Available online at www.sciencedirect.com
ScienceDirect
Journal of Molecular Graphics and Modelling 26 (2008) 813–823
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Computational alanine scanning and free energy decomposition for *E. coli* type I signal peptidase with lipopeptide inhibitor complex

Tong Li, Matheus Froeyen, Piet Herdewijn *

Laboratory for Medicinal Chemistry, Rega Institute for Medical Research, Minderbroedersstraat 10, B-3000 Leuven, Belgium

Received 16 November 2006; received in revised form 24 April 2007; accepted 28 April 2007
Available online 3 May 2007

Abstract

A thorough investigation of different roles of *Escherichia coli* type I signal peptidase residues binding to lipopeptide inhibitor has been performed by a combination of computational alanine scanning mutagenesis and free energy decomposition methods. PB and GB models are both used to evaluate the binding free energy in computational alanine scanning method and only GB model can be used to decompose the binding free energy on a per-residue basis. The regression analysis between the PB and GB model and also between the computational alanine scanning and free energy decomposition have been reported with a correlation coefficient of 0.96 and 0.83, respectively, which suggest they are both in fair agreement with each other. Moreover, the contribution components from van der Waals, electrostatic interaction, non-polar and polar energy of solvation, have been determined as well as the effects of backbones and side-chains. The results indicate that Lys145 is the most important residue for the binding but also acts as a general base, activating Ser90 to increase its nucleophility, recognizing and stabilizing the binding of lipopeptide inhibitor to the *E. coli* SPase. The hydroxyl group of Ser88 plays a key role for the binding of the inhibitor. Ser90 contributes more to the intramolecular interaction than to the intermolecular interaction. Tyr143 and Phe84 contribute larger van der Waals interaction energies, indicating that these residues can be important for the selection based on the shape of the inhibitors. The contributions from other several interfacial residues of the *E. coli* SPase are also analyzed. This study can be a guide for the optimization of lipopeptide inhibitors and future design of new therapeutic agents for the treatment of bacterial infections.

Keywords: Type I signal peptidase; Molecular dynamics; MM-PBSA/MM-GBSA; Structure-based drug design; Computational alanine scanning; Free energy decomposition

1. Introduction

Type I signal peptidase (SPase I) is a membrane-bound endopeptidase that functions by cleaving off the N-terminal signal peptides from secretory proteins in both prokaryotic and eukaryotic organisms [1,2]. The *Escherichia coli* SPase, a Gram-negative SPase, has been the best-known enzyme in this family. It contains two N-terminal transmembrane segments, a small cytoplasmic region and a C-terminal periplasmic catalytic region and belongs to a novel class of serine protease that utilizes the Ser-Lys dyad mechanism for peptide hydrolysis where Ser serves as the nucleophile and Lys serves as the general base [1–5].

Bacterial SPase I has become one of most attractive targets for the design of novel antimicrobial compounds. Inhibition of this enzyme leads to an accumulation of secretory proteins in the cell membrane and eventual cell death [1,2]. The first effective inhibitor of bacterial SPase I is a β-lactam, which binds covalently to the γ-oxygen of Ser90 and acts as time dependent irreversible inhibitor [6]. The unique non-covalently bound inhibitor, Arylomycin A_2, has been discovered only a few years ago. Arylomycin A_2 is a lipohexapeptide (D-MeSer-D-Ala-Gly-L-MeHpg-L-Ala-L-Tyr) with a 12-carbon atom branched fatty acid (iso-C12) attached via an amide bond to the amino terminus (Fig. 1) [7,8]. The X-ray structures of the catalytically active fragment of *E. coli* SPase apo-enzyme, acyl-enzyme and the lipopeptide bound enzyme have been solved recently, which gives us good templates to develop new inhibitors based on their structures [6,7,9].

Computational methods provide a molecular view of the structural and energetic consequences of mutations, and can address the origin of binding in terms of contributions from electrostatic and van der Waals interactions and changes in

* Corresponding author.
E-mail address: piet.herdewijn@rega.kuleuven.be (P. Herdewijn).

1093-3263/$ – see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2007.04.007

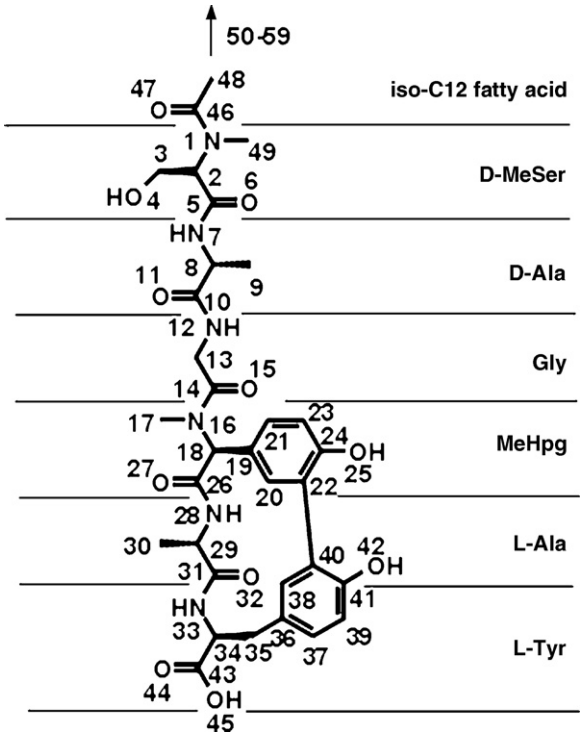

Fig. 1. Structure of Arylomycin A_2. MeSer is methylated Ser and MeHpg is *N*-methyl-4-hydroxyphenylglycine.

solvation [10]. The Molecular Mechanics Poisson Boltzmann Surface Area (MM-PBSA)/Molecular Mechanics Generalized Born Surface Area (MM-GBSA) approaches have recently become of interest in drug discovery for calculating binding affinities of biomolecular complexes, based on molecular dynamics (MD) of the given protein–ligand complex in explicit solvent [10–18]. Computational alanine scanning involving the MM-PBSA method [18–23] and free energy decomposition involving the MM-GBSA method [24,25] have been developed to investigate the binding modes in detail at the atomic level and also to estimate protein stabilities [26]. In the MM-PBSA approach, the binding free energy is estimated as the sum of the gas-phase energies, solvation free energies and entropic contributions, averaged over a series of snapshots from MD trajectories. The electrostatic contribution to the solvation term is calculated by solving the Poisson-Boltzmann (PB) equation. If the PB model is replaced by a generalized Born (GB) model, there comes the MM-GBSA method. The difference of the binding free energy can be obtained after the calculation of binding free energy of the wild-type and of the alanine mutants based on the MM-PBSA/MM-GBSA separately, which is called computation alanine scanning method. The GB model makes this variant attractive because it is much faster than the PB model and allows the decomposition of the electrostatic solvation free energy into atomic contributions in a straightforward manner. This model allows an easy and rapid free energy decomposition for the wild-type [27]. Computational alanine scanning and free energy decomposition methods complement each other and using them together may bypass their weakness, i.e. possible perturbations of the systems during virtual mutation and less accuracy to decompose the binding free energy by the MM-GBSA method.

In this work, we have analyzed the interaction between the *E. coli* SPase and Arylomycin A_2 using these two methods to discern the hotspot residues. Only relative binding energy (without entropy) was used to evaluate the relative importance of different residues, because the major thermodynamic factor is the reaction enthalpy presumably driven by hydrogen bonding interactions between the inhibitor and the signal peptidase binding site [7]. At the same time, the entropic contribution can be canceled when the relative binding free energies are calculated between the wild-type and mutants in binding to the same receptors, which has been demonstrated by Massova and Kollman [19].

