Insights into structure and function of SHIP2-SH2: homology modeling, docking, and molecular dynamics study

Uzma Saqib · Mohammad Imran Siddiqi

Received: 29 September 2010 / Accepted: 27 January 2011 / Published online: 12 February 2011
© Springer-Verlag 2011

Abstract SRC homology 2 (SH2)-containing inositol 5'-phosphatase protein (SHIP2) is a potential target for type 2 diabetes. Its ability to dephosphorylate the lipid messenger phosphatidylinositol 3,4,5-trisphosphate [PtdIns(3,4,5)P₃], important for insulin signaling, makes it an important target against type 2 diabetes. The insulin-induced SHIP2 interaction with Shc is very important for the membrane localization and functioning of SHIP2. There is a bidentate relationship between the two proteins where two domains each from SHIP2 and Shc are involved in mutual binding. However in the present study, the SHIP2-SH2 domain binding with the phosphorylated tyrosine 317 on the collagen-homology (CH) domain of Shc, has been studied due to the indispensability of this interaction in SHIP2 localization. In the absence of the crystal structure of SHIP2-SH2, its structural model was developed followed by tracking its molecular interactions with Shc through molecular docking and dynamics studies. This study revealed much about the structural interactions between the SHIP2-SH2 and Shc-CH. Finally, docking study of a nonpeptide inhibitor into the SHIP2-SH2 domain further confirmed the structural interactions involved in ligand binding and also proposed the inhibitor as a major starting point against SHIP2-SH2 inhibition. The insights gained from the current study should prove useful in the design of more potent inhibitors against type 2 diabetes.

Keywords Docking · Molecular dynamics · Shc · SHIP2-SH2 · Type 2 diabetes

Introduction

Insulin-stimulated production of phosphatidylinositol 3,4,5-trisphosphate [PtdIns(3,4,5)P₃] by phosphatidylinositol 3-kinase is important for the glucose homeostasis pathway. PtdIns(3,4,5)P₃ plays a very critical step in proceeding the insulin-signaling cascade [8, 34, 41]. However, the dephosphorylation of PtdIns(3,4,5)P₃ molecule into PtdIns(3,4)P₂ by the 5' phosphatase-SRC homology 2 (SH2)-containing inositol 5'-phosphatase protein (SHIP2) leads to attenuation of this signaling pathway. The level of PtdIns(3,4,5)P₃ is sharply regulated by SHIP proteins. SHIP proteins are SRC homology 2 (SH2)-containing inositol 5'-phosphatase proteins with isozymes, SHIP1 and SHIP2. SHIP1 is present largely in haematopoietic cells, while SHIP2, which can thus be assumed as a major regulator, is found in skeletal muscles, among other insulin-sensitive tissues [27, 37]. Studies indicate that mice deficient of SHIP2, are hypoglycemic and sensitized to insulin suggesting that SHIP2 may have an important role in negative insulin signaling [6]. Thus, targeting SHIP2 for consistent insulin signaling and glucose homeostasis can be a plausible solution to this problem.

A change of cellular localization from cytosol to membrane is crucial for SHIP2 functioning so that it might get access to its substrate [28]. This membrane targeting of SHIP2 is aided by Shc which is a member of a group of cytoplasmic-signaling proteins [17]. Shc gets tyrosine phosphorylated at its central collagen-homology (CH) domain after being stimulated with a wide variety of growth factors and cytokines [4, 7]. It has been observed

U. Saqib · M. I. Siddiqi (✉)
Molecular and Structural Biology Division,
Central Drug Research Institute, CSIR India,
Lucknow 226 001, India
e-mail: imsididi@yahoo.com

M. I. Siddiqi
e-mail: mi_siddiqi@cdri.res.in

that phosphorylation of Shc on Tyr-317 located at the CH domain creates a binding site for the SH2 domain of SHIP2 [21]. Besides the CH domain, it also contains an N-terminal phosphotyrosine [Tyr(P)] binding (PTB) domain [38, 40] which in turn binds with phosphorylated 987-Tyr residue in the NPXY motif present at the C-terminus of SHIP2 [42]. This can also be described in terms of SHIP2 which possess two potential sites that associate with Shc, one via its SH2 domain and another by its NPXY motif at C-terminal domain. The interaction taking place between the SH2 domain and phosphorylated tyrosine 317 (pTyr³¹⁷) of Shc-CH domain might be mediated via a conserved Arg47 residue of the former [9]. Studies indicate the high indispensability of this interaction with regard to SHIP2 functioning [16]. Hence, the association of SHIP2 with Shc is a key to triggering an efficient inhibitory effect on insulin cascade and this prompted us to probe the SHIP2(SH2)-Shc interaction so that further light can be thrown on their structural and molecular interaction profile.

