J Mol Model (2014) 20:2446
DOI 10.1007/s00894-014-2446-7

ORIGINAL PAPER

Combined docking, molecular dynamics simulations and spectroscopic studies for the rational design of a dipeptide ligand for affinity chromatography separation of human serum albumin

Elham Aghaee · Jahan B. Ghasemi · Firouzeh Manouchehri · Saeed Balalaie

Received: 7 April 2014 / Accepted: 26 August 2014 / Published online: 16 September 2014

Abstract A computational approach to designing a peptide-based ligand for the purification of human serum albumin (HSA) was undertaken using molecular docking and molecular dynamics (MD) simulation. A three-step procedure was performed to design a specific ligand for HSA. Based on the candidate pocket structure of HSA (warfarin binding site), a peptide library was built. These peptides were then docked into the pocket of HSA using the GOLD program. The GOLDscore values were used to determine the affinity of peptides for HSA. Consequently, the dipeptide Trp–Trp, which shows a high GOLDscore value, was selected and linked to a spacer arm of Lys[$CO(CH_2)_5NH$] on the surface of ECH-lysine sepharose 4 gel. For further evaluation, the Autodock Vina program was used to dock the linked compound into the pocket of HSA. The docking simulation was performed to obtain a first guess of the binding structure of the spacer–Trp–Trp–HSA complex and subsequently analyzed by MD simulations to assess the reliability of the docking results. These MD simulations indicated that the ligand–HSA complex remains stable, and water molecules can bridge between the ligand and the protein by hydrogen bonds. Finally, absorption spectroscopic studies were performed to illustrate the appropriateness of the binding affinity of the designed ligand toward HSA. These studies demonstrate that the designed dipeptide can bind preferentially to the warfarin binding site.

E. Aghaee · J. B. Ghasemi (✉) · F. Manouchehri
Drug Design in Silico Laboratory, Chemistry Faculty, KN Toosi University of Technology, 15875-4416, 470 Mirdamad Ave. West, 19697 Tehran, Iran
e-mail: jahan.ghasemi@gmail.com

S. Balalaie
Peptide research Center, Chemistry Faculty, K N Toosi University of Technology, Tehran, Iran

Keywords Docking · Molecular dynamics · Peptide ligand · Rational design · Affinity chromatography · Human serum albumin

Introduction

Analysis of biological fluids such as serum, plasma and urine is complicated due to the enormous dynamic range of protein concentrations existing in these complex mixtures [1–4]. Human serum albumin (HSA) generally makes up more than half of the total plasma or serum proteins, so it often interferes with detection, determination and purification of low abundance proteins that can be indicators or biomarkers for various diseases. Therefore, decreasing the concentration of high abundance proteins, especially albumin, is an essential step in the detection of low abundance proteins in serum [5–7]. On the other hand, HSA, which is commonly used for therapeutic purposes, must be of high purity. Thus, selective isolation of HSA from blood plasma can be very helpful in therapeutic procedures.

Affinity chromatography is a highly selective separation technique usually employed for the purification and isolation of biological macromolecules [8, 9]. This technique separates molecules on the basis of a reversible interaction between a target macromolecule and a specific ligand coupled to a chromatographic matrix. Affinity chromatography is a highly selective technique that offers high resolution as well as having a high capacity for the molecule(s) of interest [10]. Researchers are now focusing on the design of affinity ligands for the specific separation of target molecules [11, 12]. For example, enzymes, antibodies, antigens, small nucleic acids or peptides can be used as affinity ligands to purify their corresponding binding partners. Commonly, to identify an appropriate affinity medium, different synthetic ligands are immobilized onto a

matrix, and the more efficient medium with higher selectivity and resolution for the target molecule(s) is selected by chromatographic experiments; however, this method is time-consuming and laborious.

Recently, molecular modeling methods have been used to identify specific ligands [11–18]. To predict a compound that can bind strongly to the key regions of target molecules (with known three-dimensional structure), computer-aided structure-based design and high throughput docking-based screening are used as fundamental discovery tools [19–24]. A three-step procedure is generally used to design a ligand with high affinity toward the target protein [8, 25–27]. Building a library of several candidate molecules is the first step. Docking these molecules into the pocket of target protein is the second step. This step can be performed using a docking program, such as DOCK [28, 29], FlexX [30, 31], AutoDock [32, 33] and GOLD [34]. In the third step, these molecules are ranked by docking scores [35]; however, the predictive accuracy of the scoring functions of common docking methods is sometimes low. Docking methods and protocols usually consider the proteins as rigid entities and they routinely omit water molecules, or employ only discrete water molecules, and neglect the effect of explicit solvent on the complex structure. Molecular dynamics (MD) simulations resolve both deficiencies, namely solvent effects and the flexibility of protein [36, 37], but require time-consuming computational calculations, limiting applications of such simulations to a few ligand–protein complexes. These issues can be resolved by splitting the procedure into two steps. In the first step, a fast docking method is applied to screen large compound libraries and, in the second step, more valid but time-consuming MD simulations are performed for just a selected ligand [8, 38–40].

This work involves the design of a peptide ligand for HSA by molecular simulations, aiming at the development of a novel affinity chromatography matrix for HSA purification. This rational design includes three steps: (1) building a library of several di- and tri-peptides, (2) docking of these peptides into the pocket of the target protein and using scoring functions to evaluate the affinity of the peptides to the target protein, and (3) investigating the interactions between the selected peptide ligand coupled to a spacer arm and the target protein by MD simulations. We used this procedure to design a dipeptide of high affinity toward HSA. The dipeptide was synthesized and its affinity to HSA was evaluated by UV-vis spectroscopy to ensure that this ligand has enough affinity for the purification of HSA by affinity chromatography.

Materials and methods

Materials

HSA (free fatty acid fraction V, purity >97 %) and warfarin were purchased from Sigma (St. Louis, MO). The dipeptide (H–Trp–Trp–OH) was synthesized (in the solution phase) by the peptide research center at KN Toosi University of Technology [41–44]. Methanol as a solvent of the ligand was purchased from Merck (Darmstadt, Germany). All salts used for buffer preparation were of analytical grade and were dissolved in double distilled water. All experiments were carried out in a 10 mM phosphate buffer at pH 7.2.

Preparation of stock solutions

Stock solutions of HSA (2.4×10^{-4} mol L^{-1}) and warfarin (4.8×10^{-4} mol L^{-1}) were prepared by dissolving them in 10 mM phosphate buffer (pH 7.2). Stock solution of Trp–Trp dipeptide was prepared by dissolving an appropriate amount of the peptide in methanol, and the peptide solution was then diluted with phosphate buffer to a final concentration of 0.64 mM. All solutions used were freshly prepared.

Absorption spectroscopy

Absorption spectra were recorded at room temperature with an Agilent 8,453 single-beam spectrophotometer in the range of 200–400 nm using a quartz cell with 1.0 mm path length. The UV absorption profiles of HSA in the presence and absence of dipeptide ligand solution were recorded at pH 7.2 by keeping the concentration of HSA constant (0.012 mM), while varying the concentration of the ligand from 0.005 to 0.048 mM. The binding constant of the ligand–HSA complex was calculated according to a previously reported method [45, 46]. It was assumed that the interaction between the protein P and the ligand L results in a single complex PL (1:1). The plot of $1/\Delta A$ versus 1/[L] is linear and the binding constant can be estimated from the ratio of intercept to the slope of the line. ΔA is the difference between the initial absorbance of the free HSA at 278 nm and the recorded absorbance at different ligand concentrations, and [L] is the ligand concentration. Binding location study of the dipeptide ligand in the presence of the warfarin site marker was performed by preparing 1:1 (and 1:2) mixtures of warfarin and HSA solutions and then adding Trp–Trp solution to the mixture to achieve 1:1:1 (and 1:2:1) mixtures.

Molecular docking

The crystal structure of HSA (1H9Z) was taken from RCSB protein databank [47] (http://www.pdb.org). The candidate binding pocket (site) of HSA was identified by Discovery Studio 2.5 [48] (Accelrys, San Diego, CA) on the basis of its original ligand (warfarin). The pocket was defined as all residues within 7.7 Å of warfarin. For the preparation step of albumin, the CHARMm force field was applied, hydrogen atoms were added, all water molecules were removed, and the

pH of protein was adjusted to almost neutral, 7.4, using the protein preparation protocol.

