

Molecular dynamics of *Mycobacterium tuberculosis* KasA: implications for inhibitor and substrate binding and consequences for drug design

Benjamin Schaefer · Caroline Kisker ·
Christoph A. Sotriffer

Received: 12 July 2011 / Accepted: 2 November 2011 / Published online: 11 November 2011
© Springer Science+Business Media B.V. 2011

Abstract Inhibition of the production of fatty acids as essential components of the mycobacterial cell wall has been an established way of fighting tuberculosis for decades. However, increasing resistances and an outdated medical treatment call for the validation of new targets involved in this crucial pathway. In this regard, the β -ketoacyl ACP synthase KasA is a promising enzyme. In this study, three molecular dynamics simulations based on the wildtype crystal structures of inhibitor bound and unbound KasA were performed in order to investigate the flexibility and conformational space of this target. We present an exhaustive analysis of the binding-site flexibility and representative pocket conformations that may serve as new starting points for structure-based drug design. We also revealed a mechanism which may account for the comparatively low binding affinity of thiolactomycin. Furthermore, we examined the behavior of water molecules within the binding pocket and provide recommendations how to handle them in the drug design process. Finally, we analyzed the dynamics of a channel that accommodates the long-chain fatty acid substrates and, thereby, propose a mechanism of substrate access to this channel and how products are most likely released.

Keywords *Mycobacterium tuberculosis* · KasA · Molecular dynamics simulations · Drug design

Abbreviations

| | |
|------|-----------------------------------|
| ACP | Acyl carrier protein |
| FAS | Fatty acid synthase |
| INH | Isoniazid |
| KasA | β -Ketoacyl ACP synthase |
| MD | Molecular dynamics |
| MTB | <i>Mycobacterium tuberculosis</i> |
| PDB | Protein data bank |
| PEG | Polyethylene glycol |
| TB | Tuberculosis |
| TLM | Thiolactomycin |

Introduction

Mycobacterium tuberculosis (MTB) presents the most prominent causative agent of tuberculosis (TB). With more than 9 million new infections and about 2 million deaths per annum, TB belongs to the world's most lethal infectious diseases [1, 2]. It is estimated that 9.8 million new infections occurred in 2010, more than in any other year in history [3, 4]. A tedious and outdated first-line drug treatment from the 1960s, as well as the increasing emergence of multi-drug-resistant and extensively drug-resistant tuberculosis strains call for the identification and addressing of new targets to retrieve comprehensive susceptibility [5].

Mycolic acids are essential for MTB. They present a major part of the thick, waxy cell wall and, thus, provide the pathogen with high stability and impermeability to many antibiotics [6]. Mycolic acids are long fatty acid molecules that are synthesized by fatty acid synthase (FAS) systems. These may either be FAS-I systems, in which each enzymatic activity resides on a single polypeptide chain of a very large multifunctional enzyme [7, 8], or

B. Schaefer · C. A. Sotriffer (✉)
Institute of Pharmacy and Food Chemistry,
University of Würzburg, 97074 Würzburg, Germany
e-mail: sottriffer@uni-wuerzburg.de

C. Kisker
Rudolf Virchow Center for Experimental Biomedicine,
Institute for Structural Biology, University of Würzburg,
97080 Würzburg, Germany

FAS-II systems, in which each of the enzymatic activities corresponds to an individual polypeptide. Some bacteria, including mycobacteria, possess both an FAS-I and an FAS-II system, the latter preferring C₁₆ fatty-acid substrates and extending them up to C₅₆ chain lengths as precursors of mycolic acids [8].

Inhibition of InhA, an enzyme of the FAS-II pathway, by the first-line drug isoniazid (INH) has been an established way of fighting TB for decades [9, 10]. However, since resistance related to its special covalent binding mode is increasing [11, 12], it has become necessary to seek for alternative targets involved in this crucial pathway. In this context, the β -ketoacyl ACP synthase KasA is a promising enzyme. Induced cell lysis upon depletion shows that this enzyme is essential in mycobacteria [13]. Furthermore, the known natural product inhibitors thiolactomycin (TLM, Fig. 1), cerulenin, and platensimycin prove the general druggability of this target [14–17]. KasA is active as a homodimer and catalyzes the condensation between a malonyl functionality bound to the MTB-specific acyl carrier protein (ACP), and the growing acyl chain [18]. Initially, the acyl chain is transferred to the active site Cys171 to build an acylated KasA intermediate. This step is followed by the actual condensation reaction of malonyl-ACP with KasA, which elongates the acyl chain by two further carbon atoms [19].

