Bioorganic & Medicinal Chemistry 20 (2012) 3446–3453

Contents lists available at SciVerse ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Free energy calculations on the binding of novel thiolactomycin derivatives to *E. coli* fatty acid synthase I

Thomas Steinbrecher [a,*], David A. Case [b], Andreas Labahn [c]

[a] Institut für Physikalische Chemie, Abteilung Theoretische Chemische Biologie, Universität Karlsruhe, KIT, Kaiserstr. 12, 76131 Karlsruhe, Germany
[b] BioMaPS Institute and Department of Chemistry & Chemical Biology, Rutgers University, 174 Frelinghuysen Road, Piscataway, NJ 08854-8087, USA
[c] Institut für Physikalische Chemie, Albertstr. 23a, Universität Freiburg, Germany

ARTICLE INFO

Article history:
Received 7 February 2012
Revised 3 April 2012
Accepted 7 April 2012
Available online 16 April 2012

Keywords:
Molecular dynamics
Free energy calculations
Drug design
Antibiotics
Fatty acid synthesis
Protein–ligand interactions

ABSTRACT

Finding novel antibiotics to combat the rise of drug resistance in harmful bacteria is of enormous importance for human health. Computational drug design can be employed to aid synthetic chemists in the search for new potent inhibitors. In recent years, molecular dynamics based free energy calculations have emerged as a useful tool to accurately calculate receptor binding affinities of novel or modified ligands. While being significantly more demanding in computational resources than simpler docking algorithms, they can be employed to obtain reliable estimates of the effect individual functional groups have on protein–ligand complex binding constants.

Beta-ketoacyl [acyl carrier protein] synthase I, KAS I, facilitates a critical chain elongation step in the fatty acid synthesis pathway. Since the bacterial type II lipid synthesis system is fundamentally different from the mammalian type I multi-enzyme complex, this enzyme represents a promising target for the design of specific antibiotics. In this work, we study the binding of several recently synthesized derivatives of the natural KAS I inhibitor thiolactomycin in detail based on atomistic modeling. From extensive thermodynamic integration calculations the effect of changing functional groups on the thiolactone scaffold was determined. Four ligand modifications were predicted to show improved binding to the *E. coli* enzyme, pointing the way towards the design of thiolactomycin derivatives with binding constants in the nanomolar range.

1. Introduction

The increasing resistance of pathogenic bacteria to widely used antibiotics is a major challenge of modern medicine that has attracted a great deal of attention far beyond the academic world.[1–4] Finding novel antibiotics that affect previously untargeted biochemical pathways is of utmost interest to solve this problem. In this context, natural compounds are a promising starting point for the discovery of structurally new agents.[5,6] Modern drug design studies have benefited greatly from the increased use of computational tools for structure and binding strength predictions. Especially molecular dynamics simulation based free energy calculations have received considerable attention here over the last years.[7–15]

The bacterial fatty acid synthesis pathway is considered a promising candidate for the development of specific inhibitors with new modes of operation.[16] One reason for this is the markedly different organization of the mammalian fatty acid synthesis apparatus compared to that of bacteria. The type I, or associated, fatty acid synthesis machinery humans share with other higher organisms is comprised of a large (ca. 500 kDa) covalently bound multifunctional enzyme complex, while the type II system of plants and bacteria is made up of several distinct soluble enzymes, each catalyzing one particular reaction of the fatty acid cycle.

Structural information is available for several enzymes of the type II fatty acid synthase system from different organisms.[17–21] The chain elongation enzymes KAS I and II (or FabB and FabF), that add two more carbon atoms to a growing fatty acid chain, are among the most interesting targets, as well as the enzyme FabH that catalyzes the initial step of fatty acid synthesis. KAS I inhibitors of moderate binding strengths are known and X-ray crystal structures of cocrystallized inhibitor enzyme complexes exist,[22] but up to now, no new compound with an inhibition constant in the low nanomolar range has been found. In contrast, FabH inhibitors of nanomolar strength have been found recently.[16]

Improving known KAS I inhibitors is an interesting test bed for the use of emerging computational chemistry methods in rational drug design. Thiolactomycin, a natural compound extracted from the bacterium *Nocardia* sp., is effective against several pathogenic

Abbreviations: TI, thermodynamic integration; vdW, van der Waals; MD, molecular dynamics.

* Corresponding author. Tel.: +49 721 608 45067; fax: +49 721 608 45710.
E-mail address: thomas.steinbrecher@kit.edu (T. Steinbrecher).

0968-0896/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.bmc.2012.04.019

bacteria and is known to inhibit *Escherichia coli* KAS I by mimicking the substrate Malonyl-ACP.[23] Several thiolactomycin derivatives have been studied in the past (see Refs. [24,25]). A recently published study contained a new organic synthesis scheme providing convenient access to novel thiolactomycin derivatives that carry modified substituents at the thiolactone ring, but could not ascertain if any of them will show improved binding affinities compared to the natural compound.[26]

We continue the study of this family of compounds here by performing Thermodynamic Integration free energy calculations, in order to accurately determine their binding free energies with respect to thiolactomycin.

2. Models and methods

Thermodynamic Integration calculations are a method to compute free energy differences between two arbitrary chemical states A and B. This is accomplished by coupling the two states via a parameter λ so that a transition from zero to one along the nonphysical coordinate λ is equivalent to changing state A (the initial state) into state B (the final state). The two potential functions V(A) and V(B) are combined into a mixed, λ-dependent potential function V(λ). Since A and B can consist of different types and numbers of atoms, these calculations are sometimes referred to as 'computational alchemy'. It can easily be derived from the statistical thermodynamics expression for the free energy of a system that the free energy difference between the two λ-coupled states according to the TI formalism is:

$$\Delta G^0_{TI} = \int_0^1 \left\langle \frac{\partial V(\lambda)}{\partial \lambda} \right\rangle_\lambda d\lambda \tag{1}$$

where the angular brackets denote a Boltzmann-weighted ensemble average, using V(λ) as the potential. The integration in Eq. 1 cannot typically be performed analytically, and in practice a numerical integration scheme is used, based on simulations at a number of fixed λ-values.