Molecular docking simulations were carried out with Discovery Studio 2.5. The GOLD program was used to dock peptides into the protein [49]. GOLD is a genetic algorithm for the docking of flexible ligands into protein binding sites that accounts for the side chain flexibility for the target protein. Peptides were sketched in ChemBioDraw Ultra [50] (http://www.cambridgesoft.com) and the preparation steps were carried out in SYBYL 7.3 molecular modeling package (Tripos, St. Louis, MO) running on a Redhat Linux workstation 4.7. The resulting structures were imported into Discovery Studio, and typed with CHARMm force field and partial charges were calculated using the Momany-Rone option [51]. Charges were then minimized with Smart Minimizer, which performs 1,000 steps of steepest descent with a RMS gradient tolerance of 3, followed by conjugate gradient minimization. Other parameters were set as the default protocol settings. The structure of this protein as presented in PDB format is not complexed with any of the peptides under investigation, so, in the docking step, its original ligand was removed and the designed peptides in the library were then docked sequentially in the active site of HSA. Finally, Goldscore was used to evaluate the affinity of peptide ligands toward the candidate pocket of HSA. The dipeptide with the largest GOLDscore value was selected and the ECH spacer was linked to it. The GOLD program does not have the ability to consider the spacer rigid, so AutoDock vina was used to resolve this deficiency [52–54]. AutoDock is an automated procedure for the docking of flexible ligands to flexible receptors. AutoDock Vina is a new generation of docking software from the Molecular Graphics Laboratory at the Scripps Research Institute (La Jolla, CA). It achieves significant improvements in the average accuracy of the binding mode predictions, while being two orders of magnitude faster than AutoDock 4. Three grid boxes, all centered on the binding site with a grid spacing of 0.375 Å were examined: the larger grid box was set to 60 Å×60 Å×60 Å, the medium box was 40 Å×40 Å×40 Å and the smaller box was set to 26 Å×18 Å×16 Å, which is bigger than the size of the dipeptide. An exhaustiveness option of 8 (average accuracy) was used in the Vina calculation. AutoDockTools [55] was applied for calculation of atomic partial charges by the Gasteiger method.

Molecular surface analysis

Molecular surface analysis was performed with MOLCAD (Tripos) (http://www.tripos.com) to create a Connolly surface [72] around the selected ligand or protein The electrostatic potential energy of this surface was calculated and shown in images of the molecular surface by different colors.

Molecular dynamics simulation

To evaluate the reliability of the docking results and also to investigate the changes of the complex over time, MD simulations were performed with the GROMACS 4.5.3 simulation package [56] using the Amber99sb-ildn force field [57]. The 3D structures and the partial charges of the ligand and spacer were created by the program Chimera 1.8.1 [58]. The molecular topology file was generated using ANTECHAMBER—a tool provided with the AMBER 9 package, a commonly used package for molecular mechanics [59, 60]. There are some missing atoms and residues in the crystal structure of HSA (1H9Z); therefore the pdb file was corrected by MOE 2009.10 [61]. The ligand–protein complex structure was solvated in a cubic box of size 2,219.20 nm^3 with simple point charge (SPC) water molecules [62]. The minimum distance between the complex and the edge of the box was set to 2 nm. In order to neutralize the charge of the system, 16 Na^+ ions were added. An energy minimization step was performed for the full system without constraints using the steepest descent integrator for 50,000 steps, until a tolerance of 10 kJ mol^{-1}. This was followed by a short (200 ps) position restrained equilibration simulation at 300 K. Finally, MD simulations were performed for a period of 20 ns with a time step of 2 fs. The LINCS algorithm (Linear Constraint Solver), which is three to four times faster than the SHAKE algorithm, was used to constrain the length of covalent bonds [63]. The particle-mesh Ewald (PME) summation technique was used to compute long-ranged electrostatic interactions [64]. The Coulomb and van der Waal's cut-offs were set to 1.0 and 1.4 nm, respectively. To build a surface-like structure and to consider this fact that the spacer is covalently attached to the sepharose surface, the end of spacer should be fixed in the simulation box. Thus, positional restraint energy in the form of;

$$E_i^{\text{restraint}} = K(\Delta x_i)^2 \quad (1)$$

was used to atom i, where K is the weight of positional restraint energy (+1.0 kcal mol$Å^{-1}$) and Δx is the difference between the present and reference position of the restrained atom i [65]. The terminal spacer atom that would be covalently attached to the surface and the next two atoms were restrained throughout the whole simulation by the restrained energy mentioned above. The constant temperature and volume (NVT) ensemble at 300 K with periodic boundary conditions was used to perform the simulation. Initial velocities were assigned from Maxwell distribution at 300 K. The Berendsen thermostat was applied to keep the temperature constant.

Results and discussion

Designing a peptide ligand by molecular docking

On the basis of the candidate pocket structure of HSA, only dipeptides can be docked perfectly into the pocket because tripeptides are bigger than the pocket size. A random dipeptide library has 20^2 (= 400) unique dipeptides. This library can be built by linking two amino acids to each other. It is difficult to build the peptide library manually due to the large number of peptides. So, de novo design can be used along with molecular docking method [66] to determine the high-affinity fragments (amino acids) from the pocket structure. These fragments are then linked to form new molecules. These new molecules are docked to the target protein to evaluate their affinity.

The candidate pocket of HSA (warfarin binding site) is composed of two sites: the interior (P1), and the exterior (P2) sites (Fig. 1a). In order to know which amino acid would interact with these sites with high affinity, two dipeptide series were prepared. In the first series, alanine was kept constant and linked to 20 different amino acids. The second series was composed of 19 dipeptides in which alanine was kept constant as the second amino acid. All 39 resulting dipeptides were docked into the candidate pocket of HSA by the GOLD program. The GOLDscore was used to evaluate the affinity of the dipeptides toward HSA. It was concluded that the four amino acids tryptophan (Trp), tyrosine (Tyr), histidine (His), and phenylalanine (Phe) have favorable interactions with site P1 and the three amino acids tryptophan (Trp), tyrosine (Tyr), and phenylalanine (Phe) interact favorably with site P2. These seven amino acids were linked sequentially to form 12 (4×3) dipeptides. All 12 of these dipeptides were docked successfully into the pocket of HSA. Figure 1b shows Trp–Trp dipeptide docked into the pocket of HSA. The values of the GOLDscore obtained are listed in Table 1.

Since the Trp–Trp dipeptide had the largest GOLDscore value, it was selected for subsequent steps. As can be seen from the MOLCAD output, Trp–Trp is shorter than the pocket depth, hinting that the N-terminal of Trp is embedded in the pocket. The decrease in affinity by the attachment of the N-terminus to a solid phase can be avoided by lengthening the peptide and considering a handle arm. Therefore, lysine was linked to the N-terminus of the Trp–Trp dipeptide. This tripeptide ligand should be connected by a spacer arm to the chromatographic matrix. Some studies have shown that hydrophobic spacers are not suitable as they can provide a site for nonspecific adsorption and reduce the specificity of the affinity systems [67, 68]. The shorter or more hydrophilic spacer arms can diminish these undesirable hydrophobic interactions. Busini et al. [69] demonstrated that hydrophilic and rigid spacers provide an extended spacer and solvated ligand, while increasing the spacer length by CH_2 groups makes it more hydrophobic and flexible, which corresponds to adsorption of the ligand onto an agarose matrix. They found that if

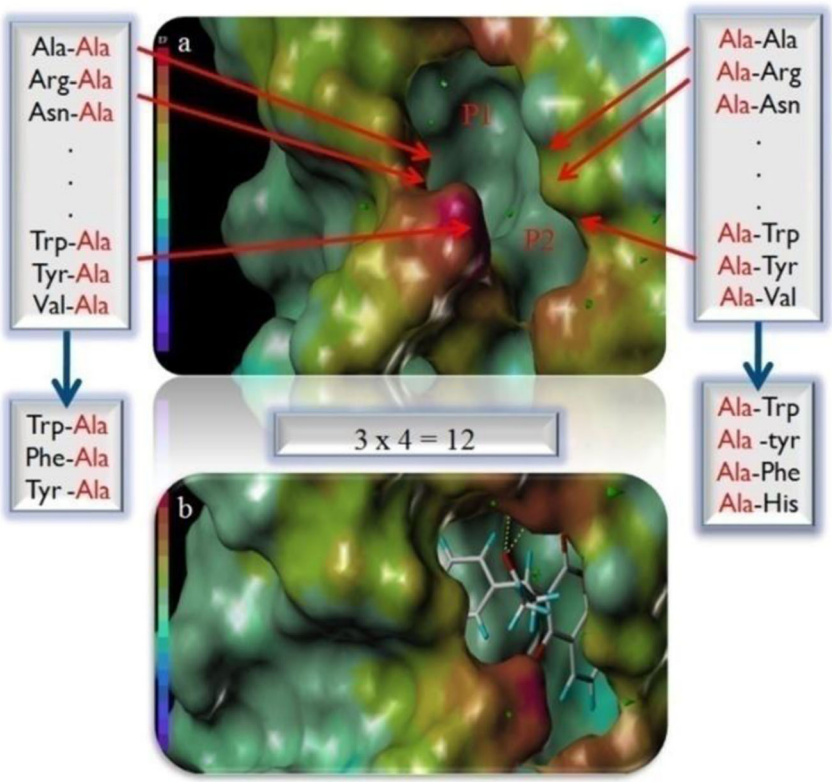

Fig. 1 **a** Structure of the candidate pocket of human serum albumin (HSA) shown as Connolly [72] created with the MOLCAD program [color-coded according to electrostatic potential which ranges from *red* (most positive) to *blue* (most negative)]. All possible dipeptides were tested for docking into the pocket. **b** The best conformation of the Trp–Trp dipeptide within the pocket of HSA. Color coding for the amino acids: *gray* carbon, *red* oxygen, *blue* nitrogen, *green* hydrogen

