano indole ring (R1) has an orientation parallel to the C-loop residues H1130, R1131, and D1132. On the other hand, the 5-cyano indole ring (R2) is stacked against the αC-helix, where it is held by the interactions with the residues F1044, E1047, and M1051 (Figure 3B). Although the binding pocket residues are fully conserved, the conformations of the binding pocket residues likely result in differences in inhibitor binding and its configurations. To understand the stability and variability of the allosteric inhibitor in complex with each kinase domain, we studied the conformational dynamics of the inhibitor using all-atom MD simulations. First, we assessed the RMSD of the protein backbone atoms and the non-hydrogen atoms of the inhibitor. The RMSD for IGF1RK (2–4.5 Å) suggests higher flexibility in its backbone than the IRK (1.5–3.5 Å) domain (Figure S3A,B). However, the RMSD for the inhibitor in the IGF1RK allosteric pocket is lower (1.5–2.5 Å) than in the IRK pocket (2.0–3.5 Å), thereby suggesting higher deviations of the inhibitor from its initial conformation in the IRK pocket than in the IGF1RK pocket (Figure S3C,D).

To further characterize the flexibility of individual residues in each protein, we calculated their RMSF values. While the overall trends of per-residue fluctuations in each protein were similar, a loop region corresponding to the residues 1093 through 1115 showed higher fluctuations in IGF1RK than IRK (Figure 4). This region includes the modeled loop (loop 1) missing in the crystal structure of IGF1RK (Figure S1). The amplitudes of the fluctuations corresponding to the RMSF values of the αC-helix residues, C-loop residues, and A-loop residues are comparable in both proteins. These critical regions in each kinase domain also form the allosteric binding pocket and some residues are actively involved in inhibitor binding. As noted above, both proteins have a high degree of conservation for the residues in these regions (Figure 1A). Moreover, slightly higher fluctuations are observed for each residue in the binding pocket of IGF1RK than IRK. Specifically, the αC-helix region, spanning the residues 1040 through

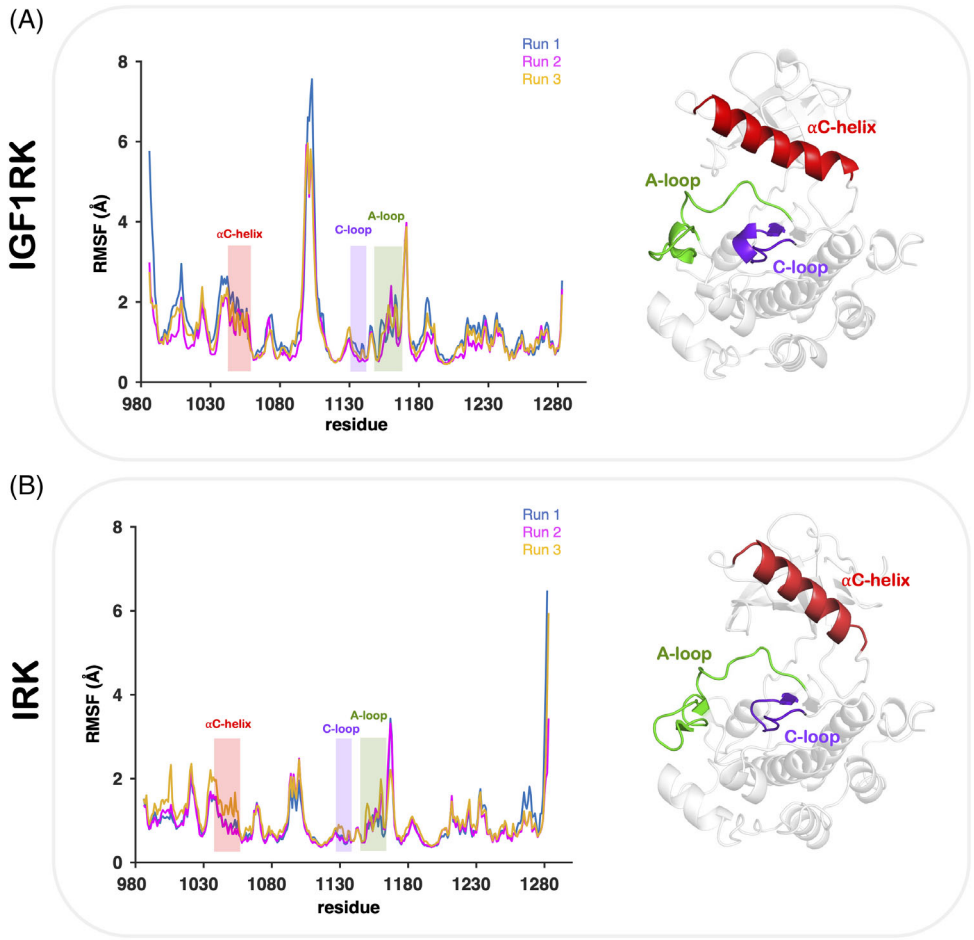

FIGURE 4 Root-mean squared fluctuation (RMSF) data for the C_α atoms for each kinase. The RMSF traces from three independent molecular dynamics (MD) simulations are shown in blue, magenta, and yellow. The RMSF data are shown for each inhibitor-bound complex: (A) IGF1RK and (B) IRK. The αC-helix, A-loop, and the C-loop regions of the respective kinases are highlighted in cartoon representations and uniquely colored red, green, and purple, respectively.

1056 in IGF1RK, showed higher fluctuations relative to the corresponding residues (1037–1053) in IRK, reflecting more flexibility of the αC-helix in IGF1RK. The αC-helix residues of IGF1RK/IRK, F1047/F1044, E1050/E1047, and M1054/M1051 are involved in the interactions with the R2 indole ring of the allosteric inhibitor in the binding sites of both proteins, suggesting that the flexibility in these residues will directly impact the inhibitor binding. The A-loop in IGF1RK spans the residues G1152 through L1174 (corresponding to IRK residues, G1149-L1170), with higher fluctuations observed for the modeled residues 1169–1171. The average RMSF value for the DFG motif in both kinases is ~1.2 Å, reflecting a lower flexibility in this region and indicating a stable DFG conformation in the binding pocket of both IGF1RK/IRK.

3.3 | Principal component analysis of protein/inhibitor complexes

Based upon MD trajectories, we performed essential dynamics analysis to understand the principal conformational motions in each protein–inhibitor complex. We assessed the lowest energy conformational ensembles of the protein–inhibitor complexes by projecting an FES along the first two principal modes. We observed several minima in the FES plot corresponding to dominant conformational populations (Figure 5). Specifically, we observed four minima for IGF1RK and three minima for IRK (designated as S1, S2, S3, and S4 in their respective FES plots). The representative conformation of the inhibitor corresponding to each free energy minimum is shown in Figure 5. We evaluated the RMSD of these inhibitor conformations relative to their initial positions in the binding pocket of each kinase domain (Figure S5). The conformational states are classified and labeled based on their RMSD values; for example, the inhibitor conformation with the lowest RMSD is denoted as S1, and so on. The inhibitor conformations for the IGF1R-inhibitor complex have an RMSD between 2 and 3 Å (Figure S5). The RMSD values for the inhibitor conformations corresponding to the free energy minima for the IRK-inhibitor complex vary between 4 and 12 Å (Figure S5).