The easiest way to couple the two endpoint potential functions into V(λ) is called linear mixing:

$$V(\lambda) = \lambda \cdot V_1 + [1-\lambda] \cdot V_0 \tag{2}$$

A multitude of other mixing schemes exist and extensive comparisons have been performed between them. Linear mixing in most cases yields stable and easily integratable free energy curves if a soft core potential is used that replaces the vdW interactions of appearing or disappearing atoms (i.e., atoms that only exist in states B or A, respectively) with a λ-dependent, modified LJ-equation.[27] In this study, the shape of the soft core pair potential used was:

$$V_{vdW,A} = 4\varepsilon(1-\lambda)\left(\frac{1}{[\alpha\lambda + (\frac{r}{\sigma})^6]^2} - \frac{1}{\alpha\lambda + (\frac{r}{\sigma})^6}\right) \tag{3}$$

$$V_{vdW,B} = 4\varepsilon\lambda\left(\frac{1}{[\alpha(1-\lambda) + (\frac{r}{\sigma})^6]^2} - \frac{1}{\alpha(1-\lambda) + (\frac{r}{\sigma})^6}\right) \tag{4}$$

where Eqs. 3 and 4 apply to the interactions of atoms present in both states A and B with those unique to states A or B, respectively; ε and σ are the common LJ parameters, r is the interatomic distance and α is an adjustable constant set to 0.5 in this study. The details of the implementation of this soft core algorithm have been described elsewhere.[28]

To allow transformations to be conducted in a single step, instead of the common breakdown into separate Coulomb and vdW transformation steps, electrostatic interactions between atoms unique to states A and B and the rest of the system were calculated using a modified Coulomb pair potential equation:

$$V_{Coulomb,A} = (1-\lambda)\frac{q_i q_j}{4\pi\varepsilon_0\sqrt{\beta\lambda + r^2}} \tag{5}$$

$$V_{Coulomb,B} = \lambda\frac{q_i q_j}{4\pi\varepsilon_0\sqrt{\beta(1-\lambda) + r^2}} \tag{6}$$

where Eqs. 5 and 6 apply to the interactions of atoms present in both states A and B with those atoms unique to states A or B, respectively; q_i and q_j are the atomic partial charges, r is the interatomic distance and β is an adjustable constant set to 24 Å^2 in this study. The details of the implementation of this one-step TI algorithm have been described elsewhere.[29]

The TI module of the AMBER modeling suite used here allows a full dual-topology approach, meaning that no atom types needed to be changed. Instead, for every changing functional group, an instance of the group before and after the transformation is present in the system. These two groups do not interact with each other and as λ increases from zero to one, the nonbonded interactions of the original group with its surroundings are switched off, while the nonbonded interactions of the new functional group are simultaneously switched on. The bonded interactions of both groups, both internal and with the remainder of the system are unchanged by the λ-transformation (and do not contribute to the free energy result) to keep the appearing or disappearing functional groups in place and restrict them to chemically meaningful conformations.

Since TI calculations give free energy differences stemming from nonphysical transformations (e.g., appearing or disappearing atoms), they have to be considered as part of a thermodynamic cycle connecting them to observable free energy changes to give interpretable results. In the context of protein–ligand binding, such a thermodynamic cycle could consist of simulating the transformation of ligand A into B both while bound to a receptor and while solvated in water. The difference in the free energy changes from the two TI transformations is then equal to the difference in the binding free energies of both ligands:

$$\Delta\Delta G^0_{Bind} = \Delta G^0_{Bind,B} - \Delta G^0_{Bind,A} = \Delta G^0_{TI,bound} - \Delta G^0_{TI,solvated} \tag{7}$$

Note that if ligand B binds stronger to the receptor than ligand A, Eq. 7 will give a negative result.

The ensemble average in Eq. 1 was approximated by performing MD simulations at given constant λ values using the AMBER11 molecular modelling suite.[30,31] System temperatures were set using a Langevin thermostat[32] with a virtual collision frequency of 2 ps^{-1} and in the case of constant pressure simulations, a weak-coupling Berendsen type algorithm based on isotropic position scaling with a relaxation time of 2 ps was used to maintain a reference pressure of 1 bar. The SHAKE algorithm[33] was used to keep the lengths of bonds involving hydrogen atoms fixed and an integration timestep of 2 fs was used. Periodic boundary conditions were applied using a particle mesh Ewald treatment[34–36] for long-range electrostatics, with a 9 Å cutoff for non-bonded interactions. An analytical, isotropic, long range van der Waals correction for dispersion interactions beyond the cutoff was applied.

The studied ligands were sketched by hand, transformed into 3D-structures and assigned atom types from the gaff forcefield.[37] In cases of missing forcefield parameters, values were taken from the most similar chemical group for which parameters were available. Ligand partial charges were generated by performing RESP fitting[38] of the electrostatic potential computed at the HF6–31G* level using GAUSSIAN03.[39] The protein parts of the system were described using the AMBER ff03 forcefield[40] and the TIP3P model was used for water molecules.[41] Visual analysis of structures was performed using VMD,[42] figures were prepared using povray and ChemDraw.

The starting structure for the protein–ligand complex was taken from the 2.35 Å X-ray crystal structure (pdb code 1FJ4) of *E. coli*

KAS I with cocrystallized thiolactomycin.[22] We have chosen to use an X-ray structure of the *E. coli* enzyme because additional experimental data are most likely to become available for this enzyme. Furthermore, to avoid binding site deformation from other cocrystallized ligands, only crystal structures containing thiolactomycin were considered. The newer structure by Pappenberger et al. (pdb code 2VB8)[43] could have been used alternatively, since both structures are of sufficiently high resolution to be used as MD starting structures. A structural comparison showed the 1FJ4 and 2VB8 X-ray structures to be similar enough (rmsd below 0.5 Å) to be expected to yield very similar MD conformational ensembles.