Table 1 Results of the best dipeptides docked into the candidate pocket of human serum albumin (HSA)

Peptide	Gold score fitness	Peptide	Gold score fitness
Trp-Trp	−70.667	Phe-Phe	−62.425
Trp-Tyr	−63.075	Phe-His	−60.269
Trp-Phe	−65.426	Tyr-Trp	−65.665
Trp-His	−63.169	Tyr-Tyr	−59.256
Phe-Trp	−66.676	Tyr-Phe	−61.530
Phe-Tyr	−60.165	Tyr-His	−59.358

the spacer is long and hydrophilic, multiple conformations are possible.

In this work, lys[$CO(CH_2)_5NH$] was selected as a spacer of sufficient length and hydrophilicity to allow the ligand to reach the pocket of HSA. The N-terminus of the dipeptide was coupled with the carboxylic group of lysine at the end of the spacer arm of lys[$CO(CH_2)_5NH$] of ECH-lysine sepharose 4 gel (Fig. 2). The attaching of the ligand to a solid matrix decreases the flexibility of the ligand and thus results in decreasing the conformational space of the ligand. Thus, the spacer arm will reduce the affinity of the peptidic ligand for HSA. Therefore, it is important to include the spacer in the simulation procedure. Finally, the designed ligand was coupled to the spacer arm of Lys-ECH and docked into the candidate pocket of HSA. The result of the docking process was used for the first guess of the binding structure of the HSA-[Trp-Trp-Lys-$CO(CH_2)_5NH$] complex for subsequent MD simulation.

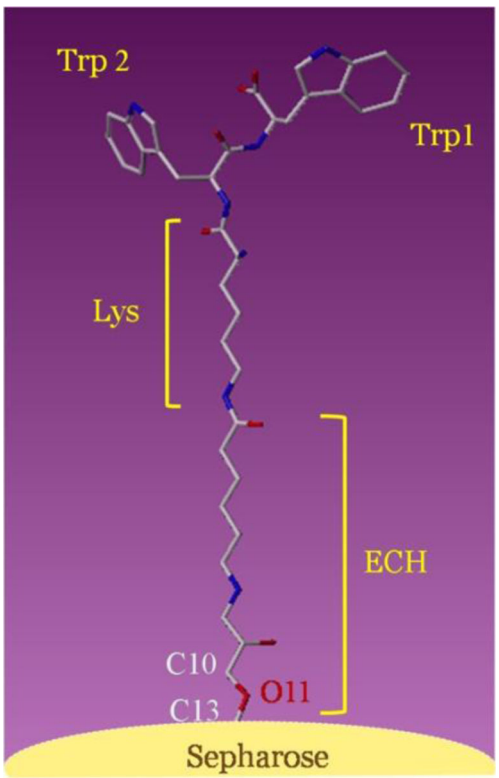

Fig. 2 Hypothesized structure of solid phase

Analyzing protein-ligand interactions

Molecular docking

The main docking procedure was performed with the GOLD program and the Trp–Trp dipeptide with the largest GOLDscore value was selected and the ECH spacer was linked to it. The GOLD program does not have the ability to consider the spacer rigid, so AutoDock vina was used to resolve this deficiency. The aggregate of dipeptide and spacer was thus defined as the ligand. The ECH spacer was considered rigid, while the Trp–Trp–Lys tripeptide was left free to rotate around the sp^3 bonds. This strategy was followed because if the ECH group is left flexible, the spacer takes the orientation toward the interior parts of the protein rather than going out of the protein surface for immobilization on a solid matrix. Three grid boxes, all centered on the binding site, were examined. The larger grid box was set to 60 Å×60 Å×60 Å, the medium box was 40 Å×40 Å×40 Å and the smaller box was set to 26 Å×18 Å×16 Å, which is bigger than the size of the Trp–Trp dipeptide. In the case of bigger grid boxes, the program considers the whole [spacer–Trp–Trp] as a ligand and docks the center of [spacer–Trp–Trp] in the pocket instead of docking Trp–Trp in the pocket and the spacer was oriented toward the interior parts of the protein rather than going out of the protein surface. Therefore the smaller grid box was selected for furthur investigation. The docking energies of the best poses of the smaller box are reported in Table 2; it can be noted that, in the first and second poses, the spacer is oriented toward the inner part of the protein while it needs to be turned toward the solvent. Furthermore, poses in which the Trp–Trp ligand is located near the candidate pocket are selected. In the first and second poses, the spacer is located in the candidate pocket. For this reason, first and second poses were excluded and the third pose was selected as the first guess of binding structures for the following MD runs (Fig. 3).

The designed Trp–Trp dipeptide shows specific interactions with HSA and can bind specifically to HSA binding pocket residues. Figure 4 shows that the hydroxyl and amino groups of the synthetic ligand can form four hydrogen (H)-

Table 2 Docking energies (kcal mol^{-1}) of HSA-[Trp-Trp-Lys-$CO(CH_2)_5NH$] complex evaluated by AutoDock Vina

Pose	Affinity
1	−9.6
2	−8.2
3	−7.6
4	−7.4
5	−7.3
6	−6.8

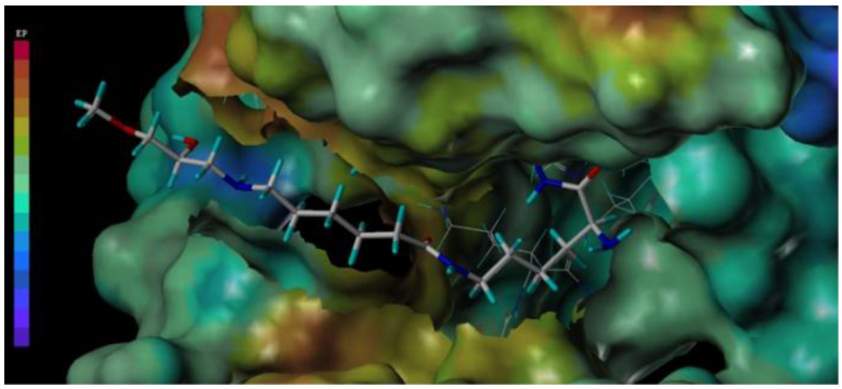

Fig. 3 The designed ligand [Trp–Trp–Lys–$CO(CH_2)_5NH$] docked into the candidate pocket of HSA by AutoDock Vina. The surface of protein is shown as Connolly created with the MOLCAD program (color-coded according to electrostatic potential as in Fig. 1)

bonds with Arg 257 of the HSA pocket, while the indole functional groups of Trp–Trp ligand have π-σ and π-cation interaction with Leu 238, Arg 222 and 218, respectively.

Affinity analysis using MOLCAD

A molecular surface analysis program was used to analyze the interactions between the ligand and protein. Figures 1 and 2 show the MOLCAD surface structures representing the electrostatic potential of HSA. The electrostatic potential color ramp ranges from blue (most electronegative potential values) to red (most electropositive potential values). These figures indicate that the pocket of HSA is a hydrophobic pocket (indicated in green), so hydrophobic amino acids with phenyl groups show high affinity toward the binding site. Therefore, the Trp–Trp dipeptide with two indole group is a good suggestion as a ligand for this pocket as it can make π–π interactions with hydrophobic residues of the pocket.