For the IGF1RK-inhibitor complex, the minima S1 and S2 emerged from a single populated ensemble with a low transition barrier (~2 kcal/mol). While examining the inhibitor conformations corresponding to these minima (S1 and S2), a similar configuration of the inhibitor was observed with a difference in the orientation of the carbonyl oxygen connecting the R1 indole group (Figure 5A). In the S3 configuration of the inhibitor, the conformation of the butyl chain and the R1 indole ring was distinct from the initial binding pose (Figure 5A and Figure S5). The minimum S3 has an energy barrier of ~4 kcal/mol relative to the minima S1 and S2. On the other hand, the minimum S4 is well separated with a comparatively higher energy barrier of

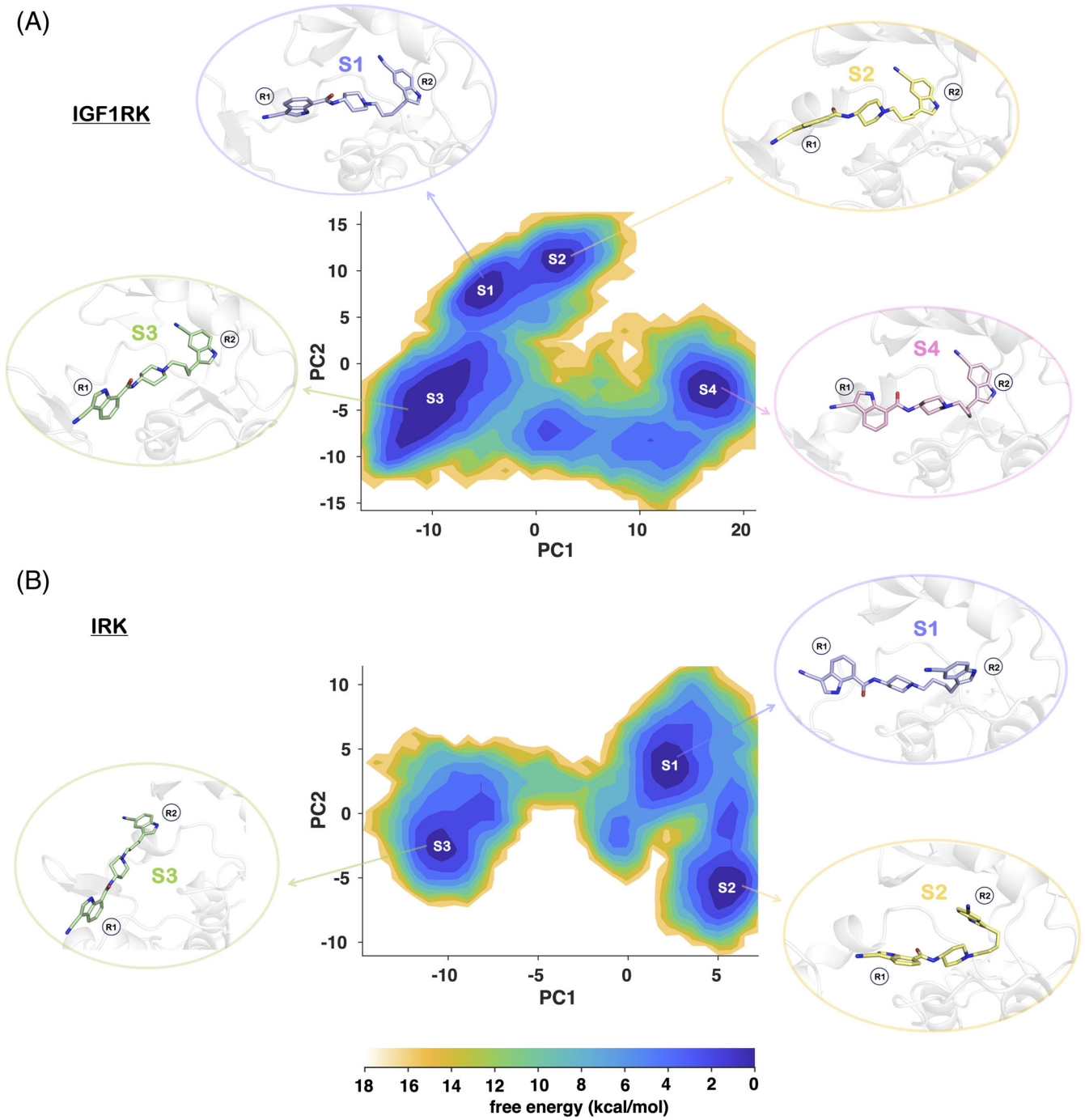

FIGURE 5 Projections of the free-energy surface (FES) along two principal components (PC1 and PC2) for each kinase. All minima in the FES are labeled and highlighted with the corresponding conformations of the inhibitor in the respective allosteric pocket of each kinase. The free energy is given in kcal/mol and indicated by the color palette ranging from lower (blue) to higher (yellow) values.

~8 kcal/mol relative to the minima S1, S2, and S3 in the FES. The conformational states of the inhibitor in the IGF1RK pocket revealed that the significant conformational transitions occur for the indole ring (R1) with the rotation along the carbonyl oxygen, consequently affecting the position of the butyl chain.

The FES plot derived from MD simulations for the IRK-inhibitor complex showed three dominant conformational ensembles (Figure 5B). All three conformational basins have distinct configurations of the inhibitor. The minimum S3 has a separate conformational basin with an energy barrier of ~10 kcal/mol, while the minima S1 and S2 are more easily accessible (~6 kcal/mol) to each other. The inhibitor conformations corresponding to the free energy minima S1 and S2 in the IRK-inhibitor complex are distinct concerning the positions and the orientations of the indole rings (R1 and R2) as well as the n-butyl chain (Figure 5B and Figure S4). Whereas the minimum S3 represents a conformation of the inhibitor where it has moved away from the binding pocket. Hence, the transition barrier between various conformational states for the IRK-inhibitor complex is higher

than that in the IGF1RK-inhibitor complex. Overall, the dominant clusters of the inhibitor conformations imply that the inhibitor has rather comparable configurations in the IGF1RK allosteric pocket region.

Further, we analyzed the protein conformations corresponding to the free-energy minima and evaluated the dominant conformational changes occurring in their structures relative to their initial configurations (Figure 6 and Figure S6). For IGF1RK, an outward shift of the