The crystal structures of the KasA apo-enzyme and of the complex with TLM have recently been solved [18] and serve as the starting point of our study. We present, for the first time, a thorough analysis of the dynamics of this protein by means of three molecular dynamics (MD) simulations—one for the complex and two using different unliganded states of the protein. Our studies reveal a dual character of the binding site with respect to its flexibility. By means of hierarchical cluster analysis, we analyzed relevant alternative binding-pocket conformations which may serve as starting points for structure-based ligand design. We further revealed a mechanism that on one hand may explain why TLM binds only moderately to wildtype KasA and, on the other, shows the importance of a water molecule in determining the accessibility of the long acyl-binding channel.

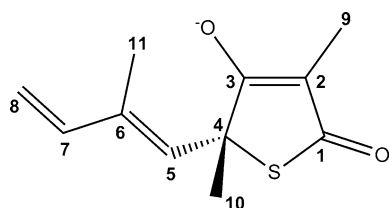


Fig. 1 Structural formula of thiolactomycin (TLM), with numbering of the carbon atoms as used in the text

Materials and methods

Protein and ligand preparation

A total of three MD simulations were initiated: one for the unaltered wildtype structure complexed with TLM (PDB code 2WGE), a second for the uncomplexed state using the apo structure as starting point (PDB code 2WGD); and a third using the complex structure, but removing the ligand prior to the simulation. The latter will be referred to as the ‘perturbed simulation’. As both the complexed and the uncomplexed wildtype state crystallized with one monomer in the asymmetric unit [18], the respective second monomer of the homodimer was created by applying the corresponding symmetry operations using PyMOL [20] (Fig. 2a). Supported by visual inspection, the binding site was defined as all amino acids within a distance of 4 Å to the ligand TLM, giving a total of fifteen residues as shown in Fig. 2b.

Although both the complexed and the unliganded protein were crystallized with similar resolutions in the same space group and under identical experimental conditions, the amount of water molecules resolved in the structures differs significantly, being 183 for apo KasA and 281 for complexed KasA. We therefore did not preserve all crystallographic water molecules, but included only the well-resolved crystallographic water molecules located within the binding-site region and the acyl-binding channel (comprising residues 115–147 as defined by Luckner et al. [18]). With respect to the binding-site region, this led to the inclusion of buried water molecules within 3.2 Å of the binding-site heavy atoms. In addition, the water molecule coordinated to the hydroxyl group of Ser346 and the carbonyl oxygen of Gly344 was retained, as it was deemed important due to the direct proximity of the two amino acids to the catalytic residue His345. Finally, we replaced a glycerol molecule that is present in the acyl-binding channels of the crystal structures by placing three additional water molecules on the coordinates of the glycerol oxygens. Such a replacement procedure was not performed for TLM after extracting the ligand in the perturbed simulation in order to examine the effect on a binding pocket which shows an artificially low hydration state. In total, 54 water molecules were kept for the apo state as well as 44 for the complex and the perturbed simulation. Additional molecules present from the crystallization conditions (e.g. isopropanol) were not included in the calculations.

The tleap module of AMBER10 [21] was used for adding hydrogens and assigning the ff99SB force field parameters. Histidines were protonated at the ϵ -nitrogen, except for His311 of the catalytic triad which was protonated at the δ -nitrogen. Hybrid quantum mechanics/molecular mechanics (QM/MM) calculations by Lee et al.