Molecular dynamics simulation

The ligand–HSA complex was further analyzed by MD simulations in order to evaluate the reliability of the docking results. The first guess at the binding structure of the HSA–[Trp–Trp–Lys–$CO(CH_2)_5NH$] complex structures was obtained from docking simulations. This structure was simulated for 20 ns using the GROMACS MD package. As discussed in Materials and methods, the above described restrained energy was applied to atoms C10, O11 and C13 of the spacer through the whole simulation (Fig. 2).

The root mean square deviation (RMSD) of the trajectories from their initial structures was used to examine whether the structure of the ligand–HSA complex remained stable under the simulation conditions or not. The RMSD values of the protein backbone and ligand ranged from 0.13 to 0.35 nm and 0.09 to 0.18 nm, respectively, as shown in Fig. 5a. The RMSD of the ligand reached about 0.13 nm after 1,500 ps, and was maintained at almost this level throughout the simulation process. This indicates that the docked complex structure is stable throughout the MD simulation. However, the whole

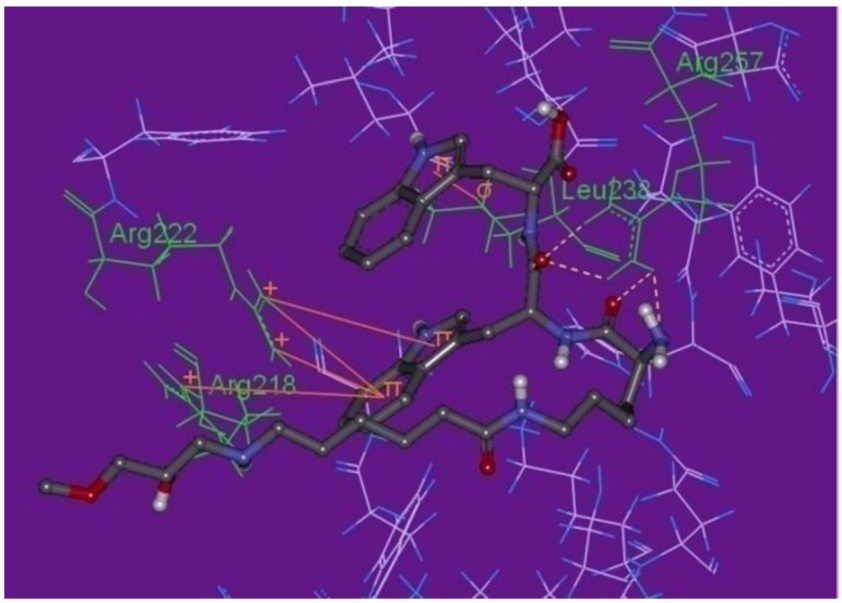

Fig. 4 The best docked ligand in the pocket of HSA results in four hydrogen (H)-bonds and three π-σ and π-cation interactions. Only residues of HSA close to the pocket are shown. *Line drawing* Protein residues *3D representation* ligand, *dotted yellow lines* H-bonds, *orange lines* π interactions

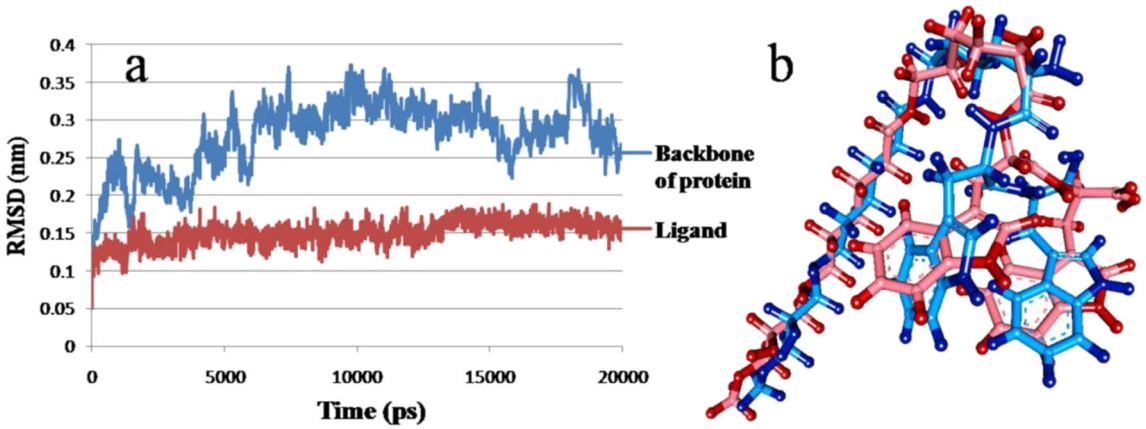

Fig. 5a,b Molecular dynamics (MD) simulation results. **a** The root mean square deviation (RMSD) of the protein (*blue*) and docked ligand (*red*) versus MD simulation time (picoseconds). **b** View of superimposed backbone atoms of the average structure of the last 200 ps of the MD simulation (*blue*) and the initial structure (*red*)

protein has a larger RMSD value of about 0.15 nm. This demonstrates that the structure of the protein shows much greater fluctuations, while the ligand exhibits smaller conformational changes. These results are in parallel with the high affinity of the ligand and the residues of the protein forming a stable complex with low flexibility.The superimposition of the average structures of the ensemble for the last 200 ps of trajectory and the docked structure is shown in Fig. 5b.

H-bond interactions between the ligand and the protein are shown in Fig. 6. Lots of H-bond interactions are identified in MD simulations, and are represented in Table 3. It can be seen that some of these H-bonds are unstable. The movements of the ligand and protein atoms are responsible for the existence of these transient H-bonds. These labile H-bonds may make a small contribution to the ligand–protein affinity. It can be observed from Fig. 6 that three H-bonds exist during the most of the simulation time and that these are very stable. These H-bonds play a key role in ligand–protein binding, and are formed between the carbonyl groups (Trp and Lys residues) of the ligand and the NH group of Arg257 as well as the hydrogen atom of the amino group of Gln196 and the oxygen atom of the ECH spacer. These stable H-bonds are also observed in the docking simulation.

Water molecules are usually removed before docking [56]. However, in the solvation and desolvation processes, water molecules are the main portion of the simulation that can affect ligand–protein binding and can bridge between the ligand and protein [70]. The final snapshot was used to study the important effect of water molecules. Figure 7 shows four water-mediated H-bonds. Three H-bonds are formed between the two nitrogen atoms of the side chain of Arg257 and the oxygen atom of the carbonyl group of Trp by mediation of the SOL4678 water molecule. The SOL41water molecule gives a H-bond donor to the oxygen atom of the carboxyl group of Leu234 and accepts the H-bond atom from the nitrogen of the indole group of Trp of the ligand. In addition, the SOL4428 water molecule bridges the side chain amino group of Lys199 and the carbonyl group of Lys of the peptide ligand by H-bonds. Finally, SOL2595 bridges between the carboxyl group of Asp451 and the nitrogen of the ECH spacer. It can be noted

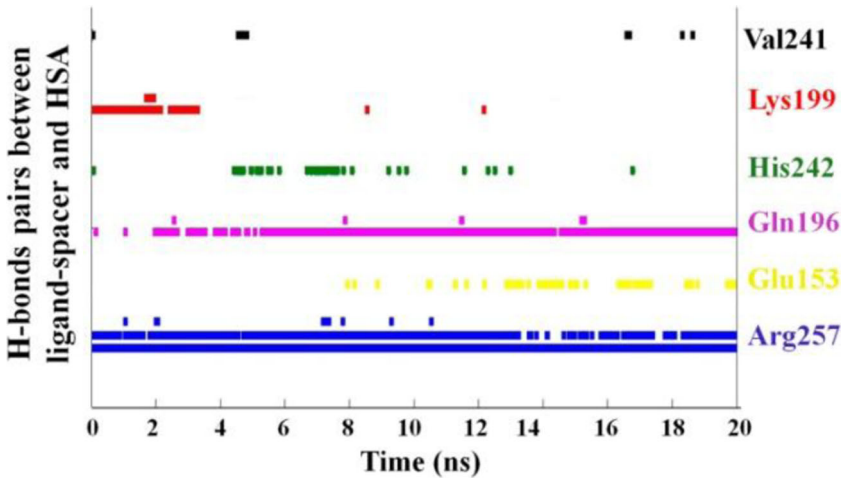

Fig. 6 H-bond interactions between the ligand [Trp–Trp–Lys–$CO(CH_2)_5NH$] and the candidate pocket of HSA observed in the 20 ns MD simulations. The presence of H-bond interactions at any time is represented by a colored mark according to protein residue

Table 3 Residues and atom types involved in intermolecular H-bonds illustrated from the 10–20 ns snapshot

Ligand residue	Atom type	Receptor residue	Atom type
Trp 1	O of the hydroxyl group of backbone	Arg 257	H of the NH1 group of side chain
Trp 1	O of the C = O group of backbone	Arg 257	H of the NE group of side chain
Trp 2	O of the C = O group of backbone	Arg 257	H of the NH1 group of side chain
Trp 2	N of the indole group of side chain	Lys 199	H of the NH group of side chain
Trp 2	H of the NH_2 group of backbone	His 242	N of the imidazole group of side chain
Lys	N of the NH_2 group of backbone	Arg 257	H of the NH1 group of side chain
Lys	O of the C = O group of backbone	Arg 257	H of the NH1 group of side chain
Lys	H of the NH_2 group of backbone	Val 241	O of the C = O group of backbone
Lys	H of the NH group of side chain	Glu 153	O of the C = O group of side chain
Lys	H of the NH group of side chain	Glu 153	O of the hydroxyl group of side chain
ECH	O of the C = O group	Gln 196	H of the NH_2 group of side chain

that these water-mediated H-bonds are stable during the entire simulation and play an important role in stabilization of the protein–ligand complex.