2. Material and methods

2.1. Molecular dynamics simulations

The starting coordinates for SPase and lipopeptide inhibitor were obtained from the X-ray crystal structure of the *E. coli* type I SPase in complex with a lipopeptide inhibitor (Protein Data Bank [28] code 1T7D), determined at 2.47 Å resolution. This file includes the coordinates of the SPase dimer, MD was performed only on the first monomer which has 208 amino acids. All simulations were conducted by using the AMBER (version 8.0) program [29,30]. The ligand was prepared by using the antechamber suite [31] in the AMBER package. Atomic charges were derived with the AM1-BCC charge method [32]. Two parameter sets were used, the biomolecular force field ff03 [33] for the protein and general AMBER force field (GAFF) [34] for the inhibitor. The complex was soaked in a truncated octahedron box of TIP3P [34,35] water molecules with a margin of 10 Å along each dimension. Four Na^+ ions were added to neutralize the system. In summary, the system consists of 1 SPase monomer, 1 lipopeptide inhibitor, 4 Na^+ ions and 6928 water molecules. The system was minimized by 200 steps of steepest descent followed by 800 steps of conjugate gradient to remove the bad contacts in the crystal structure. During the minimization, a 12 Å non-bonded cutoff was applied.

The system was then heated from 0 to 300 K in 40 ps. A subsequent 1 ns production run was performed at a constant temperature of 300 K and a constant pressure of 1 atm. The time interval was set to 2 fs. The Particle Mesh Ewald (PME) method [36] was applied to calculate long-range electrostatics interactions. The SHAKE method [37] was applied to constrain all of the covalent bonds involving hydrogen atoms. Periodic boundary conditions were applied to all dimensions. No constraint was applied to either the protein or the ligand during MD simulation. Coordinates were saved every 0.4 ps for a total of 2500 snapshots.

2.2. MM-PBSA/MM-GBSA calculation

The binding free energies were calculated using the MM-PBSA/MM-GBSA method [38]. A total number of 250

snapshots were taken from the last 500 ps of the MD trajectory with an interval of 2 ps. The MM-PBSA method can be conceptually summarized as:

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

$$G = E_{\text{gas}} + G_{\text{sol}} - TS \quad (2)$$

$$E_{\text{gas}} = E_{\text{int}} + E_{\text{vdW}} + E_{\text{ele}} \quad (3)$$

$$G_{\text{sol}} = G_{\text{PB}} + G_{\text{np}} \quad (4)$$

$\Delta G_{\text{complex}}$, $\Delta G_{\text{protein}}$ and ΔG_{ligand} are the free energies of the complex, the protein and the ligand, respectively. Each of them is calculated by summing an internal energy in gas phase (E_{gas}), a solvation free energy (G_{sol}) and a vibrational entropy term ($-TS$) in Eq. (2). E_{gas} is standard force field energy, including internal energy as well as non-covalent van der Waals and electrostatic energies (Eq. (3)). The solvation free energy, G_{sol}, is calculated with a PB/SA model, which dissects solvation free energy as the sum of an electrostatic component (G_{PB}) and a non-polar component (G_{np}). The electrostatic component is calculated using the PBSA program with the default cavity radii from the AMBER prmtop file. The dielectric constant was set to 1 for the interior solute and 80 for the surrounding solvent. The LCPO method [39] was used to calculate the Solvent Accessible Surface Area (SASA) for the estimation of the non-polar solvation free energy (ΔG_{np}) using Eq. (5) with $\gamma = 0.00542$ kcal/mol Å^{-2} and $\beta = 0.92$ kcal/mol [40].

$$\Delta G_{\text{np}} = \gamma \text{SASA} + \beta \quad (5)$$

For the MM-GBSA method, G_{PB} is replaced by G_{GB}. The Hawkins, Cramer and Truhlar pairwise generalized Born model [41,42] is used with parameters described by Tsui and Case [43]. The LCPO method was used to calculate the SASA with $\gamma = 0.005$ kcal/mol Å^{-2} and $\beta = 0.00$ kcal/mol [40].

We have used a single molecular dynamics trajectory protocol, which can qualitatively estimate the free energy consequences of many mutations [19]. As a consequence, the contribution of internal energy to the binding energy is equal to zero. $-T\Delta S$ term has been neglected according to the reasons mentioned in Section 1. Therefore, we only calculated subtotal binding free energy ($\Delta G_{\text{subtotal}} = E_{\text{gas}} + \Delta G_{\text{sol}}$).

2.3. Computational alanine scanning

The computational alanine scanning method was applied to estimate the relative binding affinity of different SPase variants to the lipopeptide inhibitor. The alanine mutant structures were generated by altering the coordinates of the wild-type trajectory. This method involved deleting atoms and truncating the mutated residue at C_γ by replacing with a hydrogen atom [19]. All parameters in the topology files for the mutated residues were accordingly replaced by the alanine residue parameters. The same 250 snapshots taken from the last 500 ps of the MD trajectory with the time interval of 2 ps were used.

Key residues of the SPase, Glu82, Phe84, Gln85, Ile86, Ser88, Ser90, Met91, Ile101, Val132, Asp142, Tyr143, Ile144, Lys145 and Asp280, were chosen from the putative active site [6,9]. Proline residues were not mutated since their backbone conformations differ significantly from the alanine residue [19]. As a result, the Pro83 and Pro87 which belong to the active site are not selected.

Explicit MD simulations for the mutated proteins F84A, S88A, Y143A, I144A and L145A with the lipopeptide inhibitor have been carried out separately to conform whether there is a global conformation change after mutation which can cause the largest changes in binding free energy for the five residues. The starting coordinates of alanine mutant structures were also generated by altering the coordinates of the wild-type trajectory. The parameters of the MD are the same as those of the wild-type MD simulation.

2.4. Binding free energy decomposition

Any virtual mutation may lead to perturbations that transcend specific localized interactions in the binding interface [24]. Thus, evaluation of the contribution of each residue to the binding free energy has been made at the atomic level by means of free energy decomposition. Although, it is possible to decompose the energy through PB calculation, it is rather time-consuming and computationally expensive. Alternatively, a generalized born model has been introduced to determine the electrostatic contribution to the solvation energy [24,43,44,45].

$$E_{\text{pol}}^{\text{GB}} = -\frac{1}{2}\left(1 - \frac{e^{-\kappa f_{\text{GB}}}}{\varepsilon_\omega}\right)\sum_{ij}\frac{q_i q_j}{f_{\text{GB}}} \quad (6)$$

q_i and q_j are atomic partial charges, ε_ω the solvent dielectric constant, κ the Debye-Huckel screening parameter and the double sum runs over all pairs of atoms. Here, ε_ω and κ have been set to 80 and 0, respectively. f_{GB} has been defined as follows:

$$f_{\text{GB}} = \left[r_{ij}^2 + \alpha_i\alpha_j \exp\left(\frac{-r_{ij}^2}{4\alpha_i\alpha_j}\right)\right]^{1/2} \quad (7)$$

r_{ij} is the distance between atom i and j, α_i and α_j are the effective Born radius of atom i and j. The contribution of atom i to the electrostatic free energy is obtained by:

$$E_{\text{elec}}^{i} = -\frac{1}{2}\sum_{j}\left(1 - \frac{e^{-\kappa f_{\text{GB}}}}{\varepsilon_\omega}\right)\frac{q_i q_j}{f_{ij}^{\text{GB}}(r_{ij})} + \frac{1}{2}\sum_{j\neq i}\frac{q_i q_j}{r_{ij}} \quad (8)$$

The calculation of internal energies has been canceled because ΔE_{int} is equal to zero under the assumption of single trajectory simulation. The SASA per atom was estimated with a recursive algorithm and in every recursion step, each triangular face of the polyhedron is divided into four pieces of equal size, hence a better approximation of a sphere is obtained [24]. The corresponding non-polar solvation energy per atom can be obtained based on the corresponding SASA. Moreover, one half of the pairwise van der Waals interaction energy is added to E^i to which the atom belongs that is part of the interaction pair [24]. Again, entropy terms are also neglected [24]. The total relative binding free energy of a

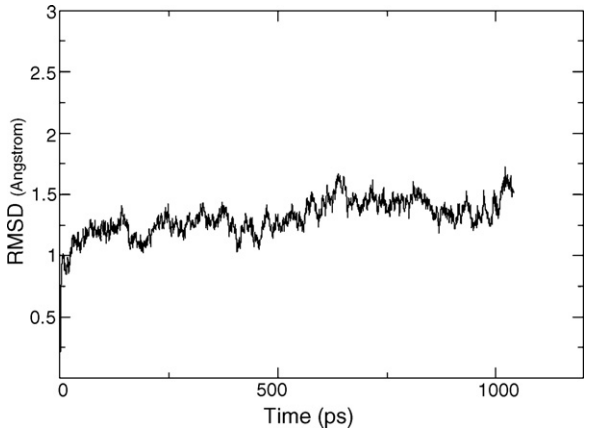

Fig. 2. RMSDs for the backbone atoms during the MD simulations of the *E. coli* type I signal peptidase with lipopeptide inhibitor complex.

given residue can be obtained by summing the contribution of each atom of this residue. The separate contribution of backbones and side-chains can be organized from the relevant atoms. In all, 250 snapshots which were the same as those in computational alanine scanning method were used to average the energy term.