Computational studies on macromolecule interactions are being routinely done these days. These studies provide a robust way of probing the key residues involved in a three-dimensional (3D) protein–ligand interaction system. There have been many recent publications where the results of computational analyses either helped in further narrowing down the studies or validated the already performed studies at the molecular level. Our understanding about the protein structure and function is mostly dependent upon X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy [5]. However, many proteins are difficult to crystallize, e.g., membrane proteins, or are too large to be studied with NMR, and both approaches are very time intensive. With continuous emergence of new sequences of proteins involved in various diseases, our desire to predict their structures in order to understand their function is greatly elevated. Hence, quick and accurate strategies for computational modeling have been in use in past couple of years [12, 24, 29, 43]. Recently, a computational study has been done which sheds light on how a protein GEF mediates the activation of a small G protein Arf1 [25]. Also, a computational study on PDZ binding to the BAR domain of PICK1 elucidated by molecular dynamics has been done which details the molecular mechanisms of their binding [13]. Other similar computational studies on protein–ligand interactions have been reported earlier [1, 2, 22].

The present work details the construction of a 3D model for the human SHIP2-SH2 domain due to the unavailability of its crystal structure in Protein Data Bank (PDB), followed by its docking and molecular dynamics study with Shc-CH domain, its “partner” in membrane localization. Docking studies were then performed using a non-peptide inhibitor, AP22650 [36] of src-SH2 domain. AP22650 though initially designed for src-SH2 domain was non-specific for the same.

However, due to its potency along with highly similar active site of shc-SH2 with that of SHIP2-SH2 prompted us to use it against the latter. As expected, AP22650 showed quite reasonable binding conformation in the SHIP2-SH2 active site along with displaying the major binding interactions to be further qualified as a SHIP2-SH2 inhibitor. The interactions of SHIP2-SH2 with Shc-CH and with AP22650 provided some important clues for the design of novel and potent inhibitors against SHIP2-SH2.

Materials and methods

Computational resources

Molecular modeling was carried out on SGI Origin 300 workstation equipped with 4×600 MHz R12000 processors workstation. For sequence alignment ClustalW 1.8 in InsightII/homology module [Insight II Program, Accelrys Inc. (2001) San Diego, CA <http://accelrys.com/>] were used. Comparative modeling was performed using InsightII/MODELER. Energy minimization and molecular dynamic simulations were carried out using InsightII/CharmM module. Ramachandran plot was generated through structure analysis and verification server (SAVS) while other statistics Prostat and Profile-3D were generated in InsightII/HOMOLOGY. Molecular docking of Shc-CH was done by PatchDock protein–protein docking server while that of AP22650 was done using FlexX program interfaced with Sybyl 7.1 (Tripos Inc., 1699 South Hanley Road, St. Louis, MO, 63144, USA. <http://www.tripos.com/>).

Sequence alignment and comparative modeling

The amino acid sequence of the target protein SHIP2-SH2 (accession no. O15357), was extracted from the Swiss-Prot protein sequence database (<http://www.expasy.ch/sprot/>). The SHIP2-SH2 sequence is composed of 97 residues. Sequence alignment was derived with the ClustalW 1.8 package and Align123 in InsightII/Homology module using the BLOSUM matrices [14] for scoring the alignments. Procuring a high quality multiple sequence alignment of the target with the templates is a central step in protein homology modeling in view of the fact that only small local errors made in alignment can be corrected in the subsequent steps. The average NMR structure and X-ray crystal structure of template proteins SHIP1-SH2 and SAP-SH2 both from homosapiens, PDB Id: 2YSX and PDB Id: 1D4W [31], respectively, were retrieved from Protein Data Bank. The resulting alignment was quite straightforward as there were no loop insertions or deletions during the entire length of the alignment. Finally, the alignment was used as an input for the automated homology modeling program

MODELER interfaced with the InsightII software suite. MODELER is used for homology or comparative modeling of protein three-dimensional structures. Out of the three models generated, the model with best stereochemical quality as determined by Ramachandran plot was selected for further analysis. Hydrogen atoms were added to it followed by energy minimization, employing the InsightII/Discover module in a stepwise fashion following a standard procedure consisting of 500 steps of steepest descent and 1,000 steps of conjugate gradient minimization with an RMS gradient of the potential energy of 0.001 kcal/mol Å in each step.

