J Mol Model (2013) 19:3749–3766
DOI 10.1007/s00894-013-1900-2

ORIGINAL PAPER

Effects of organic solvents and substrate binding on trypsin in acetonitrile and hexane media

Yanyan Meng · Yuan Yuan · Yanyan Zhu · Yanzhi Guo · Menglong Li · Zhimeng Wang · Xuemei Pu · Lin Jiang

Received: 1 March 2013 / Accepted: 27 May 2013 / Published online: 23 June 2013

Abstract In this work, we used molecular dynamic (MD) simulation to study trypsin with and without a six-amino-acid peptide bound in three different solvents (water, acetonitrile and hexane) in order to provide molecular information for well understanding the structure and function of enzymes in non-aqueous media. The results show that the enzyme is more compact and less native-like in hexane than in the other two polar solvents. The substrate could stabilize the native protein structure in the two polar media, but not in the non-polar hexane. There are no significant differences in the conformation of the S1 pocket upon the substrate binding in water and acetonitrile media while a reverse behavior is observed in hexane media, implying a possible induced fit binding mechanism in the non-polar media. The substrate binding enhances the stability of catalytic H-bond network since it could expel the solvent molecules from the active site. The enzyme and the substrate appear to be more appropriate to the reactive conformation in the organic solvents compared with aqueous solution. There is much greater substrate binding strength in hexane media than the water and acetonitrile ones since the polar solvent significantly weakens electrostatic interactions, which are observed to be the main driving force to the binding. In addition, some residues of the S1 pocket could remain favorable contribution to the binding despite the solvent change, but with differences in the contribution extent, the number and the type of residues between the three media.

Keywords Enzyme · Molecular dynamics simulation · Non-aqueous media · Substrate binding

Electronic supplementary material The online version of this article (doi:10.1007/s00894-013-1900-2) contains supplementary material, which is available to authorized users.

Y. Meng · Y. Zhu · Y. Guo · M. Li · Z. Wang · X. Pu (✉) · L. Jiang (✉)
Faculty of Chemistry, Sichuan University, Chengdu 610064, People's Republic of China
e-mail: xmpuscu@scu.edu.cn
e-mail: jianglin203@scu.edu.cn

Y. Yuan
College of Management, Southwest University for Nationalities, Chengdu 610041, People's Republic of China

Introduction

Non-aqueous enzymatic catalysis has been one of the most active areas in enzymology in recent years and attracted considerable research interests [1–5]. The interests mainly stem from its synthetic and processing advantages with respect to aqueous environment, for example, lower side reactions with water, higher selectivity and thermostability. However, as known, the enzymes have very low activities universally in organic media and there are still widely different or even contradictive opinions on explaining the drop of the activity. Some research works reported that the organic solvent decreases the flexibility of enzyme, and reduces its activity accordingly [6–9]. However, some investigations indicated that the enzymatic activity is not directly associated with its flexibility [5, 10]. In addition, some studies suggested that the penetration of organic solvents into the active site of enzymes should play an important role in reducing the activity of enzymes [2, 11]. In fact, these different opinions are mainly attributed to the complexity of the system composition and the structure of proteins, which lead to difficulty in experimentally detecting subtle conformational variations induced by the organic solvents. To date, no experimental technique has been able to provide a detailed structural-dynamics description of this phenomenon. Thus, the structure and function of enzymes in non-aqueous media is still not fully understood, especially at the molecular level, which limits its further application. To survey this issue, a number of MD simulation studies of proteins in water/organic solvent mixtures have been performed to supplement experimental investigations [12–22] since the

molecular dynamics simulation could provide a molecular level view of the solute-solvent interactions and reveal the dynamic behavior of proteins [23–26]. The organic solvent induced changes in the structure and dynamics of enzymes, solvent distribution and hydration mechanism were discussed in these MD works. However, the previous MD studies almost all centered on apo-enzyme systems without substrate bound. Little information was reported for the complex system of enzyme and substrate. Although the observations from the free enzyme systems provide valuable information for understanding the structure and function of enzymes in non-aqueous media, the detailed aspects of the dynamic processes of substrate binding almost remained unstudied in the field. Similarly, the information has been very limited on experiments since a Michaelis complex would instantly lead to the product state under the regular experimental conditions, resulting in difficulty in experimentally detecting the dynamics process of the substrate binding. However, the dynamic process is crucial in the catalytic function of enzymes [27–29]. Thus, it is necessary to obtain the information at the atomic level using MD method.

On the basis of these considerations above, we, in the work, carried out MD simulations on trypsin with one peptide substrate bound in three media (viz., aqueous, acetonitrile and hexane). Trypsin is a typical representative of serine proteases, which is one of the most extensively studied families of enzymes. As accepted, its catalytic mechanism mainly depends on three amino acid residues [30]. (viz., His57, Asp102, Ser195), which form a hydrogen bond network (see Fig. 1). By means of the network, the catalytic reaction is performed [30, 31]. The acetonitrile [32] and hexane [33] solvents were chosen in the work due to the fact that the catalytic activities of enzymes are significantly associated with the polarity of organic solvents used [5, 10, 34–36]. For example, subtilisin was exhibited to be higher activity in polar tert-Amyl alcohol and acetone media than non-polar hexane one [34]. However, α-chymotrypsin was observed to have higher activity in non-polar octane than highly polar acetonitrile [10, 35]. Similarly, it was reported that the transesterification activity of cutinase is higher in n-hexane than that in acetonitrile [36]. To gain insight into the effects of polar and non-polar organic solvents on the structure and the function of trypsin, we selected acetonitrile and hexane as representative polar and non-polar organic media in the work. The two solvents were often used in nonaqueous enzymatic catalysis [32, 33].

It is noted that the two organic media used in the work include the crystal waters in the trypsin structure due to the importance of essential water in maintaining catalytic activities of enzymes in nonaqueous media [1–5]. The organic media with inclusion of the crystal waters may be served as microhydration non-aqueous environment. Experimental research already revealed that bulk water is not absolutely necessary and enzymes can function in media with very little water content [37]. In some cases, as little as tens of molecules of water on a protein are sufficient for observable catalytic function [38, 39].

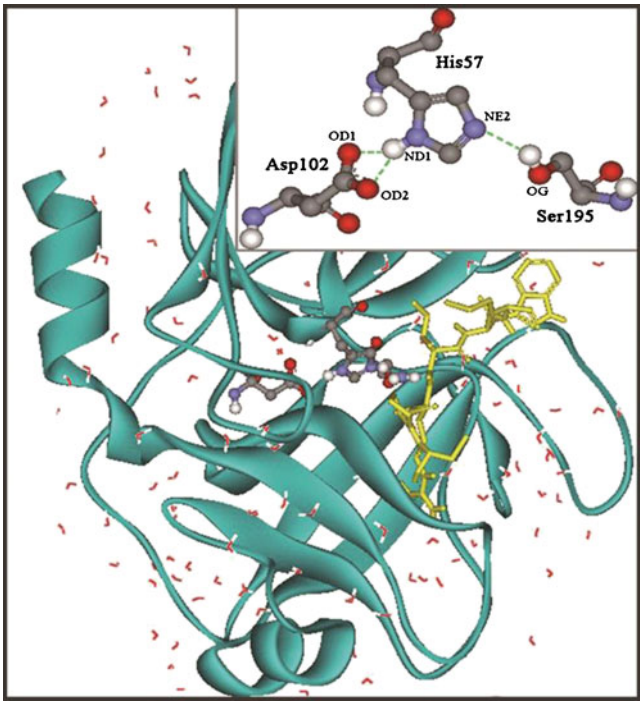

Fig. 1 Illustration of the secondary structure of trypsin with its catalytic triad residues (ball and stick) and substrate (*yellow stick*), derived from the crystal structure. The *green dot lines* denote hydrogen bond

In addition, the trypsin without the substrate (named as free enzyme), as a reference, was studied in the work. By comparison of the free enzyme system with the complex one in three media, we could examine the solvent effects on the enzymatic structure, solvent distribution and substrate binding as well as changes in these properties induced by the substrate binding. Indeed, some different observations from the previous studies on free enzyme systems were obtained in this work. Thus, the work could provide new insight adding to previous studies so that we could well understand the structure and function of enzymes in organic media.

Methods and materials

The initial coordinate of trypsin-substrate system was obtained from the X-ray crystal structure of trypsin with an inhibitor bound (PDB code 1MCT [40]), which contains 223 residues and 135 crystal waters. The sequence of the substrate was taken from the inhibitor at the active site (Cys3-Pro4-Arg5-Ile6-Trp7-Met8, this serial number came from the original inhibitor sequence). The scissile bond of the substrate is located between the Arg5 and Ile6 residues. The complex system constructed above was successfully used to study catalytic factors in aqueous solution [30].

In this work, the initial structure of free trypsin without the substrate was obtained mainly through removing the inhibitor from the crystal structure of the 1MCT for the following reasons. Firstly, the identical structure of the enzyme between the complex system and the free one is necessary in the study in order to more reliably observe the substrate binding induced differences in the calculated properties (for example, structures, solvent distributions and so on). However, there is no crystal structure of the free enzyme available, which has the sequence that is identical to the enzyme in 1MCT structure. Secondly, some experimental and calculated studies [30, 31] on trypsin in aqueous solution indicated that no observable conformation changes occur when trypsin is bound to substrates or inhibitors. In fact, the following MD results derived from aqueous environment in this work also confirmed the finding. Furthermore, we also used MD to simulate the crystal structure of the free trypsin (PDB code 2PTN [41]) using the method described latter. The 2PTN also contains 223 residues with the inclusion of 82 crystal waters. But, 42 out of the 223 residues are different from the enzyme in the 1MCT complex and the 42 residues are located far from the active site and S1 pocket. The results derived from the 2PTN are similar to those derived from the free enzyme constructed above (see supporting information Tables S1–S2), further confirming that it should be reasonable to use the free enzyme constructed through removing the substrate to study the effects of the solvent and the substrate. To neutralize the system, six chloride ions were added into the models.

Molecular dynamics (MD) simulations were performed using the program AMBER11 [42] with parm99SB force field [43]. Three solvent environments were considered in this work, viz., aqueous solution (labeled as WAT) as a reference, acetonitrile media (labeled as ACN), hexane media (labeled as HEX). Water was represented by the TIP3P [44] model. Two organic solvent molecules were parameterized. The geometric parameters were derived by ab initio geometry optimization on the HF/6-31G* level using Gaussian 09 [45] in the gas phase. The partial charge was calculated using the RESP program [46] of AMBER11 to the electrostatic potential. In organic solutions, we still retained all 135 crystal waters in 1MCT structure. Extra water and organic solvent molecules were added using the LEAP utility. The rectangle periodic box was set up so that any solute atom is at least 10 Å, 12 Å, 14 Å from any box edges for the aqueous, acetonitrile and hexane systems, respectively. As a result, the number of water molecules in WAT is 8610 for the complex system and 8587 for the free system. The number of acetonitrile molecules in ACN media is 2147 for the complex system and 1987 for the free system. The complex and free systems in hexane media (viz., HEX) contain 1441 and 1440 hexane molecules, respectively. The calculated densities of 1.02 (water) g cm^{-3}, 0.80 (acetonitrile) and 0.68 g cm^{-3} (hexane) were in good agreement with corresponding experimental values [47] of 1.00 g cm^{-3}, 0.79 g cm^{-3} and 0.66 g cm^{-3}, respectively.