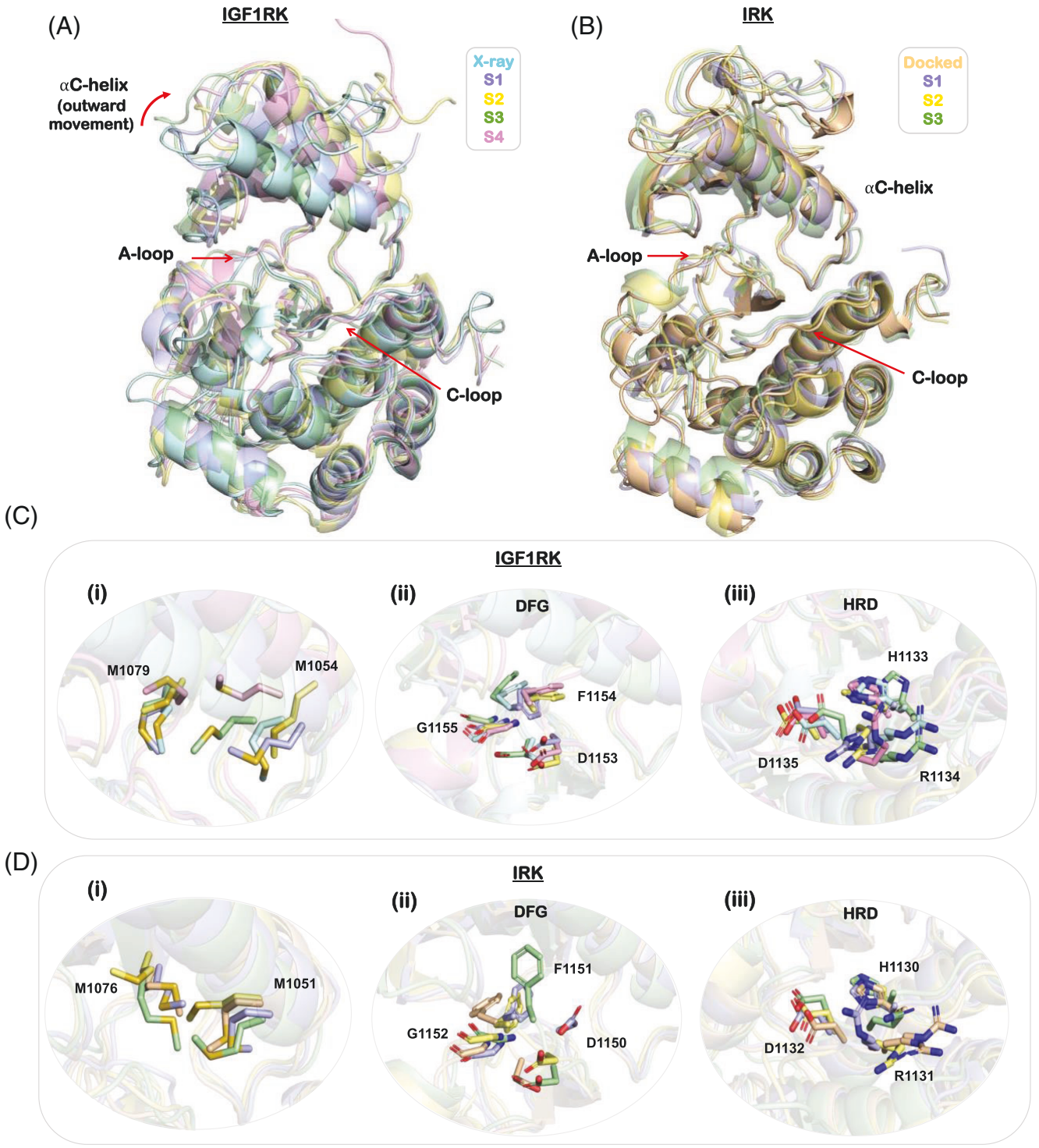

FIGURE 6 Structural superimposition of the IGF1R and IR kinase domain conformations obtained from the free-energy surface (FES; cf. Figure 5). (A) IGF1RK conformations: S1 (purple), S2 (yellow), S3 (green), and S4 (pink) superimposed on the initial x-ray conformation (cyan). (B) IRK conformations: S1 (purple), S2 (yellow), and S3 (green) superimposed on the conformation obtained after docking (wheat). (C) A comparison of the conformations of the residue sidechains for IGF1RK: (I) residue pairs M1054 and M1079, (ii) aspartic acid-phenylalanine-glycine (DFG) (D1153, F1154, and G1155) motif, and (iii) histidine-arginine-aspartic acid (HRD) (H1133, R1134, and D1135) motif. (D) A comparison of the conformations of the residue sidechains for IRK: (i) residue pairs M1051 and M1076, (ii) DFG (D1150, F1151, and G1152) motif, and (iii) HRD (H1130, R1131, and D1132) motif (see also Figure S6).

αC-helix was observed in conformations corresponding to all minima (Figure 6A). The different sidechain conformations of residues M1079 and M1054 in all states reflect their flexibility in accommodating the 5-cyano indole (R2) moiety of the allosteric inhibitor (Figure 6Ci). The corresponding residues M1076 and M1051 in IRK have relatively less conformational variability and attain a conformation in each state, which would partially block the binding pocket for the inhibitor R2 ring (Figure 6Di). The DFG motif in IGF1RK has no major conformational shift and remains in the "D-in/F-out" conformation (Figure 6Cii) to stabilize the butyl and the piperidine group of the inhibitor. Whereas in IRK, the DFG motif obtains slightly varied conformation in each of the dominant populated states. The phenylalanine ring of residue F1151 is populated in such a configuration in one of the conformations (corresponding to minimum S3) that it causes steric hindrance and blocks the pocket for the allosteric inhibitor (Figure 6Dii). As a result, the inhibitor was observed to move out of the binding pocket region in conformations corresponding to the minimum S3 (Figure 5B). This conformation of F1151 in the DFG motif also resembles the one observed in the apo IGF1RK structure (Figure S2). The residue D1150 stays in the "D-in" configuration, with a slightly more inward conformation (toward the allosteric pocket) affecting the position of the inhibitor in the IRK pocket. In the histidine-arginine-aspartic acid (HRD) motif of the C-loop, the relatively longer sidechains of the residues R1134/R1131 occur in multiple conformations in both kinases, influencing the 3-cyano indole ring of the inhibitor (Figure 6Ciii and Diii). We also carried out the comparison of the dominant conformations of the apo (unliganded) IGF1RK (PDB ID 1P4O) and IRK (PDB ID 1IRK) structures (Figure S7). The apo IRK structure also shows a higher population of some residue conformations that would likely disfavor the binding of the allosteric inhibitor (Figure S8A). Specifically, the side chains of residues M1051 and M1076 in the apo IRK crystal structure (Figure 2) retain similar configurations in the MD-derived dominant conformation of apo IRK (Figure S8A). As a result, these sidechains are orientated differently in apo IRK in comparison to apo IGF1RK and therefore partially block the IRK binding site for the inhibitor's indole ring R2 (where V1060 of IRK is inaccessible for forming a hydrogen bond with the NH9 atom in the indole ring R2), similar to what was observed in the other conformations (docked, S1, S2, and S3) of the IRK (Figure S8A, ii). As a result, both apo IRK and docked IRK have similar overall binding pocket shapes. Additionally, we compared the apo IGF1RK crystal structure conformation (PDB ID 1P4O) to the dominant conformation from the apo IGF1RK simulation (Figure S8B). A conformational shift in the A-loop and the DFG motif ("D-in" and "F-in") reflects a potential role of conformational selection and induced fit in the binding of the allosteric inhibitor in the IGF1RK pocket. Overall, these results suggest that the conformations of the sidechains of residues in the αC-helix, DFG motif, and the C-loop HRD motif significantly influence the conformation of the inhibitor in the allosteric pocket of both proteins. The conformational changes in the residues within the binding pocket are responsible for the specific configuration of the 5-cyano indole group (R2), the butyl chain, and the piperidine ring observed in all populated conformations of the IGF1RK-inhibitor complex. To understand the relative importance of individual residue sidechains, we quantified their energetic contributions toward the binding affinity and selectivity of the inhibitor.

3.4 | Energetic basis of inhibitor interactions in each kinase/inhibitor complex

To gain an understanding of the molecular interactions and energetics involved in the binding of the allosteric inhibitor with IGF1RK and IRK, a thorough analysis of the binding free energy was carried out using the MMGBSA method (see Methods section 2.4 and Table S1). The estimates on the energetics were averages over three independent all-atom MD simulations of each complex (Figure 7A). Among the nonbonded interaction energies, the contributions favoring the binding of the inhibitor are the van der Waals interaction energy followed by the electrostatic interaction energy. The total binding free energy (ΔG_{bind}) values of the allosteric inhibitor with IGF1RK and IRK are −42.75 kcal/mol and −39.44 kcal/mol, respectively. The contribution of the van der Waals energy (E_{vdw}) to the total binding free energy for both complexes is comparable, that is, −54.13 kcal/mol for IGF1RK and −52.29 kcal/mol for IRK. However, the higher binding affinity of the allosteric inhibitor toward IGF1RK in comparison to IRK is by virtue of the electrostatic energy contribution. The electrostatic energy (E_{ele}) contributions for the IGF1RK and IRK complexes are −31.99 kcal/mol and −21.19 kcal/mol, respectively. Hence, a significant energy contribution and difference is observed between these complexes predominantly involving polar and charged interactions.

To characterize these interactions, we estimated the number of hydrogen bonds between the inhibitor and protein for each kinase-inhibitor complex. The allosteric inhibitor typically creates about 1–2 average hydrogen bonds in each kinase domain. We estimated the occupancy of the hydrogen bonds to understand their significance in the interactions of the allosteric inhibitor in the binding site. The IGF1RK/IRK residues participating in the hydrogen bond formation are K1033/K1030, V1063/V1060, V1132/V1129, H1133/H1130, R1134/R1131, D1135/D1132, K1171/K1168, and L1173/L1170, also shown in Figure 7B with their respective occupancy values. All these residues have hydrogen bond occupancy (H_0) < 25% except for the residue V1063 in the IGF1RK binding site, which has an H_0 value of about 65%. The allosteric inhibitor at the binding site forms a stable hydrogen bond with V1063, which has also been observed in the crystal structure of the IGF1RK inhibitor complex (Figure 7C,D). The hydrogen bond is formed between the NH9 atom of the R2 indole ring of the inhibitor and the main chain carbonyl oxygen atom of V1063 in the IGF1RK domain. However, this interaction is not attained by the corresponding residue V1060 in the IRK pocket because of the distinct orientation of the R2 indole ring of the allosteric inhibitor (Figure 7C). Hence, this unique hydrogen bonding interaction at the IGF1RK allosteric pocket is critical and likely contributes toward the high selectivity of the allosteric inhibitor for the IGF1RK.