Model quality assessment

The quality of the final refined model was assessed by subjecting it to a series of tests for its internal consistency and reliability. The InsightII program Prostat was used to analyze the properties of bonds, angles, and torsions while the Profile-3D program was used to check the structure and sequence compatibility. The cut-off used was 5 standard deviations from the reference value. The online SAVS server's PROCHECK suite of programs was used for assessing the "stereochemical quality" of the modeled protein structure.

SHIP2(SH2)-Shc(CH) docking

For the purpose of complex protein formation, the 3D structure of Shc (PDB Id: 1WCP) was truncated to carry the central collagen-homology domain, CH domain (residues: 211–368) and docked onto the SHIP2-SH2 domain using PatchDock. To perform docking, the set of amino acid residues that, according to published data [31], are important for SH2-Shc interactions was specified as a possible binding site on the SHIP2-SH2 domain surface. This set included Arg28, Arg47, Ser49, Glu50, Ser51, Arg70, and Thr83. PatchDock is a geometry-based molecular docking algorithm, which aims in finding docking transformations that yield good molecular shape complementarity.

Molecular dynamics of the binary complex

This resulting binary complex structure was further refined by energy minimization by 500 steps of steepest descent followed by 2,000 steps of conjugate descent molecular dynamics (MD) simulation with CHARMM program using CHARMM force field [3]. Then, the complex was embedded in a 10 Å layer of water molecules to imitate aqueous solvent conditions. This solvated complex was subjected to energy minimization first by 500 steps of steepest descent followed by 2,000 steps of conjugate gradient method. The system was heated from 50 K to 300 K over a period of 50 ps with a time step of 1 fs and the velocities being reassigned in the system every 0.05 ps.

The system was further equilibrated with a 1 fs time step for 100 ps so that the energy of the system achieves complete stability. Production runs were performed at 300 K and carried out under constant number of particles, volume, and temperature conditions for 500 ps with a 1 fs time step. All the bonds involving hydrogen atom were constrained using SHAKE algorithm in all simulations [33]. The molecular trajectory for the systems generated by the molecular dynamics simulations were analyzed using the visual molecular dynamics (VMD) [15] software.

Docking and molecular dynamic studies of AP22650

Molecular docking using FlexX program interfaced with Sybyl 7.1 was carried out of the AP22650 inhibitor on SHIP2-SH2 domain. Docking was done using the standard parameters of the FlexX program [18, 32] as implemented in Sybyl 7.1. All the flexibilities of its rotatable bonds were considered in the process of docking for identifying the best binding conformation. The refined model was chosen as the receptor for molecular docking. The ligand-binding site in SHIP2-SH2 was defined based on the available information from the literature. The inhibitor was docked into the protein binding-site with hydrogens present and formal charges were assigned by FlexX. The best conformation was chosen based on the best Dock score and binding pose among the resulting 30 docking solutions. The molecular dynamics studies of the resulting binary complex of SHIP2-SH2 with docked AP22650 were done exactly similar to that done above for SHIP2-SH2-Shc-CH complex.

Results and discussion

3D model building of SHIP2-SH2 domain

Comparative modeling is used to develop a 3D structure of the target protein making use of the structural information from the previously resolved protein structures with substantial similarity. To date, several NMR and crystal structures of SH2 domains from various sources have been reported by different research groups [10, 23, 26, 30]. However, recently the NMR structure of SH2 domain of SHIP1, an isozyme of SHIP2-SH2 has been deposited in Protein Data Bank by Kasai et al. (data not yet published). The SHIP1-SH2 (PDB: 2YSX) shares 54% identity with the SHIP2-SH2 domain. Another SH2 domain from SAP protein in complex with SLAM phosphopeptide (PDB: 1D4W) showing an identity of 42% was also used as a template, as that can be useful in guiding the corresponding active site residues. The alignment of the deduced amino acid sequence of the SHIP2-SH2 with the sequences from SHIP1 and SAP obtained from ClustalW 1.8 [39] is shown in Fig. 1. As can

Fig. 1 Multiple sequence alignment using ClustalW. Residues highlighted in red correspond to identical/conserved residues, while residues in red text are similar in the three proteins. Secondary structural elements are shown for SHIP1 (2YSX)

be observed, the alignment between the three proteins is highly conserved assuring superior model quality. The superimposed structures of all the three proteins are shown in Fig. 2, which besides displaying that the SHIP2-SH2 has the characteristic SH2 fold including a central β sheet with α helices packed against either side also show that the homology modeling has been quite accurate as the three proteins superimpose very well.