Before production run, all systems were prepared by a series of combined energy minimizations (the steepest descent method for the first 3000 steps, the conjugate gradient algorithm for the second 2000 steps). After the minimizations, the systems were heated gradually from 0 to 300 K within 120 ps. Then 5 ns dynamics simulations were carried out with periodic boundary conditions in the NVT ensemble at 300 K using the Berendsen temperature coupling [48]. Finally, 95 ns NPT simulation (T=300 K and P=1 atm) was performed in the canonical ensemble. For all MD simulations, a time step of 2 fs was utilized. The SHAKE algorithm [49] was applied to constrain all the bonds involving a hydrogen atom with a tolerance of 1.0×10^{-5} Å. Nonbond interactions were handled with a 12 Å atom-based cutoff. The particle-mesh-Ewald (PME) method [50, 51] was applied to treat the long-range electrostatic interactions. The trajectory was saved every 2 ps for analysis. All of the MD results were analyzed using the analysis module of AMBER11 and VMD [52] as well as some other developed specific trajectory analysis software.

Results and discussion

Root-mean-square deviations (RMSD)

The root mean square deviation (RMSD) is for deviation from the crystal structure of trypsin. The RMSD values of backbone atoms of the enzyme were plotted as a function of time in Fig. 2. As shown in Fig. 2, after 50 ns, the RMSD values in the three media present minor fluctuations and approach equilibration, suggesting that all systems were well relaxed and stable within the scale of 100 ns simulation time. The average RMSD values over the last 20 ns simulation in the three media are listed in Table 1. It is apparent that the RMSD values in three solvents follow the order of HEX>ACN>WAT for the free and complex systems, indicating that the enzyme structures in the organic solvents have a larger deviation from the crystal structure. Moreover, the deviation in the non-polar solvent is higher than that in the polar solvent, consistent with observations from previous research works [17, 21, 53–55]. A comparison of the free enzyme with the complexed one (see Table 1 and Fig. 2) shows that the trypsin with the substrate bound exhibits smaller RMSD values than the free one in the two polar media (viz., acetonitrile and water) while a reverse trend is presented in the non-polar hexane. The observation implies that the six-amino-acid peptide substrate would stabilize the trypsin native structure in the polar media, but destabilize it in the non-polar one.

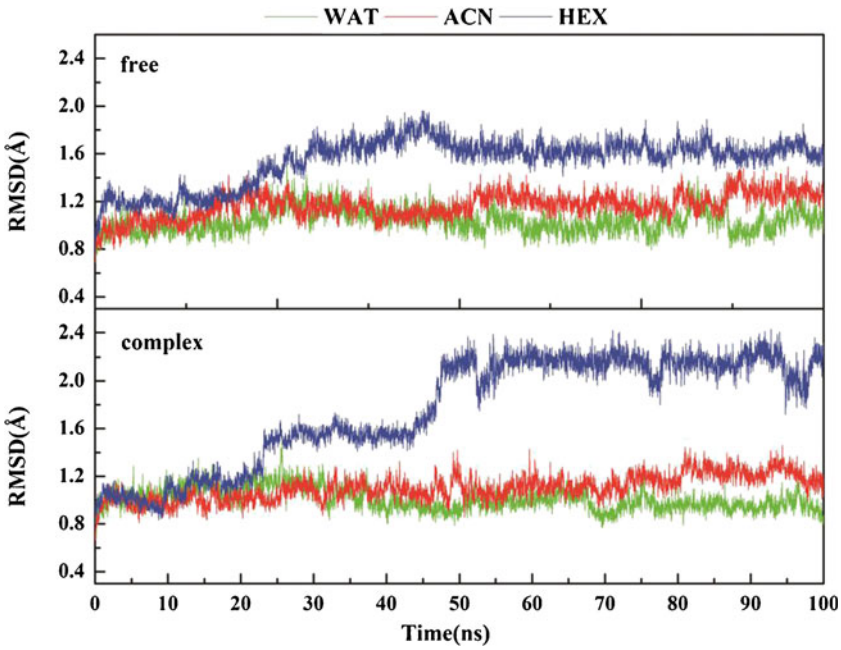

Fig. 2 The RMSD values of backbone atoms of trypsin with respect to the simulation time in aqueous (labeled as WAT), acetonitrile (labeled as ACN) and hexane (labeled as HEX) media for the free enzyme (*top*) and the complex system (*bottom*). The RMSD is for deviation from the crystal structure

In addition, we also calculated the RMSD values of each residue, as shown in Fig. 3 and Fig. 4. It can be seen from Fig. 3 that the RMSD values of most residues follow the order of HEX>ACN>WAT, either for the free or the complexed enzyme, consistent with the variation trend of the total RMSD values above. As shown in Fig. 4, the RMSD values of most residues of the free enzyme are greater than those of the complexed one in the polar water and acetonitrile media. The difference is more pronounced in acetonitrile solvent than that in aqueous solution. As a result, the average RMSD value of the free enzyme is larger than that of the complexed one in the two polar solvents, further confirming the stabilization effect of the substrate on the enzyme native structure in the polar media. In contrast, most residues of the free enzyme have larger deviations from the crystal structure than those of the complexed one in hexane media (see Fig. 4), exhibiting unfavorable effects of the substrate on stabilization the enzyme native structure in the non-polar media.

The specificity of serine proteases is usually associated with the S1 pocket, which is adjacent to the surface and

Table 1 Summary of the RMSD values (in Å unit) of backbone atoms, radius of gyration (R_g), the percentage of the α-helix and β-sheet, SASA and the percentage of hydrophobic and hydrophilic SASA (%) and their standard deviation over last 20 ns trajectories for aqueous solution (WAT), acetonitrile (ACN) and hexane (HEX) media

System	Free[a]			Complex[b]		
Media	WAT	ACN	HEX	WAT	ACN	HEX
RMSD of protein						
backbone	1.047±0.104	1.254±0.085	1.624±0.067	0.928±0.060	1.218±0.076	2.158±0.113
R_g and the percentage of the helix and sheet						
R_g (Å)	16.53±0.05	16.48±0.05	16.26±0.04	16.42±0.05	16.34±0.03	16.27±0.03
α-Helix%	6.0±0.2	5.6±0.2	5.0±0.2	7.2±0.3	4.7±0.2	4.4±0.2
β-Sheet%	32.8±0.5	32.4±0.4	32.6±0.5	32.6±0.5	32.7±0.5	32.0±0.4
Total SASA, hydrophobic and hydrophilic residues SASA/total SASA						
SASA(×10^3 Å)	10.54±0.15	10.15±0.14	9.69±0.11	10.32±0.15	9.96±0.11	9.74±0.12
Hydrophilic%	75.95±0.54	72.59±0.69	71.39±0.73	72.57±0.87	71.85±0.50	68.63±0.62
Hydrophobic%	24.05±0.54	27.41±0.69	28.61±0.73	27.43±0.87	28.15±0.50	31.37±0.62

The RMSD values are for deviations from the crystal structure

[a] Free denotes the free enzyme system

[b] Complex denotes the enzyme-substrate complex system

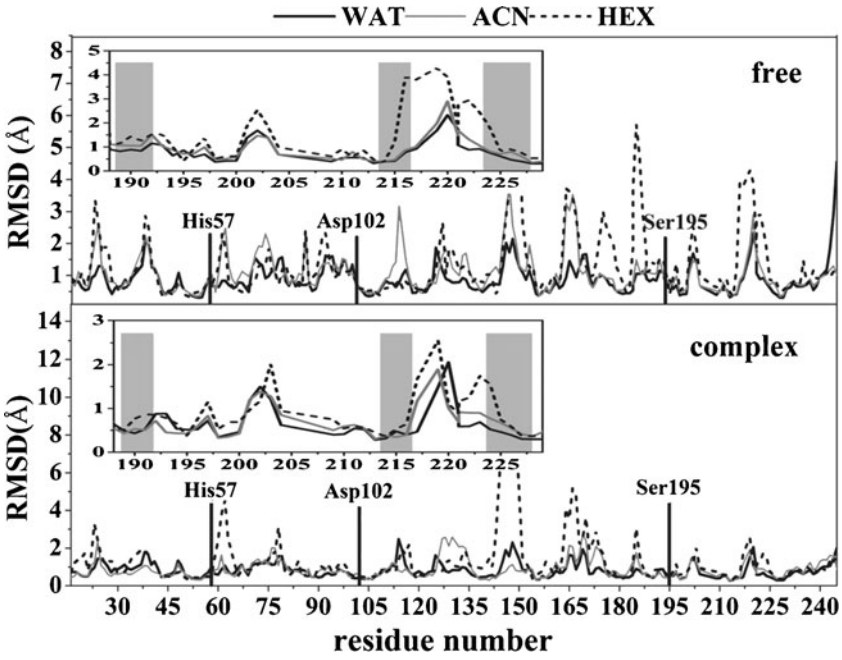

Fig. 3 A comparison of average RMSD values of per-residue of the protein between aqueous solution (WAT), acetonitrile (ACN) and hexane (HEX) media over the last 20 ns trajectories for the free enzyme (*top*) and the complex (*bottom*) systems. The RMSD is for deviation from the crystal structure. The residues that correspond to the S1 pocket are highlighted in gray and the three catalytic residues are highlighted in *black lines*

bordered with the side chains of residues 189–192, 214–216, and 224–228 [31]. Thus, we, herein, especially concern the changes in the S1 pocket induced by the solvents and the substrate binding. Similar to observations above, the RMSD values of the S1 pocket residues are slightly larger in acetonitrile media than those in aqueous solution (see Fig. 3) for the free enzyme system. However, significantly larger deviations of the S1 pocket from the native protein are presented in the non-polar hexane solvent for the free enzyme. The observation implies a large variation of the pocket shape in the non-polar solvent, which is probably associated with the specificity change to some extent. To more clearly evaluate the effect of organic solvent on the S1 pocket, we used VMD to show its structure in the three media, based on the average protein structure over the last 20 ns trajectories (see Fig. 5). As can be seen from Fig. 5, there are no large differences in the S1 pocket structure between the aqueous solution and the acetonitrile media for the free enzyme. However, the S1 pocket exhibits a relatively closed state in the acetonitrile media compared with the aqueous solution. Interestingly, there are significant differences in the S1 pocket structure between the hexane media and the two polar solvents for the free enzyme. The S1 pocket is almost fully closed in hexane media compared with aqueous solution and acetonitrile ones, indicating that the non-polar solvent could induce larger changes in the S1 pocket conformation than the polar organic one. The observations derived from Fig. 5 are consistent with the calculated RMSD values above, which could rationalize experimental findings that the organic solvent could change substrate specificity. Different from the free enzyme system, the RMSD values of the S1 pocket in the complex systems do not present significant differences between the three solvents, as revealed by Fig. 3. The observation also displays the role of the substrate in stabilizing the S1 pocket structure. Compared with the free enzyme, the substrate binding only induces minor conformational changes in the S1 pocket in aqueous solution and polar acetonitrile media, as shown in Fig. 5. The observation is `consistent with the experimental finding in aqueous solution [31] that there are no observable conformation changes when trypsin is bound to substrates or inhibitors. However, a comparison of the free enzyme with the complexed one in hexane media shows that the non-polar solvent could cause large changes in the S1 pocket conformation (see Fig. 4 and Fig. 5), implying that a classical induced fit binding mechanism may be operative in the non-polar media. However, the behavior has been not reported for chymotrypsin and trypsin in aqueous solution.