Since KAS I is dimeric in solution,[44,45] only chains A and B from the crystal structure were used. Protons were added using the *xleap* tool in the AMBER modelling suite. The two active side histidines 298 and 333 in each of the two active sites were built as ε-protonated, to enable the hydrogen bonds between these residues and the thiolactone-O of thiolactomycin observed in the crystal structure to form in the MD simulations. The two cocrystallized thiolactomycin molecules were kept in the binding sites. This complex was embedded into a rectangular box of water molecules so that the minimum distance between any solute atom and the box edge was 12 Å and neutralized by adding sodium ions (*E. coli* KAS I is an acidic enzyme with an isoelectric point of 5.2). Water molecules that were present in the crystal structure were retained in the MD model.

3. Results

3.1. System preparation and stability

An initial MD equilibration of the complex was performed to test for the general stability of ligand KAS I complexes under the simulation conditions used and to generate a well equilibrated solvated complex structure as a starting point for TI calculations. The system was subjected to the following initial equilibration procedure: A steepest descent minimization was carried out for 200 steps to remove bad contacts. Then, a 10 ps MD run under NVT conditions was performed during which the system temperature was raised to 298 K using an increased collision frequency of 10 ps^{-1} for the Langevin thermostat. Simulation conditions were switched to the NPT ensemble and a 100 ps density equilibration run was performed. During the first 10 ps the system density reached ca. 1.01 g/cm^3 for all simulations and stabilized at this value for the rest of the MD simulation. An unrestrained 5 ns equilibration run was then performed with the temperature and density equilibrated complex in the NPT ensemble at 298 K. During this run, the complex total kinetic and potential energy as well as density fluctuated less than 0.5%. The complex rmsd was monitored during this final equilibration (Fig. 1) for 1000 MD snapshots 5 ps apart and shows that the protein structure stabilized at ca. 1.2 Å Cα rmsd compared to the X-ray starting structure. Also, the ligand positions within the binding sites exhibited stable rmsd values in the range of 0.9–1.3 Å.

The final binding pose adopted by the ligands is very similar to their starting conformation, but a slight change in hydrogen bonding patterns was observed. In the X-ray crystal structure, both His 298 and 333 form hydrogen bonds to the thiolactone-O of thiolactomycin. During the MD equilibration, the O—His 333 H-bond was observed in 76% of trajectory snapshots and the O—His 298 H-bond was present in 75% of snapshots. In addition a thiolactomycin-S—His 298 H-bond involving the thiolactone-S was present in 46% of the trajectory snapshots. The changed hydrogen bonding pattern emerged due to small translational movements of His 298 and rotational movement of His 333, as well as a small repositioning of the ligand in the binding site (Fig. 1). However, the overall thiolactomycin binding mode observed during the MD equilibration was very similar to that of the original X-ray crystal structure, as the low rmsd values indicate.

Preliminary tests were performed in which only a single monomer of KAS I was solvated and simulated using the same protocol as above. This led to an MD trajectory in which the protein structure underwent larger conformational changes and did not stabilize at a constant rmsd after several nanosceonds of MD simulation time. No further analysis of the monomeric KAS I model was performed. Another preliminary test used a dimeric KAS I model, but removed all water molecules from the X-ray crystal structure before solvating the complex. An equilibration of this system for 500 ps showed a stable protein conformation, but one of the thiolactomycin ligands showed major conformational movements in the active site. All further simulations were therefore based on the dimeric model including the crystal structure water molecules.

3.2. Selected ligand transformations and TI setup

Starting from the previously equilibrated complexes, an initial set of six TI ligand transformations were set up so that the starting state always corresponded to the natural compound thiolactomycin **1** (Fig. 2). The thiolactomycin ligand located in chain A was selected to be transformed, while the ligand in the chain B binding site was left unchanged. The derivatives were chosen because they have recently been synthesized and experiments regarding their antibacterial properties are underway.[24–26,46] Four of the compounds, **2**, **3**, **4** and **5**, add or remove methyl groups at the thiolactone headgroup. **6** replaces the lipophilic sidechain with a methyl group and **7** exchanges the terminal ethenyl group with a ethinyl group. Table 1 gives an overview of what functional groups were changed in the ligands. For **5**, both the keto- and enol tautomers of the ligand were built. Additionally, a 'null-transformation' that transforms the lipophilic diene sidechain of **1** into an identical copy of itself was set up to test for the accuracy of the results.

Models of the ligands both bound to KAS I in solution as well as solvated in water are needed for a TI calculation of relative binding energies. For the ligand protein complexes, the first thiolactomycin ligand in the preequilibrated structure of KAS I was replaced by the new compound while the second was kept in place. The new compounds were placed into the binding site by keeping the atomic position of all atoms that occur both in thiolactomycin and the new compound unchanged. For models of the ligands in water, individual ligand molecules were embedded in a 12 Å deep water layer in a rectangular simulation box. For these systems, containing thiolactomycin derivatives, an additional equilibration run was performed before free energy data was collected. The equilibration consisted of 500 steps of steepest descent minimization to remove bad vdW contacts introduced by the newly placed ligand and 10 ps of NVT heating, followed by 50 ps of constant volume equilibration at 298 K as a final equilibration step. This 50 ps equilibration proved to be sufficient to reach a temperature of 298 K and density of ca. 1 g/cm^3. The density and temperature equilibrated sets of protein–ligand complexes and solvated ligands served as the starting structures for several consecutive rounds of 500 ps length TI simulations. All MD production runs were performed at 298 K in the NPT ensemble.

All TI free energy changes reported were calculated from using nineteen λ-points in order to construct the free energy curve by numerical evaluation of the integral in Eq. 1. The λ-points chosen were 0.05, 0.1, 0.15, ... 0.95 and the trapezoidal rule was used for the integration. Model systems were equilibrated at their respective λ-value for each simulation window before data collection started.