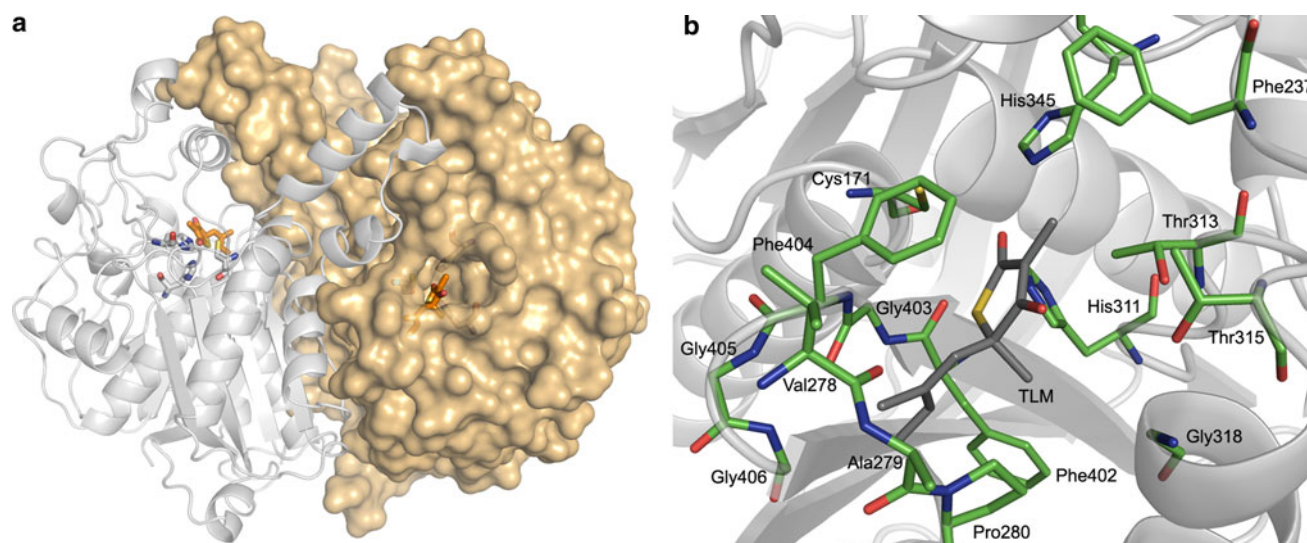


Fig. 2 **a** Dimer of wildtype KasA in complex with TLM (orange carbons). One monomer of the homodimer is shown in cartoon, the other in surface representation. Residues of the catalytic triad are

depicted as sticks with gray carbons. **b** Depiction of the immediate binding pocket. Active-site residues (green carbons) and the ligand (dark gray carbons) are shown in stick representation

[22] revealed that the active site of KasA is able to easily switch between a zwitterionic state with Cys171 deprotonated and His311 positively charged, and a neutral state with His311 protonated at the δ -nitrogen. In both cases, an energetically favorable hydrogen bond is established between the δ -NH group and a water molecule (cf. Fig. 6). For our calculations, the neutral state was chosen as the starting point. Atomic charges of the ligand were calculated using the RESP [23] methodology at the HF/6-31G* level in Gaussian 03 [24]. Missing force field parameters of the ligand were determined using the parmchk [25] module in AMBER10 according to the GAFF force field [26]. The correct GAFF atom and bond types of the ligand were determined with the antechamber program [25]. After a short energy minimization of 200 steps using a generalized Born implicit solvent model [27, 28], the systems were again loaded into tleap and solvated in a box of TIP3P water molecules [29]. This resulted in simulation box sizes of $\sim 84 \text{ \AA} \times 89 \text{ \AA} \times 103 \text{ \AA}$ and $\sim 19,600$ water molecules.

Molecular dynamics simulations

All three MD simulations were performed by means of the SANDER module of AMBER10 [21] with the *ff99SB* force field. After the minimization and solvation step, the solvent molecules were disordered in a constant-volume simulation, keeping the solute fixed with the *ibelly* option. For this purpose, the system was heated from 100 to 300 K for 20 ps and then cooled back to 100 K over a period of 5 ps, using the Berendsen weak coupling algorithm with a 0.5 ps time constant. After this procedure, the whole system (i.e.