In addition, we also analyzed the RMSD values of the three active residues (viz., His57, Asp102 and Ser195) since the conformation of the active site should play an important role in maintaining optimal activity of the enzyme. As shown in Fig. 3, the three active residues of the enzyme with or without the substrate bound exhibit much lower RMSD values than most residues of the protein in the three media, displaying small mobility. The results are consistent with previous observations derived from free chymotrypsin in hexane [53] and acyl-chymotrypsin in aqueous solution [56]. In order to provide a visual view for the solvent induced conformation deviation from the crystal structure, Fig. 6 presents a superposition of the three catalytic residues of the last frame structures between the three media and the crystal one. It can be seen from Fig. 6 that there are only minor conformation displacements between the three media and the crystal structure for the active residues of the free

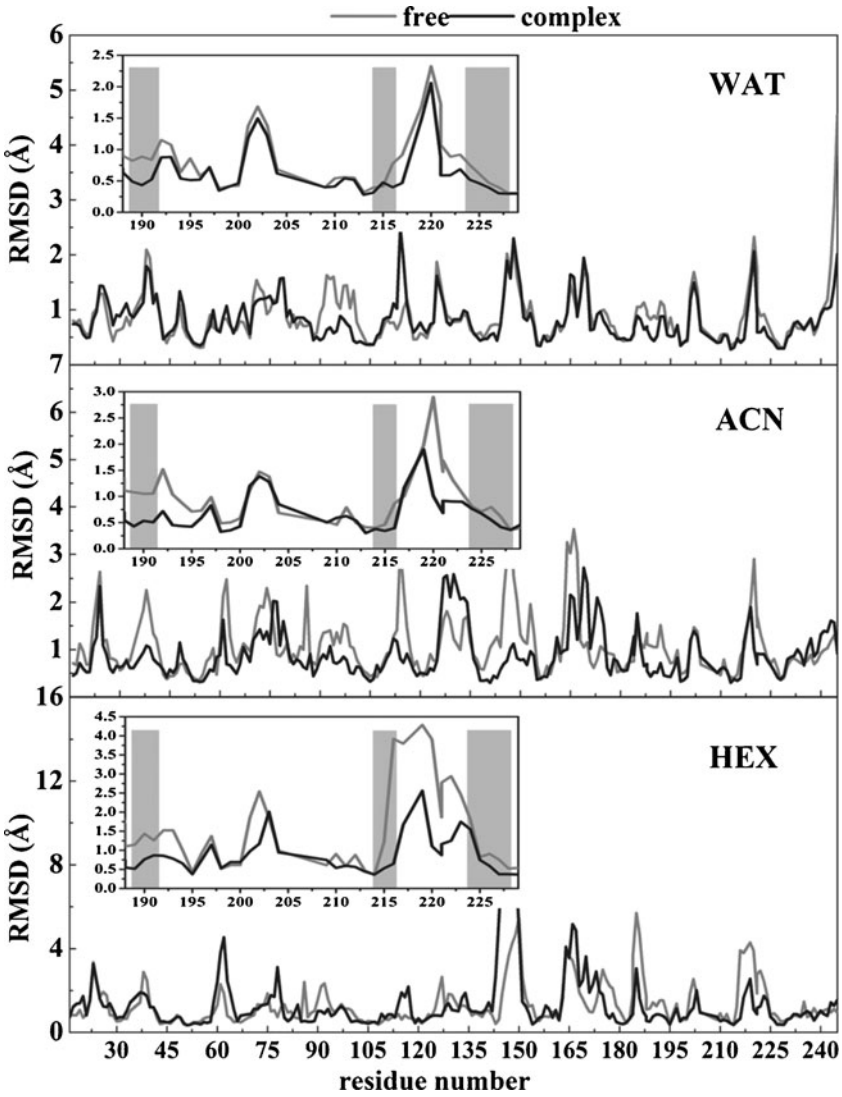

Fig. 4 A comparison of average RMSD values of per-residue between the complex and free enzyme over the last 20 ns trajectories in aqueous (WAT), acetonitrile (ACN) and hexane (HEX). The RMSD is for deviation from the crystal structure. The residues that correspond to the S1 pocket are highlighted in *gray*

enzyme and the complexed one, further displaying the relative rigidity of the active site. The relative rigidity of the active site residues should contribute to the fact that enzymes could remain catalytic activity in the polar and non-polar medium with low water-content [57–59]. In addition, the RMSD values of the three residues are similar between the complex system and the free enzyme one, indicating that the substrate binding does not play an observable role in influencing the structures of the active residues.

Radius of gyration and secondary structure

The radius of gyration is defined as the mass-weighted RMS distance of a collection of atoms from their common center

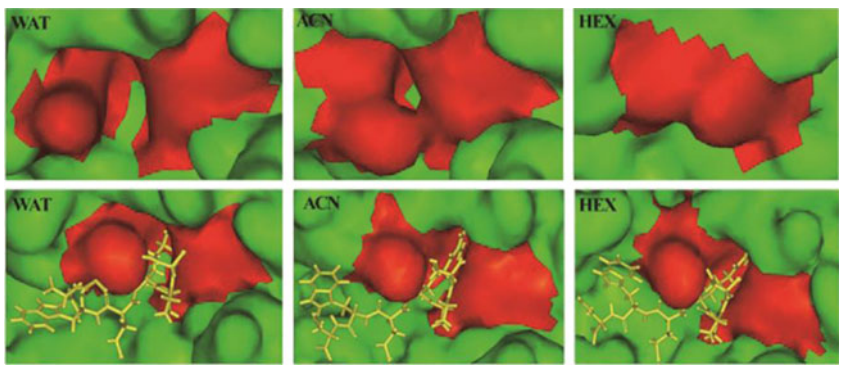

Fig. 5 Conformations of S1 pockets of the free enzyme (*top*) and the complexed enzyme (bottom) in aqueous solution (WAT), acetonitrile (ACN) and hexane (HEX) media, derived from the average structure over last 20 ns trajectories. Red, yellow and green denotes S1 pocket, substrate and the other residues around S1 pocket, respectively

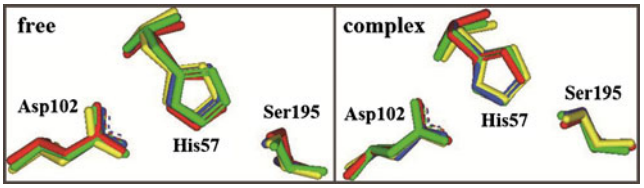

Fig. 6 A superimposition of the three catalytic residues of the free (free) and complexed (complex) enzymes between the crystal structure and the last frame structure in three media. Hydrogen atoms are not shown. Color code: *blue*, *crystal structure*; *red*, *aqueous solution*; *green*, *acetonitrile solution*; *yellow*, *hexane solution*

of mass [60]. Hence, this analysis could give us insight into the overall dimensions of the protein. The radius of gyration (R_g) is commonly estimated in terms of Eq. 1 [61]:

$$R_g = \sqrt{\frac{\sum_{i=i}^{N} m_1 \left(\overrightarrow{r_i} - \overrightarrow{r_{com}}\right)^2}{\sum_{i=1}^{N} m_i}}, \quad (1)$$

where N is the number of atoms and r_{com} is the center-of-mass position of the protein, while m_i and r_i are the mass and position of atom i, respectively. Table 1 lists the average radius of gyration calculated over the last 20 ns in the three media. In the two enzyme states, the radius of gyrations are in the order of WAT>ACN>HEX, displaying a positive dependence on the solvent polarity. Namely, lower the polarity of solvent, more compact the enzyme structure. On a whole, the gyration radiuses of the enzyme with or without the substrate bound are similar between the three media (nearly 1.6 nm), indicating that the solvent changes and the substrate binding only play minor role in influencing the overall structure of the enzyme.

To gain more insight into the effect of the solvents and the substrate on the trypsin conformation, we also calculated the secondary structures of the protein in the three media using DSSP algorithm implemented in Amber program [62] and the calculated results are listed in Table 1. The data in Table 1 shows that the solvent and the substrate hardly influence the β-sheet structure of the enzyme. However, it is slightly easier for the α-helix structure to be influenced by the solvent and the substrate, but still with small extent. As can be seen from Table 1, the organic solvent could reduce the content of α-helix. Although the substrate is far away from the α-helix regions, its binding still induces a slight increase in the α-helix content in aqueous solution and a slight drop in acetonitrile and hexane media, as revealed by a comparison of the free enzyme system with the complexed one.

SASA (solvent-accessible surface area)

Solvent accessible surface areas (SASA) were calculated using the program VMD defined by rolling a probe of given size (1.4 Å) around a van der Waals surface of the protein, which could provide information of how different parts of a protein can be affected by media. Table 1 lists the calculated SASA and the relative hydrophobic as well as the hydrophilic surface percentage for 1000 snapshots, which were selected from the last 20 ns trajectories at the interval of 20 ps.

An inspection of Table 1 reveals that the total SASA values follow the order of HEX<ACN<WAT in both the complex and the free enzyme systems, in agreement with the variation trend of radius of gyration. A comparison of the total SASA between the complexed enzyme and the free enzyme exhibits slight differences, suggesting that the substrate binding only plays a minor role in influencing the total SASA. To gain more insight into the variation of SASA, we analyzed SASA variations of hydrophobic and hydrophilic residues. The percentage of hydrophilic SASA and hydrophobic SASA relative to the total SASA (SASA %) are calculated in the three media and listed in Table 1. Similar to the variation trend of the total SASA, the percentage of hydrophilic SASA follows the order of HEX<ACN<WAT while a reverse trend is observed for the percentage of hydrophobic SASA. A comparison of the free enzyme and the complexed one reveals that the substrate binding induces an increase in the hydrophobic SASA and a decrease in the hydrophilic SASA. The variation is more pronounced in strong polar water environment and non-polar hexane media compared with polar acetonitrile solvent. The observation indicates that the solvent and the substrate make some hydrophobic and hydrophilic side chains reorient themselves on the protein surface. And some of the polar side chains, which are originally accessible to the solvent in the free enzyme state, are buried into the interior of the trypsin, and some non-polar side chains become to be exposed to the solvents. As a result, the hydrophobic surface area is increased with decreasing solvent polarity or upon the substrate binding while a reverse trend is observed for the hydrophilic surface area.

Distribution of solvents and crystal water

In order to gain insight into the distribution of water and the organic solvents around the protein, the spatial probability density distribution of solvents were calculated by binning atoms positions from rms coordinate fit frames over all protein atoms at 20 ps intervals into $(0.5)^3$ m grids over the last 20 ns trajectories.

The number densities of the bulk water, acetonitrile and hexane solvents are calculated to be 0.03344, 0.0116 and 0.0046 atoms/Å^3. The contour level cutoff values in Fig. 7

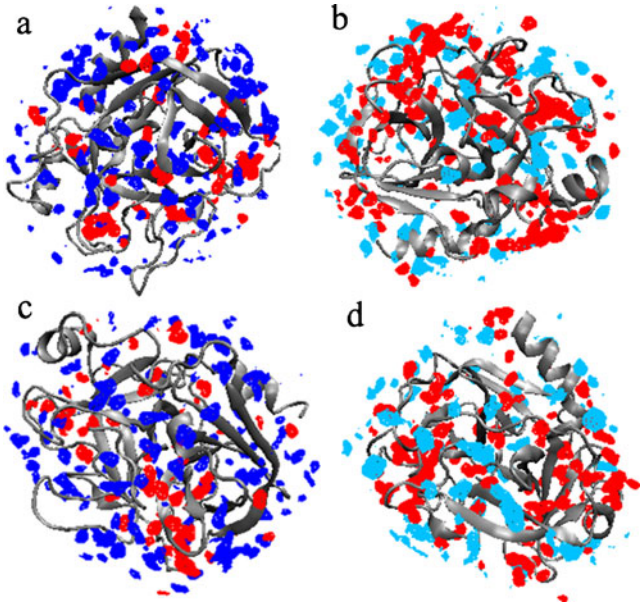

Fig. 7 The spatial distribution of water and organic molecules around the protein for the free enzyme (**a**, **b**) and the complexed one (**c**, **d**). The protein corresponds to the average structure of trypsin over the last 20 ns equilibrium trajectory. The contours enclose regions with a probability density above five times the average density for water (red contour), 14 times for acetonitrile (blue contour) and 36 times for hexane (cyan contour). The average structure of the enzyme was drawn according to its second structure

are 5, 14 and 36 times average density of bulk water, acetonitrile and hexane, which could provide a clear picture for the distribution of the two organic solvents and the crystal water around the protein. As shown in Fig. 7, the spatial contours enclosing high probability regions of the crystal waters are mainly located at the protein surface and a fraction in the protein interior, not diffusing into the bulk organic solvent. In addition, the density of crystal water molecules around the protein in hexane media is significantly higher than that in acetonitrile one both for the complex system and the free enzyme one. To obtain quantitative information of the solvent distribution around the protein, we also calculated the number of solvent molecules within 2.5 Å distance from the protein surface (see Table 2) since the 2.5 Å distance is generally considered as the first hydration layer interacting with the protein [17]. As show in Table 2, most of the crystal water molecules in the organic solvents are located in the 2.5 Å region, especially in hexane media, where the number of water molecules in the region is much more than that in acetonitrile media, either for the free enzyme or the complexed one. The quantitative observations are consistent with Fig. 7, suggesting that the acetonitrile has stronger ability to strip off water from the protein than the hexane. Similarly, there is little difference in the distribution of the water and organic solvent around the protein between the free enzyme and the complex systems, judging from Fig. 7 and Table 2.