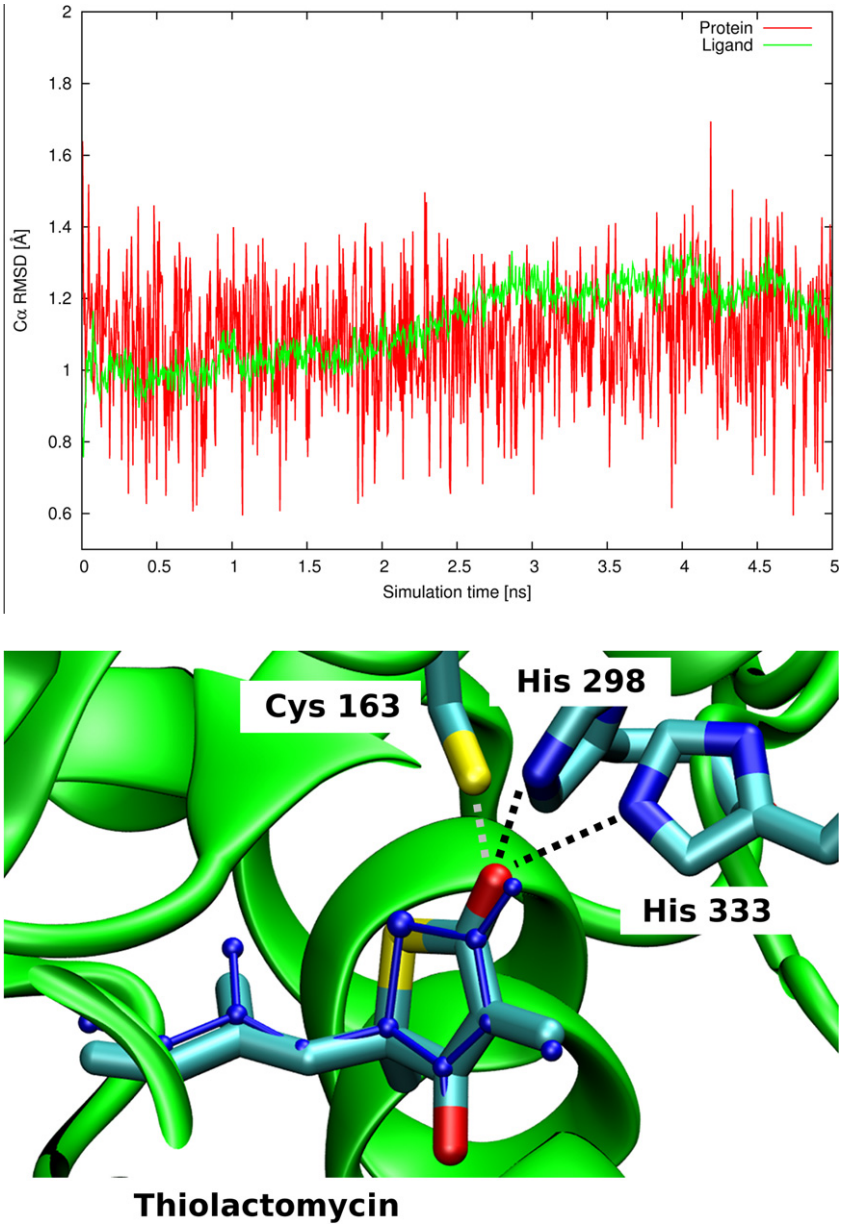

Figure 1. Structural changes in the KAS I thiolactomycine complex during five nanoseconds of equilibration. (Top) rmsd values in Å for the protein backbone α-carbon atoms and the bound thiolactomycin ligand involved in the TI transformations. The second ligand showed a similarly stable trajectory (not displayed). (Bottom) Comparison of the binding pose of thiolactomycin (colored by element) after 5 ns of equilibration with respect to the X-ray crystal structure (blue). The active site amino acid residues of the catalytic triad are displayed (from left to right Cys 163, His 298 and His 333, remainder of the protein in green). Conformational adjustments results in very small changes in ligand binding pose. The two H-bonds to the catalytic triad His residues are stable (black dashes) and an H-bond to Cys 163 is also seen (gray dashes).

3.3. Sampling and error estimation

In order to close the thermodynamic cycle that allows the calculation of relative binding free energies according to Eq. 7, the ligand transformations need to be simulated twice, once in complex and once with the ligand solvated in water, for every transformation shown in Figure 2. Absolute free energy changes for either transformation step ranged from −20 to +150 kcal/mol and are not reported here, because they cannot be connected to any observable. Only the free energy differences reported below are physically meaningful. Data collection was conducted from a total simulation time of 3 ns per λ-window for the protein–ligand complex transformations (for a total of 57 ns) and of 5 ns per λ-window for the solvated ligand transformation (for a total of 95 ns). Longer simulation times were used for the computationally much cheaper solvated ligand transformation steps to reduce statistical errors from this substep.

Errors result from the sampling error due to the finite simulations length, inaccuracies of the molecular mechanics model and numerical integration errors. Among these, only the first error contribution can be reliably estimated: For each individual simulation window, $\partial V/\partial\lambda$-values showed standard deviations in the range of 3–5 kcal/mol. The standard error of the mean for a TI transformation can be estimated by determining the number of independent samples from the $\partial V/\partial\lambda$-autocorrelation time, as described in Ref. [29]. However, such an analysis can severely underestimate errors, since the common assumption of single exponentially decaying $\partial V/\partial\lambda$ autocorrelation functions may not be true for the complex potential energy surfaces of protein–