solute and solvent) was allowed to move and gradually brought to 300 K over 25 ps. Subsequently, the constant-volume periodic boundary conditions were switched to constant pressure, and the simulation was continued over 15,050 ps. Pressure was adjusted by isotropic position scaling, with a default reference pressure of 1 bar and a pressure relaxation time of 1 ps. The temperature was kept constant by setting the time constant of the coupled heat bath to 1 ps. Covalent bonds involving hydrogen atoms were constrained by the SHAKE algorithm [30], and a time step of 0.002 ps was used. An 8 Å cutoff was applied for the van der Waals interactions, while the electrostatics were treated by the Particle Mesh Ewald method [31]. After the first 50 ps, trajectory data were collected for further analysis to give a total time of 15,000 ps per dimer. Coordinates were saved at intervals of 1 ps. Analyses were performed by means of the ptraj module of AMBER10 and *carnal* of the AMBER7 suite, while for visualization VMD [32] and PyMOL [20] were used. The hierarchical cluster analysis was performed with SYBYL-X [33].

Results and discussion

KasA is a homodimeric protein comprising 416 amino acids per subunit (Fig. 2a). An illustration of the binding site is provided in Fig. 2b. The isoprenoid chain and the methyl group C¹⁰ of TLM, as well as C⁴ and the sulfur of the thiolactone ring are accommodated by Val278, Ala279, Pro280, and Gly318 on one side and Phe402, Gly403, Phe404, Gly405, plus Gly406 on the other. The methyl

group C⁹ is located in a sub-pocket formed by Phe237, Thr313, and Thr315. The carbonyl oxygen of the thiolactone ring points towards the residues of the catalytic triad, Cys171, His311, and His345, where a hydrogen bond to the protonated ϵ -nitrogen of His345 seems geometrically favorable.

For all three simulations, a high stability of the overall fold was observed. The root-mean-square deviation (RMSD) of the C ^{α} -atoms from the X-ray structure is only 1.21 ± 0.07 Å for the complex and equally low for the apo structure (1.16 ± 0.07 Å). For the perturbed simulation, an RMSD of 1.37 ± 0.13 Å was calculated; when compared to the unperturbed simulations, a trend towards higher values can be observed until ~ 12 ns, when a plateau is reached at ~ 1.6 Å (cf. Fig. 3). The fact that the value for the perturbed simulation is rather low and comparable to the values obtained for the unperturbed trajectories suggests that mainly local conformational changes take place. Higher values are observed when all protein atoms are taken into account for RMSD calculation instead of only the C ^{α} -trace: 1.90 ± 0.11 Å for the complex, 1.86 ± 0.08 Å for the apo, and 2.04 ± 0.15 Å for the perturbed simulation, respectively. This provides an indication that structural changes of the protein are to a considerable extent rather due to re-orientations of side chains than to alterations of the overall fold. Furthermore, only moderate overall fluctuations seem to occur, as indicated by the low standard deviations.

Dynamics of the binding pocket

Considering only the fifteen binding-site residues (as defined in Materials and Methods), higher standard

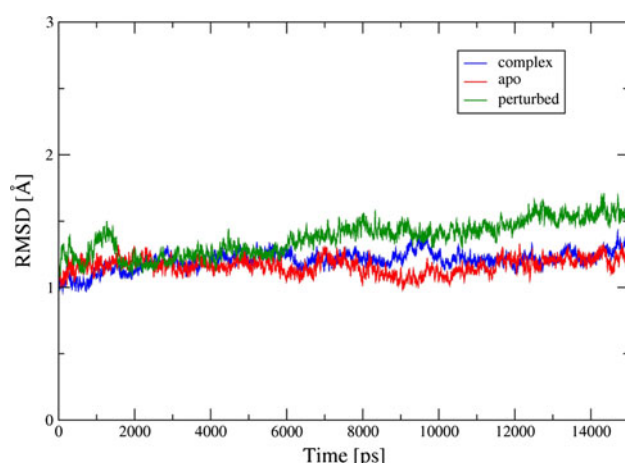


Fig. 3 RMSD values of all C ^{α} -atoms with respect to the starting X-ray structure for each simulation as a function of simulation time