Protein structure validation

The quality of the initial model was improved by subjecting it to a crude energy minimization protocol as detailed in the “Materials and methods” section. These minimizations helped relieve any steric clashes or improper geometries in the protein structure to produce a model with correct bond lengths and bond angles and where individual atoms are not too close together. The refined structure was evaluated for overall quality using available analysis procedures. These analyses compare specific properties of

the model with those for known high quality protein structures. For this purpose, three protein analysis programs: PROCHECK [20], Prostat, and Profile-3D were utilized. Prostat was used to assess the stereochemical quality of the model. This program verifies the accuracy of parameters such as bond lengths, bond angles, and correctness of amino acid chirality. No spurious angle or bond length was detected in our model. The results are listed at the bottom of Table 1.

Another important indicator of the stereochemical quality of the model is the distribution of the main chain torsion angles phi and psi which may be examined in a Ramachandran plot. The Ramachandran plot of the phi–psi plots is shown in Fig. 3 while the detailed results are listed in Table 1. The plot clearly shows that all the residues are either in most favored or additional allowed regions and none in generously allowed or disallowed regions, suggesting high model quality. Finally, the 3D homology model was verified using the Profile-3D program in InsightII software, shown in Fig. 4. Profile-3D is a program based on algorithms that measure the compatibility of an amino acid sequence with a three-dimensional structure by reducing the structure to a one-dimensional representation, known as the

Fig. 2 Superimposed structures of SHIP2-SH2 (magenta), SHIP1-SH2 (blue), and SAP-SH2 (red)

Table 1 Results of protein structure check by PROCHECK and Prostat

Residues in most favored regions	72 (86.7%)
Residues in additional allowed regions	11 (13.3%)
Residues in generously allowed regions	0 (0.0%)
Residues in disallowed regions	0 (0.0%)
Number of non-glycine and non-proline residues	83
Number of end-residues (excl. Gly and Pro)	2
Number of glycine residues (shown as triangles)	8
Number of proline residues	4
Total number of residues	97
Number of bond distances with significant deviationsb (by Prostat)	0
Number of bond angles with significant deviations (by Prostat)	0

Fig. 3 Ramachandran plot of the homology-modeled structure of SHIP2-SH2. The different colored areas indicate “disallowed” (white), “generously allowed” (light yellow), “additional allowed” (yellow), and “most favored” (red) regions (refer to Table 1)

3D profile, which can be aligned with the sequence. Hence, the resulting alignment score is a measure of the compatibility of the sequence with the structure. A smoothing window size of ten residues was used. The analysis yielded an overall score of 41.19 similar to the typical score of 43.74 for a native protein of equivalent size and well above 19.68, a score that would indicate an incorrect structure. In summary, the above mentioned analyses indicate that the model structure is consistent with our current understanding of the protein structure.

Fig. 4 The evaluation of the SHIP2-SH2-modeled structure by Profile-3D program

Docking of SHC into SHIP2-SH2

The binding of one protein to the active site of another protein is typically associated with local and global structural rearrangement of the receptor (induced-fit behavior). As a result, protein–protein interaction studies and structure-based drug design preferentially relies on the structures of protein–protein complexes in which the second protein behaves like a ligand. Keeping this in mind, the next step was to develop a protein–protein complex of SHIP2-SH2 with Shc-CH that would offer a more detailed and accurate picture of the interactions and structural complementarities between the active sites of both. Information from such a protein–protein complex may be used appropriately to design inhibitors and may provide more meaningful structural models. Consequently, the protein–protein complex was developed from the initially generated homology model of SHIP2-SH2 domain and PDB deposited homology model of Shc-CH domain through docking by PatchDock server [11, 35], by including information about the putative binding sites of both. The structure of human Shc modeled through homology modeling by Suenaga et al. deposited in the PDB was our structure of choice among other crystal structures of Shc because the Shc-CH domain carrying the highly conserved pTyr³¹⁷ (YVNV motif) residue responsible for interacting with SHIP2-SH2 is not resolved in other crystal structures of Shc. The docked Shc-CH domain into the SHIP2-SH2 domain is shown in Fig. 5.