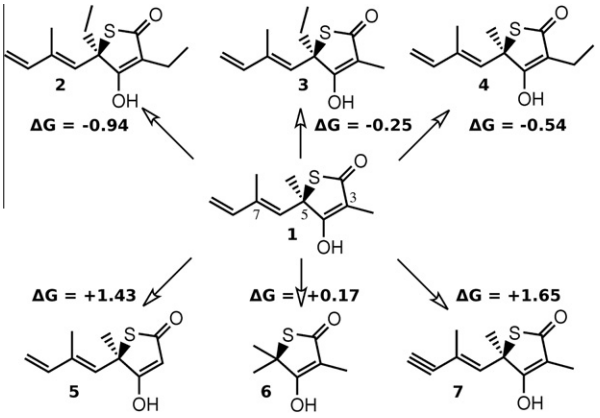

Figure 2. Six thiolactomycin derivatives studied by TI calculations. Every arrow symbolizes a binding free energy prediction, performed as two TI transformations for the complexed and solvated ligands respectively, each using 19 λ-windows per substep. All ligand transformations used the natural compound **1** as the initial state. Compounds **2** to **5** change the thiolactone ring substitution pattern, **6** and **7** represent changes in the ligand diene sidechain. $\Delta\Delta G$ gives the predicted changes in binding free energy for the different compounds in kcal/mol (see text, Table 1 for details).

ligand complexes. Additionally, other contributions may add significantly to the total error.[47]

Therefore, we have selected a simpler but more robust error estimation method here: Each TI transformation was broken down into consecutive 500 ps length windows and free energy changes were computed for each. Final results were obtained from averaging the free energy change for each substep (ligand transformation in water and in complex) and the standard error of the mean was estimated from GAUSSIAN error propagation. For the error values reported in Table 1, the contributions to the standard error from the complex transformation were typically ca. 2–3 times higher than those from the solvated ligand transformation.

When comparing this batch averaging scheme with error estimations from the $\partial V/\partial\lambda$-autocorrelation time as described above, we find the later yielded error values about four times smaller than those reported in Table 1.

To get another estimate of the expected accuracy of the simulation results, a 'null transformation' was set up, transforming the (2-methyl)-butadiene sidechain of thiolactomycin into an identical copy of itself. Since this transformation cannot result in a change of free energy, any deviation of the simulation result from zero must be due to insufficient sampling during the MD simulation. The two individual simulation results gave free energy changes of -0.28 ± 0.12 kcal/mol for the complex transformations and -0.02 ± 0.07 kcal/mol for the transformation in water, both values close to zero.

As expected, the transformation of the solvated ligand in isotropic surroundings gives results closer to the ideal value of zero. Combining those numbers yields a total free energy change for the 'null transformation' of -0.26 ± 0.13 kcal/mol. Since the (2-methyl)-butadiene sidechain is significantly larger than most functional group changes in the following, this can be considered a conservative error estimate and is indeed about twice as large as most standard errors from batch averaging in Table 1.

To estimate the degree to which insufficient sampling influences the results, we have conducted a cycle closure reverse transformation for the **4** to **1** transformation, starting from the final structure of the complex after 3 ns and the solvated ligands after 5 ns simulation time. Apart from the starting structures and direction of the transformation, the same protocol as for all other simulations was used. A total relative binding free energy of $+0.76 \pm 0.19$ kcal/mol was obtained, in good agreement to the forward transformation result of -0.54 ± 0.16 kcal/mol (the error estimate ranges overlap when the reversed sign is considered). We therefore conclude that the transformation direction and hysteresis effects have only little influence on our results.

Since determining appropriate atomic partial charges for novel compounds is a major problem in computational drug design,[48–50] we have investigated the effect that alternative partial charge derivation schemes have on our results. The **1** to **4** ligand transformation was repeated using ligands parameterized with AM1-BCC,[51,52] Mulliken[53] and Gasteiger[54] charges in addition to the RESP charges used in the remainder of this work. To reduce computational effort, we did calculate the relative solvation free energy of the compounds instead of their relative binding free energy (by conducting two TI transformations in water and in vacuum). The $\Delta\Delta G^0_{solv}$ averaged over 5 ns simulation time was 0.71 ± 0.07 kcal/mol for the RESP case, indicating that the additional methyl group slightly reduces compound solubility. In comparison, a relative solvation free energy of 0.54 ± 0.05 kcal/mol was found using AM1-BCC charges, while 0.31 ± 0.15 kcal/mol were computed with Mulliken charges and 0.38 ± 0.02 kcal/mol with Gasteiger charges. While the partial charge derivation method clearly influences the result, the widely used RESP and AM1-BCC charge sets give values in reasonable agreement, while Mulliken and Gasteiger charges, which were not designed specifically for solvated MD simulations, deviate slightly more.

3.4. Transformations of KAS I bound ligands

TI calculations for the six ligand transformations detailed in Figure 2 were conducted starting from structures preequilibrated as described above. For the 3 ns length ligand-protein complex simulations, system stability was monitored by computing the

Table 1
Functional groups changed during the TI transformations. Disappearing and appearing groups are unique to the initial and final states of the transformations. See Figures 2 and 3 for ligand structures. The **1** to **1** transformation represents a 'null transformations' used to estimate errors. The **1** to **5** transformation used the enol form of **5**. The $\Delta\Delta G_{Bind}$ columns give the binding free energy of the end state ligand relative to thiolactomycin combining data from 19×3 ns of complex and 19×5 ns of solvated ligand simulations. Negative values indicate the derivative to be a stronger binder. See Section 3.5. for the transformations involving **8** and **9**. See text for details

Transformation	Disappearing funct. groups	Appearing funct. groups	$\Delta\Delta G_{Bind}$ (kcal/mol)
1 → **1**	5-(2-Methyl)-butadienyl	5-(2-Methyl)-butadienyl	-0.26 ± 0.13
1 → **2**	3-Methyl, 5-methyl	3-Ethyl, 5-ethyl	-0.94 ± 0.15
1 → **3**	5-Methyl	5-Ethyl	-0.25 ± 0.10
1 → **4**	3-Methyl	3-Ethyl	-0.54 ± 0.16
1 → **5**	5-Methyl	5-H	$+1.43 \pm 0.12$
1 → **6**	5-(2-Methyl)-butadienyl	5-Methyl	$+0.17 \pm 0.34$
1 → **7**	7-Ethenyl	7-Ethinyl	$+1.65 \pm 0.14$
4 → **1**	3-Ethyl	3-Methyl	$+0.76 \pm 0.19$
4 → **8**	3-Ethyl	3-Isopropyl	$+0.52 \pm 0.13$
4 → **9**	3-Ethyl	3-Tertbutyl	-0.27 ± 0.11

ligand and protein rmsd values with respect to their starting positions for all simulation windows. For the protein Cα-rmsd values of ca. 1.1 Å were found during all simulations, indicating no significant conformational changes in the receptor. Ligand rmsd values uniformly rose to values in the 1.5–2.0 Å range and did not further increase after that.