3. Result and discussion

3.1. Molecular stability and structure from MD

To assess the stability of the MD trajectories, the backbone atoms root-mean-square deviations (RMSDs) from the X-ray structure of the *E. coli* SPase with lipopeptide inhibitor complex have been plotted in Fig. 2. During the first 40 ps, a sharp rise was observed and then the function keeps stable until a little increase occurs at about 500 ps. After the first 500 ps of production run, the RMSD values converge to a lower value about 1.3 Å which reflects the stability of the dynamics. Fig. 3 shows the superimposition of the average structure of the last 500 ps snapshots with the X-ray structure of this complex. Only the interfacial amino acids of the enzyme have been displayed. This figure shows that the SPase and inhibitor have almost the same conformation as the X-ray structure. The simulated structure generally reproduces the experimentally determined one. Based on the observation, we selected the last 500 ps snapshots to be used in the computational alanine scanning and free energy decomposition analysis.

3.2. Computational alanine scanning of SPase binding-site residues

In order to validate that there are no global conformation changes upon Ala mutation, explicit MD simulations for the mutated proteins F84A, S88A, Y143A, I144A and L145A with the lipopeptide inhibitor have been carried out. The backbone atoms RMSDs of the five complexes compared to the starting X-ray structure have been shown in Fig. 4. As can be seen, the RMSD from the X-ray structure is around 1.75 Å for the Ile144Ala system and the rest four RMSDs are between 1.2 and

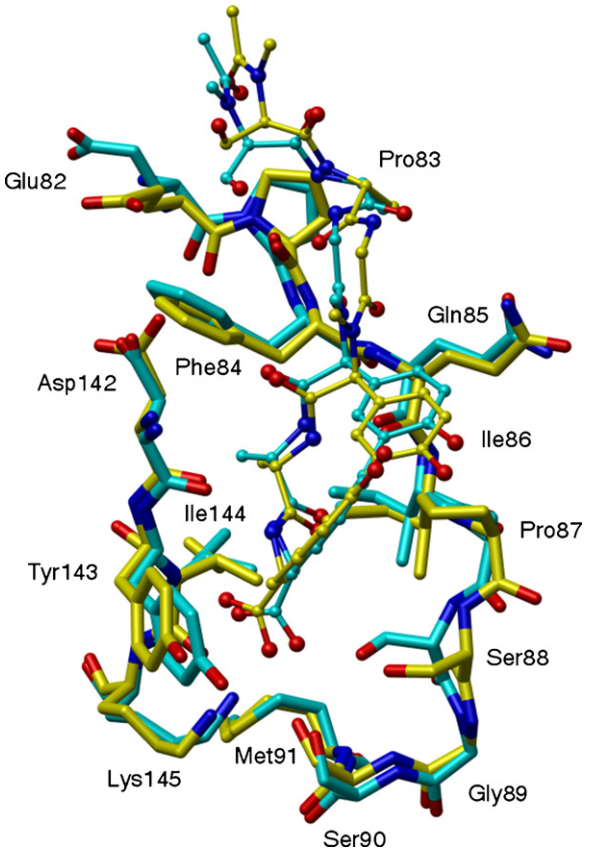

Fig. 3. The average structure of the last 500 ps molecule dynamics snapshots of the *E. coli* SPase with the lipopeptide inhibitor complex superimposed with its X-ray structure. The inhibitors are shown in ball-and-stick representation. The residues from the enzyme are shown in stick. Carbon atoms in the X-ray structure and MD snapshots are colored in yellow and cyan, respectively. Oxygen is red and Nitrogen is blue. Figure is created using Chimera [46].

1.5 Å. Thus, the five mutation systems show a satisfying stability during the MD simulation.

The calculated relative binding free energy and contributions of vdW, electrostatic interaction and solvation energy using the single trajectory MM-PBSA/MM-GBSA method are listed in Table 1. The relative binding free energy ($\Delta\Delta G_{\text{subtotal}}$) is −53.68 kcal/mol by MM-PBSA calculation ($\Delta\Delta G_{\text{subtotal,PB}}$)

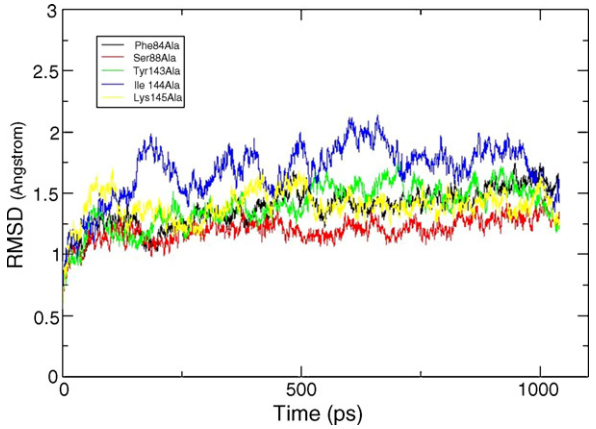

Fig. 4. RMSDs of the backbone atoms of the five simulated models compared to the starting X-ray structure.

Table 1
Energy components and free energy of binding (kcal/mol) for the *E. coli* SPase with lipopeptide inhibitor complex

	Complex		Receptor		Ligand		Delta	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
$E_{electrostatic}$	−5525.58	45.21	−5434.07	45.12	−8.54	4.54	−82.98	9.12
E_{vdW}	−892.80	23.81	−863.16	23.03	18.56	3.30	−48.20	3.52
E_{gas}	−1734.74	55.45	−1743.36	55.54	139.79	6.93	−131.18	8.95
$G_{non\text{-}polar,PB}$	53.59	0.80	54.17	0.81	4.45	0.04	−5.03	0.12
G_{PB}	−2013.37	41.11	−1959.38	42.79	−136.52	3.95	82.53	8.63
$G_{subtotal,PB}$	−3694.52	43.46	−3648.57	42.87	7.72	5.99	−53.68	6.64
$G_{non\text{-}polar,GB}$	48.59	0.74	49.12	0.75	3.25	0.04	−3.79	0.11
G_{GB}	−2123.36	37.26	−2067.33	37.95	−133.66	3.62	77.63	6.51
$G_{subtotal,GB}$	−3809.52	41.17	−3761.57	40.38	9.39	6.05	−57.34	3.78

$G_{subtotal,PB/GB} = E_{gas} + G_{solvation}$, $G_{solvation} = G_{non\text{-}polar,PB/GB} + G_{PB/GB}$.

and −57.34 kcal/mol by MM-GBSA calculation ($\Delta\Delta G_{subtotal,GB}$). There is a 3.66 kcal/mol difference in $\Delta\Delta G_{subtotal}$ between the two methods. The difference completely originates from the calculation of the contributions to the solvation energy which is slightly higher in the MM-PBSA calculation than in the MM-GBSA calculation.

To elucidate the key residues in the binding site of the enzyme and the most favorable interaction modes, computational alanine scanning was performed. This method depends on the assumption that local changes of the protein do not influence the whole conformation of the complex significantly. The two important S1 and S3 pockets have been proposed previously, which are formed by the amino acids: Ile86, Pro87, Ser88, Ser90, Met91, Leu95, Tyr143, Ile144, Lys145 and Phe84, Ile86, Ile101, Val132, Asp142, Ile144, respectively [6,9]. All the residues of the S1 and S3 pockets and other interfacial residues of SPase have been mutated to alanine, except for the proline residue whose backbone is remarkably different from that of alanine. Table 2 shows the results of the computational alanine scanning mutagenesis for 14 residues of SPase. Because the binding free energy difference was calculated according to the following equation: $\Delta\Delta G = \Delta G_{wild\text{-}type} - \Delta G_{mutant}$, the positive and negative values indicate the unfavorable and favorable contributions, respectively.