As can be seen from Fig. 7, there are some water and organic solvent molecules penetrating into the protein interior. To gain further insight into the distribution of solvents in the active site, we calculated the average resident number of acetonitrile, hexane and water in the active pocket, defined as the space within 5 Å centered at the atom OG@Ser195 owing to its crucial role in the catalytic reaction (see Table 2). The data in Table 2 shows that there are 2∼10 water molecules and 1∼3 organic solvent molecules resided in the active site for the free enzyme, with more water molecules in the hexane media and more organic solvent molecules in the acetonitrile one. For the complex system, there is nearly no water molecule or organic solvent molecules penetrating into the 5.0 Å region in acetonitrile and hexnae media, but there are about two water molecules located in the active site in aqueous solution. The observations reveal that the peptide substrate could expel the solvent molecule from the active site and further confirm the observation above that the acetonitrile has stronger ability to strip off water molecule than the hexane.

H-bond analysis in the active site

As revealed, the hydrogen bond network between the three active residues (viz., His57, Asp102 and Ser195) plays an important role in the catalytic ability of serine proteases [31, 63]. Thus, we, in the part, carried out a detailed analysis of H-bonding between the three residues (see Table 3). Herein, the criteria used to define hydrogen bond are solely geometric. We consider one hydrogen bond to be presented if both the distance between hydrogen atom-acceptor and hydrogen atom-donor is less than 3.5 Å and the angle between the hydrogen bond donor atom, hydrogen atom, and acceptor atom is more than 120° criteria are satisfied simultaneously. A hydrogen bond is considered to be stable if it exists for>90 % of the trajectory time.

As can be seen from Table 3, the H-bonding between the Asp102 and the His57 residues has been stably existed over more than 95 % life time, suggesting that the H-bonding is very stable and insensitive to the solvent changes and the substrate binding. However, the solvent and the substrate should induce minor changes in the H-bond strength, deduced from the differences in the H-bond distances between the systems under study. In contrast, the solvent and the substrate could significantly influence the H-bonding between the His57 and Ser195 residues. For the enzyme without the substrate bound, the H-bond is broken in most of the simulation time in water and acetonitrile media but could stably exist in non-polar hexane media with lifetime of 95 %. We also calculated the H-bonding between the Ser195 residue and the water molecule (see Table 3). In the free enzyme systems, the calculated result shows that there is H-bonding between the Ser195 and the water molecule over more than half of the

Table 2 Average number of water molecules (water) and organic solvent molecules (organic) residing in the surface of the protein and the active site over the last 20 ns trajectories for aqueous solution (WAT), acetonitrile (ACN) and hexane (HEX) media

System	Free			Complex		
	WAT	ACN	HEX	WAT	ACN	HEX
No. of solvent molecules residing within 2.5 Å region from the surface of protein						
Water molecule	418±12	85±6	121±3	463±12	82±4	126±3
Organic molecule	/	152±9	110±6	/	159±8	109±6
No. of solvent molecules siding in the active site[a] (5 Å)						
Water molecule	8.3±1.4	1.5±1.1	3.2±0.5	1.6±1.2	0±0	0±0
Organic molecule	/	2.4±1.1	1.3±0.6	/	0±0	0±0

[a] Average number of solvent molecules in the active site, defined as the space with 5.0 Å centered at the atom OG@Ser195

"/" denotes no organic molecule in aqueous solution

simulation time in aqueous solution and acetonitrile media while a low life time (15 %) is observed in hexane media. The observations should partly contribute to the instability of the H-bond in water and acetonitrile media and the stability in hexane solvent for the free enzyme. Different from the free enzyme system, the substrate binding significantly enhances the H-bonding between the His57 and Ser195 residues, as demonstrated by higher lifetime in the complex system relative to the free one (see Table 3). As revealed above, there is no water molecule penetrating into the active site to form H-bond with the Ser195 residue for the complexed enzyme in acetonitrile and hexane media. While for the complex system in aqueous solution, only very occasional H-bond with lifetime of 7 % is formed between water and the Ser195 residue, which should play to some extent role in perturbing the H-bonding between the His57 and Ser195 residues. As a result, the life time of the H-bond between the two residues is ~53 % for the complex system in aqueous solution, significantly lower than those in the other two organic media.

The key distances between the substrate and the catalytic residues

As generally accepted in the catalytic mechanism of serine proteases, the Ser195 residue attacks the carbonyl of the peptide substrate, assisted by the His57 residue acting as a general base to withdraw the proton from the Ser195, to yield a tetrahedral intermediate in the acylation reaction step [31]. Therefore, we focus on the effects of the solvents on the distance between the OG@Ser195 atom and the carbonyl carbon atom of Agr5 of the peptide substrate (viz., C@Arg5). Figure 8 plots the distance between the OG@Ser195 and the C@Arg5 atoms *vs* time in the three media. It is apparent that the distance is slightly fluctuated around 3.0 Å in the organic solvents in most of the simulation time, close to the reactive distance. However, it displays large fluctuations in aqueous solution, larger than 3.5 Å in more than half of the simulation time.

As observed above, in aqueous solution, there have been 1~2 water molecules located in the active site during the whole simulation time. Therefore, it is reasonable to assume if the water molecule resided in the active site results in the large distance between the OG@Ser195 atom and the C@Arg5 one in aqueous solution. We analyzed the distance between the OG@Ser195 atom and the nearest water molecule (viz., O@H_2O) to it. As shown in Fig. 8, there is a negative correlation between the two distances (viz., OG@Ser195…C@Arg5 and OG@Ser195…O@H_2O). Namely, the larger the distance of OG@Ser195…O@H_2O, the smaller the distance of OG@Ser195…C@Arg5, vice versa. The observation confirms the assumption above that the perturbation of the water molecules penetrating into the active site leads to larger fluctuations in the distance of OG@Ser195…C@Arg5 in

Table 3 Average distances of hydrogen bonds (in Å) and their percentage occupation (%) in aqueous solution (WAT), acetonitrile (ACN) and hexane (HEX) media over the last 20 ns trajectories

System	Free (%/Å)			Complex (%/Å)		
Media	WAT	ACN	HEX	WAT	ACN	HEX
NE2@His57···OG@Ser195	32/2.90	31/3.04	95/2.89	53/2.90	100/2.87	100/2.88
OD2@Asp102···ND1@His57	100/2.88	100/2.83	99/2.91	100/2.84	98/2.89	96/2.88
O@H_2O···OG@Ser195	79 %	59 %	15 %	7 %	0 %	0 %

NE2@His57 depicts OG atom of His57 residue. OG@Ser195 depicts OG atom of Ser195 residue. OD2@Asp102 depicts OD2 atom of Asp102 residue. ND1@His57 depicts ND1 atom of His57 residue. O@H_2O depicts O atom of the water molecule

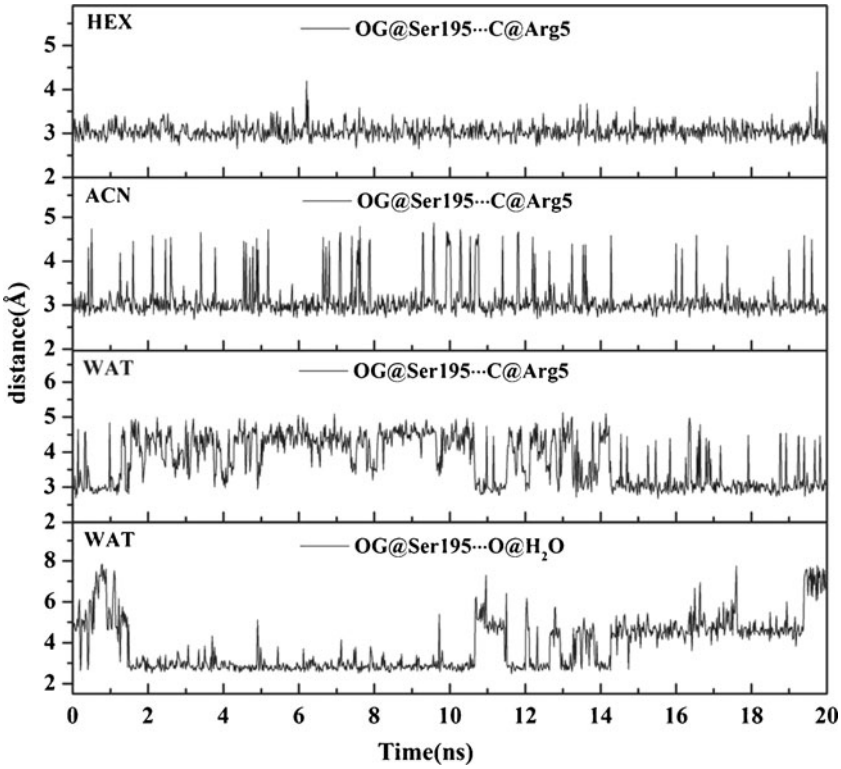

Fig. 8 The time-dependent distances between the OG@Ser195 atom and the C@Arg5 atom in aqueous (WAT), acetonitrile (ACN) and hexane (HEX) solution and the distances between OG@Ser195 atom and oxygen atom of the water molecule (O@H_2O) closest to it in aqueous solution (WAT), derived from the last 20 ns trajectories

aqueous solution than the other two organic solvents. As revealed above, there are nearly no solvent molecules penetrating into the active site for the complex system in the two organic media. It should be the absence of the perturbation of solvent molecule that partly contributes to smaller OG@Ser195…C@Arg5 distance in the organic media than that in aqueous solution for the complex system.

As revealed by experiments, the catalytic reaction mainly depends on the attack of OG@Ser195 to C@Arg5 and the assistance of the H-bond network among the Ser195, His57 and Asp102 residues. Thus, the effects of solvents on these distances should be associated with the variation of the enzyme activity. Figure 9 summarizes the three distances of NE2@His57…OG@Ser195, ND1@His57…OD2@Asp102 and OG@Ser195…C@Agr5 in 1000 snapshots selected from the last 20 ns trajectories at the interval of 20 ps. The results show that the three distances in 989 of 1000 snapshots are almost smaller than 3.5 Å in hexane media. In acetonitrile media, the three distances are also smaller than the value of 3.5 Å in 933 of 1000 snapshots, slightly less than hexane media. It seems that the enzyme and the substrate are appropriate to the reactive conformation in the two organic solvents, which may be associated with the fact that the enzyme could retain activity in micro-hydrated organic media. The result should be attributed to more compact structure of the enzyme induced by the organic solvent and the absence of perturbation of solvent molecules in the active site. In contrast, the three distances approach to the reactive conformation only in less than half of 1000 snapshots (viz., 431 snapshots), which should be attributed to the perturbation of water molecules in the active site. However, in fact, the enzymatic activity is higher in several orders of magnitude in aqueous media than in organic media [1, 5, 10, 64], suggesting that these key distances presented in the initial binding step seem to play no dominant role in the enzymatic activity.