An additional measure of complex stability was the continuing presence of the thiolactone-O His 333 H-bond which was found in >70% of all MD snapshots in all simulations. The H-bond between the thiolactone-O and His 297 was observed in ca. 40% of all MD snapshots, a slightly lower frequency of occurrence when compared to the initial MD simulations. The thiolactone-S—His 297 H-bond was also observed in ca. 40% of all MD snapshots, comparable to the frequency observed in the initial MD simulations of the thiolactomycin KAS I complex.

The free energy changes from each of these TI calculations, as well as for the null transformation of **1** into itself, are given in Table 1. These values include the correction for the ligand transformations in water and therefore represent the predicted differences in binding free energy to KAS I for thiolactomycin and each studied derivative. Note that a negative result indicates stronger binding for the ligand that **1** is transformed into. The calculated free energy changes range from −0.94 ± 0.15 to +1.65 ± 0.14 kcal/mol, equivalent to more than a two order of magnitudes spread in binding constants.

For the transformations of **1** into **5** or **7**, unfavorable relative binding free energies of more than +1 kcal/mol were found, while for **6** a small decrease in affinity, within the margin of error, is predicted. This indicates that these three thiolactone derivatives will show diminished or at best nearly unchanged binding strength to KAS I compared to the natural compound. Removal of the lipophilic sidechain or the 5-methyl group as well as the seemingly small chemical change of the terminal ethenyl into an ethinyl group are all predicted to negatively influence the ligand's binding strength.

The **1** to **5** transformation was conducted with the enol tautomer of the ligands. While it is assumed that the enol form is the stable one in aqueous solution, no such information for a ligand protein complex is available. Therefore, an additional TI calculation was set up in which the enol form of the ligand was transformed into its keto tautomer both in water and as a KAS I complex. Using the same equilibration and data collection process as detailed above, a relative binding free energy for the two ligand tautomers of +1.4 ± 0.16 kcal/mol was found. The free energy change for the keto/enol transformation cannot be directly compared to experimental results of the two compound's relative stability (because a molecular mechanics forcefield is not suitable to describe such chemical changes), but a thermodynamic cycle similar to the one above shows that the relative binding free energy of the two tautomers is equal to the difference in their relative stability solvated in water and bound to the protein:

$$\begin{aligned}\Delta\Delta G^0_{keto/enol} &= \Delta G^0_{TI,keto\rightarrow enol,bound} - \Delta G^0_{TI,keto\rightarrow enol,solvated} \\ &= \Delta\Delta G^0_{Bind} = \Delta G^0_{Bind,keto} - \Delta G^0_{Bind,enol} \end{aligned} \quad (8)$$

The result indicates that the binding site does not stabilize the ligand in its keto form, which would be an even weaker binder to KAS I than the enol tautomer. Therefore, the keto form was not considered further here.

The cases of adding methyl groups to the ligand (by changing methyl substituents into ethyl substituents) produce more favorable relative binding free energies, predicting increased binding strengths for three of the new ligands. For the **1** to **2** transformation a relative binding free energy of −0.94 ± 0.15 kcal/mol is calculated, predicting an increase in binding strength significantly larger than the estimated margin of error. For the **1** to **3** transformation, a relative binding free energy of −0.25 ± 0.10 kcal/mol is found, indicating a slightly higher binding strength for the **3** derivative than for the natural compound but the free energy difference computed is close to the estimated error threshold.

For the transformations into **2** and **4** a marked increase in the ligand binding strength of 0.54 ± 0.16 and 0.94 ± 0.15 kcal/mol respectively is predicted. This free energy change is significantly larger than the estimated error in both cases. The two ligand for which increased binding strength is predicted both increase the size of the functional group in the 3-position from a methyl to an ethyl group.

The chemical change to generate ligand **2** (change both the 3- and 5-methyl into ethyl groups) is the sum of the changes to generate ligands **3** and **4** (change the 5- or the 3-methyl into ethyl groups, respectively) and the resulting relative binding free energies computed for the latter two transformations approximately add up to the result for the former. This is not necessarily expected, since **3** and **4** will have different equilibrium conformations (induced by adding additional sterical load at different ligand positions) and the effect of changing one part of a ligand can in principle be strongly influenced by previous changes to it. In this case however, such cooperative effects seem to be small and ligand optimization could be conducted on both the 3- and 5-positions of **1**.

3.5. Potential new ligands

The results from the previous chapter suggest that increasing the size of the substituent in the 3-position (and to a smaller extent the 5-position as well) of the thiolactone ring leads to compounds that show improved binding strength to KAS I. Therefore two new ligands not yet synthesized were created, **8** and **9** which replace the 3-methyl with an isopropyl or *t*-butyl group, respectively (Fig. 3). Two TI calculations were set up in which **4** was transformed into **8** and **9** respectively. Starting structures for the protein–ligand complexes were taken from the final structures of the **1** to **4** transformation. As before, 5 ns length MD simulations were performed for the ligand transformations in water and 3 ns MD were performed for the transformation of protein bound ligands in both cases. The protein and ligand rmsd values were measured during the MD simulations, with protein rmsd values staying constant at ca. 1.1 Å and stable ligand rmsd values in the range of 1.5 Å with respect to the starting structure.