The warfarin-binding site is a hydrophobic pocket, therefore hydrophobic interactions could be an important interaction in the association of ligand and the pocket of HSA. The most relevant hydrophobic interaction can be observed between the hydrophobic residues of the binding site (Ala291, Leu238, and Phe223) and the Trp–Trp dipeptide. The minimum distance between these hydrophobic residues and the Trp–Trp dipeptide was studied by MD simulations; no notable changes were observed in the distances between these residues and the dipeptide during the last 10 ns of MD simulation (Fig. 8). However, compared with Ala291 and Leu238, Phe223 has a larger distance of about 0.1 nm from the dipeptide and also fluctuates more widely, demonstrating that the hydrophobic Ala291-dipeptide and Leu238-dipeptide interactions are important and form a stable complex with only small conformational changes.

Absorption spectroscopy

The Trp–Trp dipeptide was designed by means of docking and MD simulations. To examine whether the designed ligand has sufficient affinity to HSA or not, an absorption spectroscopy study was carried out. Figure 9a shows that an increase in the dipeptide concentration resulted in an increase in absorption at 278 nm that can be related to complex formation. The double reciprocal plot of $1/(A - A_0)$ versus 1/(ligand concentration) is linear and the binding constant (K) can be estimated from the

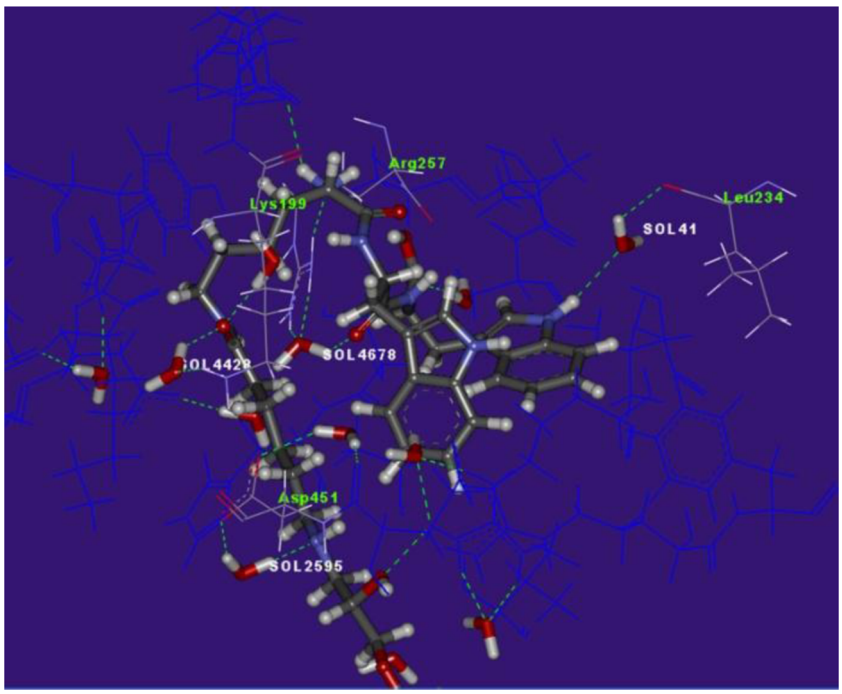

Fig. 7 Snapshot of H-bond interactions between the protein as well as ligand and four water molecules (SOL41, SOL4678, SOL2595 and SOL4428) found near the candidate pocket of HSA observed in the final conformation of the 20 ns MD simulations. Residues of the candidate pocket that interact with the water molecules are labeled (Lys199, Arg257, Leu234 and Asp451). *Line drawing* Protein residues, *3D representation* ligand and water molecules, *dotted-green lines* H-bonds. Atoms: *red* oxygen, *blue* nitrogen, *white* hydrogen, *gray* carbon. The structures were plotted using VMD software (http://www.ks.uiuc.edu/Research/vmd/)

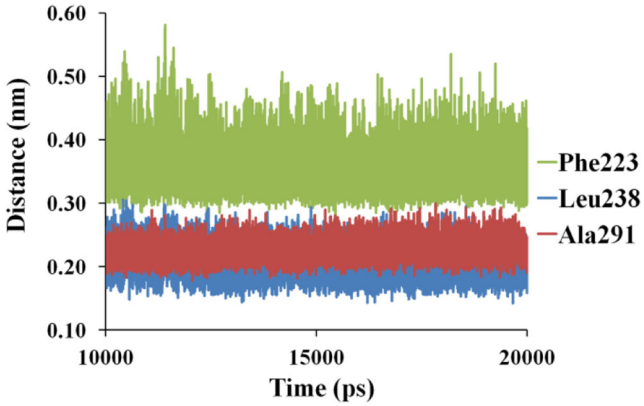

Fig. 8 Changes in distances of the most relevant hydrophobic interactions between hydrophobic residues of the binding site [Ala291 (*red*), Leu238 (*blue*), and Phe223 (*green*)] and the Trp–Trp dipeptide during the last 10 ns of MD simulation

ratio of the intercept to the slope as described above (Fig. 9b). A_0 is the initial absorbance of the free HSA at 278 nm and A is the measured absorbance at different ligand concentrations. A binding constant of $3.01(\pm 0.09)\times 10^3$ M^{-1} was determined from this experiment. The relationship between the binding constant, $K_{binding}$, and the binding free energy between protein and ligand, $\Delta G_{binding}$, is

$$\Delta G_{binding} = -RT \ln K_{binding} \quad (2)$$

where R is the gas constant, 1.987 cal K^{-1} mol^{-1}, and T is the absolute temperature, 298.15 K. An experimental binding free energy between protein and ligand of −4.7 kcal mol^{-1} was obtained, confirming that this ligand has enough affinity for the purification of HSA by affinity chromatography.

HSA has two major and selective binding sites: I and II. Warfarin is known as the site I marker [71]. In order to investigate whether the Trp–Trp dipeptide, which was designed for the warfarin binding site (site I), interacts with the target site, binding location studies between the Trp–Trp dipeptide and HSA in the presence of the warfarin site marker were performed. As can be seen from Fig. 10, warfarin shows a band at approximately 309 nm. Upon formation of the HSA–warfarin complex, the band intensity decreased slightly due to the decrease in free warfarin concentration. Conversely, when adding Trp–Trp dipeptide to the HSA–warfarin