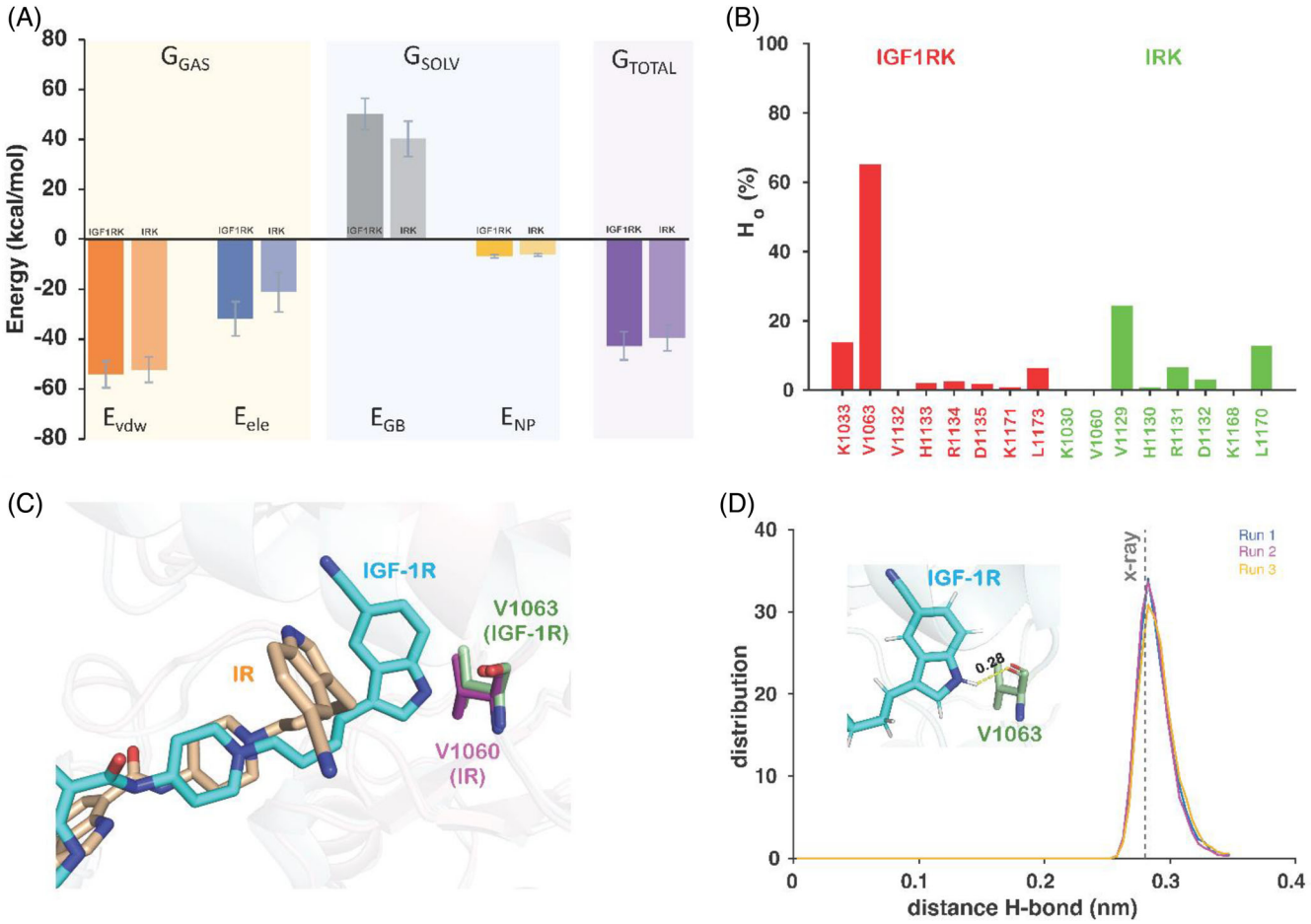

FIGURE 7 Binding free energy of the inhibitor and the hydrogen bond analysis for each kinase complex. (A) The binding free energy of the inhibitor decomposed into individual components for the IGF1RK and IRK complexes. The lighter and darker shades for each energy component represent energies corresponding to the IRK and IGF1RK complexes, respectively. The error bars represent the standard deviation for each component (see also Table S1). (B) Hydrogen bond occupancy (H_O) for each kinase/inhibitor complex. The IGF1RK residues and IRK residues are labeled and shown in red and green color, respectively. The residues with no bar reflect 0% H-bond occupancy. (C) The superimposed conformations of the 5-cyano indole ring of the inhibitor for the IGF1RK (cyan) and IRK (wheat) is shown with the neighboring valine residues, V1063 (green; IGF1RK) and V1060 (magenta; IRK). (D) The distributions of a distance characterizing the hydrogen bond between the NH atom in the indole ring of the inhibitor and the main chain carbonyl oxygen atom of V1063 in IGF1RK. The distribution data are based on three independent all-atom MD simulations. The vertical line in distributions marks the value of the distance characterizing the hydrogen bond (2.8 Å) in the x-ray crystal structure.

3.5 | Interactions governing the binding of allosteric inhibitor to IGF1RK and IRK

To investigate the interactions governing the inhibitor conformation in the binding pocket of each kinase, we analyzed the lowest energy conformational state, S1, which has the least RMSD of the allosteric inhibitor from its initial conformation in each complex (Figure 8). In the binding pocket of IGF1RK, the R2 indole ring of the inhibitor forms a stable hydrogen bond with V1063 and the hydrophobic contacts with the residues L1065, M1079, M1054, and G1152 (Figure 8C). The residue D1153 from the DFG motif also interacts with the inhibitor, establishing nonbonded contacts. Furthermore, the conformation of the R1 indole ring of the inhibitor is primarily guided by the cation-π interactions with R1134 and the hydrophobic contacts formed by L1173. Correspondingly, the R1 indole ring is held by forming cation-π and hydrophobic interactions with the residues R1131 and L1170 in the IRK pocket (Figure 8D). The piperidine ring of the inhibitor at the IRK binding pocket is supported by the C-loop residues H1130 and D1132 along with the A-loop residues D1150 and G1149. The R2 group of the inhibitor in the IRK-bound complex predominantly interacts with the nonpolar and aromatic residues L1065, M1051, F1054, and F1128. The aromatic residue F1128 makes π-stacking interactions with the R2 indole ring influencing the inhibitor conformation in the IRK pocket. These observations indicate that the allosteric inhibitor in both binding pockets is held predominantly by hydrophobic contacts, with a single hydrogen bond established with the inhibitor in the IGF1RK pocket.

Next, we performed residue-level decomposition of the binding free energy to quantify the energetic contribution of each binding site residue with the allosteric inhibitor. In Figure 9A, we show the ΔG_{bind}