Table 1 Simulation averages ($t = 15$ ns) of the RMSD values (Å) with respect to the starting structure for the entire KasA dimer and the binding pockets

| | Complex structure | Apo structure | Perturbed structure |
|--|-------------------|-----------------|---------------------|
| Protein (C ^{α}) | 1.21 ± 0.07 | 1.16 ± 0.07 | 1.37 ± 0.13 |
| Protein (all atoms) | 1.90 ± 0.11 | 1.86 ± 0.08 | 2.04 ± 0.15 |
| Pocket A (C ^{α}) | 0.82 ± 0.10 | 1.53 ± 0.27 | 1.35 ± 0.29 |
| Pocket A (all atoms) | 1.41 ± 0.14 | 2.11 ± 0.34 | 1.94 ± 0.26 |
| Pocket B (C ^{α}) | 1.11 ± 0.13 | 0.98 ± 0.21 | 1.04 ± 0.27 |
| Pocket B (all atoms) | 1.78 ± 0.20 | 1.67 ± 0.43 | 1.41 ± 0.21 |

deviations of the respective RMSD values are observed in all three simulations (Table 1), suggesting a larger degree of mobility compared to the overall structure. Furthermore, a stabilizing effect of the ligand on the pocket conformation is revealed, as a distinctly more flexible behavior is monitored for the active-site residues when TLM is not included in the simulations. Besides this enhanced degree of mobility, conformational changes are observed that only occur in the absence of the ligand.

2D-RMSD calculations

To illustrate the conformational space sampled in the three 15-ns simulations, a combined 2D-RMSD plot was recorded. Snapshots were taken at regular intervals of 50 ps for each monomer, resulting in a total of 1,800 different sets of coordinates. The fifteen binding-site residues were superimposed and then used to calculate mutual RMSD values. The resulting plot is shown in Fig. 4. In the following, one pocket of the KasA homodimer will be referred to as pocket A, while the second will be called pocket B. This naming will be kept for the remaining sections.

When considering each simulation separately, maximum deviations of 2.47 Å (pocket A) and 2.90 Å (pocket B) are calculated for the complex (bottom left third of Fig. 4), while the values are 2.84 Å/2.49 Å for the apo (central four squares of Fig. 4), and 3.19 Å/2.58 Å for the perturbed structure (top right third), respectively. For the combined 2D-RMSD calculations, deviations of up to 3.54 Å are found. The plot clearly shows that the conformational states obtained for both pocket A and pocket B of the complex stay close to the starting structure throughout the simulation, which again proves the stabilizing effect of the ligand on the binding site. The conformations of the bound structure, however, are evidently different from the majority of the conformations that occur when the ligand is not present. Also, and expectedly, both uncomplexed structures bear a significantly more diverse set of conformations.

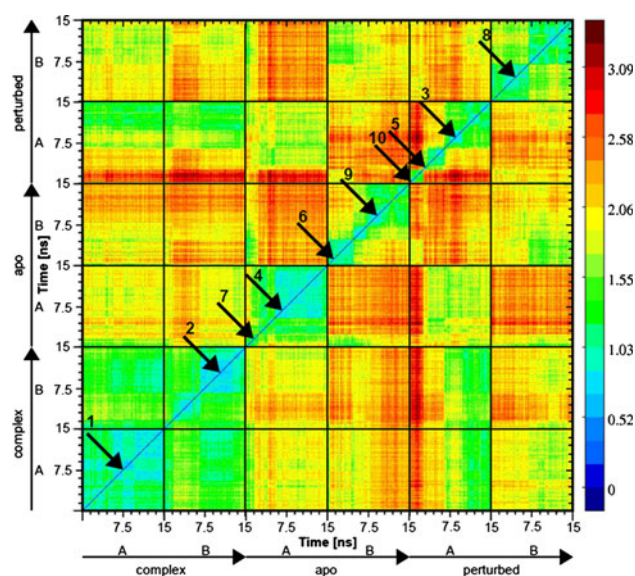


Fig. 4 Combined 2D-RMSD plot for the binding-site residues of all three simulations. Pockets are delineated as *black-framed squares*. Calculations are based on an all-atom fit of the binding-site residues. *Numbered arrows* in the plot indicate snapshots that were chosen as distinct representatives after hierarchical cluster analysis

The 2D-RMSD plot further shows that the sampled binding-pocket conformations are not equally distributed between the two subunits of each homodimer on the simulation time scale. Accordingly, one would need to extend the simulation times to accomplish a more exhaustive coverage of conformational space. Nevertheless, with a timescale of 90 ns in total, the three trajectories already allow valuable insights into the dynamics of this system and, not least, represent an appropriate compromise in terms of computational efforts.