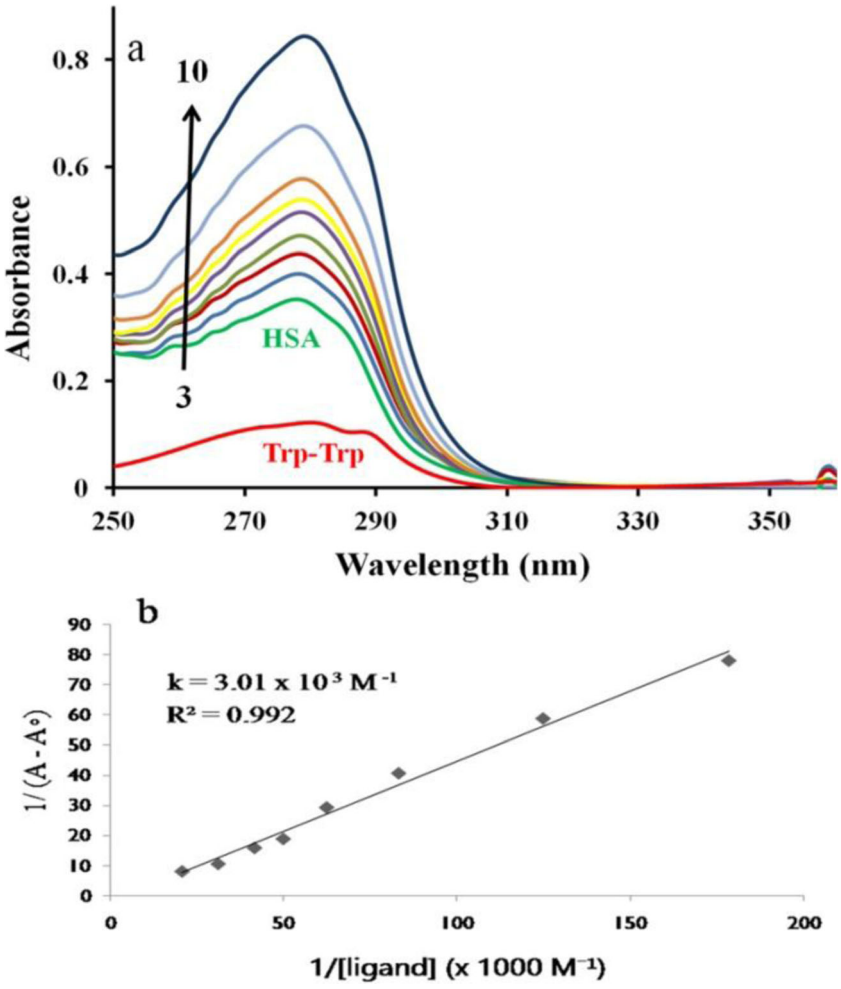

Fig. 9a,b Absorption spectra of free HSA and its Trp–Trp dipeptide complexes in aqueous solution. **a** Spectra of *1* free dipeptide (0.012 mM); *2* free HSA (0.012 mM); *3–10* dipeptide–HSA complexes: *3* 0.005 mM, *4* 0.008 mM, *5* 0.012 mM, *6* 0.016 mM, *7* 0.02 mM, *8* 0.024 mM, *9* 0.032 mM and *10* 0.048 mM. **b** $1/(A - A_0)$ versus (1/ligand concentration), where A_0 is the initial absorbance of HSA (278 nm) and A is the measured absorbance at different ligand concentrations (0.005–0.048 mM) with constant HSA concentration of 0.012 mM at pH 7.2

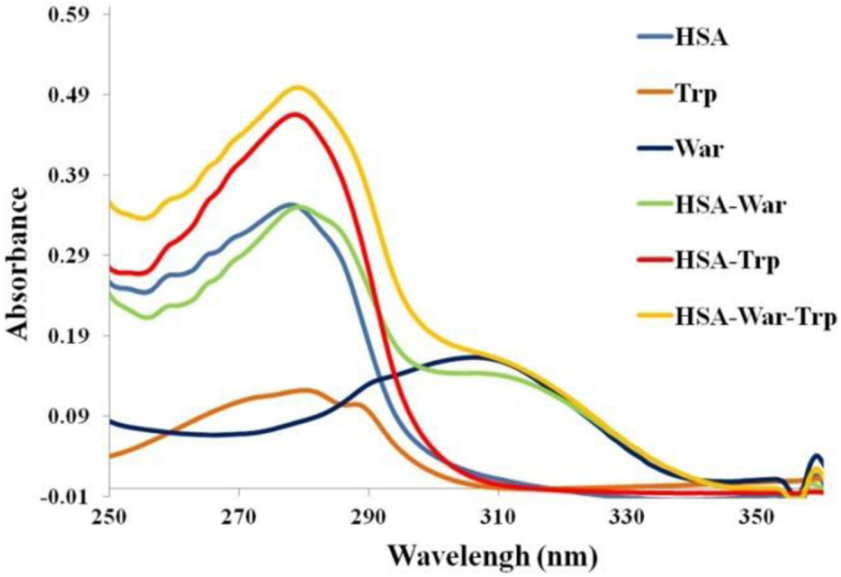

Fig. 10 Absorption spectra of free HSA (0.012 mM), warfarin (War; 0.012 mM), Trp–Trp dipeptide (0.012 mM) and their 1:1 complexes in aqueous solution. Spectra of free HSA (*blue*), free dipeptide (*brown*), free War (*indigo*), 1:1 HSA:War solution (*green*), 1:1 HSA:dipeptide solution (red), and 1:1:1 HSA:War:dipeptide solution (*yellow*) are shown

complex, the observed 309 nm band remains unchanged. This could verify the specific interaction between the dipeptide ligand and HSA by releasing warfarin and consequently increasing its concentration in the solution. It is clear that the binding of warfarin to HSA is responsible for the shift of the HSA band with λ_{max} at 278 nm towards a higher wavelength (i.e., red shift) of the spectrum. No similar effect was observed after addition of the Trp–Trp dipeptide to HSA-warfarin complex solution, suggesting that the dipeptide ligand can displace warfarin in a competitive manner to bind to the HSA site.

The absorption band at 278 nm is due mainly to the presence of the Trp214 residue of HSA. Basically, although

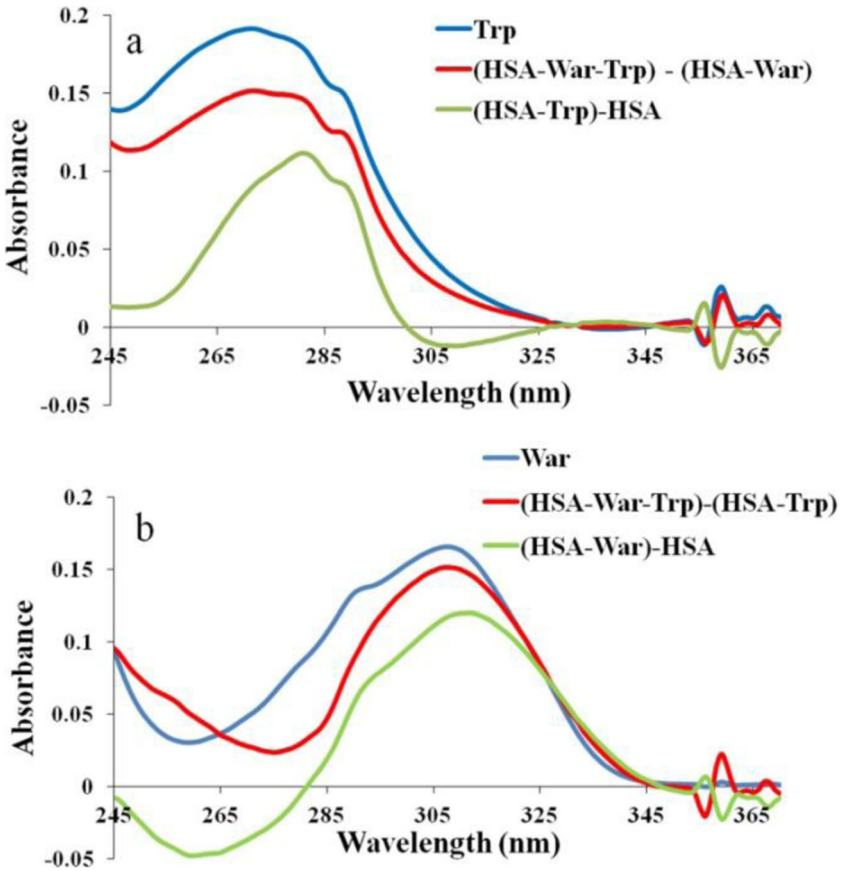

Fig. 11 Absorption spectra of **a** free Trp-Trp dipeptide (0.012 mM), **b** free warfarin (War, 0.012 mM) and their HSA complexes with difference spectra obtained at the concentration of 0.012 mM for all components at pH 7.2

there is no significant spectral shifting upon interaction between the dipeptide ligand and HSA, its intensity exhibits considerable variations. To identify and investigate these variations, difference spectra were obtained as follows. In Fig. 11a, the spectra of the 1:1:1 mixture of HSA:warfarin:Trp–Trp was subtracted from the 1:1 HSA-warfarin complex to certify the existence of the free dipeptide in solution. The difference spectrum has lower intensity than the dipeptide solution spectrum with the same concentration. This observation indicates that some of dipeptides can interact with HSA in the presence of warfarin. The difference spectrum between the 1:1 mixture of HSA:dipeptide and HSA solution was smaller in intensity than the two spectra of Trp–Trp and HSA:War:Trp–Trp. This result suggests a reduction in free dipeptide concentration in the solution due to the interaction of dipeptide with protein. As indicated in Fig. 11b, difference spectra that show the presence of free warfarin in solution can be represented. It can be seen that the interaction between warfarin and HSA decreases in the presence of the Trp–Trp dipeptide in solution. However, if the concentration of warfarin is increased two-fold (1:2:1 HSA:warfarin:dipeptide mixture), the dipeptide no longer has any chance to compete with warfarin to interact with the binding site. These results are in agreement with the higher binding constant of warfarin and HSA.