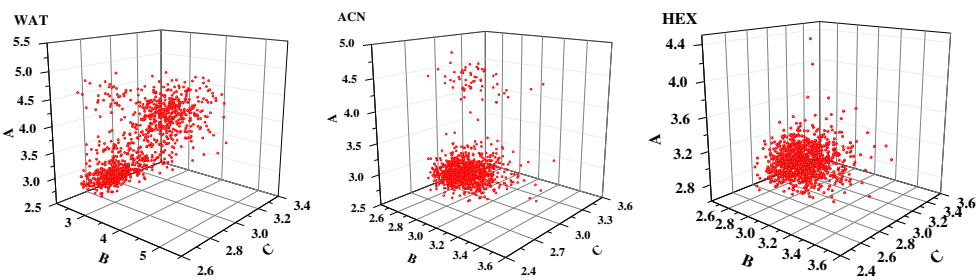

Fig. 9 Plot for three key distances (in Å) of NE2@His57···OG@Ser195 (labeled as A), ND1@His57···OD2@Asp102 (labeled as B), and OG@Ser195···C@Arg5 (labeled as C) in aqueous (WAT), acetonitrile (ACN) and hexane (HEX) media

Analysis on MM/GBSA binding free energy

To study the thermodynamics of the interaction between the substrate and the enzyme in the three media, binding free energy calculations were performed using molecular mechanics generalized Born/surface area (MM/GBSA) method [65]. The binding free energy $\Delta G_{\text{binding}}$ was calculated using Eq. 2

$$\Delta G_{binding} = G_{complex} - \left(G_{enzyme} + G_{substrate}\right). \quad (2)$$

Where G_{complex}, G_{enzyme} and $G_{\text{substrate}}$ are the free energies of the complex, enzyme, and substrate, respectively. The free energies are estimated as a sum of four terms (see Eq. 3).

$$G = E_{gas} + G_{psolv} + G_{npsolv} - TS \quad (3)$$

Where E_{gas} is the molecular mechanics energy of the molecule expressed as the sum of the internal energy of the molecule plus the electrostatic and van der Waals interactions, G_{psolv} and G_{npsolv} are the polar and nonpolar contribution to solvation energy of the molecule, respectively. T is the absolute temperature and S is the molecule entropy. Similar to many investigations [66–73], the change in solute entropy was not included in the free energy calculations in the work since we mainly focus on the relative order of binding affinities.

The GB model (IGB=2) was used to calculate the polar contribution to the solvation energy. The nonpolar contribution was determined on the basis of solvent-accessible surface area (SASA) using the LCPO method (see Eq. 4).

$$\Delta G_{npsolv} = 0.0072 \times \Delta SASA \quad (4)$$

The value of 1 was used for the interior dielectric constant, while values of 80, 37.5, and 1.88 were used for the exterior solvent dielectric constant for water, acetonitrile and hexane, respectively. Snapshots without the water and organic molecules as well as chloride ions were extracted from the MD trajectories in the last 20 ns at 20 ps intervals for the energy analysis. Table 4 lists the calculated energy and its components in the three solvents. Negative $\Delta G_{\text{binding}}$ values in Table 4 mean that the binding between the enzyme and the substrate is spontaneous in the three media. The binding strength in acetonitrile media is slightly weaker by 2.8 kcal mol^{-1} than that in aqueous solution. However, the $\Delta G_{\text{binding}}$ value in hexane media is high up to −195.3 kcal mol^{-1}, much higher than those in water and acetonitrile solvents. Rashidi determined Michaelis–Menten kinetic values for the oxidation of phenanthridine by aldehyde oxidase in the absence and presence of a volume of the water-miscible organic solvents (including N-N-dimethylformamide, acetonitrile, tetrahydrofuran, 1-propanol, 2-propanol, ethanol, pyridine, dioxane, and methanol solvents) that produces 30 % inhibition [74]. Their results predicted that the substrate affinity to the active site of the enzyme should be increased in the organic solvent-containing medium since the kinetic K_{m} values determined are smaller in the organic solvents than aqueous solution. In addition, kinetic parameters of rice peroxidase determined in 40 % organic solvents were observed to be lower in low-polar chloroform than that in polar ethanol [75]. In terms of the relationship between the binding strength and the kinetic K_{m} value, it can be assumed that the binding strength between the peroxidase and the substrate in ref. [75] is greater in the low-polar chloroform than that in polar ethanol. Our calculation results provide a support for these experimental observations above. As indicated by some experimental studies, higher catalytic activities were often observed for some enzymes in nonpolar solvents compared with the polar ones [2–5]. Laane [76] studied the rules for optimization of biocatalysis in organic solvents and concluded that the biocatalysis in organic solvents is low in polar solvents and high in non-polar ones. Feng [21] studied the activity of CALB in organic solvent using MD simulation and indicated that the activation energy of CALB was lower in non-polar solvents than that in polar solvents. It seems that the higher binding energy in non-polar solvent revealed by the work may rationalize the previous findings to some extent.

To gain insight into the driving force for the binding strength, we also analyzed the binding free energy components. Data in Table 4 suggest that the Coulombic interaction (ΔE_{ele}) and van der Waals interaction (ΔE_{vwd}) are the main driving force to the binding strength in the three media, in particular for the Coulombic interaction, which is highly attractive.

Table 4 Energetic analysis (in kcal mol^{-1}) of trypsin-substrate complex and their standard deviations in aqueous solution (WAT), acetonitrile (ACN) and hexane (HEX) media

Media	WAT	ACN	HEX
ΔE_{ele} [a]	−117.7±15.7	−103.6±13.0	−220.4±15.3
ΔE_{vdw} [b]	−62.7±4.9	−61.1±4.2	−70.6±5.0
ΔE_{gas} [c]	−180.4±15.3	−164.7±12.2	−291.1±14.8
ΔG_{npsolv} [d]	−8.6±0.4	−8.1±0.3	−9.7±0.3
ΔG_{psolv} [e]	131.7±13.5	118.2±10.3	105.5±5.4
ΔG_{solv} [f]	123.1±13.4	110.3±10.3	95.8±5.4
ΔG_{ele} [g]	14.1±5.7	14.7±5.2	−114.9±10.7
$\Delta G_{\text{binding}}$ [h]	−57.3±4.8	−54.5±4.7	−195.3±10.0

[a] Non-bonded electrostatic energy calculated by the MM force field

[b] Non-bonded van der walls contribution from MM force field

[c] Absolute free energy in gas phase

[d] Nonpolar contribution to the solvation free energy

[e] Polar contribution to the solvation free energy calculated

[f] Solvation free energy

[g] Total electrostatic energy contribution to the binding energy

[h] Binding energy

$\Delta E_{\text{gas}} = \Delta E_{\text{ele}} + \Delta E_{\text{vdw}}$, $\Delta G_{\text{solv}} = \Delta G_{\text{npsolv}} + \Delta G_{\text{psolv}}$

$\Delta G_{\text{ele}} = \Delta E_{\text{ele}} + \Delta G_{\text{psolv}}$, $\Delta G_{\text{binding}} = \Delta E_{\text{gas}} + \Delta G_{\text{solv}}$

Moreover, the Coulombic interaction ΔE_{ele} in hexane media is high up to −220.4 kcal mol^{-1}, much larger than those in aqueous solution (−117.7 kcal mol^{-1}) and acetonitrile media (−103.6 kcal mol^{-1}). The strong Coulombic attraction in the three media is largely offset by the unfavorable electrostatic solvation free energy (ΔG_{psolv}), as shown in Table 4. The result is expected since the Coulombic interaction is generally anti-correlated with the electrostatic solvation free energy ΔG_{psolv} [77]. An inspection of Table 4 shows that the net electrostatic contributions ΔG_{ele} disfavors the binding in aqueous solution and acetonitrile media, consistent with previous studies in aqueous solution [77–81]. The observation suggests that the polar solvent would disfavor the binding between the enzyme and the substrate. In contrast to the finding in the polar solvent, the electrostatic contribution is strongly favorable to the binding in non-polar hexane media. Indeed, our previous work, which studied effects of solvents with different polarity on the H-bond strength between formamide (as a model of protein) and the water molecule using continuum dielectric model, already revealed that the binding strength is decreased with increasing solvent polarity through reducing the electrostatic interaction [82], which is consistent with the observation from this work.

Important residues contributed to the substrate binding

Interfacial residues in the complex can be identified by considering the inter-atomic distance between the enzyme and the substrate [83, 84]. Residues across the interface with distances below a cut-off are considered as interacting. The cut-off distance is computed as the sum of van der Waal's radii of interacting atoms plus 0.5 Å [83, 84]. PSAIA software was used to identify the binding interface residues. Figure 10 representatively shows the binding interface of the crystal structure. Table 5 lists the calculated results in the three media, derived from the average structure over the last 20 ns trajectories. As expected, in aqueous solution, the residues of the S1 pocket are almost located in the binding interface and the interfacial residues also include oxygen hole residues (Gly193-Ser195). In acetonitrile media, the interfacial residues are similar to those in aqueous solution, along with minor variations in the number and the type of residues. However, significant changes in the number and the type of interfacial residues are observed in hexane media (see Table 5), compared with water and acetonitrile media. These observations further exhibit the solvent induced changes in the conformation of the enzyme, especially in the non-polar hexane solvent. However, in the three media, most of the S1 pocket residues are preserved in the interface, for example, the Asp189, Gly216 and Gly226 residues, which account for trypsin's specificity for the substrate containing arginine [31]. In addition, the oxygen hole residues (Gly193-Ser195) are also preserved in the interface in the three media.

In order to further identify the important residues contributed to the binding, we also decomposed the binding energy on a per-residue basis to obtain the contribution from structural elements, in terms of Eq. 5.

$$\Delta G_{i-binding} = G_{i-trypsin.substrate} - G_{i-trypsin} \quad (5)$$

Where $G_{\text{i-trypsin.substrate}}$ and $G_{\text{i-trypsine}}$ represent the energy of the ith residue in the complexed and the free trypsin, respectively. The energy for the residue i can be calculated using Eq. 6.

$$G_{i-A} = G_{i-gas} + G_{i-solv} \quad (6)$$

Where $G_{\text{i-A}}$ denotes $G_{\text{i-trypsin.substrate}}$ or $G_{\text{i-trypsin}}$ above; $G_{\text{i-gas}}$ is the gas phase internal energy of the residue i computed by AMBER force field; $G_{\text{i-solv}}$ is the solvation energy of the ith residue, including contribution from both polar ($G_{\text{i-polar}}$) and nonplar ($G_{\text{i-nonpolar}}$). The polar contribution was calculated with GB model. The nonpolar contribution arising from cavity formation and van der Waals interactions between the solute and the solvent was estimated using Eq. 4. Similar to the binding free energy calculation, snapshots without the water and organic molecules and chloride ions were extracted from the last 20 ns MD trajectories at 20 ps intervals for the calculation of binding energy decomposition on a per-residue. On the basis of calculated $\Delta G_{\text{i-binding}}$, we could identify the residues with significant impact on the trypsin-substrate binding (see Fig. 11). A comparison of Fig. 11 with Table 5 reveals that the interaction between the trypsin and the substrate is dominated by some high energy residues located in the binding interface. The rest of the binding energy comes from low