The free energy changes computed from the TI calculations resulted in a relative binding free energy of +0.52 ± 0.13 kcal/mol for **8** and −0.27 ± 0.10 kcal/mol for **9**. This predicts **8** to be an inhibitor with a lower binding strength than **4**, but comparable to **1** (with a difference in binding strength for **1** and **8** below the estimated error) and **9** an inhibitor with slightly increased binding strength compared to **4**, making it also a stronger binder than the natural compound thiolactomycin.

Our calculations resulted in an unexpected pattern of changes in the binding free energy for the sequence of ligands **1**, **4**, **8** and **9**. These differ in a steadily increasing size of their 3-substituent

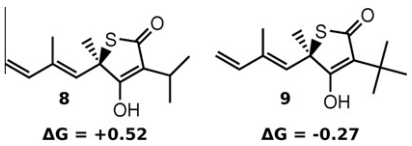

Figure 3. Two proposed new ligands which increase the size of the thiolactone ring 3-substituent. The simulations set up for **8** and **9** changed the 3-ethyl group of **4** into an iso-propyl or *t*-butyl group respectively. $\Delta\Delta G$ gives the predicted changes in binding free energy for the different compounds in kcal/mol (see text, Table 1 for details).

from methyl to ethyl, *iso*-propyl and *t*-butyl, but the TI calculations order the compounds with respect to their binding strength as **8** ≈ **1** < **4** < **9**. Nevertheless, a slightly larger residue at the 3-position of **1** appears beneficial for the compound binding strength. Organic synthesis of the proposed new ligands can now be undertaken to test if this predicted order of binding strengths is sustained by experimental data.

4. Discussion

MD simulations of thiolactone ligand KAS I complexes yielded stable complexes that provide insights into the ligand binding modes on an atomic level. The critical interaction of thiolactomycin with both active side histidines was reproduced in the simulations and consistently predicted for its derivatives as well. The change in hydrogen bonding pattern that produced a hydrogen bond involving the thiolactone sulphur atom resulted only in minor changes of the ligand binding pose. This hydrogen bond might be caused by an artificial preference of the force field for hydrogen bonds involving sulphur atoms.

The TI calculations in this study were used to compute the binding free energy differences between 8 new potential inhibitors of the bacterial fatty acid synthase KAS I and the natural inhibitor thiolactomycin. Four inhibitors studied, **2, 3, 4** and **9**, were predicted to bind stronger to the receptor than thiolactomycin. The most promising of these compounds exchanged the 3-methyl substituent of thiolactomycin with a bulkier residue (an ethyl or *t*-butyl group). Furthermore, a thiolactomycin derivative with a slightly larger substituent at the thiolactones 5-position (**2** and **3**) was also predicted to have a slightly higher binding affinity to the receptor. However, all changes in the diene sidechain were predicted to diminish the binding affinity of the ligand. Several previous studies deal with thiolactomycin derivatives with modified diene sidechains, but variations of the thiolactone ring substitution pattern have received less attention so far, mainly because no efficient and flexible synthesis for such derivatives existed until recently. This study highlights the potential of thiolactone ring variations in the design of thiolactomycin derivatives with improved binding strength.

Despite their high computational expense free energy calculations are regarded as useful tools to accurately compute relative binding free energies and our results support this. The binding strength differences observed here are small, below 1 kcal/mol, which is not uncommon for ligand optimization studies. Simpler binding energy estimation tools like Ligand Docking and MM-PBSA would not be capable of reliably predicting such small changes in affinity. This finding is in good agreement to previous studies.[55–57]

From the 'null-transformation' of thiolactomycin into itself described in Section 3.3, an error of 0.26 kcal/mol was estimated for the TI calculations. The simulation showed that the error mainly stems from the simulation of the protein bound ligand transformation.

This work suggests that for similar applications of TI ligand optimizations, MD simulations in the low nanosecond range (with less than a microsecond of total complex simulation length invested here overall) can yield good free energy estimates for small molecule transformations. This points to their applicability to guide the design of new compounds by precise estimates of their binding strengths, even though more method validation is desirable.

We therefore conclude that TI free energy calculations are useful tools to study protein–ligand interactions. The increasing availability of high performance computing will make these methods more widely applied as parts of the computational drug design toolkit.[58,59] In the future, combining MD based free energy predictions with ligand docking calculations may well allow the rapid and accurate prediction of the binding properties of new compounds in silico. While the application of TI calculations to protein–ligand complexes is still a developing field, we anticipate these methods to become routine tools for computational chemists in the near future.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bmc.2012.04.019. These data include MOL files and InChiKeys of the most important compounds described in this article.