The difference spectra in Fig. 12 are the result of two consecutive subtractions or a double subtraction spectrum. In the first step, the dipeptide spectrum is subtracted from 1:1:1 HSA:warfarin:dipeptide spectrum and in the second step the warfarin spectrum is subtracted from the difference spectrum resulting from the first step (the concentration of all components is 0.012 mM). The double subtraction spectrum shows a considerable intensity in the range of 300–350 nm with regard to the HSA spectrum. This can be assigned to the presence of warfarin in the binding site of HSA. This observation can be considered as probable confirmation that the dipeptide ligand can more or less separate HSA from the HSA–warfarin complex or any HSA–drug complex. This finding must be proved by further computational and experimental studies.

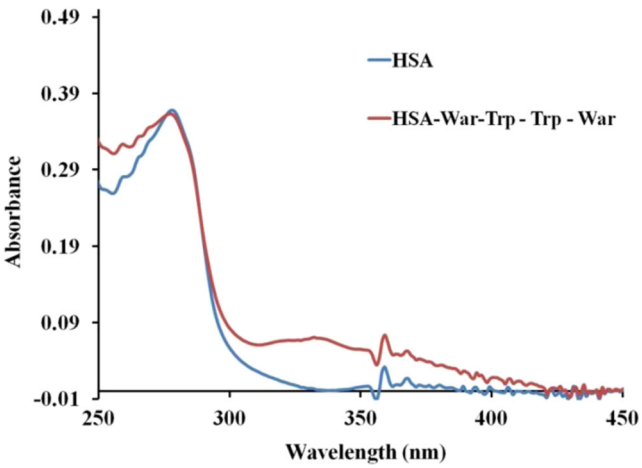

Fig. 12 Absorption spectra of free HSA (0.012 mM, *blue*) and the difference spectrum between the 1:1:1 HSA:War:dipeptide solution and dipeptide solution and then subtracted again from warfarin solution obtained at the concentration of 0.012 mM (*red*) for all components at pH 7.2

The results of these spectroscopic studies show that the designed ligand has enough affinity for purification of HSA by affinity chromatography. However, the dipeptide ligand needs to be attached to ECH-lysine sepharose gel to see whether or not this solid phase can absorb the HSA efficiently and specifically from the serum. Furthermore, the binding mechanism of the peptide ligand and HSA should be studied in more detail in subsequent research.

Conclusions

This article introduces a computational approach to the design of a dipeptide ligand for HSA purification by exploiting structure-based docking and MD simulation. In the first step, a peptide library of 39 dipeptides was built. Afterwards, these peptides were docked into the candidate pocket (warfarin binding site) of HSA. It was found that four amino acids have high affinity with the interior site and three amino acids show high affinity interactions with the exterior site. Hence, a second dipeptide library was built by linking these amino acids sequentially. After docking simulation, a Trp–Trp dipeptide with the highest GOLDscore value was selected. In the third step, this dipeptide was connected to a spacer arm of ECH-lysine sepharose gel in order to simulate the chromatographic separation process. Afterwards, the ligand-spacer set was docked into the candidate pocket of HSA to obtain a first guess binding structure of the HSA–[Trp–Trp–Lys–$CO(CH_2)_5NH$] complex structure for MD simulation. To identify the interactions between the designed ligand and HSA, 20-ns MD simulations were performed. It was found that water molecules play an important role in stabilization of the protein–ligand complex by bridging the ligand and the protein by H-bonds. In addition, all the H-bonds found in the docking were confirmed by MD simulations. Finally, high affinity of the dipeptide ligand toward HSA was examined by spectroscopic analysis. It was affirmed that the designed dipeptide can bind specifically to the warfarin binding site. The present work has demonstrated that a combination of docking and MD simulations is a reliable strategy for the rational design of new affinity ligands on the basis of protein structure.

References

1. Brgles M, Clifton J, Walsh R, Huang F, Rucevic M, Cao L, Hixson D, Müller E, Dj J (2011) Selectivity of monolithic supports under overloading conditions and their use for separation of human plasma and isolation of low abundance proteins. J Chromatogr A 1218: 2389–2395
2. Nice EC, Rothacker J, Weinstock J, Lim L, Catimel B (2007) Use of multidimensional separation protocols for the purification of trace components in complex biological samples for proteomics analysis. J Chromatogr A 1168:190–210
3. Anderson NL, Anderson NG (2002) The human plasma proteome. Mol Cell Proteomics 1:845–867
4. Issaq HJ (2001) The role of separation science in proteomics research. Electrophoresis 22:3629–3638
5. Zhu G, Zhao P, Deng N, Tao D, Sun L, Liang Z, Zhang L, Zhang Y (2012) Single chain variable fragment displaying M13 phage library functionalized magnetic microsphere-based protein equalizer for human serum protein analysis. Anal Chem 84: 7633–7637
6. Anderson LN, Polanski M, Pieper R et al (2004) The human plasma proteome. Mol Cell Proteomics 3:311–326
7. Ray S, Reddy PJ, Jain R, Gollapalli K, Moiyadi A, Srivastava S (2011) Proteomic technologies for the identification of disease biomarkers in serum: advances and challenges ahead. Proteomics 11: 2139–2161. doi:10.1002/pmic.201000460
8. Liu FF, Dong XY, Wang T, Sun Y (2007) Rational design of peptide ligand for affinity chromatography of tissue-type plasminogen activator by the combination of docking and molecular dynamics simulations. J Chromatogr A 1175:249–258
9. Ren D, Penner NA, Slentz BE et al (2004) Contributions of commercial sorbents to the selectivity in immobilized metal affinity chromatography with Cu(II). J Chromatogr A 1031:87–92
10. Affinity chromatography Handbook; principles and methods. GE Healthcare, product code 18-1022-29.
11. Qiao Y, Li P, Chen Y, Feng J, Wang J, Wang W, Ma Y, Sun P, Yuan Z (2010) Design, optimization and evaluation of specific affinity adsorbent for oligopeptides. J Chromatogr A 1217:7539–7546
12. Qiao Y, Zhao J, Li P, Wang J, Feng J, Wang W, Sun H, Ma Y, Yuan Z (2010) Adsorbents with high selectivity for uremic middle molecular peptides containing the Asp-Phe-Leu-Ala-Glu sequence. Langmuir 26:7181–7187
13. Labrou NE, Eliopoulos E, Clonis YD (1999) Molecular modeling for the design of biomimetic chimeric ligand. Aplication to the purification of bovine heart L-lactate dehydrogenase. Biotechnol Bioeng 63: 322–332
14. Liu FF, Wang T, Dong XY, Sun Y (2007) Rational design of affinity peptide ligand by flexible docking simulation. J Chromatogr A 1146: 41–50
15. Platis D, Sotriffer CA, Clonis Y, Labrou NE (2006) Rational design of affinity peptide ligand by flexible docking simulation. J Chromatogr A 1128:138–151
16. Ghasemi JB, Tavakoli H (2012) Improvement of the prediction power of the CoMFA and CoMSIA models on Histamine H3 antagonists by different variable selection methods. Sci Pharm 80:547–566. doi:10.3797/scipharm.1204-19
17. Ghasemi JB, Meftahi N, Pirhadi S, Tavakoli H (2013) Docking and pharmacophore-based alignment comparative molecular field analysis three-dimensional quantitative structure–activity relationship analysis of dihydrofolate reductase inhibitors by linear and nonlinear calibration methods. J Chemometr 27:287–296
18. Ghasemi JB, Aghaee E, Jabbari A (2013) Docking, CoMFA and CoMSIA studies of a series of N-benzoylated phenoxazines and phenothiazines derivatives as antiproliferative agents. Bull Korean Chem Soc 34:899–906
19. Budin N, Majeux N, Tenette-Souaille C, Caflisch A (2001) Structure-based ligand design by a build-up approach and genetic algorithm search in conformational space. J Comput Chem 22:1956–1970
20. Pirhadi S, Shiri F, Ghasemi JB (2013) Methods and applications of structure based pharmacophores in drug discovery. Curr Top Med Chem 13:1036–1047
21. Ardakani A, Ghasemi JB (2013) Identification of novel inhibitors of HIV-1 integrase using pharmacophore-based virtual screening combined with molecular docking strategies. Med Chem Res 22:5545–5556
22. Warren GL, Andrews CW, Capelli AM et al (2006) A critical assessment of docking programs and scoring functions. J Med Chem 49: 5912–5931
23. Sousa SF, Fernandes PA, Ramos MJ (2006) Protein-ligand docking: current status and future challenges. Proteins 65:15–26
24. Rupasinghe CN, Spaller MR (2006) The interplay between structure-based design and combinatorial chemistry. Curr Opin Chem Biol 10: 188–193
25. Kroemer RT (2007) Structure-based drug design: docking and scoring. Curr Protein Pept Sc 8:312–328
26. Hunter WN (2009) Structure-based ligand design and the promise held for antiprotozoan drug discovery. J Biol Chem 284:11749–11753
27. Stahl M, Rarey M (2001) Detailed analysis of scoring functions for virtual screening. J Med Chem 44:1035–1042
28. Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE (1982) A geometric approach to macromolecule-ligand interactions. J Mol Biol 161:269–288
29. Ewing TJA, Kuntz ID (1997) Critical evaluation of search algorithms for automated molecular docking and database screening. J Comput Chem 18:1176–1189
30. Rarey M, Kramer B, Lengauer T, Klebe G (1996) A fast flexible docking method using an incremental construction algorithm. J Mol Biol 261:470–489
31. Rarey M, Kramer B, Lengauer T (1999) The particle concept: placing discrete water molecules during protein-ligand docking predictions. Proteins 34:17–28
32. Morris GM, Goodsell DS, Halliday RS et al (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J Comp Chem 19:1639–1662
33. Goodsell DS, Morris GM, Olson AJ (1996) Automated docking of flexible ligands: applications of autodock. J Mol Recognit 9:1–5
34. Jones G, Willett P, Glen RC, Leach AR, Taylor R (1997) Development and validation of a genetic algorithm for flexible docking. J Mol Biol 267:727–748
35. Kitchen DB, Decornez H, Furr JR, Bajorath J (2004) Docking and scoring in virtual screening for drug discovery: methods and applications. Nat Rev Drug Discov 3:935–949. doi:10.1038/nrd1549
36. Alonso H, Bliznyuk AA, Gready JE (2006) Combining docking and molecular dynamic simulations in drug design. Med Res Rev 26: 531–568
37. Ghasemi JB, Hooshmand S (2013) 3D-QSAR, docking and molecular dynamics for factor Xa inhibitors as anticoagulant agents. Mol Simul 39:453–471
38. Manetti F, Locatelli GA, Maga G et al (2006) A combination of docking/dynamics simulations and pharmacophoric modeling to discover new dual c-Src/Abl kinase inhibitors. J Med Chem 49:3278–3286
39. Li L, Wei DQ, Wang JF, Chou KC (2007) Computational studies of the binding mechanism of calmodulin with chrysin. Biochem Biophys Res Commun 358:1102–1107
40. Kesner TF, Elcock AH (2006) Computational sampling of a cryptic drug binding site in a protein receptor: explicit solvent molecular dynamics and inhibitor docking to p38 map kinase. J Mol Biol 359: 202–214