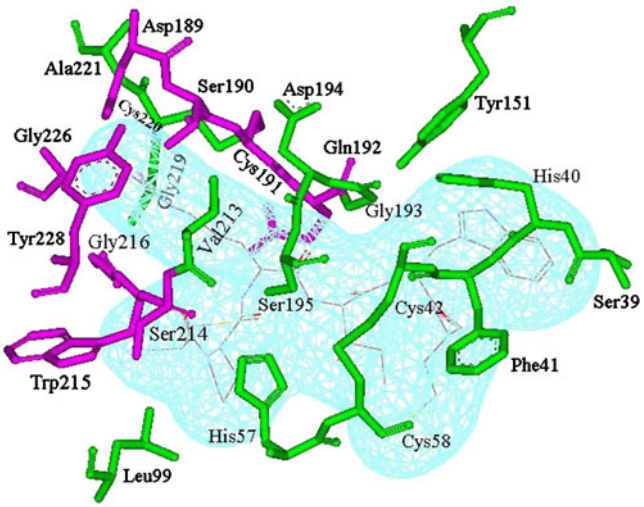

Fig. 10 Residues in the binding interface between the enzyme and the substrate, calculated using PSAIA software for the crystal structure. Pink and green denote the S1 pocket residues and the other residues in the binding interface, respectively. Substrate is shown by a light blue wire mesh

Table 5 The binding interface residues of the complex system in aqueous solution (WAT), acetonitrile (ACN) and hexane (HEX) media, derived from the average structure over the last 20 ns trajectories

	Name and serial number of amino acid		
	WAT	ACN	HEX
S1 pocket	Asp189 Ser190 Gln192 Ser214 Try215 Gly216 Gly226 Val227 Tyr228	Asp189 Ser190 Gln192 Ser214 Trp215 Gly21 Gly226 Val227	Asp189 Ser190 Gln192 Trp215 Gly216 Gly226
other	Phe41 Leu99 Tyr151 Gly193 Asp194 Ser195 Val213 Tyr217 Pro225	Phe41 Lys60 Leu99 Try151 Gly193 Ser195 Val213 Tyr217 Gly218 Ala221	Phe41 Lys60 Arg62 Leu99 Asn143 LYS145 Ser146 Gly148 Ser149 Try151 Gly193 Asp194 Ser195 Val213 Gly219Ala221

energy residues with a broad spatial distribution in the protein, which may be considered as a background interaction.

In aqueous solution, the His57, Tyr151, Asp189, Ser190, Gln192, Trp215 and Gly216 residues are observed to significantly contribute to the binding energy, as revealed by Fig. 11. Further analysis indicates that the Asp189, Ser190, and Gly216 residues favor the binding mainly via electrostatic interactions (see Fig. 12 and supporting information Table S3) and the backbone atoms of these residues form hydrogen bonds with the substrate, as illustrated in Fig. 13. The other two S1 pocket residues (viz., Gln192 and Trp215), the His57 and Tyr151 residues contribute their substrate affinity mainly through van der Waals interactions. Indeed, the Tyr151 and Gln192 residues also form hydrogen bonds with the substrate (see Fig. 13), which significantly favor the Coulomb binding energy by −9.7 and −14.2 kcal mol^{-1} (see supporting information Table S3), respectively. However, the favorable contribution derived from the H-bonding is largely offset by unfavorable polar solvation energy (9.3 kcal mol^{-1} for Tyr151 and 14.1 kcal mol^{-1} for Gln192). As a net result, the van der Waals interactions

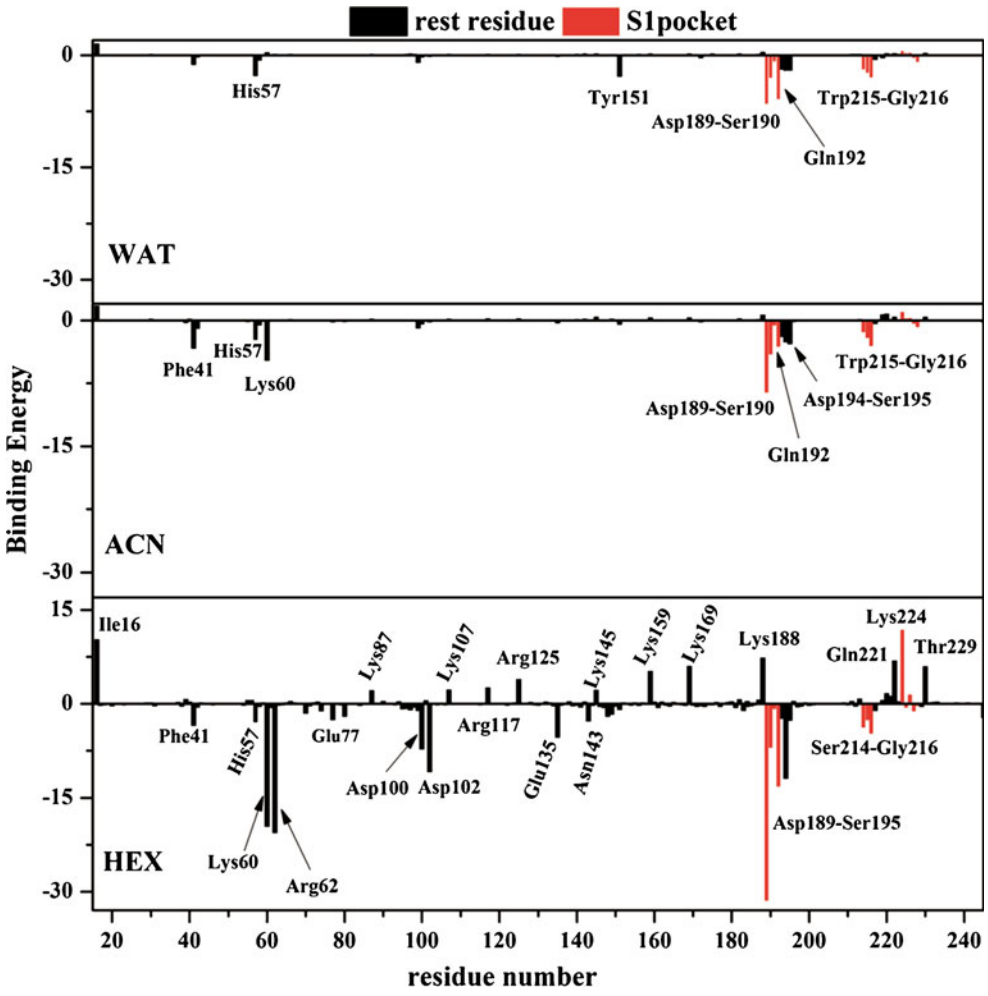

Fig. 11 Per-residue binding energy ΔGi for the substrate binding in aqueous (WAT), acetonitrile (ACN) and hexane (HEX) media and only high energy residues ($|\Delta Gi| \geq 2$ kcal mol^{-1}) are shown

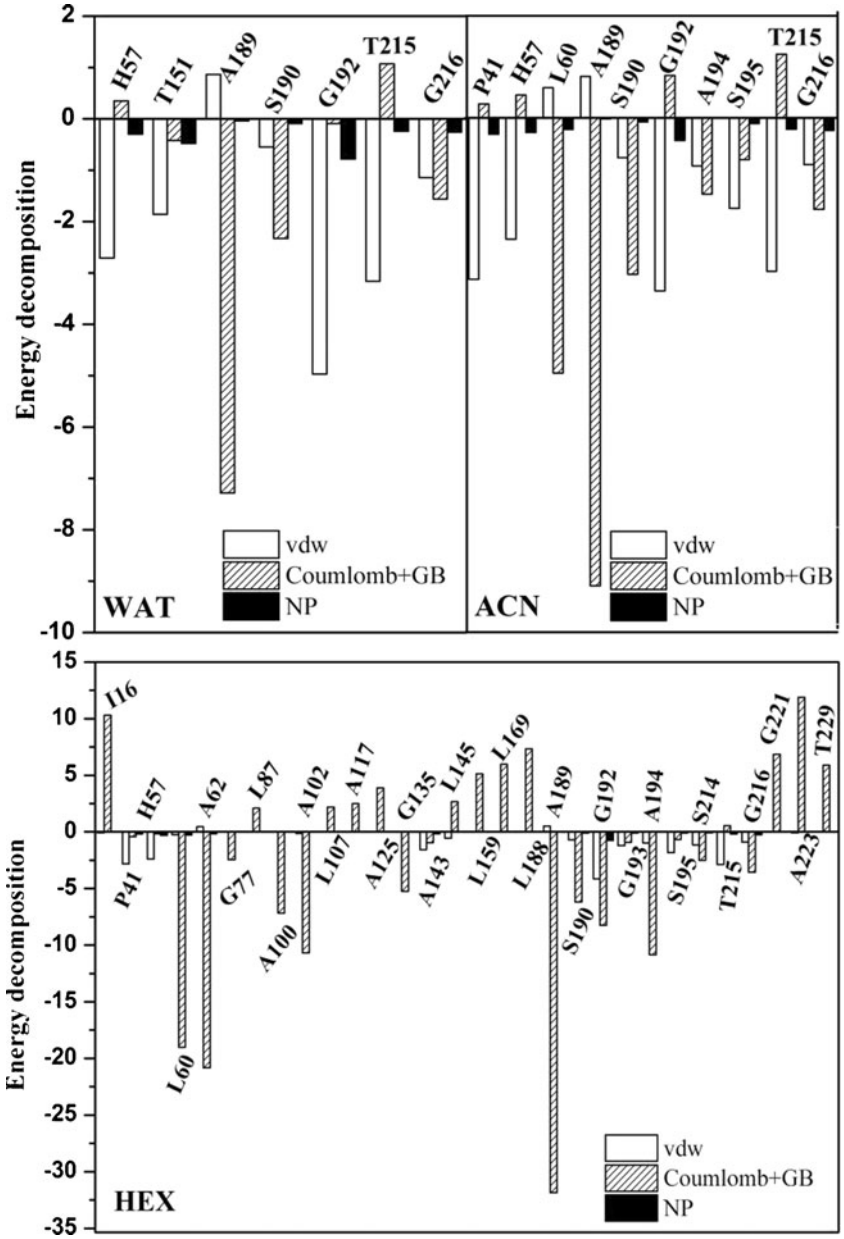

Fig. 12 Energy decomposition into contributions from van der Waals energy (vdw), the sum of Coulombic interaction and the polar solvation energy (Coulomb+GB), and the nonpolar term of solvation energy (NP) for the residues whose absolute value of binding energy was greater than 2.0 kcal mol^{-1}. WAT denotes aqueous solution. ACN denotes acetonitrile media and HEX denotes hexane media

between the two residues and the substrate instead display dominant contributions to the binding.

In acetonitrile media, those residues, which are observed to favor the binding in aqueous solution, still retain their favorable contribution (see Fig. 11 and Fig. 13). However there are some changes in their contribution extents, as reflected by the differences in these energy values between aqueous solution and acetonitrile media (see Fig. 11 and supporting information Table S4). Besides these residues above, the residues Phe41, Lys60 and the two residues located in the oxyanion hole (viz., Asp194 and Ser195) also significantly exhibit favorable contribution to the binding energy in acetonitrile media, which are observed to only play a minor role in favoring the binding in aqueous solution. The Phe41 and Ser195 residues favor the interaction mainly through van der Waals interactions while the Lys60 and Asp194 residues contribute to the binding affinity mainly through the electrostatic interactions (see Fig. 12). In addition, significantly favorable contribution from the Tyr151 residue observed in aqueous solution is largely weakened by acetonitrile media. These observations further confirm that the residues in the binding interface occur to some extent reorientations in acetonitrile media, compared with aqueous solution.