The binding affinity is decreased by more than 3 kcal/mol when four residues—Ser88, Tyr143, Ile144 and Lys145 are mutated to alanine according to the MM-PBSA calculation and three residues—Phe84, Tyr143 and Lys145 according to the MM-GBSA calculation. Fig. 5 shows the regression between the calculated $\Delta\Delta G_{subtotal,PB}$ and $\Delta\Delta G_{subtotal,GB}$ for the 14 residues obtained by the computational alanine scanning. As can be seen, the correlation between the two models is quite good. The correlation coefficient is 0.96. But there is still some striking disagreement between the $\Delta\Delta G_{subtotal,PB}$ and $\Delta\Delta G_{subtotal,GB}$ of Phe84Ala, Ile144Ala. Larger standard deviations have been obtained by the PB than by the GB model. For example, the standard deviations of $\Delta\Delta G_{subtotal,PB}$ for Lys145A and Ile144A are 6.45 and 3.44 kcal/mol, only 1.62 and 0.47 kcal/mol of $\Delta\Delta G_{subtotal,GB}$. It seems that the PB model is more sensitive to the atomic coordinates than the GB model during the calculation of contributions to solvation energy. In the following, we analyze some SPase residues that are successfully predicted to be important for the binding of lipopeptide inhibitor based on the computational alanine scanning. Fig. 6 shows the interaction between the lipopeptide and SPase from the average structure of the MD. Analysis of intermolecular and intramolecular hydrogen bonds of the last 500 ps trajectories has been revealed in Table 3.

Phe84. The side-chain phenyl ring of this residue makes a favorable −3.67 kcal/mol contribution to the binding free energy, coming from the vdW contact with the lipopeptide inhibitor (Table 2). There is some disagreement between $\Delta\Delta G_{PB}$ and $\Delta\Delta G_{GB}$, which suggests the PB model is more sensitive to the atomic coordinates than the GB model. In fact, the direction and position of the side-chain of Phe84 have proved quite variable in different situations [7], which may be another reason. The larger unfavorable $\Delta\Delta G_{PB}$ can compensate the favorable electrostatic and vdW interaction which are only partially compensated by the less unfavorable $\Delta\Delta G_{GB}$. However, it is clear that vdW contact of this residue with the lipopeptide inhibitor plays an important role for the binding.

Ser88. This residue has a significant effect on the binding with a −4.28 kcal/mol reduction in $\Delta\Delta G_{subtotal,PB}$ and −2.47 kcal/mol reduction in $\Delta\Delta G_{subtotal,GB}$ (Table 2). Taking the standard deviation into account, they are in reasonable agreement. The favorable contribution is provided mainly by electrostatic interaction. The hydrogen bond of Ser88-OGH···O45-Arylomycin A2 with a high occupancy 82.88% and close average distance 2.627 Å (Table 3) is interrupted after alanine mutation, which can explain most of the loss of the binding free energy. The intramolecular interaction between Ser88 and Ser90 is also important because the hydrogen bond of Ser88-O···HO-Ser90 has a 52.40% occupancy (Table 3). The contribution of the side-chain Ser88 has been investigated through in vivo and in vitro characterization of Ala-, Cys- and Thr-substituted mutants [4]. In their work, Pro-OmpA Nuclease A (PONA) has been used as a substrate of SPase. The experimental $\Delta\Delta G_{binding}$ of Ser88Ala is −5.2 kcal/mol. They confirmed the hydroxyl group of Ser88 plays a key role in the binding of substrate, which is also found by the computation alanine scanning method in our present work.

Ser90. This residue does not contribute substantially to the binding free energy. The duration of the hydrogen bond of Ser90-OGH···O45-Arylomycin A2 is only 3.44%. But the experimental Ser90Ala have very low activity [47]. According

Table 2
Computational alanine scanning results (kcal/mol) for the *E. coli* SPase with lipopeptide inhibitor complex

Contribution	Glu82Ala	Phe84Ala	Gln85Ala	Ile86Ala	Ser88Ala
$\Delta\Delta E_{electrostatic}$	16.73(2.36)	−1.53(0.90)	1.22(0.76)	−0.44(0.18)	−17.11(4.33)
$\Delta\Delta E_{vdW}$	−1.42(0.42)	−3.67(0.56)	−1.56(0.43)	−1.20(0.44)	2.19(1.66)
$\Delta\Delta E_{gas}$	15.31(2.47)	−5.21(1.10)	−0.34(0.86)	−1.64(0.53)	−14.92(3.37)
$\Delta\Delta G_{non\text{-}polar,PB}$	−0.06(0.02)	−0.22(0.03)	−0.12(0.04)	−0.04(0.01)	−0.05(0.03)
$\Delta\Delta G_{PB}$	−16.00(2.67)	5.15(2.51)	−0.35(0.97)	0.72(1.34)	10.69(2.50)
$\Delta\Delta G_{subtotal,PB}$	−0.76(2.05)	−0.28(2.18)	−0.81(0.97)	−0.95(1.45)	**−4.28(2.32)**
$\Delta\Delta G_{non\text{-}polar,GB}$	−0.06(0.02)	−0.20(0.03)	−0.11(0.04)	−0.03(0.01)	−0.04(0.03)
$\Delta\Delta G_{GB}$	−15.81(2.38)	2.02(0.95)	−0.55(0.64)	0.15(0.19)	12.49(2.23)
$\Delta\Delta G_{subtotal,GB}$	−0.56(0.48)	**−3.39(0.67)**	−1.01(0.33)	−1.51(0.53)	−2.47(1.40)

Contribution	Ser90Ala	Met91Ala	Leu95Ala	Ile101Ala	Val132Ala
$\Delta\Delta E_{electrostatic}$	−0.40(3.12)	−0.12(0.53)	−0.01(0.13)	−0.15(0.06)	1.74(0.15)
$\Delta\Delta E_{vdW}$	−0.27(0.22)	−0.16(0.05)	−0.06(0.02)	−0.24(0.06)	−0.55(0.15)
$\Delta\Delta E_{gas}$	−0.67(2.99)	−0.28(0.52)	−0.07(0.13)	−0.40(0.10)	1.20(0.17)
$\Delta\Delta G_{non\text{-}polar,PB}$	−0.03(0.01)	0.00(0.00)	0.01(0.01)	0.01(0.01)	−0.01(0.01)
$\Delta\Delta G_{PB}$	1.25(3.09)	−0.47(1.56)	−2.61(1.33)	−0.46(1.65)	−0.41(1.97)
$\Delta\Delta G_{subtotal,PB}$	0.56(1.27)	−0.75(1.49)	−2.68(1.35)	−0.86(1.65)	0.78(1.98)
$\Delta\Delta G_{non\text{-}polar,GB}$	−0.02(0.01)	0.00(0.00)	0.00(0.00)	0.01(0.01)	0.00(0.01)
$\Delta\Delta G_{GB}$	1.67(2.43)	0.22(0.51)	0.01(0.12)	0.14(0.06)	−1.72(0.16)
$\Delta\Delta G_{subtotal,GB}$	0.98(0.72)	−0.05(0.14)	−0.06(0.02)	−0.25(0.06)	0.52(0.15)

Contribution	Asp142Ala	Tyr143Ala	Ile144Ala	Lys145Ala
$\Delta\Delta E_{electrostatic}$	33.33(2.10)	0.75(1.20)	1.05(0.29)	−96.69(4.88)
$\Delta\Delta E_{vdW}$	−0.65(0.18)	−3.12(0.51)	−1.50(0.35)	1.06(1.43)
$\Delta\Delta E_{gas}$	32.68(2.04)	−2.37(1.47)	−0.45(0.46)	−95.63(3.97)
$\Delta\Delta G_{non\text{-}polar,PB}$	−0.04(0.02)	−0.13(0.03)	−0.01(0.02)	−0.02(0.03)
$\Delta\Delta G_{PB}$	−31.02(2.14)	−2.40(1.75)	−2.94(3.37)	67.70(6.04)
$\Delta\Delta G_{subtotal,PB}$	1.62(1.89)	**−4.91(1.82)**	**−3.40(3.44)**	**−27.96(6.45)**
$\Delta\Delta G_{non\text{-}polar,GB}$	−0.04(0.02)	−0.12(0.03)	−0.01(0.02)	−0.02(0.03)
$\Delta\Delta G_{GB}$	−31.80(2.02)	−0.57(0.86)	−0.85(0.21)	81.89(3.07)
$\Delta\Delta G_{subtotal,GB}$	0.85(0.35)	**−3.07(0.83)**	−1.31(0.47)	**−13.76(1.62)**

$|\Delta G_{subtotal,PB/GB}| \geq 3$ kcal/mol values are displayed in bold and underlined.

to hydrogen bond analysis of the MD (Table 3), the occupancy of the hydrogen bonds of Ser88-O···HO-Ser90 and Lys145-NZH···O-Ser90 is 52.40% and 80.08%, respectively, which suggests that Ser90 contributes more to the intramolecular interaction and not to the intermolecular interaction.