Classical SH2 domains bind phosphopeptides in a “two-pronged” fashion; the phosphorylated tyrosine residue binds in a pocket on one side of the central sheet, and the three to five residues C-terminal to it bind in a hydrophobic pocket or groove on the opposite side [19]. The phosphotyrosine-binding pocket of SHIP2-SH2 contains

Fig. 5 Ribbon representation of docked CH domain (blue) of Shc into SH2 domain of SHIP2 (red)

Fig. 6 Interactions of pTyr³¹⁷ of Shc-CH domain into the phosphotyrosine-binding pocket of SHIP2-SH2. The highly conserved interaction between pTyr³¹⁷ and Arg47 are shown in thicker sticks as compared other residues. The hydrogen bond interaction between pTyr³¹⁷ and Arg28 is mediated by water (red sphere)

some highly conserved residues like Arg28, Arg47, Ser49, Glu50, Ser51, Arg70, etc. Active site of the enzyme was defined by all these residues which constitute conserved binding site regions in SH2 domains. The Shc-CH retains these general binding features, with pTyr³¹⁷ extending into the phosphotyrosine-binding pocket (Fig. 6). In the docked complex, pTyr³¹⁷ is positioned in the phosphotyrosine-binding BC loop. The pTyr³¹⁷ is coordinated in a manner like that has been observed in other typical SH2 complexes. The universally conserved Arg47 forms the expected bidentate hydrogen bonds with phosphate oxygens. Besides, the other conserved Arg28, which is little farther away from the phosphate oxygens is making a hydrogen bond via a conserved water molecule. Interactions with Ser49, Glu50, Ser51, and Arg70 further stabilize the SH2-CH interaction. Though the importance of N-terminal pY-1 and pY-3 positions (with respect to pTyr³¹⁷) of SH2 in protein binding has been described in one of the templates SAP, however since the corresponding residues are not conserved in SHIP2-SH2 hence only the well conserved residues at the C-terminus of pTyr³¹⁷ seemed to be important for further binding. In order to validate the docking study and to confirm the stability of the protein–protein complex, further minimization and dynamics study was performed.

Molecular dynamics

Energy minimization was carried out on the resulting protein–protein complex in an aqueous environment, as detailed in the “Materials and methods” section. Care was taken to maintain the overall geometry of the complex during the refinement. The stability of the model structure was finally investigated by using an unconstrained MD simulation. After the dynamics studies were over, the stability of the binary complex was evaluated based on the analyses of the root mean square deviation (RMSD) of the geometry and the total energy of the protein. Both the analyses have revealed that the system was stabilized thus, confirming that the simulation is reliable with no artifacts (Fig. 7a, b). The total energy did not deviate much after attaining a stable state, also the RMSD between the initial

Fig. 7 Graph showing energy (a), RMSD (in Å) of protein geometry (b) and RMSD of the hydrogen bond distance between the phosphate oxygen of the pTyr³¹⁷ and the –NH group of Arg47 (c) observed with respect to time throughout the trajectory of SHIP2-SH2-Shc-CH binary complex MD simulation analysis

Fig. 8 pTyr³¹⁷VNV motif of the SHC-CH into the phosphotyrosine binding site and the hydrophobic pocket (Val + 3 site) of SH2 domain

and final structures was also quite low, suggesting that the two proteins were reasonable in their energy as well as geometry and their relative movements from their presumed position was not too large. The highly conserved pTyr³¹⁷ and Arg47 interaction, which has been well documented in the literature as the most important interaction, has also been studied during the entire dynamic simulation. As shown in Fig. 7c, the RMSD of the hydrogen bond distance formed between the phosphate oxygen of the pTyr³¹⁷ and the –NH group of Arg47 residue with respect to the average structure, oscillates very narrowly at a fixed position during the entire simulation confirming that this is one of the most stable and uniform interaction between the two proteins. Other important interactions have also been noticed during the simulation and they also follow the same pattern as of Arg47 (data not shown). Hence, all these studies confirm that the system remained in equilibrium during the entire simulation. Figure 8 shows the position of pTyr³¹⁷VNV motif of the Shc-CH into the phosphotyrosine binding site and the hydrophobic pocket (Val + 3 site) of SH2 domain. The pTyr³¹⁷ retains all the conserved interactions as observed during the docking studies. The pY + 1 to pY + 3 positions following pTyr³¹⁷ belonging to the conserved motif YVNV are coordinated by interactions with the residues in the hydrophobic cleft of the Val + 3 site. Comparison of the initial unbound SHIP2-SH2 domain structure with the Shc-CH bound form after molecular dynamics shows small overall deviations. Importantly, in the unbound form, the EF and BG loops fold inward to close the hydrophobic +3 binding groove, however due to the insertion of Val-320

(pTyr³¹⁷ + 3 position) the hydrophobic cleft can be observed in its open conformation (Fig. 9).