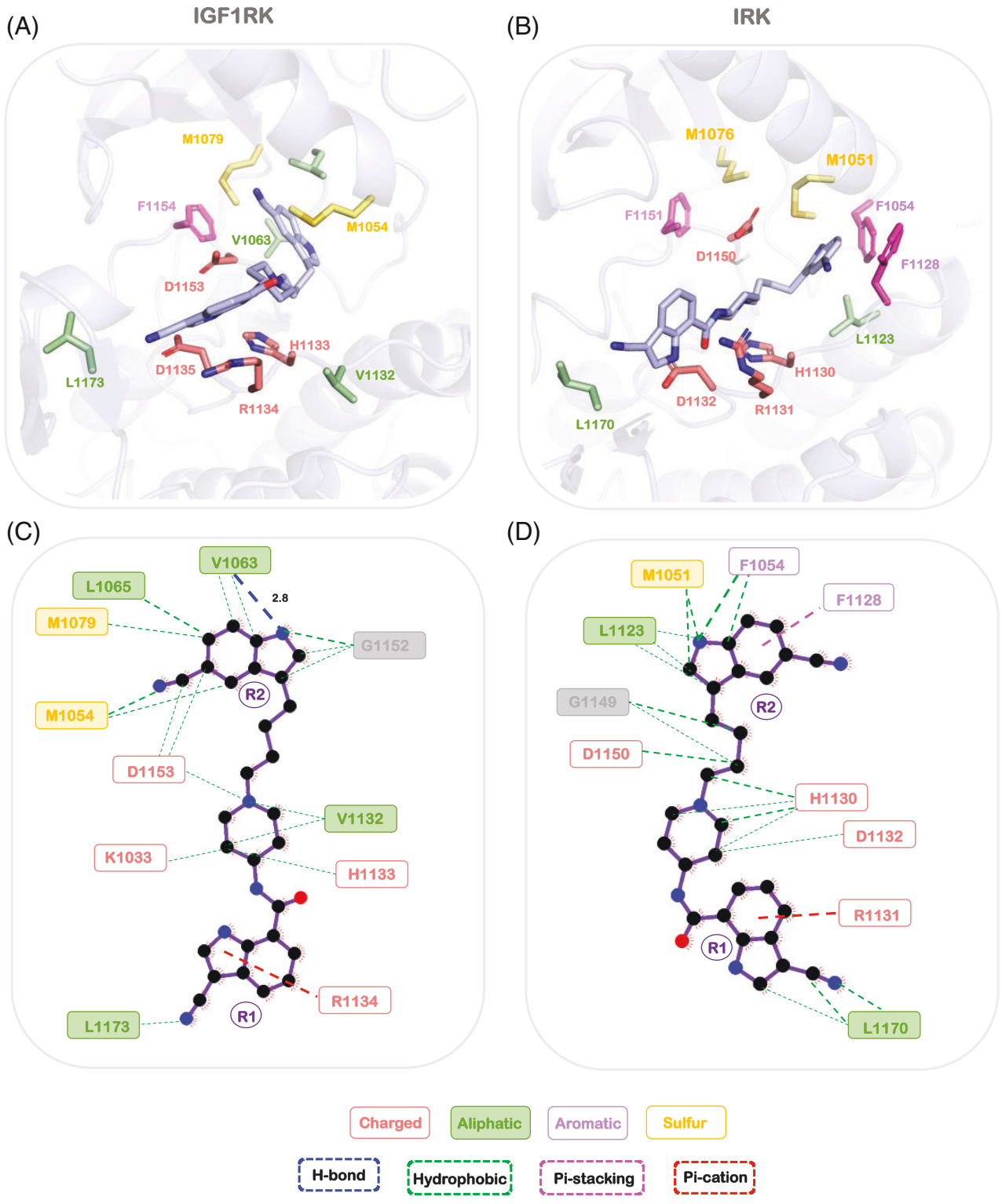

FIGURE 8 Inhibitor interactions in the IGF1RK and IRK binding pockets. The bound conformations of the inhibitor are shown for IGF1RK and IRK, respectively. For each inhibitor, the conformation corresponds to the free-energy minimum state S1 (cf. Figure 5). The binding pocket residues involved in the interactions are shown in stick representations. For each complex, a 2D representation of the interactions between the allosteric inhibitor and the binding site residues is shown in panel (C) IGF1RK and panel (D) IRK. The residues and the type of interactions are categorized and colored uniquely. The thickness of the dashed lines correspond to the number of nonbonded contacts. A thin dashed line represents only a single nonbonded atom-to-atom contact, and a thick line represents more than one nonbonded contact with the residue.

values for those amino acid residues which are identical in the IGF1RK and IRK inhibitor-bound complexes. The snapshots representing the contrasting conformational differences for the residue sidechains are shown in Figure 9B. The αC-helix residues M1054 (IGF1RK) and M1051 (IRK) in both complexes energetically favor the binding of the inhibitor with a contribution of −2.08 kcal/mol and −1.22 kcal/mol, respectively. It has the maximum van der Waals energy contribution compared to the other αC-helix residues (Figure S9). The orientations of the sidechains of M1054 and M1051 are shown in Figure 9B, i,ii, respectively. The other methionine residue supporting the R2 indole group of the inhibitor in the IGF1RK is M1079. This residue has a contribution of about −1.00 kcal/mol to the binding free energy in the IGF1RK pocket. The energy contribution of M1079 is entirely derived from the van der Waals energy arising from its interaction with the indole ring of the inhibitor. On the other hand, M1076 has an energy contribution of only −0.08 kcal/mol in the IRK complex. As a result, the contribution of this methionine residue (M1079/M1076) to the binding energy in the IGF1RK and IRK domain differs significantly. This difference arises from the conformation of the inhibitor's aromatic ring and the orientation of the sulfur-containing side chain (Figure 9B, v,vi).

Further, the IGF1RK residue V1063 significantly contributes to inhibitor binding with a ΔG_{bind} value of −1.35 kcal/mol, whereas the corresponding residue V1060 at IRK has no contribution to the binding of the inhibitor. A stable hydrogen bond between the two moieties, the backbone carbonyl oxygen atom of V1063 and the NH9 atom of the inhibitor results in an electrostatically favored interaction. Hence, V1063 has the highest electrostatic energy contribution with an energy

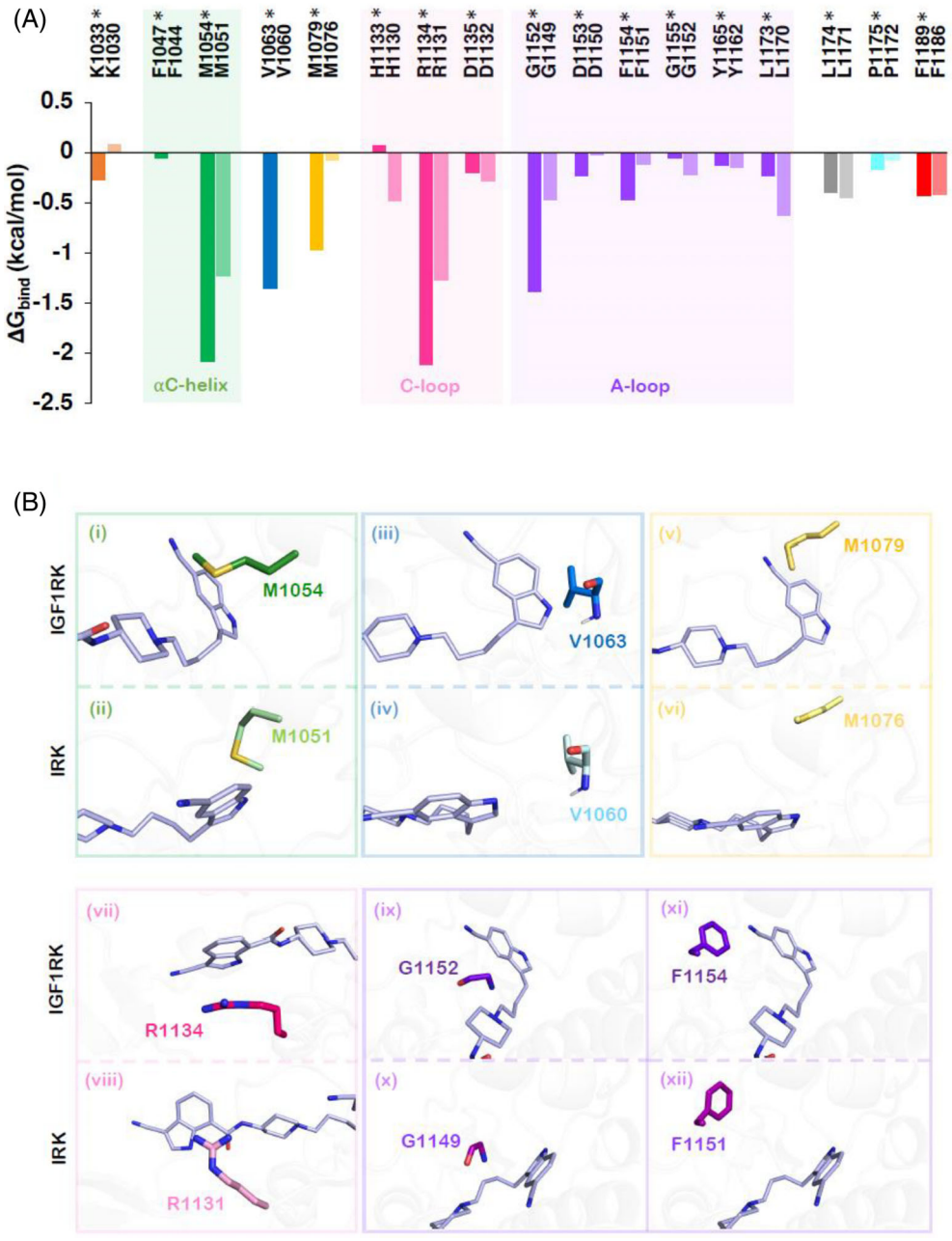

FIGURE 9 Data on residue-level interactions between the allosteric inhibitor and each kinase domain. (A) The ΔG_{bind} for each residue involved in the allosteric inhibitor binding in both kinase domains. The lighter and darker shades represent the contribution of the corresponding residue of IRK and IGF1RK, respectively. IGF1RK residues are marked with "*". The free energy data for residues corresponding to the αC-helix, A-loop, and C-loop regions are shown and highlighted in green, magenta, and purple bars, respectively. (B) The conformations of the allosteric inhibitor and selected residues are shown for IGF1RK and IRK complexes. Snapshots are taken from the minimum state S1 of each complex (cf. Figure 5). The color scheme for each residue in panel B is consistent with data shown in panel A. The panels are labeled for IGF1RK and IRK complexes, respectively.