Tyr143. The largest part to the binding free energy of Tyr143 is the vdW interaction (Table 2). This indicates vdW interaction between Tyr143 and the inhibitor plays an important role and

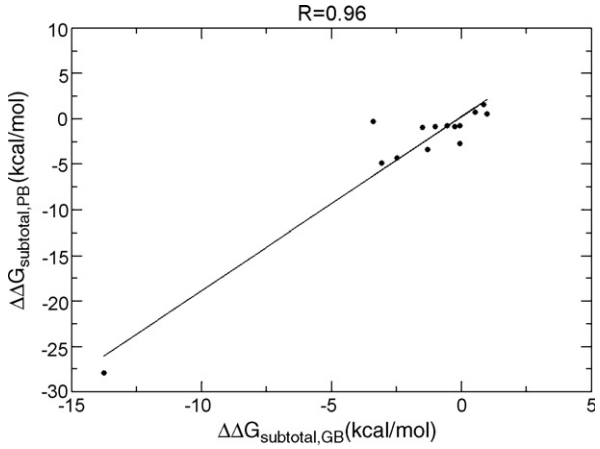

Fig. 5. Regression between the calculated $\Delta\Delta G_{subtotal,PB}$ and $\Delta\Delta G_{subtotal,GB}$ values for the 14 residues obtained by the computational alanine scanning.

may be one of the reasons why a large cyclic ring of arylomycin A_2 is preferred.

Ile144. Contribution to favorable binding energy of Ile144 originates mainly from polar solvation free energy and vdW interaction. There is a 3.37 kcal/mol standard deviation of $\Delta\Delta G_{PB}$, which again suggests the PB model is more sensitive to the atomic coordinates than the GB model with only 0.21 kcal/mol standard deviation of $\Delta\Delta G_{GB}$ (Table 2). During the computational alanine scanning method, it is impossible to find the contribution from the backbone atoms which are not changed. The hydrogen bond of Ile144-NH···O44-Arylomycin A2 with a 99.44% occupancy (Table 3) is not broken during the computational alanine scanning, which may explain there is no large difference in $\Delta\Delta E_{electrostatic}$. It has been found that when the Ile144 is mutated to Ala, the SPase still keeps 50% activity relative to the WT SPase (PONA as substrate) [48]. It seems that the backbone atoms of Ile144 play an important role in sustaining the function of SPase, which has also been suggested by the free energy decomposition results with almost double contributions from the backbone than from the sidechain (Table 5).

Lys145. Obviously, this residue is the most important residue with a −27.96 kcal/mol reduction in $\Delta\Delta G_{subtotal,PB}$ and −13.76 kcal/mol reduction in $\Delta\Delta G_{subtotal,GB}$ (Table 2). Most of the binding energy difference of the L145A mutation comes

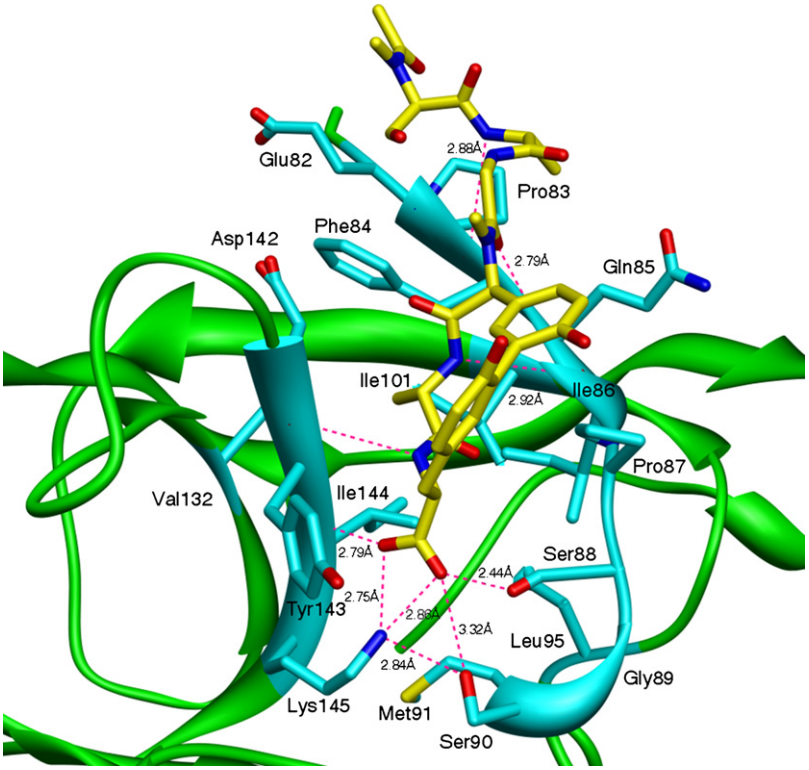

Fig. 6. Interactions between lipopeptide inhibitor and *E. coli* SPase based on the average structure of this complex from the MD. Carbon atoms of the inhibitor are colored in yellow. The backbones and side-chains of the interfacial residues from the peptidase are shown in cyan ribbon and cyan stick, respectively. The rest of the residues of the peptidase are shown in green ribbon. Hydrogen bonds are displayed in magenta dashed line with the corresponding distances. Red is for oxygen and blue is for nitrogen. Figure is created using Chimera [46].

Table 3
Intermolecular and intramolecular hydrogen bonds of the last 500 ps trajectories

Hydrogen bond[a]			% Occupied[b]	Distance
SPase residues		Arylomycin A2		
Pro83-O	···	HN7	95.65	2.897(0.12)
Gln85-O	···	HN28	90.40	2.930(0.12)
Gln85-NH	···	O15	99.68	2.837(0.10)
Ser88-OGH[c]	···	O45	82.88	2.627(0.11)
Ser88-NH	···	O33	13.84	2.718(0.14)
Ser90-OGH[c]	···	O45	3.44	2.848(0.17)
Asp142-O	···	HN33	94.96	2.936(0.11)
Ile144-NH	···	O44	99.44	2.806(0.10)
Lys145-NZH[d]	···	O44	55.92	2.806(0.11)
Lys145-NZH[d]	···	O45	61.28	2.904(0.13)
Hydrogen bond[a]			% Occupied[b]	Distance
SPase residues		SPase residues		
Ser88-O	···	HO-Ser90	52.40	2.790(0.13)
Lys145-NZH[d]	···	GO-Ser90	80.08	2.898(0.12)

[a] Hydrogen bonds were assigned when the distance of two heavy atoms (O or N) is less than 3.2 Å and the angle (heavy atom-hydrogen-heavy atom) is larger than 120°.
[b] The percentage of hydrogen bond populating over the last 500 ps MD.
[c] Hydrogen of the side-chain hydroxyl of Ser88 and Ser90 residues.
[d] Hydrogen of the side-chain amine of Lys145 residue.

from the great difference between $\Delta\Delta E_{\text{electrostatic}}$ and $\Delta\Delta G_{\text{PB}}$/$\Delta\Delta G_{\text{GB}}$. A large desolvation energy is paid, but it can not offset the larger electrostatic interaction, which leads to the large loss in binding energy. Lys145 N-ζ can form three strong hydrogen bonds with the carboxylate O45 and O44 of arylomycin A_2 and Ser90 O-γ with 2.904, 2.806 and 2.898 Å average distances separately (Table 3). The hydrogen bond of Lys145-NZH···O-Ser90 exists in the 80.08% time of the last 500 ps MD, which suggests the intramolecular interaction is critical for keeping the enzymatic function of SPase. It has been shown that SPase is inactive in vitro and vivo when Lys145 is mutated to alanine [49]. More and more evidence reveals that *E. coli* SPase uses a Ser90/Lys145 dyad mechanism, where Lys145 acts as the general base and Ser90 acts as nucleophile [4–7,50,51]. It is conformed that Lys145 has two functions: one is for activating Ser90 to increase the nucleophility and the other is to recognize and stabilize the binding of substrates to the *E. coli* SPase [6,7,9]. These two functions have been observed also in our present work.