Docking and molecular dynamic studies of AP22650 with SHIP2-SH2

In order to further gain insight in the SHIP2-SH2 interactions with interacting partners, we performed docking studies of an inhibitor, AP22650 (Fig. 10a). The inhibitor though initially designed for targeting osteoclasts carrying the src SH2 domain was non-selective. Hence, docking studies were done in order to find out whether the same could bind to SHIP2-SH2, too. Though shc-SH2 is just 28.8% identical to SHIP2-SH2 in overall sequence length, however, there is considerable identity/similarity in their binding site residues with the exception of few, yet important residues. Interestingly, AP22650 docked quite well with a FlexX docking energy of –18.067 kcal/mol and a reasonable conformation into the SHIP2-SH2 active site as shown in Fig. 10b. It has been observed that docked AP22650 acquire the same position and displayed the same interactions in the SHIP2-SH2 as observed for pTyr³¹⁷VNV motif of Shc-CH. The phosphate group of the inhibitor snugly fits into the phosphotyrosine binding site, while the preceeding hydrophobic groups mimic the +1 and +3 positions of the pTyr³¹⁷VNV motif of Shc-CH, thereby entering into the hydrophobic pocket between the BG and EF loop around the hydrophobically conserved Thr 83. In order to further affirm this binding conformation along with stability of various important binding interactions, we

Fig. 9 The open (cyan) and closed (magenta) conformation of SHIP2-SH2 domain

Fig. 10 AP22650 (**a**) docked into the SHIP2-SH2 active site (**b**)

carried out molecular dynamics simulation of the docked AP22650 in the SHIP2-SH2 binding site just exactly similar to that done for SHIP2-SH2-Shc binary complex. Again, the stability of the binary complex was evaluated based on the analyses of the total energy and RMSD of the geometry of the complex Fig. 11a, b, respectively. The total energy and RMSD of the complex remained quite consistent throughout the simulation. Also, all the important interactions involved in ligand binding have been carefully checked for their consistency, which too shown minute deviations from the initial starting structure. Thus, the docking studies combined with the molecular dynamics suggest AP22650 to be a potential prototype for further modification so that highly specific inhibitor against SHIP2-SH2 can be designed based on this. Hence, the differences between the active sites of src-SH2 and SHIP2-SH2 have to be taken into account in order to come up with more specific and highly selective SHIP2-SH2 inhibitors.

After a thorough comparison of binding modes of both Shc and AP22650, several conclusions can be drawn with respect to SH2 binding-site interactions. First, SH2 domains have a well-conserved phosphotyrosine-binding site on one side of the central sheet where phosphorylated moieties which can be either the phosphotyrosine residues of Shc-CH domain (natural ligand) or the 4'-phosphono-Phe group of

the AP22650 (a non-peptide inhibitor) bind with stable ionic and hydrogen bond interactions via their oxygen to both protein backbone and side-chain atoms of important active-site residues like Arg28, Arg47, Ser49, Glu50, Ser51, Arg70, etc. The 4'-phosphonate moiety of AP22650 also penetrates deeply into the phosphotyrosine-binding site and was predicted to form similar set of molecular interactions as that of phosphotyrosine and hence is expected to be functionally equivalent to it. On the other hand, the seven-membered ring of AP22650 fits well in the hydrophobic groove formed by the side chains of Tyr69 and Arg70 making hydrophobic interactions with them similar to that observed for the side-chain group of Val318 from pTyr³¹⁷VNV motif of Shc. The benzamide carbonyl forms a hydrogen bond with the backbone NH of R70 which is also shown by the backbone peptide moiety of pYVNV motif of Shc. Finally, the cyclohexyl group of AP22650 extends into the hydrophobic pY+3 pocket of the SH2 binding site, just like the Val320 of pYVNV motif where both make majority of hydrophobic interactions with non-polar residues including the important interaction with Thr83. Thus, the non-peptide inhibitor AP22650 exhibits very similar interactions as well as mode of binding to the classical pYNV motif of Shc. This might be attributed to

Fig. 11 Graph showing energy (**a**), RMSD (in Å) of protein geometry (**b**) observed with respect to time throughout the trajectory of SHIP2-SH2-AP22650 MD simulation analysis

their highly similar geometry and conformation in the protein binding site, which in turn is again due to almost similar positioning of functional groups in both. Hence, more potent inhibitors might be developed based upon AP22650 binding, where activity and specificity can be enhanced by judicious modification of functional groups at its various critical interacting positions.