value of −2.83 kcal/mol (Figure S9A). The conformation of this residue and its vicinity with the inhibitor in the binding pocket of IGF1RK and IRK differs significantly (Figure 9Biii,iv). This results in the formation of key backbone interactions at the IGF1RK pocket contributing toward inhibitor selectivity. The C-loop HRD motif, H1133/H1130, R1134/R1131, and D1135/D1132 have considerable energy contributions in both IGF1RK and IRK. It is worth pointing out that the IRK residue H1130 has a favorable contribution toward the binding of the inhibitor with a ΔG_{bind} value of −0.48 kcal/mol whereas, an unfavorable contribution is observed for H1133 in IGF1RK (0.07 kcal/mol). On the other hand, the C-loop residue R1134 in IGF1RK and R1131 in IRK have major contributions to the total binding free energy. For IGF1RK, this charged residue has a ΔG_{bind} value of −2.11 kcal/mol with the maximum contribution arising from the van der Waals energy. The corresponding residue R1131 for IRK has the ΔG_{bind} value of −1.27 kcal/mol coming from both the electrostatic and van der Waals energy. Notably, the relatively longer sidechain of this residue (R1134/R1131) has varying conformations in both pockets, and it is also influenced by the orientation of the R1 indole ring of the inhibitor and vice versa (Figure 9Bvii, viii). Additionally, the adjacent residue D1135/D1132 also contributes positively with a ΔG_{bind} value of −0.19 kcal/mol in IGF1RK and −0.28 kcal/mol for IRK. Indeed, the allosteric inhibitor is predominantly stabilized in the binding pocket of each kinase through interactions with the charged residues.

Further, the DFG motif and the other A-loop residues such as G1152/G1149, Y1165/1162, and L1173/L1170 also have favorable energy contributions. The IGF1RK/IRK residue G1152/G1149 adjacent to the DFG motif has a ΔG_{bind} value of −1.38 kcal/mol and −0.47 kcal/mol, respectively. However, in the IGF1RK binding site, G1152 stabilizes the inhibitor with relatively equal contribution from the van der Waals and the electrostatic energy (Figure S9).

The phenylalanine residue from the DFG motif, F1154 favors higher binding energy contribution with a ΔG_{bind} value of −0.46 kcal/mol at IGF1RK than F1151 at IRK, which contributes a ΔG_{bind} value of −0.12 kcal/mol. Similarly, the residue D1153 at the IGF1RK has a higher energetic contribution toward inhibitor binding than the corresponding residue D1150 in IRK (Figure 9A). Hence, the conformations of the D and F residue sidechains of the DFG motif with respect to the inhibitor also influence their interaction energy contributions (Figure 9Bxi,xii). Furthermore, the residues L1170 and L1171 in the IRK pocket have more favorable energy contributions with the ΔG_{bind} values of −0.62 kcal/mol and −0.44 kcal/mol, respectively. These interactions together contribute to the binding affinity differences for the allosteric inhibitor toward IGF1RK and IRK. However, the higher energy contribution of the sulfur-containing sidechains of M1054 and M1079, the charged side chain arm of R1134, and the DFG motif, along with the favorable electrostatic energy contribution by V1063 are altogether responsible for a higher binding affinity of the inhibitor toward IGF1RK.

4 | DISCUSSION

Kinases are crucial mediators in signaling pathways and have been identified as prospective therapeutic targets. However, achieving inhibitor selectivity among kinases is a significant challenge. Given that the selectivity barriers are frequently associated with orthosteric type I and type II inhibitors, developing allosteric inhibitors (type III and type IV) is one of the strategies for overcoming the selectivity hurdle in the kinase drug discovery.[32] One of the prominent causes of clinical trial failures is off-target inhibitor toxicity due to the structurally identical kinase drug binding pockets.[78] Hence, selectivity should be a key consideration in designing both orthosteric and allosteric kinase inhibitors. The discovery of selective inhibitors for Abl, CK2, PI3K, MEK, Akt, JAK, IGF1R, and other kinases has demonstrated that selectivity can be achieved among structurally similar kinases.[12] Moreover, insights into the binding mechanisms of kinase inhibitors are necessary for developing novel inhibitors with adequate selectivity and specificity. Furthermore, there is a significant gap between the availability of experimental bioactivity data and the structural knowledge of the mechanisms of action for most allosteric kinase inhibitors.[79]

In this work, we investigated the molecular level details of the allosteric binding (type III) of an IGF1RK inhibitor (MSC1609119A-1) to the homologous IRK domain. Although the binding mode of this inhibitor for IGF1RK was resolved using the x-ray crystallography method, there is no structural information on its interaction with IRK and therefore, a lack of understanding of its decreased affinity and activity.[39] Heinrich et al.[39] describe the kinase selectivity in terms of the IC_{50} values measured in a biochemical screening assay. In their experimental investigation, they report that the allosteric inhibitor MSC1609119A-1 is potent and selective for IGF1RK, with an IC_{50} value of 0.4 μM for IGF1RK, and 6.9 μM for IRK. Given that the amino acid residues directly interacting with the inhibitor in IGF1RK are fully conserved in IRK, the selectivity of the allosteric inhibitor for IGF1RK is challenging to rationalize. The structural spectrum of kinases is diverse, with multiple states that include the αC-helix "inward" and "outward" conformations as well as DFG motif's "in" and "out" conformations.[78,80] Our initial structural comparison of the IGF1RK (apo and inhibitor co-crystalized) structures and the apo IRK structure suggest that there is a potential role of conformational selection, especially for the residues M1054 and M1079 in IGF1RK because the corresponding residues M1051 and M1076 in IRK have steric overlaps if the inhibitor were to take a conformation similar to what is observed for IGF1RK (Figure 1C). However, the x-ray structures may fail to capture the precise dynamic local conformational attributes in the binding pocket regions, which is critical in kinase-selective drug discovery. Therefore, we quantified the structural and energetic contributions of the IGF1RK/IRK residues toward the binding of the allosteric inhibitor MSC1609119A-1 for rationalizing its selectivity for IGF1RK over IRK.

Our findings imply that the allosteric inhibitor in IGF1RK has a stable configuration, with conformational changes limited to the 3-cyano indole ring (R1) (Figure 5A). This provides an explanation for the most probable arrangement of the R1 ring of the inhibitor in IGF1RK that was not resolved in the electron density map of the cocrystallized complex. We show that the binding configuration of the inhibitor in both allosteric pockets could be largely influenced by the conformations of the binding pocket residues. Adding to this, our findings revealed that

the outward movement of the αC-helix and the conformational flexibility of the residues M1054 and M1079 in the IGF1RK pocket accommodates the 5-cyano indole (R2) substituent of the inhibitor (Figure 6A,C). In contrast, in IRK, there is no conformational shift in the position of the αC-helix, and the M1051 and M1076 residues also achieve conformations that would block the allosteric pocket in IRK (Figure 6B,D). Consequently, this leads to a sterically hindered conformation of the allosteric inhibitor in IRK. IGF1RK has an intrinsic ability to stabilize the "D-in" and "F-out" conformation of the DFG motif in the presence of the inhibitor, whereas in IRK this conformation is not stable. In IRK, the DFG motif is populated in multiple conformations, leading to an altered pocket conformation affecting inhibitor binding (Figure 6D). Furthermore, the DFG-"F-in" conformation in the unliganded apo IGF1RK suggests that both conformational selection and induced fit play a role in the binding mechanism (Figure S8B).[81] Hence, the distinct residue conformations could be an explanation for the preferential binding of the allosteric inhibitor (MSC1609119A-1) to IGF1RK.