3.3. Decomposition of free energy on a per-residue basis

Free energy decomposition serves as a faster alternative for computational alanine scanning mutagenesis. The per-atom contributions can be summed over atom groups such as residues, backbones and side-chains, to obtain their contributions to the total binding free energy [25]. Thus, contributions

Table 4
Total binding free energy decomposition (kcal/mol) to *E. coli* SPase and Arylomycin A_2

	ΔE_{vdW}	$\Delta E_{electrostatic}$	$\Delta G_{non\text{-}polar}$	ΔG_{GB}	$\Delta G_{subtotal}$
Arylomycin A_2	−24.10(1.76)	−41.49(4.56)	−2.78(0.07)	49.37(3.58)	−19.00(1.98)
E. coli SPase	−24.10(1.76)	−41.49(4.56)	−1.97(0.06)	28.26(3.09)	−39.30(2.10)
Complex	−48.20(3.52)	−82.98(9.12)	−4.75(0.11)	77.63(6.51)	−58.30(3.79)

due to ''non-mutatable'' functional groups have been calculated at the atomic level using the MM-GBSA method [24].

Table 4 reports the decomposition of the binding energy into contributions from vdW energy, electrostatic interaction energy, non-polar solvation free energy, polar solvation free energy for *E. coli* SPase and Arylomycin A_2 separately. Table 5 depicts the decomposition of $\Delta G_{gas+solv}$ on a per residue basis for *E. coli* SPase.

The total relative binding energy of *E. coli* SPase and Arylomycin A_2 complex is −58.30 kcal/mol (Table 4), about 4.5 kcal/mol lower than the value −53.68 kcal/mol calculated via the MM-PBSA method (Table 1). Here, *E. coli* SPase and Arylomycin A_2 contribute each −39.30 and −19.00 kcal/mol, respectively. The summed binding energies (−21.36 and −12.16 kcal/mol, respectively) from the S1 and S3 pocket residues occupies about 54.48% and 31.01% of the *E. coli* SPase contribution (−39.21 kcal/mol) to the binding and the remaining 15% of the binding energy comes from the residues of the other pockets. Lys145 has a contribution of −7.84 kcal/mol, most from the side-chain (−7.99 kcal/mol) and essentially through electrostatic interaction, its backbone (0.15 kcal/mol) is a little unfavorable to the binding. Tyr143 and Ile144 contribute −4.49 and −3.46 kcal/mol to the binding energy in good agreement with the alanine scanning data. Interestingly, Ser90 also appears to disfavor the binding, but only to a lesser extent. Complementarily, Pro83 which are not be mutated to alanine in the computation alanine scanning method and Gln85 have been recognized as critical residues. In this case, the contribution of Pro83 and Gln85 arises more from the backbone atoms than the side-chains (Table 5). The backbone oxygen of Pro83 forms a strong hydrogen bond with N7 of Arylomycin A_2 at a average distance of 2.897 Å. This hydrogen bond has a high occupancy of 95.65% (Table 3). Similarly, there are two strong hydrogen bonds, Gln85-O···HN28-Arylomycin A2 and Gln85-NH···O15-Arylomycin A2 (Table 3). Except for electrostatic interaction, van der Waals contacts from Pro83 and Gln85 contribute substantially to the binding free energy (Table 3).

Phe84 contributes about −4.66 kcal/mol to the binding affinity, more than a half of which originates from the side-chain according to the free energy decomposition calculation. Ser88 has only a −1.31 kcal/mol contribution to the binding energy while it has −2.47 kcal/mol contribution ($\Delta\Delta G_{subtotal,GB}$) to the binding energy after being mutated to alanine. Taking account of the approximations of these models and the dependency upon parameters, they are in fair agreement. But, generally, the results from the binding energy decomposition agree well with those from the alanine scanning based on GB models.

3.4. Comparison between computation alanine scanning and free energy decomposition

Regression between the calculated $\Delta\Delta G_{subtotal,GB}$ obtained by the computational alanine scanning and $\Delta G_{subtotal}$ obtained by free energy decomposition for the 14 residues has been plotted in Fig. 7. The calculations of binding free energy are

Table 5
Decomposition of $\Delta G_{gas+solv}$ (kcal/mol) on a per-residue basis into contributions from van der Waals energy (E_{vdW}), electrostatic interaction energy ($E_{electrostatic}$), non-polar solvation free energy ($G_{non\text{-}polar}$), polar solvation free energy (G_{GB})

	ΔE_{vdW}	$\Delta E_{electrostatic}$	$\Delta G_{non\text{-}polar}$	ΔG_{GB}	$\Delta G_{subtotal}$	$B\Delta G_{subtotal}$	$S\Delta G_{subtotal}$
E82	−1.76(0.28)	8.30(0.97)	−0.19(0.02)	−7.61(1.13)	−1.26(0.41)	−0.47(0.19)	−0.80(0.31)
P83	−1.74(0.60)	−3.88(0.82)	−0.20(0.02)	2.45(0.42)	**−3.37(0.57)**	−2.27(0.53)	−1.10(0.25)
F84	−3.82(0.37)	−3.27(0.63)	−0.24(0.03)	2.67(0.49)	**−4.66(0.47)**	−2.02(0.27)	−2.64(0.38)
Q85	−2.41(0.63)	−3.52(0.95)	−0.27(0.03)	1.70(0.49)	**−4.50(0.47)**	−3.57(0.43)	−0.93(0.19)
I86	−1.34(0.33)	−2.21(0.39)	−0.02(0.01)	1.71(0.24)	−1.85(0.55)	−1.06(0.30)	−0.79(0.28)
P87	−2.25(0.41)	−0.90(0.71)	−0.28(0.03)	0.74(0.62)	−2.69(0.39)	−0.39(0.08)	−2.30(0.38)
S88	0.27(0.88)	−10.28(2.6)	−0.07(0.02)	8.78(1.64)	−1.31(0.77)	−0.53(0.31)	−0.78(0.68)
G89	−0.06(0.02)	−0.75(0.26)	0.00(0.00)	0.76(0.25)	−0.06(0.03)	−0.01(0.02)	−0.05(0.01)
S90	−0.32(0.10)	−2.32(1.66)	−0.01(0.01)	3.05(1.39)	0.40(0.32)	−0.24(0.09)	0.64(0.33)
M91	−0.15(0.04)	−0.97(0.39)	0.00(0.00)	1.05(0.36)	−0.07(0.05)	0.02(0.02)	−0.09(0.04)
L95	−0.06(0.01)	0.03(0.21)	0.00(0.00)	−0.02(0.20)	−0.05(0.01)	0.01(0.00)	−0.06(0.01)
I101	−0.23(0.04)	−0.15(0.16)	0.00(0.00)	0.22(0.16)	−0.16(0.04)	0.00(0.02)	−0.16(0.03)
V132	−0.44(0.10)	0.00(0.05)	−0.22(0.01)	0.01(0.05)	−0.45(0.11)	−0.08(0.02)	−0.36(0.09)
D142	−2.00(0.42)	14.98(1.22)	−0.23(0.03)	−14.32(1.04)	−1.58(0.38)	−1.57(0.29)	−0.01(0.25)
Y143	−3.05(0.40)	−4.70(0.76)	−0.13(0.02)	3.39(0.45)	**−4.49(0.53)**	−3.12(0.35)	−1.38(0.43)
I144	−1.29(0.61)	−0.61(0.97)	−0.06(0.01)	−1.49(0.42)	**−3.46(0.44)**	−2.29(0.36)	−1.16(0.24)
K145	0.42(0.71)	−47.89(2.45)	−0.02(0.01)	39.65(1.42)	**−7.84(0.82)**	0.15(0.04)	−7.99(0.84)

$BG_{subtotal}$ stands for the backbone $\Delta G_{gas+solv}$ and $S\Delta G_{subtotal}$ stands for side-chain $\Delta G_{gas+solv}$. $|\Delta G_{subtotal}| \geq 3$ kcal/mol values are displayed in bold and underlined.