Conclusions

In summary, the 3D structure of SHIP2-SH2 was developed using the PDB deposited 3D structures of SHIP1 and SAP proteins. Further, the model was refined by energy minimization and validated by various methods. Next, the protein–protein complex was obtained by docking Shc-CH domain into the SHIP2-SH2 domain. This docked complex was then minimized in order to relieve any uneven geometries and a molecular dynamics study was performed. The resulting complex was analyzed for its structural consistency in terms of its final energy, RMSD, and molecular interactions with reference to available literature. It can be used as a significant tool to augment our understanding of interactions of the SHIP2-SH2 domain with the Shc-CH domain at the residue level. This information could also aid in the understanding of its catalytic mechanism, thus providing a prototype for inhibitor binding studies. Finally, a known src-SH2 inhibitor AP22650 was used for docking studies in order to further take the study to a next level. AP22650 seemed to be quite reasonable in terms of its Dock score and interactions. The small, yet important sequence differences between the src-SH2 and SHIP2-SH2 can be exploited for further design of specific inhibitors based on AP22650.

The two structural models of SHIP2-SH2 with Shc-CH and with AP22650 should prove useful in the design and development of inhibitors as potential novel therapeutic agents against diabetes by either de novo drug design or virtual screening of large chemical databases. Although the probability of non-selectivity and non-specificity of the SH2 domain inhibitors cannot be ruled out due to highly conserved nature of SH2 domains, though, small differences at sequence level among the various SH2 domains from various sources could be potentially utilized in designing highly specific inhibitors. Besides, further work has to be done in order to probe the overall SHIP2-Shc bidentate partnership, in which the Shc PTB domain interactions with phosphorylated NPxY motifs of SHIP2; concomitantly with the SHIP2-SH2 domain binding with the Shc-CH domain carrying the pTyr³¹⁷VNV motif could be evaluated as a single interaction system in a more detailed context.

Acknowledgments This manuscript is CDRI communication number 7901. This work was supported by the grants from Council of Scientific and Industrial Research (CSIR India) funded network project NWP0034 (validation of identified screening models and development of new alternative models for evaluation of new drug entities). US thank CSIR for fellowship.

References

1. Bakan A, Bahar I (2011) *Pac Symp Biocomput* 2011:181–192
2. Bodian DL, Radmer RJ, Holbert S, Klein TE (2011) *Pac Symp Biocomput* 193–204
3. Brooks BR, Brucoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) *J Comp Chem* 4:187–217
4. Burns LA, Karnitz LM, Sutor SL, Abraham RT (1993) *J Biol Chem* 268:17659–17661
5. Chasse GA, Rodriguez AM, Mak ML, Deretey E, Perczel A, Sosa CP, Enriz RD, Csizmadia IG (2001) *J Mol Struct* 537:319–361
6. Clement S, Krause U, Desmedt F, Tanti JF, Behrends J, Pesesse X, Sasaki T, Penninger J, Doherty M, Malaisse W, Dumont JE, Marchand-Brustel YL, Erneux C, Hue L, Schurmans S (2001) *Nature* 409:92–97
7. Crowe AJ, McGlade J, Pawson T, Hayman MJ (1994) *Oncogene* 9:537–544
8. Czech MP, Corvera S (1999) *J Biol Chem* 274:1865–1868
9. Damen JE, Liu L, Rosten P, Humphries RK, Jefferson AB, Majerus PW, Krystal G (1996) *Proc Natl Acad Sci USA* 93:1689–1693
10. Donaldson LW, Gish G, Pawson T, Kay LE, Forman-Kay JD (2002) *Proc Natl Acad Sci USA* 99:14053–14058
11. Duhovny D, Nussinov R, Wolfson HJ (2002) *Proceedings of the 2nd Workshop on Algorithms in Bioinformatics (WABI)* Rome, Italy, *Lecture Notes in Computer Science*. 2452:185–200
12. Eswar N, Eramian D, Webb B, Shen M, Sali A (2008) *Methods Mol Biol* 426:145–159
13. He Y, Liwo A, Weinstein H, Scheraga HA (2011) *J Mol Biol* 405:298–314
14. Henikoff S, Henikoff JG (1992) *Proc Natl Acad Sci USA* 89:10915–10919
15. Humphrey W, Dalke A, Schulten K (1996) *J Mol Graph* 14:33–38
16. Ishihara H, Sasaoka T, Ishiki M, Wada T, Hori H, Kagawa S, Kobayashi M (2002) *Mol Endocrinol* 16:2371–2381
17. Koch CA, Anderson D, Moran MF, Ellis C, Pawson T (1991) *Science* 252:668–674
18. Kramer B, Rarey M, Lengauer T (1999) *Proteins Struct Funct Genet* 37:228–241
19. Kuriyan J, Cowburn D (1997) *Annu Rev Biophys Biomol Struct* 26:259–288
20. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) *J Appl Crystallogr* 26:283–291
21. Liu L, Damen JE, Hughes MR, Babic I, Jirik FR, Krystal G (1997) *J Biol Chem* 272:8983–8988
22. Lukman S, Grant BJ, Gorfe AA, Grant GH, McCammon JA (2010) *PLoS Comput Biol* 6:e1000922
23. Mallis RJ, Brazin KN, Fulton DB, Andreotti AH (2002) *Nat Struct Biol* 9:900–905
24. Marti-Renom M, Pieper U, Madhusudhan M, Rossi A, Eswar N, Davis F, Al-Shahrour F, Dopazo J, Sali A (2007) *Nucleic Acids Res* 35:W393–W397
25. Meriam BH, Charles HR (2010) *PLoS ONE* 5:e9142
26. Nolte RT, Eck MJ, Schlessinger J, Shoelson SE, Harrison SC (1996) *Nat Struct Biol* 3:364
27. Pesesse X, Deleu S, Smedt FD, Drayer L, Erneux C (1997) *Biochem Biophys Res Commun* 239:697–700