Consistent with the experimentally observed activity,[39] the allosteric inhibitor showed different binding free energies for the IGF1RK and IRK domains. The energy contribution from the electrostatic components (E_{ele}) is significantly higher in the IGF1RK-inhibitor complex relative to the IRK-inhibitor complex (Figure 7A). Furthermore, in the IGF1RK pocket, the allosteric inhibitor forms a stable hydrogen bond with the residue V1063. This hydrogen bonding interaction is not observed in the IRK pocket and the corresponding residue V1060 has no energetic contribution to inhibitor binding. Therefore, these results suggest that unique interactions at the kinase binding pockets may contribute toward achieving inhibitor selectivity. Our study demonstrates an excellent example of the role of rigidity and flexibility of the allosteric ligand binding sites critical to protein–ligand interactions,[82] thereby providing a molecular basis for designing kinase inhibitors. This study reflects that in kinase drug discovery, it is critical to identify inhibitor-accessible local characteristics in the binding pockets of one kinase that are not observed in its isoform/homolog or other kinases. The structural basis of inhibitor selectivity established in this work may benefit kinase drug discovery toward enhanced isoform selectivity in the insulin receptor family and other RTKs.

5 | CONCLUSION

In this work, we report molecular details of binding of an allosteric inhibitor to the kinase domains of IGF1R and IR. The data from MD simulations and binding free energy calculations provide a direction for discovering unresolved molecular factors responsible for the binding specificity of the inhibitor. Our findings suggest that the conformations of the residues M1054 and M1079 coupled with the outward movement of the αC-helix and the stable DFG (D-in and F-out) motif conformation, favor the selectivity of MSC1609119A-1 toward IGF1RK. Furthermore, we postulate that the selectivity attained is a result of the differences in the electrostatic interaction energy and the formation of a unique hydrogen bond in the IGF1RK pocket. The hydrogen bond between the indole ring of the inhibitor and Val1063 of IGF1RK, which is absent with the corresponding Val1060 residue of IRK, provides directionality and specificity to the allosteric inhibitor, thereby making it selective for IGF1RK over IRK. Our study suggests that the conformations of the allosteric pocket residues that lead to inherent differences in the binding affinity are responsible for the selectivity of the allosteric inhibitor for IGF1RK. Overall, the findings from this study enable a better understanding of the IGF1RK/IRK allosteric inhibitor binding mechanism and may potentially aid in developing selective kinase inhibitors.

AUTHOR CONTRIBUTIONS

Jyoti Verma: Conceptualization; investigation; writing – original draft; methodology; visualization; software; data curation; formal analysis; validation. **Harish Vashisth:** Conceptualization; funding acquisition; writing – review and editing; project administration; supervision; resources.

ACKNOWLEDGMENTS

We acknowledge the financial support provided by the National Institute of Health (NIH) through Grant No. R35GM138217. We are grateful for computational support through the following resources: Premise, a central shared HPC cluster at UNH supported by the Research Computing Center, and Biomade, a heterogeneous CPU/GPU cluster supported by the NSF EPSCoR award (OIA-1757371).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the supplementary material of this article.

ORCID

Harish Vashisth https://orcid.org/0000-0002-2087-2880

REFERENCES

1. Pawson T, Scott JD. Protein phosphorylation in signaling – 50 years and counting. *Trends Biochem Sci*. 2005;30(6):286-290.
2. Pang K, Wang W, Qin JX, et al. Role of protein phosphorylation in cell signaling, disease, and the intervention therapy. *MedComm*. 2022; 3(4):e175.
3. Bauman AL, Scott JD. Kinase- and phosphatase-anchoring proteins: harnessing the dynamic duo. *Nat Cell Biol*. 2002;4(8):E203-E206.
4. Adams JA. Kinetic and catalytic mechanisms of protein kinases. *Chem Rev*. 2001;101(8):2271-2290.
5. Cicenas J, Zalyte E, Bairoch A, Gaudet P. Kinases and cancer. *Cancers*. 2018;10(3):63.
6. Bhullar KS, Lagarón NO, McGowan EM, et al. Kinase-targeted cancer therapies: progress, challenges and future directions. *Mol Cancer*. 2018;17(1):1-20.
7. Sato S, Sanjo H, Takeda K, et al. Essential function for the kinase TAK1 in innate and adaptive immune responses. *Nat Immunol*. 2005; 6(11):1087-1095.
8. Chong ZZ, Shang YC, Wang S, Maiese K. A critical kinase cascade in neurological disorders: PI3K, Akt, and mTOR. *Future Neurol*. 2012; 7(6):733-748.
9. Tabit CE, Shenouda SM, Holbrook M, et al. Protein kinase C-β contributes to impaired endothelial insulin signaling in humans with diabetes mellitus. *Circulation*. 2013;127(1):86-95.

10. Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. *Science*. 2002; 298(5600):1912-1934.
11. Modi V, Dunbrack RL. A structurally-validated multiple sequence alignment of 497 human protein kinase domains. *Sci Rep*. 2019;9(1): 1-16.
12. Norman RA, Toader D, Ferguson AD. Structural approaches to obtain kinase selectivity. *Trends Pharmacol Sci*. 2012;33(5):273-278.
13. Davis MI, Hunt JP, Herrgard S, et al. Comprehensive analysis of kinase inhibitor selectivity. *Nat Biotechnol*. 2011;29(11):1046-1051.
14. Kiselyov A, Balakin K, Tkachenko S, Savchuk N. Recent progress in development of non-ATP competitive small-molecule inhibitors of protein kinases. *Mini Rev Med Chem*. 2006;6(6):711-717.
15. Liu Y, Gray NS. Rational design of inhibitors that bind to inactive kinase conformations. *Nat Chem Biol*. 2006;2(7):358-364.
16. Johnson LN. Protein kinase inhibitors: contributions from structure to clinical compounds. *Q Rev Biophys*. 2009;42(1):1-40.
17. Wang W, Krosky D, Ahn K. Discovery of inactive conformation-selective kinase inhibitors by utilizing cascade assays. *Biochemistry*. 2017;56(34):4449-4456.
18. Schindler T, Bornmann W, Pellicena P, Miller WT, Clarkson B, Kuriyan J. Structural mechanism for STI-571 inhibition of abelson tyrosine kinase. *Science*. 2000;289(5486):1938-1942.
19. Aleksandrov A, Simonson T. Molecular dynamics simulations show that conformational selection governs the binding preferences of imatinib for several tyrosine kinases. *J Biol Chem*. 2010;285(18):13807-13815.
20. Lin YL, Meng Y, Jiang W, Roux B. Explaining why Gleevec is a specific and potent inhibitor of Abl kinase. *Proc Natl Acad Sci U S A*. 2013; 110(5):1664-1669.
21. Agafonov RV, Wilson C, Otten R, Buosi V, Kern D. Energetic dissection of Gleevec's selectivity towards human tyrosine kinases. *Nat Struct Mol Biol*. 2014;21(10):848-853.
22. Ferguson AD, Sheth PR, Basso AD, et al. Structural basis of CX-4945 binding to human protein kinase CK2. *FEBS Lett*. 2011;585(1): 104-110.
23. Battistutta R, Cozza G, Pierre F, et al. Unprecedented selectivity and structural determinants of a new class of protein kinase CK2 inhibitors in clinical trials for the treatment of cancer. *Biochemistry*. 2011; 50(39):8478-8488.
24. Robinson DD, Sherman W, Farid R. Understanding kinase selectivity through energetic analysis of binding site waters. *ChemMedChem*. 2010;5(4):618-627.
25. Heffron TP, Wei B, Olivero A, et al. Rational design of phosphoinositide 3-kinase inhibitors that exhibit selectivity over the phosphoinositide 3-kinase isoform. *J Med Chem*. 2011;54(22):7815-7833.
26. Barelier S, Pons J, Gehring K, Lancelin JM, Krimm I. Ligand specificity in fragment-based drug design. *J Med Chem*. 2010;53(14):5256-5266.
27. Bamborough P, Brown MJ, Christopher JA, Chung CW, Mellor GW. Selectivity of kinase inhibitor fragments. *J Med Chem*. 2011;54(14): 5131-5143.
28. Erlanson DA, Arndt JW, Cancilla MT, et al. Discovery of a potent and highly selective PDK1 inhibitor via fragment-based drug discovery. *Bioorg Med Chem Lett*. 2011;21(10):3078-3083.
29. Feng L, Geisselbrecht Y, Blanck S, et al. Structurally sophisticated octahedral metal complexes as highly selective protein kinase inhibitors. *J Am Chem Soc*. 2011;133(15):5976-5986.
30. Wang B, Wu H, Hu C, et al. An overview of kinase downregulators and recent advances in discovery approaches. *Signal Transduct Target Ther*. 2021;6(1):1-19.
31. Cox KJ, Shomin CD, Ghosh I. Tinkering outside the kinase ATP box: allosteric (type IV) and bivalent (type V) inhibitors of protein kinases. *Future Med Chem*. 2010;3(1):29-43.
32. Huggins DJ, Sherman W, Tidor B. Rational approaches to improving selectivity in drug design. *J Med Chem*. 2012;55(4):1424-1444.
33. Shukla D, Meng Y, Roux B, Pande VS. Activation pathway of Src kinase reveals intermediate states as targets for drug design. *Nat Commun*. 2014;51(5):1-11.
34. Roskoski R. Classification of small molecule protein kinase inhibitors based upon the structures of their drug-enzyme complexes. *Pharmacol Res*. 2016;103:26-48.
35. Cohen P, Cross D, Jänne PA. Kinase drug discovery 20 years after imatinib: progress and future directions. *Nat Rev Drug Discov*. 2021; 20(7):551-569.
36. Wu P, Clausen MH, Nielsen TE. Allosteric small-molecule kinase inhibitors. *Pharmacol Ther*. 2015;156:59-68.
37. Wang Z, Canagarajah BJ, Boehm JC, et al. Structural basis of inhibitor selectivity in MAP kinases. *Structure*. 1998;6(9):1117-1128.
38. Quambusch L, Landel I, Depta L, et al. Covalent-allosteric inhibitors to achieve Akt isoform-selectivity. *Angew Chem Int Ed Engl*. 2019;58(52): 18823-18829.
39. Heinrich T, Grädler U, Böttcher H, Blaukat A, Shutes A. Allosteric IGF-1R inhibitors. *ACS Med Chem Lett*. 2010;1(5):199-203.
40. 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):6406.
41. Lee J, Pilch PF. The insulin receptor: structure, function, and signaling. *Am J Physiol Physiol*. 1994;266(2):C319-C334.
42. Cai W, Sakaguchi M, Kleinridders A, et al. Domain-dependent effects of insulin and IGF-1 receptors on signalling and gene expression. *Nat Commun*. 2017;8(1):1-14.
43. Belfiore A, Frasca F, Pandini G, Sciacca L, Vigneri R. Insulin receptor isoforms and insulin receptor/insulin-like growth factor receptor hybrids in physiology and disease. *Endocr Rev*. 2009; 30(6):586-623.
44. Hubbard SR. The insulin receptor: both a prototypical and atypical receptor tyrosine kinase. *Cold Spring Harb Perspect Biol*. 2013;5(3): a008946.
45. Ullrich A, Gray A, Tam AW, et al. Insulin-like growth factor I receptor primary structure: comparison with insulin receptor suggests structural determinants that define functional specificity. *EMBO J*. 1986; 5(10):2503-2512.
46. Munshi S, Kornienko M, Hall DL, et al. Crystal structure of the apo, unactivated insulin-like growth factor-1 receptor kinase: implication for inhibitor specificity. *J Biol Chem*. 2002;277(41):38797-38802.
47. Hubbard SR, Wei L, Hendrickson WA. Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature*. 1994; 372(6508):746-754.
48. Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem*. 2004;25(13):1605-1612.
49. Yang Z, Lasker K, Schneidman-Duhovny D, et al. UCSF Chimera, MODELLER, and IMP: an integrated modeling system. *J Struct Biol*. 2012;179(3):269-278.
50. Madeira F, Pearce M, Tivey ARN, et al. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res*. 2022; 50(W1):W276-W279.
51. Schrödinger, LLC. PyMOL Molecular Graphics System, Version 2.0. Schrödinger, LLC. 2015.
52. Madhavi Sastry G, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des*. 2013; 27(3):221-234.
53. Shelley JC, Cholleti A, Frye LL, Greenwood JR, Timlin MR, Uchimaya M. Epik: a software program for pK a prediction and protonation state generation for drug-like molecules. *J Comput Aided Mol Des*. 2007;21(12):681-691.