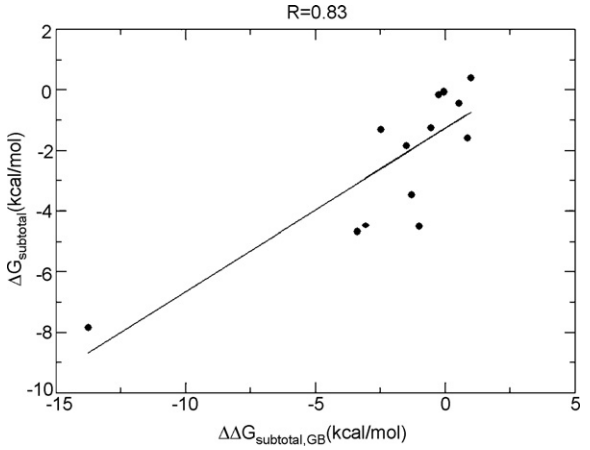

Fig. 7. Regression between the calculated $\Delta\Delta G_{\text{subtotal,GB}}$ values obtained by the computational alanine scanning and $\Delta G_{\text{subtotal}}$ values obtained by free energy decomposition for the 14 residues.

both based on the GB model. A correlation of $R = 0.83$ between $\Delta\Delta G_{\text{subtotal,GB}}$ and $\Delta G_{\text{subtotal}}$ was observed. There is a satisfying correlation between the two methods. It is remarkable that free energy decomposition has more advantage than computation alanine scanning. First, it is faster because the binding free energy is calculated only once, while this term must be recalculated for all mutants in the computational alanine scanning approach. Second, it is not necessary to consider the global conformation changes which should be validated in the computational alanine scanning approach, since the assumption of minimal conformational changes upon Ala mutation is not always justified. Third, it can decompose the binding energy into both of the backbones and sidechains while only the importance of sidechains of residues can be probed by the computational alanine scanning method.

4. Conclusion

Computational alanine scanning and free energy decomposition methods have been performed to investigate the different contribution of *E. coli* SPase binding site residues to the binding of Arylomycin A_2. The PB and GB models have been used to calculate the binding free energy during computational alanine scanning mutation, but only the GB model has been used to decompose the free energy at an atomic level. The regression analysis between the PB and GB model and also between the computational alanine scanning and free energy decomposition have been reported with a correlation coefficient of 0.96 and 0.83, respectively, which suggest they are both in fair agreement with each other. Larger standard deviations have been observed in PB model than in GB model. For example, the standard deviations of $\Delta G_{\text{subtotal,PB}}$ for Lys145A and Ile144A are 6.45 and 3.44 kcal/mol, only 1.62 and 0.47 kcal/mol of $\Delta G_{\text{subtotal,GB}}$. Moreover, there is still some discrepancy between the two methods. It appears that the MM-PBSA method is more sensitive to the atomic coordinates than the MM-GBSA method during the calculation of contributions to the solvation energy. For the system under investigation, the free energy decomposition of the wild-type system provides information similar to that of the computational alanine scanning. There are more advantages in the free energy decomposition method because it is faster and not necessary to consider the global conformation changes and also it can probe the effects of both the backbones and sidechains of residues.

We did not consider entropic contributions, because it can be neglected when the relative binding free energies are calculated between the wild-type and mutants binding to the same receptor. Enthalpy presumably driven by hydrogen bond interactions between *E. coli* SPase and Arylomycin A_2 proved to be the major thermodynamic factor. Explicit MD simulations for the mutated proteins F84A, S88A, Y143A, I144A and L145A with the lipopeptide inhibitor have been carried out to validate there are no global conformation changes upon Ala mutation. The backbone atoms RMSDs of the five simulated models compared to the starting X-ray structure show a satisfying stability during the MD simulation.

Our data confirm that Lys145 is the most important residue for the binding with strong hydrogen bonding with the carboxylate of Arylomycin A_2 and Ser90 O-γ. Its first function is to activate Ser90 to increase the nucleophility, its second function is to recognize and stabilize the binding of *E. coli* SPase with substrates. The hydroxyl group of Ser88 plays a key role for the binding of the inhibitor. Ser90 contributes more to the intramolecular interaction than to the intermolecular interaction. Again, the energy components of Tyr143 and Phe84 show that their larger van der Waals interaction energy contribute substantially to the binding, revealing that the phenyl rings of these two residues are more important in selecting the shape of the substrates. The contribution to favorable binding energy of Ile144 originates mainly from polar solvation free energy and vdW interaction. Moreover, the backbone atoms of Gln85 and Pro83 interact with L-Ala and D-Ala parts of Arylomycin A_2 via van der Waals contacts and strong hydrogen bonds. The residues from the S1 and S3 pocket occupied about 54.48% and 31.01% of the contribution of *E. coli* SPase to the binding energy.

Overall, the results from the computational alanine scanning and free energy decomposition methods agree with each other very well. These two methods provide good ways to determine rapidly the influences of single residues on the binding affinity and pinpoint the hotspot residues. In that sense, structure-based ligand design can benefit from the inexpensive computational screening.

References

[1] M. Paetzel, R.E. Dalbey, N.C. Strynadka, The structure and mechanism of bacterial type I signal peptidases: a novel antibiotic target, Pharmacol. Ther. 87 (2000) 27–49.
[2] M.T. Black, G. Bruton, Inhibitors of bacterial signal peptidases, Curr. Pharm. Des. 4 (1998) 133–154.
[3] M. Paetzel, A. Karla, N.C. Strynadka, R.E. Dalbey, Signal peptidases, Chem. Rev. 102 (2002) 4549–4580.
[4] T.B. Michael, Evidence that the catalytic activity of prokaryote leader peptidase depends upon the operation of a serine-lysine catalytic dyad, J. Bacterial. 175 (1993) 4957–4961.
[5] M. Paetzel, N.C. Strynadka, W.R. Tschantz, R. Casareno, P.R. Bullinger, R.E. Dalbey, Use of site-directed chemical modification to study an