Surprisingly, the favorable or unfavorable contributions from the residues are observed to be much larger in the non-polar hexane solvent than those in the polar water and acetonitrile media. Some residues favoring the binding in the polar solvents, for example the Asp189, Ser190, Gln192 and Gly216 residues, still preserve their favorable contribution in the non-polar hexane (see Fig. 11 and Fig. 13), but with much

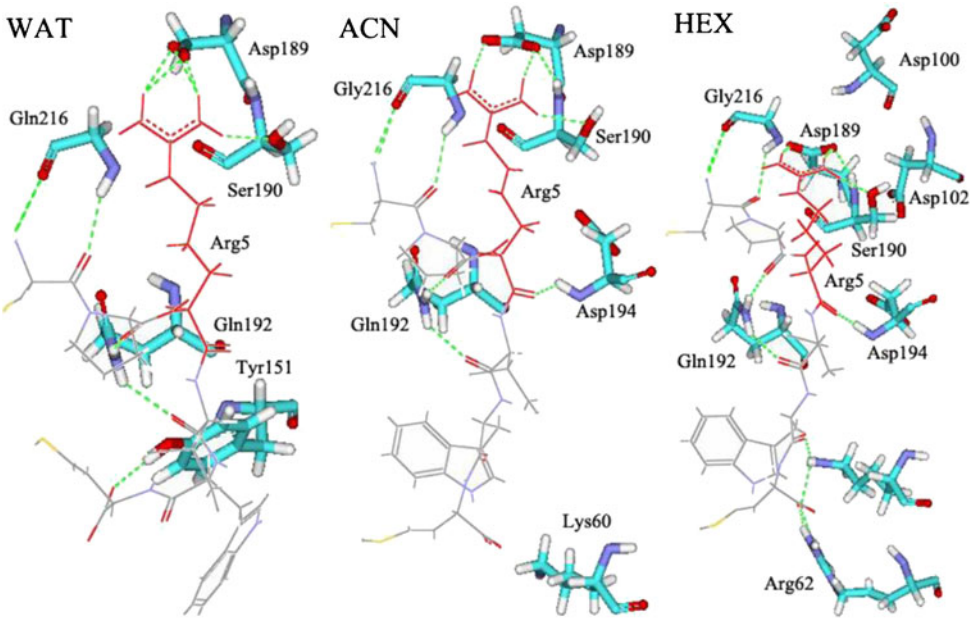

Fig. 13 Representative interactions between some higher energy residues in the binding interface ($|\Delta Gi| \geq 2$ kcal mol^{-1}) and the substrate in aqueous (WAT), acetonitrile (ACN) and hexane (HEX) solution. Line and stick denotes substrate and higher energy residues of trypsin, respectively. The green dotted line denotes hydrogen bond. The red line denotes the Arg5 residue with positive charge of the substrate

larger extent. Besides these residues, some other residues, which do not present favorable contribution to the binding in aqueous solution and acetonitrile media, are observed to strongly favor the binding in hexane media, further confirming the changes in the conformation of binding surface induced by the hexane solvent. For instance, the Arg62, Asp100 and Asp102 residues significantly contribute to the binding energy mainly through electrostatic interactions (see Fig. 12 and supporting information Table S5) in hexane media, resulting from the H-bonding of the residue Arg62 with the substrate and the Coulombic interaction between the negative charge of the Asp100 and Asp102 residues and the positive charge of the Arg5 residue of the substrate (also see Fig. 13). On a whole, in three media, the favorable residues contribute to the binding affinity mainly through electrostatic interactions, consistent with the analysis of the total binding energy above. Furthermore, the effects are much stronger in non-polar hexane media than those in the other two polar solvents. Not surprisingly, the polarity of water and acetonitrile solvent should weaken the electrostatic interactions between the favorable residues and the substrate, as confirmed by our previous work [82]. Thus, the favorable residues could exhibit much larger contribution to the binding affinity in hexane media than in the two polar solvents. As a result, the total binding energy between the enzyme and the substrate is observed to be largest in the non-polar hexane solvent.

Conclusions

Using 100 ns MD simulations, we studied the trypsin with or without a six-amino-acid peptide bound in aqueous solution and the two organic media (acetonitrile and hexane) with the inclusion of 135 crystal waters. The effects of the organic solvent and the substrate binding on the structure of enzyme, solvent distribution, intra-protein and protein-substrate interactions in non-aqueous media are studied by comparison of the results from these systems.

The results indicate that the trypsin in organic media has larger deviation from the crystal structure than that in aqueous solution either for the free state or the complexed form, and the deviation is more significant in non-polar hexane than that in polar acetonitrile media. In the free enzyme system, the S1 pocket in hexane media is significantly different from those in polar water and acetonitrile solvents, presenting a fully closed state while it exhibits open states in the two polar solvents. The observations suggest that there are to some extent changes in the specificity of the enzyme induced by the non-polar organic solvent. However, no large difference in the S1 pocket structure is observed for the complexed enzyme between the three media, which may stem from the substrate stabilization. The observation implies that the six-amino-acid peptide substrate may induce a fit binding mechanism in non-polar hexane media, which was not experimentally observed in the polar water and acetonitrile solvents. Similarly, the stabilization effect from the substrate binding is also observed for the overall structure of the enzyme in the polar solvents (water and acetonitrile). In contrast, the substrate could destabilize the native structure of the enzyme in the non-polar hexane media.

Although no unwinding or denaturation of the enzyme is observed in the organic solvents within the scale of the simulation time, the enzyme exhibits more compact structure in the organic solvents than aqueous solution, especially for hexane media. In addition, the organic solvents also induce some reorientations of the hydrophobic and hydrophilic side chains. Accordingly, a decrease in the SASA is observed in the organic media and the reduced extent is larger in hexane media than

acetonitrile one. The substrate binding and the decrease of solvent polarity lead to an increase in the hydrophobic surface while a reverse trend is observed for the hydrophilic surface area. However, the solvent and the substrate only play a minor role in the secondary structure of the enzyme. In addition, the substrate did not play an observation role in the distribution of the solvent around the protein surface. The crystal water in the organic solvents is mainly clustered on the protein surface, not diffusing into the bulk organic solvent. However, the acetontrile solvent exhibits stronger ability to strip off the bound water from the protein surface than the hexane one. The catalytic H-bond network and the key distance between the Ser195 residue and the substrate seem to appear more productive Mixchaeils complex conformation in the organic solvents (viz., acetonitrile and hexane) than aqueous solution. The results may partly stem from more compact structure of the enzyme caused by the organic solvent and the absence of perturbation of solvent molecules in the active site owing to the substrate's expelling.

The binding free energy indicates that the non-polar hexane solvent could yield a much more favorable binding between the substrate and the enzyme than the polar water and acetonitrile solvents. Further analysis reveals that the four residues of the S1 pocket (viz., Asp189, Ser190, Gln192 and Gly216) keep favorable affinity to the binding in the three media, displaying an insensitivity to the solvent variation. However, the organic solvents cause to some extent changes in the number and the type of residues favoring the binding as well as the contribution extent, especially for the non-polar hexane. The favorable residues contribute to the binding affinity mainly through electrostatic interactions while the polar solvent could significantly weaken the electrostatic interactions. As a result, the binding strength between the substrate and the enzyme is much larger in the non-polar hexane solvent than those in the polar water and acetonitrile ones. However, as confirmed by experiments [1, 5, 10, 64], the enzymatic activity is typically several orders of magnitude lower in organic media compared with aqueous solution. Based on that fact, it seems that the organic solvent induced higher stability of the catalytic H-bond network, shorter key distance associated with the catalytic reaction and stronger substrate binding strength in the initial binding steps may play no dominant role in influencing the enzymatic activity. Some other changes induced by the solvent in the initial (for example protein overall structure variations) and successive reaction steps may corporately contribute to the drop of enzymatic activity in organic solvents.

On a whole, the work provides a systematic study on the substrate induced structural and dynamics changes for the trypsin in the polar and non-polar organic media and the effects of the solvents on the substrate binding. Some new and different observations from previous free enzyme systems are obtained in the work, which could help us to better understand the structure and the function of enzyme in the initial reaction phase of non-aqueous enzymatic catalysis.

Acknowledgments This project is supported by the National Science Foundation of China (Grant No. 20973115, 21273154, U1230121) and the International Science and Technology Cooperation Foundation of Sichuan province (Grant No. 2011H0003).

References

1. Klibanov AM (2001) Improving enzymes by using them in organic solvents. Nature 409(6817):241–246
2. Halling PJ (2004) What can we learn by studying enzymes in non-aqueous media? Philosophical transactions of the Royal Society of London. Series B, Biological sciences 359(1448):1287–1297
3. Hudson EP, Eppler RK, Clark DS (2005) Biocatalysis in semi-aqueous and nearly anhydrous conditions. Curr Opin Biotechnol 16(6):637–643
4. Mansfeld J, Ulbrich-Hofmann R (2007) The stability of engineered thermostable neutral proteases from Bacillus stearothermophilus in organic solvents and detergents. Biotechnol Bioeng 97(4):672–679
5. Serdakowski AL, Dordick JS (2008) Enzyme activation for organic solvents made easy. Trends Biotechnol 26(1):48–54
6. Zaks A, Klibanov AM (1988) Enzymatic catalysis in nonaqueous solvents. J Biol Chem 263(7):3194–3201
7. Eppler RK, Komor RS, Huynh J, Dordick JS, Reimer JA, Clark DS (2006) Water dynamics and salt-activation of enzymes in organic media: Mechanistic implications revealed by NMR spectroscopy. Proc Natl Acad Sci USA 103(15):5706–5710
8. Affleck R, Xu ZF, Suzawa V, Focht K, Clark DS, Dordick JS (1992) Enzymatic catalysis and dynamics in low-water environments. Proc Natl Acad Sci USA 89(3):1100–1104
9. Watanabe K, Yoshida T, Ueji S (2004) The role of conformational flexibility of enzymes in the discrimination between amino acid and ester substrates for the subtilisin-catalyzed reaction in organic solvents. Bioorg Chem 32(6):504–515
10. Clark DS (2004) Characteristics of nearly dry enzymes in organic solvents: implications for biocatalysis in the absence of water. Philosophical Transactions Of the Royal Society Of London Series B-Biological Sciences 359(1448):1299–1307
11. Colombo G, Carrea G (2002) Modeling enzyme reactivity in organic solvents and water through computer simulations. J Biotechnol 96(1):23–33
12. Zheng YJ, Ornstein RL (1996) A molecular dynamics and quantum mechanics analysis of the effect of DMSO on enzyme structure and dynamics: Subtilisin. J Am Chem Soc 118(17):4175–4180
13. Soares CM, Teixeira VH, Baptista AM (2003) Protein structure and dynamics in nonaqueous solvents: Insights from molecular dynamics simulation studies. Biophys J 84(3):1628–1641
14. Micaelo NM, Teixeira VH, Baptista AM, Soares CM (2005) Water dependent properties of cutinase in nonaqueous solvents: a computational study of enantioselectivity. Biophys J 89(2):999–1008
15. Peters GH, vanAalten DMF, Edholm O, Toxvaerd S, Bywater R (1996) Dynamics of proteins in different solvent systems: analysis of essential motion in lipases. Biophys J 71(5):2245–2255
16. Yang L, Dordick JS, Garde S (2004) Hydration of enzyme in nonaqueous media is consistent with solvent dependence of its activity. Biophys J 87(2):812–821
17. Micaelo NM, Soares CM (2007) Modeling hydration mechanisms of enzymes in nonpolar and polar organic solvents. FEBS J 274(9):2424–2436
18. Zheng YJ, Ornstein RL (1996) Molecular dynamics of subtilisin Carlsberg in aqueous and nonaqueous solutions. Biopolymers 38(6):791–799
19. Rezaei-Ghaleh N, Amininasab M, Nemat-Gorgani M (2008) Conformational changes of alpha-chymotrypsin in a fibrillation-promoting condition: a molecular dynamics study. Biophys J 95(9):4139–4147