28. Phee H, Jacob A, Coggeshall KM (2000) *J Biol Chem* 275:19090–19097
29. Pieper U, Eswar N, Webb B, Eramian D, Kelly L, Barkan D, Carter H, Mankoo P, Karchin R, Marti-Renom M, Davis F, Sali A (2009) *Nucleic Acids Res* 37:D347–D354
30. Pletneva EV, Sundd M, Fulton DB, Andreotti AH (2006) *J Mol Biol* 357:550–561
31. Poy F, Yaffe MB, Sayos J, Saxena K, Morra M, Sumegi J, Cantley LC, Terhorst C, Eck MJ (1999) *Mol Cell* 4:555–561
32. Rarey M, Kramer B, Lengauer T, Klebe G (1996) *J Mol Biol* 261:470–489
33. Ryckaert JP, Ciccitti G, Berendsen HJC (1977) *J Comp Phys* 23:327–341
34. Saltiel AR (1996) *Am J Physiol* 270:375–385
35. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ (2005) *Nucl Acids Res* 33:363–367
36. Shakespeare W, Yang M, Bohacek R, Cerasoli F, Stebbins K, Sundaramoorthi R, Azimioara M, Vu C, Pradeepan S, Metcalf C, Haraldson C, Merry T, Dalgarno D, Narula S, Hatada M, Lu X, Schravendijk MR, Adams S, Violette S, Smith J, Guan W, Bartlett C, Herson J, Iulucci J, Weigele M, Sawyer T (2000) *Proc Natl Acad Sci USA* 97:9373–9378
37. Sleeman MW, Wortley KE, Lai KM, Gowen LC, Kintner J, Kline WO, Garcia K, Stitt TN, Yancopoulos GD, Wiegand SJ, Glass DJ (2005) *Nat Med* 11:199–205
38. Songyang Z, Shoelson SE, McGlade J, Olivier P, Pawson T, Bustelo XR, Barbacid M, Sabe H, Hanafusa H, Yi T, Ren R, Baltimore D, Ratnofsky S, Feldman RA, Cantley LC (1994) *Mol Cell Biol* 14:2777–2785
39. Thompson JD, Higgins DG, Gibson TJ (1994) *Nucleic Acids Res* 22:4673–4680
40. van der Geer P, Wiley S, Lai VKM, Olivier JP, Gish G, Stephens R, Kaplan D, Shoelson S, Pawson T (1995) *Curr Biol* 5:404–412
41. Virkamäki A, Ueki K, Kahn CR (1999) *J Clin Invest* 103:931–943
42. Wisniewski D, Strife A, Swendeman S, Erdjument-Bromage H, Geromanos S, Kavanaugh WM, Tempst P, Clarkson B (1999) *Blood* 93:2707–2720
43. Zhou H, Skolnick J (2007) *Biophys J* 93:1510–1518