41. Arabanian A, Mohammadnejad M, Balalaie S, Gross JH (2009) Synthesis of novel Gn-RH analogues using Ugi-4MCR. Bioorg Med Chem Lett 19:887–890
42. Arabanian A, Mohammadnejad M, Balalaie S (2010) A novel and efficient approach for the amidation of C-terminal peptides. J Iran Chem Soc 7:840–845
43. Rezaee Z, Arabanian A, Balalaie S, Ahmadiani A, Nasoohi S (2012) Semicarbazide substitution enhances enkephalins resistance to ace induced hydrolysis. Int J Pept Res Ther 18:305
44. Tahoori F, Sheikhnejad R, Balalaie S, Sadjadi M (2013) Synthesis of novel peptides through Ugi-ligation and their anti-cancer activities. Amino Acids 45:975–981
45. Abdi K, Nafisi S, Manouchehri F, Bonsaii M, Khalaj A (2012) Interaction of 5-Fluorouracil and its derivatives with bovine serum albumin. J Photochem Photobiol B 107:20–26
46. Fani N, Bordbar AK, Ghayeb Y (2013) Spectroscopic, docking and molecular dynamics simulation studies on the interaction of two Schiff base complexes with human serum albumin. J Lumin 141: 166–172
47. Petitpas I, Bhattacharya AA, Twine S, East M, Curry S (2001) Crystal structure analysis of warfarin binding to human serum albumin. J Biol Chem 276:22804–22809
48. Discovery Studio. Accelrys software Inc, San Diego, CA, 2009. http://www.accelrys.com.
49. Politi A, Durdagi S, Moutevelis-Minakakis P, Kokotos G, Mavromoustakos T (2010) Development of accurate binding affinity predictions of novel rennin inhibitors through molecular docking studies. J Mol Graph Model 29:425–435
50. Kerwin SM (2010) ChemBioOffice Ultra 2010 suite. J Am Chem Soc 132:2466–2467
51. Momany FA, Rone RJ (1992) Validation of the general purpose QUANTA 3.2/CHARMm force field. J Comput Chem 13:888–900
52. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. J Comput Chem 31:455–461
53. Cosconati S, Forli S, Perryman AL, Harris R, Goodsell DS, Olson AJ (2010) Virtual screening with AutoDock: theory and practice. Expert Opin Drug Discov 5:597–607
54. Forli S, Olson AJ (2012) A force field with discrete displaceable waters and desolvations entropy for hydrated ligand docking. J Med Chem 55:623–638
55. Sanner MF, Huey R, Dallakyan S, Karnati S et al (2007) AutoDockTools, Version 1.4.5, The Scripps Research Institute, La Jolla, CA
56. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC, GROMACS: fast, flexible, and free. J Comput Chem 26:1701–1718.
57. Lindorff-Larsen K, Piana S, Palmo K et al (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins 78:1950–1958
58. Pettersen EF, Goddard TD, Huang CC et al (2004) UCSF Chimera–a visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612
59. Pearlman DA, Case DA, Caldwell JW et al (1995) Amber, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. Comput Phys Commun 91:1–41
60. Case DA, Cheatham TE, Darden T et al (2005) The amber biomolecular simulation programs. J Comput Chem 26:1668–1688
61. Molecular Operating Environment (MOE) from chemical computing group Inc., Montreal, QC, Canada, 2013
62. Berendsen HIC, Postma JPM, Van Gunsteren W, Hermans J (1981) Interaction models for water in relation to protein hydration. In: Pullman B (ed) Intermolecular forces. Reidel, Dordrecht, pp 331–342
63. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) LINCS: a linear constraint solver for molecular simulations. J Comput Chem 18:1463–1472
64. Darden T, York D, Pedersen L (1993) Particle Mesh Ewald-an N.Log(N) method for Ewald sums in large systems. J Chem Phys 98:10089–10092
65. Dismer F, Hubbuch J (2010) 3D structure-based protein retention prediction for ion-exchange chromatography. J Chromatogr A 1217: 1343–1353
66. Honma T (2003) Recent advances in de novo design strategy for practical lead identification. Med Res Rev 23:606–632
67. O'Carra P, Barry S, Griffin T (1974) Spacer arms in affinity chromatography: use of hydrophilic arms to control or eliminate non-biospecific adsorption effects. FEBS Lett 43:169–175
68. Lowe CR (1977) The synthesis of several 8-substituted derivatives of adenosine 5′-monophosphate to study the effect of the nature of the spacer arm in affinity chromatography. Eur J Biochem 73:265–274. doi:10.1111/j.1432-1033.1977.tb11316.x
69. Busini V, Moiani D, Moscatelli D, Zamolo L, Cavallotti C (2006) Investigation of the influence of spacer arm on the structural evolution of affinity ligands supported on agarose. J Phys Chem B 110: 23564–23577
70. Levy Y, Onuchic JN (2006) Water mediation in protein folding and molecular recognition. Annu Rev Biophys Biomol Struct 35:389–415. doi:10.1146/annurev.biophys.35.040405.102134
71. Ni Y, Zhang X, Kokot S (2009) Spectrometric and voltammetric studies of the interaction between quercetin and bovine serum albumin using warfarin as site marker with the aid of chemometrics. Spectrochim Acta A 71:1865–1872
72. Connolly MJ (1983) J Appl Cryst 16:548–558