54. Friesner RA, Banks JL, Murphy RB, et al. Glide: a new approach for rapid, accurate docking and acoring. 1. Method and assessment of docking accuracy. *J Med Chem*. 2004;47(7):1739-1749.
55. Halgren TA, Murphy RB, Friesner RA, et al. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J Med Chem*. 2004;47(7):1750-1759.
56. Friesner RA, Murphy RB, Repasky MP, et al. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J Med Chem*. 2006;49(21):6177-6196.
57. Berendsen HJC, van der Spoel D, van Drunen R. GROMACS: a message-passing parallel molecular dynamics implementation. *Comput Phys Commun*. 1995;91(1–3):43-56.
58. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins*. 2006;65(3):712-725.
59. Wang J, Wang W, Kollman PA, Case DA. Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model*. 2006;25(2):247-260.
60. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. *J Comput Chem*. 2004; 25(9):1157-1174.
61. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys*. 1983;79(2):926-935.
62. Chester C, Friedman B, Ursell F. An extension of the method of steepest descents. *Proc Cambridge Philos Soc*. 1957;53(3):599-611.
63. Berendsen HJC, Postma JPM, Van Gunsteren WF, Dinola A, Haak JR. Molecular dynamics with coupling to an external bath. *J Chem Phys*. 1984;81(8):3684-3690.
64. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *J Chem Phys*. 2007;126(1):014101.
65. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: a linear constraint solver for molecular simulations. *J Comput Chem*. 1997;18(12): 1463-1472.
66. 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.
67. Tsui V, Case DA. Theory and applications of the generalized born solvation model in macromolecular simulations. *Biopolymers*. 2000; 56(4):275-291.
68. Chen J, Peng C, Wang J, Zhu W. Exploring molecular mechanism of allosteric inhibitor to relieve drug resistance of multiple mutations in HIV-1 protease by enhanced conformational sampling. *Proteins*. 2018;86(12):1294-1305.
69. Genheden S, Ryde U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin Drug Discov*. 2015; 10(5):449-461.
70. Wang E, Sun H, Wang J, et al. End-point binding free energy calculation with MM/PBSA and MM/GBSA: strategies and applications in drug design. *Chem Rev*. 2019;119(16):9478-9508.
71. Valdés-Tresanco MS, Valdés-Tresanco ME, Valiente PA, 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.
72. Chen J, Zeng Q, Wang W, Sun H, Hu G. Decoding the identification mechanism of an SAM-III riboswitch on ligands through multiple independent gaussian-accelerated molecular dynamics simulations. *J Chem Inf Model*. 2022;62(23):6118-6132.
73. Miller BR, McGee TD, Swails JM, Homeyer N, Gohlke H, Roitberg AE. MMPBSA.py: an efficient program for end-state free energy calculations. *J Chem Theory Comput*. 2012;8(9):3314-3321.
74. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC. GROMACS: fast, flexible, and free. *J Comput Chem*. 2005;26(16):1701-1718.
75. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph*. 1996;14(1):33-38.
76. David CC, Jacobs DJ. Principal component analysis: a method for determining the essential dynamics of proteins. *Methods Mol Biol*. 2014;1084:193-226.
77. Laskowski RA, Swindells MB. LigPlot+: multiple ligand–protein interaction diagrams for drug discovery. *J Chem Inf Model*. 2011;51(10): 2778-2786.
78. Zhang M, Liu Y, Jang H, Nussinov R. Strategy toward kinase-selective drug discovery. *J Chem Theory Comput*. 2023;19(5):1615-1628.
79. Hu H, Laufkötter O, Miljković F, Bajorath J. Data set of competitive and allosteric protein kinase inhibitors confirmed by X-ray crystallography. *Data Brief*. 2021;35:106816.
80. Taylor SS, Wu J, Bruystens JGH, et al. From structure to the dynamic regulation of a molecular switch: a journey over 3 decades. *J Biol Chem*. 2021;296:100746.
81. Morando MA, Saladino G, D'Amelio N, et al. Conformational selection and induced fit mechanisms in the binding of an anticancer drug to the c-Src kinase. *Sci Rep*. 2016;6:24439.
82. Luque I, Freire E. Structural stability of binding sites: consequences for binding affinity and allosteric effects. *Proteins*. 2000;41(S4): 63-71.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Verma J, Vashisth H. Molecular basis for differential recognition of an allosteric inhibitor by receptor tyrosine kinases. *Proteins*. 2024;1-18. doi:10.1002/prot.26685