essential lysine in *Escherichia coli* leader peptidase, J. Biol. Chem. 272 (1997) 9994–10003.
[6] M. Paetzel, R.E. Dalbey, N.C. Strynadka, Crystal structure of a bacterial signal peptidase in complex with a β-lactam inhibitor, Nature 396 (1998) 186–190.
[7] M. Paetzel, J.J. Goodall, M. Kania, R.E. Dalbey, M.G. Page, Crystallographic and biophysical analysis of a bacterial signal peptidase in complex with a lipopeptide-based inhibitor, J. Biol. Chem. 279 (2004) 30781–30790.
[8] K. Palaniappan, J.K. Adam, A.S. Mark, D.B. Matthew, A.S. Tim, C. Matthew, R.S. James, L.M. Kristina, F. Zheng, L.A. Eddie, M. Deborah, N.J. Louis, J.K. Valentine, I.N. Thalia, C.T. Richard, S.B. Peng, Novel lipoglycopeptides as inhibitors of bacterial signal peptidase I, J. Biol. Chem. 279 (2004) 36250–36258.
[9] M. Paetzel, R.E. Dalbey, N.C. Strynadka, Crystal structure of a bacterial signal peptidase apoenzyme, J. Biol. Chem. 277 (2002) 9512–9519.
[10] J.M. Swanson, R.H. Henchman, J.A. McCammon, Revisiting free energy calculations: a theoretical connection to MM/PBSA and direct calculation of the association free energy, Biophys. J. 86 (2004) 67–74.
[11] B. Kuhn, P. Gerber, T. Schulz-Gasch, M. Stahl, Validation and use of the MM-PBSA approach for drug discovery, J. Med. Chem. 48 (2005) 4040–4048.
[12] C. Luo, L. Xu, S. Zheng, X. Luo, J. Shen, H. Jiang, X. Liu, M. Zhou, Computational analysis of molecular basis of 1:1 interactions of NRG-1beta wild-type and variants with ErbB3 and ErbB4, Proteins 59 (2005) 742–756.
[13] S. Huo, J. Wang, P. Cieplak, P.A. Kollman, I.D. Kuntz, Molecular dynamics and free energy analyses of cathepsin D-inhibitor interactions: insight into structure-based ligand design, J. Med. Chem. 45 (2002) 1412–1419.
[14] Y. Xu, R.X. Wang, A computational analysis of the binding affinities of FKBP12 inhibitors using the MM-PB/SA method, Proteins 64 (2006) 1058–1068.
[15] M. Lepsik, Z. Kriz, Z. Havlas, Efficiency of a second-generation HIV-1 protease inhibitor studied by molecular dynamics and absolute binding free energy calculations, Proteins 57 (2004) 279–293.
[16] B. Kuhn, P.A. Kollman, Binding of a diverse set of ligands to avidin and streptavidin: an accurate quantitative prediction of their relative affinities by a combination of molecular mechanics and continuum solvent models, J. Med. Chem. 43 (2000) 3786–3791.
[17] J. Wang, P. Morin, W. Wang, P.A. Kollman, Use of MM-PBSA in reproducing the binding free energies to HIV-1RT of TIBO derivates and predicting the binding mode to HIV-1 RT of Efavirenz by docking and MM-PBSA, J. Am. Chem. Soc. 123 (2001) 5221–5230.
[18] H.B. Thorsteinsdottir, T. Schwede, V. Zoete, M. Meuwly, How inaccuracies in protein structure models affect estimates of protein–ligand interactions: computational analysis of HIV-1 protease inhibitor binding, Proteins 65 (2006) 407–423.
[19] I. Massova, P.A. Kollman, Computational alanine scanning to probe protein–protein interactions: a novel approach to evaluate binding free energies, J. Am. Chem. Soc. 121 (1999) 8133–8143.
[20] S. Huo, I. Massova, P.A. Kollman, Computational alanine scanning of the 1:1 human growth hormone-receptor complex, J. Comput. Chem. 23 (2002) 15–27.
[21] O. Villacanas, R.M. Jaime, Reducing CDK4/6-p16^{INK4a} interface: computational alanine scanning of a peptide bound to CDK6 protein, Proteins 63 (2006) 797–810.
[22] S.M. Irina, A.F. Pedro, J.R. Maria, Unraveling the importance of protein-protein interaction: application of a computational alanine-scanning mutagenesis to the study of the IgG1 streptococcal protein G (C2 fragment) complex, J. Phys. Chem. B 110 (2006) 10962–10969.
[23] T.C. Lillian, C.S. William, W.P. Jed, S.P. Vijay, Kinetic computational alanine scanning: application to p53 oligomerization, J. Mol. Biol. 357 (2006) 1039–1049.
[24] H. Gohlke, C. Kiel, D.A. Case, 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 (2003) 891–913.
[25] V. Zoete, M. Meuwly, M. Karplus, Study of the insulin dimerization: binding free energy calculations and per-residue free energy decomposition, Proteins 61 (2005) 79–93.
[26] V. Zoete, M. Meuwly, Importance of individual side chains for the stability of a protein fold: computational alanine scanning of the insulin monomer, J. Comput. Chem. 27 (2006) 1843–1857.
[27] V. Zoete, O. Michielin, Comparison between computational alanine scanning and per-residue binding free energy decomposition for protein–protein association using MM-GBSA: application to the TCR-p-MHC complex, Protein (March 2007) (Epub ahead of print).
[28] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, Nucleic Acids Res. 28 (2000) 235–242.
[29] D.A. Pearlman, D.A. Case, J.W. Caldwell, W.S. Ross, T.E. Cheatham III, S. DeBolt, D. Ferguson, G. Seibel, P.A. Kollman, 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 (1995) 1–41.
[30] D.A. Case, T.A. Darden, T.E. Cheatham III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, K.M. Merz, B. Wang, D.A. Pearlman, M. Crowley, S. Brozell, V. Tsui, H. Gohlke, J. Mongan, V. Hornak, G. Cui, P. Beroza, C. Schafmeister, J.W. Caldwell, W.S. Ross, P.A. Kollman, AMBER, version 8.0, University of California, 2004.
[31] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general AMBER force field, J. Comput. Chem. 25 (2004) 1157–1174.
[32] A. Jakalian, B.L. Bush, D.B. Jack, C.I. Bayly, Fast, efficient generation of high-quality atomic charges. AM1-BCC Model: I. Method, J. Comput. Chem. 21 (2000) 132–146.
[33] Y. Duan, C. Wu, S. Chowdhury, M.C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, P. Kollman, A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations, J. Comput. Chem. 24 (2003) 1999–2012.
[34] W.L. Jorgensen, Revised TIPS for simulations of liquid water and aqueous solutions, J. Chem. Phys. 77 (1982) 4156–4163.
[35] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926–935.
[36] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an NLog(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.
[37] S. Miyamoto, P.A. Kollman, SETTLE: an analytical version of the SHAKE and RATTLE algorithm for rigid water models, J. Comput. Chem. 13 (1992) 952–962.
[38] P.A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D.A. Case, T.E. Cheatham III, Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models, Acc. Chem. Res. 33 (2000) 889–897.
[39] J. Weiser, P.S. Shenkin, W.C. Still, Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO), J. Comput. Chem. 20 (1999) 217–230.
[40] D. Sitkoff, K.A. Sharp, B. Honig, Accurate calculation of hydration free energies using macroscopic solvent models, J. Phys. Chem. 98 (1994) 1978–1988.
[41] G.D. Hawkins, C.J. Cramer, D.G. Truhlar, Pairwise solute descreening of solute charges from a dielectric medium, Chem. Phys. Lett. 246 (1995) 122–129.
[42] G.D. Hawkins, C.J. Cramer, D.G. Truhlar, Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium, J. Phys. Chem. 100 (1996) 19824–19839.
[43] A. Onufriev, D. Bashford, D.A. Case, Modification of the generalized born model suitable for macromolecules, J. Phys. Chem. B 104 (2000) 3712–3720.
[44] B. Jayaram, D. Sprous, D.L. Beveridge, Solvation free energy of biomacromolecules: parameters for a modified generalized born model consis-

tent with the AMBER force field, J. Phys. Chem. B 102 (1998) 9571–9576.

[45] V. Tsui, D.A. Case, Theory and applications of the generalized Born Solvation Model in macromolecular simulations, Biopolymers 56 (2001) 257–291.

[46] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera—a visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612.

[47] M. Sung, R.E. Dalbey, Identification of potential active-site residues in the *Escherichia coli* leader peptidase, J. Biol. Chem. 267 (1992) 13154–13159.

[48] A. Karla, M. Lively, M. Paetzel, R.E. Dalbey, The identification of residues that control signal peptidase cleavage fidelity and substrate specificity, J. Biol. Chem. 280 (2005) 6731–6741.

[49] J.L. Carlos, P.A. Klenotic, M. Paetzel, N.C. Strynadka, R.E. Dalbey, Mutational evidence of transition state stabilization by Ser 88 in *Escherichia coli* Type I Signal Peptidase, Biochemistry 39 (2000) 7276–7283.

[50] W.R. Tschantz, M. Sung, V.M. Delgado-Partin, R.E. Dalbey, A serine and a lysine residue implicated in the catalytic mechanism of the *Escherichia coli* leader peptidase, J. Biol. Chem. 268 (1993) 27349–27354.

[51] M. Paetzel, N.C. Strynadka, Common protein architecture and binding sites in proteases utilizing a Ser/Lys dyad mechanism, Protein Sci. 8 (1999) 2533–2536.