References and notes

1. Antibiotic Runoff, *The New York Times*, Sep 18th 2007, Editorial.
2. ''How to Avoid, Identify And Treat Staph Germ'', *The Washington Post*, Oct 18th 2007, B.2.
3. 'Preventing overuse of antibiotics', *The Irish Times*, <http://www.irishtimes.com/newspaper/letters/2012/0103/1224309735633.html>, Jan 3rd, 2012.
4. 'Concern grows over rise of antibiotic resistant salmonella', *The Daily Mail*, <http://www.dailymail.co.uk/news/article-2078802/Concern-grows-antibiotic-resistant-salmonella.html> Dec 26th 2011.
5. Singh, S. B.; Barrett, J. F. *Biochem. Pharmacol.* **2006**, *71*, 1006.
6. Butler, M. S.; Buss, A. D. *Biochem. Pharmacol.* **2006**, *71*, 919.
7. Michel, J.; Essex, J. *J. Med. Chem.* **2008**, *51*, 6654.
8. Michel, J.; Essex, J. *J. Comput. Aided Mol. Des.* **2010**, *24*, 639.
9. Ali, A.; Bandaranayake, R.; Cai, Y.; King, N.; Kolli, M.; Mittal, S.; Murzycki, J.; Madhavi, M.; Nalivaika, E.; Özen, A.; Prabu-Jeyabalan, M.; Thayer, K.; Schier, C. *Viruses* **2010**, *2*, 2509.
10. Bortolato, A.; Moro, S. *Expert Opin. Drug Disc.* **2008**, *3*, 579.
11. Jorgensen, W. *Science* **2004**, *303*, 1813.
12. Gilson, M.; Zhou, H. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 21.
13. deAzevedo, W.; Dias, R. *Curr. Drug Targets* **2009**, *9*, 1031.
14. Klebe, G. *Drug Discovery Today* **2006**, *11*, 580.
15. Gohlke, H.; Klebe, G. *Angew. Chem., Int. Ed.* **2002**, *41*, 2644.
16. Khandekar, S. S.; Daines, R. A.; Lonsdale, J. T. *Curr. Protein Pept. Sci.* **2003**, *4*, 21.
17. Olsen, J. G.; Kadziola, A. v.; Wettstein-Knowles, P.; Siggaard-Andersen, M.; Larsen, S. *Structure* **2001**, *9*, 233.
18. Davies, C.; Heath, R. J.; White, S. W.; Rock, C. O. *Structure* **2000**, *8*, 185.
19. Moche, M.; Dehesh, K.; Edwards, P.; Lindqvist, Y. *J. Mol. Biol.* **2001**, *305*, 491.
20. Scarsdale, J. N.; Kazanina, G.; He, X.; Reynolds, K. A.; Wright, H. T. *J. Biol. Chem.* **2001**, *276*, 20516.
21. Lomakin, I. B.; Xiong, Y.; Steitz, T. A. *Cell* **2007**, *129*, 319.
22. Price, A. C.; Choi, K.; Heath, R. J.; Li, Z.; White, S. W.; Rock, C. O. *J. Biol. Chem.* **2001**, *276*, 6551.
23. Oishi, H.; Noto, T.; Sasaki, H.; Suzuki, K.; Hayashi, T.; Okazaki, H.; Ando, K.; Sawada, M. *J. Antibiot.* **1982**, *35*, 391.
24. Dormann, K. L.; Brueckner, R. *Angew. Chem.* **2007**, *119*, 1178.
25. Dormann, K. L.; Brueckner, R. *Angew. Chem., Int. Ed.* **2007**, *46*, 1160.
26. Dormann, K.L. *Ph.D. Thesis,* University of Freiburg, 2007.
27. Shirts, M. R.; Pande, V. S. *J. Chem. Phys.* **2005**, *122*, 134508.
28. Steinbrecher, T.; Mobley, D. L.; Case, D. A. *J. Chem. Phys.* **2007**, *127*, 214108.
29. Steinbrecher, T.; Joung, I.; Case, D. A. *J. Comput. Chem.* **2011**, *32*, 3253.
30. Ponder, J. W.; Case, D. A. *Adv. Prot. Chem.* **2003**, *66*, 27.
31. Case, D. A.; Darden, T. A.; Cheatham III, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossvary, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. AMBER11 University of California: San Francisco, 2010.
32. Chandrasekhar, S. *Rev. Mod. Phys.* **1943**, *15*, 1.
33. Ryckaert, J. P.; Cicotti, G.; Berendsen, H. P. *J. Comput. Phys.* **1997**, *23*, 327.
34. Ewald, P. *Ann. Phys.* **1921**, *64*, 253.
35. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
36. Darden, T.; York, D.; Pedersen, L. G. *J. Chem. Phys.* **1993**, *98*, 10089.
37. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157.
38. Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269.
39. Frisch, M. J. et al. GAUSSIAN 03, Revision C.02 (Gaussian, Inc.: Wallingford CT, 2004).
40. Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T. *J. Comput. Chem.* **2003**, *24*, 1999.
41. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
42. Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.
43. Pappenberger, G.; Schulz-Gasch, T.; Kusznir, E.; Mueller, F.; Hennig, M. *Acta Crystallogr., Sect D* **2007**, *63*, 1208.
44. D'Agnolo, G.; Rosenfeld, I. S.; Vagelos, P. R. *J. Biol. Chem* **1975**, *14*, 5289.
45. Garwin, J. L.; Klages, A. L.; Cronan, J. E. *J. Biol. Chem* **1980**, *255*, 11949.
46. Dormann, K. L.; Steinbrecher, T.; Grond, S.; Labahn, A.; Brueckner, R. in preparation.

47. Chodera, J. D.; Mobley, D. L.; Shirts, M. R.; Dixon, R. W.; Branson, K.; Pande, V. S. *Curr. Opp. Struct. Biol.* **2011**, *21*, 150.
48. Reddy, M. R.; Erion, M. D. *J. Am. Chem. Soc.* **2007**, *129*, 9296.
49. Reddy, M. R.; Singh, U. C.; Erion, M. D. *J. Am. Chem. Soc.* **2011**, *133*, 8059.
50. Rathore, R. S.; Reddy, R. N.; Kondapi, A. K.; Reddanna, P.; Reddy, M. R. *Theor. Chem. Acc.* **2012**, *131*, 1096.
51. Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2000**, *21*, 132.
52. Jakalian, A.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2002**, *23*, 1623.
53. Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833.
54. Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219.
55. Jorgensen, W. *Acc. Chem. Res.* **2009**, *42*, 724.
56. Steinbrecher, T.; Case, D. A.; Labahn, A. *J. Med. Chem.* **2006**, *49*, 1837.
57. Steinbrecher, T.; Labahn, A. *Curr. Med. Chem.* **2010**, *17*, 767.
58. Deng, Y. Q.; Roux, B. *J. Chem. Theor. Comput.* **2006**, *2*, 1255.
59. Hamelberg, D.; McCammon, J. A. *J. Am. Chem. Soc.* **2004**, *126*, 7683.