20. Diaz-Vergara N, Pineiro A (2008) Molecular dynamics study of triosephosphate isomerase from Trypanosoma cruzi in water/decane mixtures. J Phys Chem B 112(11):3529–3539
21. Li C, Tan T, Zhang H, Feng W (2010) Analysis of the conformational stability and activity of Candida antarctica lipase B in organic solvents: insight from molecular dynamics and quantum mechanics/simulations. J Biol Chem 285(37):28434–28441
22. Zhu L, Yang W, Meng YY, Xiao X, Guo Y, Pu X, Li M (2012) Effects of organic solvent and crystal water on gamma-chymotrypsin in acetonitrile media: observations from molecular dynamics simulation and DFT calculation. J Phys Chem B 116(10):3292–3304
23. Wu R, Xie H, Cao Z, Mo Y (2008) Combined quantum mechanics/molecular mechanics study on the reversible isomerization of glucose and fructose catalyzed by Pyrococcus furiosus phosphoglucose isomerase. J Am Chem Soc 130(22):7022–7031
24. Yang M-J, Pang X-Q, Zhang X, Han K-L (2011) Molecular dynamics simulation reveals preorganization of the chloroplast FtsY towards complex formation induced by GTP binding. J Struct Bio 173(1):57–66
25. Da L-T, Wang D, Huang X (2012) Dynamics of pyrophosphate ion release and its coupled trigger loop motion from closed to open state in RNA polymerase II. J Am Chem Soc 134(4):2399–2406
26. Lousa D, Baptista AM, Soares CM (2011) Structural determinants of ligand imprinting: a molecular dynamics simulation study of subtilisin in aqueous and apolar solvents. Protein Sci 20(2): 379–386
27. de Groot BL, Hayward S, van Aalten DMF, Amadei A, Berendsen HJC (1998) Domain motions in bacteriophage T4 lysozyme: a comparison between molecular dynamics and crystallographic data. Proteins 31(2):116–127
28. vanAalten DMF, Conn DA, deGroot BL, Berendsen HJC, Findlay JBC, Amadei A (1997) Protein dynamics derived from clusters of crystal structures. Biophys J 73(6):2891–2896
29. Mello LV, de Groot BL, Li SL, Jedrzejas MJ (2002) Structure and flexibility of Streptococcus agalactiae hyaluronate lyase complex with its substrate - Insights into the mechanism of processive degradation of hyaluronan. J Biol Chem 277(39):36678–36688
30. Ishida T, Kato S (2003) Theoretical perspectives on the reaction mechanism of serine proteases: the reaction free energy profiles of the acylation process. J Am Chem Soc 125(39):12035–12048
31. Hedstrom L (2002) Serine protease mechanism and specificity. Chem Rev 102(12):4501–4523
32. Lousa D, Cianci M, Helliwell JR, Halling PJ, Baptista AM, Soares CM (2012) Interaction of counterions with subtilisin in acetonitrile: insights from molecular dynamics simulations. J Phys Chem B 116(20):5838–5848
33. Wedberg R, Abildskov J, Peters GH (2012) Protein dynamics in organic media at varying water activity studied by molecular dynamics simulation. J Phys Chem B 116(8):2575–2585
34. Kim J, Clark DS, Dordick JS (2000) Intrinsic effects of solvent polarity on enzymic activation energies. Biotechnol Bioeng 67(1):112–116
35. Suzawa VM, Khmelnitsky YL, Giarto L, Dordick JS, Clark DS (1995) Suspended and immobilized chymotrypsin in organic media: structure-function relationships revealed by electron spin resonance spectroscopy. J Am Chem Soc 117(32):8435–8440
36. Garcia S, Vidinha P, Arvana H, Gomes da Silva MDR, Ferreira MO, Cabral JMS, Macedo EA, Harper N, Barreiros S (2005) Cutinase activity in supercritical and organic media: water activity, solvation and acid-base effects. J Supercrit Fluids 35(1):62–69
37. Gupta MN, Roy I (2004) Enzymes in organic media - Forms, functions and applications. Eur J Biochem 271(13):2575–2583
38. Zaks A, Klibanov AM (1988) The effect of water on enzyme action in organic media. J Biol Chem 263(17):8017–8021
39. Rupley JA, Careri G (1991) Protein hydration and function. Adv Protein Chem 41:37–172
40. Huang Q, Liu S, Tang Y (1993) Refined 1·6 Å resolution crystal structure of the complex formed between porcine β-trypsin and MCTI-A, a trypsin inhibitor of the squash family: detailed comparison with bovine β-trypsin and its complex. J Mol Biol 229(4):1022–1036
41. Walter J, Steigemann W, Singh TP, Bartunik H, Bode W, Huber R (1982) On the disordered activation domain in trypsinogen: chemical labelling and low-temperature crystallography. Acta Crystallographica Section B 38(5):1462–1472
42. Case DA D, Cheatham TE III, Simmerling CL, Wang J, Duke RE, Luo R, Crowley M, Walker RC, Zhang W, Merz KM, Wang B, Hayik S, Roitberg A, Seabra G, Kolossvary I, Wong KF, Paesani F, Vanicek J, Wu X, Brozell SR, Steinbrecher T, Gohlke H, Yang L, Tan C, Mongan J, Hornak V, Cui G, Mathews DH, Seetin MG, Sagui C, Babin V, Kollman PA (2010)Amber12; University of California: San Francisco
43. Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, Shaw DE (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins 78(8):1950–1958
44. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79(2):926–935
45. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA Jr, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas O, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) Gaussian 09 revision A01. Gaussian Inc, Wallingford, CT
46. Bayly CI, Cieplak P, Cornell W, Kollman PA (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. J Phy Chem 97(40):10269–10280
47. Speight, J. (2004) Lange's Chemistry Handbook 16th Edn. McGraw-Hill, New Tork
48. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81(8):3684–3690
49. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 23(3):327–341
50. 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(12):10089–10092
51. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A smooth patticile mesh ewald method. J Chem Phys 103(19):8577–8593
52. Humphrey W, Dalke A, Schulten K (1996) VMD: Visual molecular dynamics. J Mol Graph Model 14(1):33–38
53. Toba S, Hartsough DS, Merz KM (1996) Solvation and dynamics of chymotrypsin in hexane. J Am Chem Soc 118(27):6490–6498
54. Tejo BA, Salleh AB, Pleiss J (2004) Structure and dynamics of Candida rugosa lipase: the role of organic solvent. J Mol Model 10(5–6):358–366
55. Bordes F, Barbe S, Escalier P, Mourey L, Andre I, Marty A, Tranier S (2010) Exploring the conformational states and rearrangements of yarrowia lipolytica lipase. Biophys J 99(7):2225–2234

56. Nakagawa S, Yu HA, Karplus M, Umeyama H (1993) Active site dynamics of acyl-chymotrypsin. Proteins 16(2):172–194
57. Kisee H, Fujimoto K, Noritomi H (1988) Enzymatic reactions in aqueous-organic media. VI. Peptide synthesis by α-chymotrypsin in hydrophilic organic solvents. J Biotechnol 8(4):279–290
58. Čeěovský V, Jakubke H-D (1994) Acyl transfer reactions catalyzed by native and modified α-chymotrypsin in acetonitrile with low water content. Enzyme Microb Technol 16(7):596–601
59. Simon LM, Kotormán M, Garab G, Laczkó I (2001) Structure and activity of α-chymotrypsin and trypsin in aqueous organic media. Biochem Biophys Res Commun 280(5):1367–1371
60. Kuszewski J, Gronenborn AM, Clore GM (1999) Improving the packing and accuracy of NMR structures with a pseudopotential for the radius of gyration. J Am Chem Soc 121(10):2337–2338
61. Verde AV, Acres JM, Maranas JK (2009) Investigating the specificity of peptide adsorption on gold using molecular dynamics simulations. Biomacromolecules 10(8):2118–2128
62. Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22(12):2577–637
63. Topf M, Richards WG (2004) Theoretical studies on the deacylation step of serine protease catalysis in the gas phase, in solution, and in elastase. J Am Chem Soc 126(44):14631–14641
64. Klibanov AM (1997) Why are enzymes less active in organic solvents than in water? Trends Biotechnol 15(3):97–101
65. Zhou P, Tian F, Shang Z (2009) 2D depiction of nonbonding interactions for protein complexes. J Comput Chem 30(6):940–951
66. Lyne PD, Lamb ML, Saeh JC (2006) Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. J Med Chem 49(16):4805–4808
67. Rastelli G, Degliesposti G, Del Rio A, Sgobba M (2009) Binding estimation after refinement, a new automated procedure for the refinement and rescoring of docked ligands in virtual screening. Chem Biol Drug Des 73(3):283–286
68. Del Rio A, Baldi BF, Rastelli G (2009) Activity prediction and structural insights of extracellular signal-regulated kinase 2 inhibitors with molecular dynamics simulations. Chem Biol Drug Des 74(6):630–635
69. Lafont V, Armstrong AA, Ohtaka H, Kiso Y, Amzel LM, Freire E (2007) Compensating enthalpic and entropic changes hinder binding affinity optimization. Chem Biol Drug Des 69(6):413–422
70. Rastelli G, Del Rio A, Degliesposti G, Sgobba M (2010) Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. J Comput Chem 31(4):797–810
71. Gilson MK, Zhou H-X (2007) Calculation of protein-ligand binding affinities. Journal 36:21–42
72. Brown SP, Muchmore SW (2007) Rapid estimation of relative protein-ligand binding affinities using a high-throughput version of MM-PBSA. J Chem Inf Model 47(4):1493–1503
73. Wang JM, Morin P, Wang W, Kollman PA (2001) Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA. J Am Chem Soc 123(22):5221–5230
74. Rashidi M-R, Dehghany M, Dehghan G, Jouyban A, Faridi A (2013) Aldehyde oxidase activity and stability in water-miscible organic solvents. Appl Biochem Biotechnol 169(3):901–910
75. Singh P, Prakash R, Shah K (2012) Effect of organic solvents on peroxidases from rice and horseradish: prospects for enzyme based applications. Talanta 97:204–210
76. Laane C, Boeren S, Vos K, Veeger C (1987) Rules for optimization of biocatalysis in organic solvents. Biotechnol Bioeng 30(1):81–87
77. Deng N-J, Cieplak P (2009) Insights into affinity and specificity in the complexes of alpha-lytic protease and its inhibitor proteins: binding free energy from molecular dynamics simulation. PCCP 11(25):4968–4981
78. Gohlke H, Kiel C, Case DA (2003) Insights into protein–protein binding by binding free energy calculation and free energy decomposition for the Ras–Raf and Ras–RalGDS complexes. J Mol Biol 330(4):891–913
79. Noskov SY, Lim C (2001) Free energy decomposition of protein-protein interactions. Biophys J 81(2):737–750
80. Wang W, Kollman PA (2000) Free energy calculations on dimer stability of the HIV protease using molecular dynamics and a continuum solvent model. J Mol Biol 303(4):567–582
81. Gohlke H, Case DA (2004) Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. J Comput Chem 25(2):238–250
82. Chen A, Pu X, He S, Guo Y, Wen Z, Li M, Wong N-B, Tian A (2009) Solvent effects on isolated formamide and its monohydrated complex: observations from PCM study. New J Chem 33(8):1709–1719
83. Keskin O, Tsai CJ, Wolfson H, Nussinov R (2004) A new, structurally nonredundant, diverse data set of protein-protein interfaces and its implications. Protein Sci 13(4):1043–1055
84. Swapna LS, Bhaskara RM, Sharma J, Srinivasan N (2012) Roles of residues in the interface of transient protein-protein complexes before complexation. Sci Rep 2