Hierarchical clustering

To further probe the sampled conformational space, a hierarchical cluster analysis based on the 2D-RMSD calculations was performed. A similarity cutoff was chosen to obtain a reasonable number of ten clusters which at the same time cover the conformational space of all simulations. The respective members were used to calculate average structures and the member bearing the lowest RMSD value compared to the mean structure of a cluster was chosen as cluster representative. The snapshots thus selected are illustrated in Fig. 5 and indicated by black arrows in Fig. 4. All conformations of the complex simulation are found in the first two clusters. Cluster 1 mostly contains states from pocket A, while Cluster 2 is dominated by structures of pocket B. In addition, several conformations of pocket A of the perturbed simulation are included in these two clusters. This partial similarity between states of the complex and the perturbed structure is well

illustrated in the 2D-RMSD plot by green crosspeaks (cf. top left third or bottom right third of Fig. 4). The remaining eight clusters include conformations solely from the apo and perturbed structure, which clearly reflects the fact that a considerably enhanced conformational diversity is provided by the two uncomplexed simulations.

Binding-site flexibility

Comparison of the ten selected representative conformations, combined with the analysis of the per-residue RMSD values and the corresponding standard deviations of the binding-pocket amino acids reveals that the active site is composed of a rather rigid and a more flexible half. The stable part includes the residues of the catalytic triad, Cys171, His311, and His345, plus Phe237, which is directly adjacent. The standard deviations of the per-residue RMSD values (averaged over all three simulations) are 0.22 Å for Cys171, 0.23 Å for His311, 0.19 Å for His345, and 0.27 Å for Phe237, respectively, which clearly reflects restricted mobility within this region of the pocket. The remaining, more flexible residues show average standard deviations of up to 0.67 Å. Such a dual character of the binding site is a common finding for enzymes and appears to be important if not essential for enzymatic activity [34].

Analysis of the more rigid part reveals a complex interaction pattern of Phe237 and the catalytic triad as shown in Fig. 6. The stability of Cys171 and His311 is partly attributed to three crystallographic water molecules which are not located within, but directly adjacent to the binding pocket and whose positions are precisely retained throughout all simulations, with occupancy rates between 95 and 99%. Distance measurements reveal that one water molecule forms hydrogen bonds towards the backbone oxygens of Cys171 and Ile347, as well as the nitrogen of Gly351. A second water molecule interacts with the oxygen of Ala310 and thereby also stabilizes the neighboring backbone of His311. Further hydrogen bonds are found for this water molecule with the carboxylate moiety of Glu354 and, more importantly, the backbone nitrogen of Lys340. By immobilizing this residue, the charged side-chain nitrogen of Lys340 acts as a permanent donor to the carbonyl oxygen of His311. Further stabilization is achieved by a hydrogen bond between the His311 backbone nitrogen and the carboxylate oxygens of Glu322. Finally, the flexibility of the His311 side chain is limited by an interaction between the δ -nitrogen and a third water molecule that also interacts with the side-chain oxygen of Glu354.

For His345 and Phe237, conformational backbone stability is exclusively due to intramolecular interactions, specifically with Ala215, Ser217, Val238, and Gly240 (Fig. 6b). The Phe237 side chain is kept in place by non-polar interactions with Met216, Phe239, and His345, all of

which display only small average positional changes below 1 Å. In the complex simulation, the side chains of the catalytic triad are stabilized by the thioester oxygen of the ligand. In the uncomplexed structures, however, this position is partly taken by alternating water molecules, showing occupancy rates of 75%/59% (pocket A/B) for the apo and 66%/85% for the perturbed simulation, respectively.

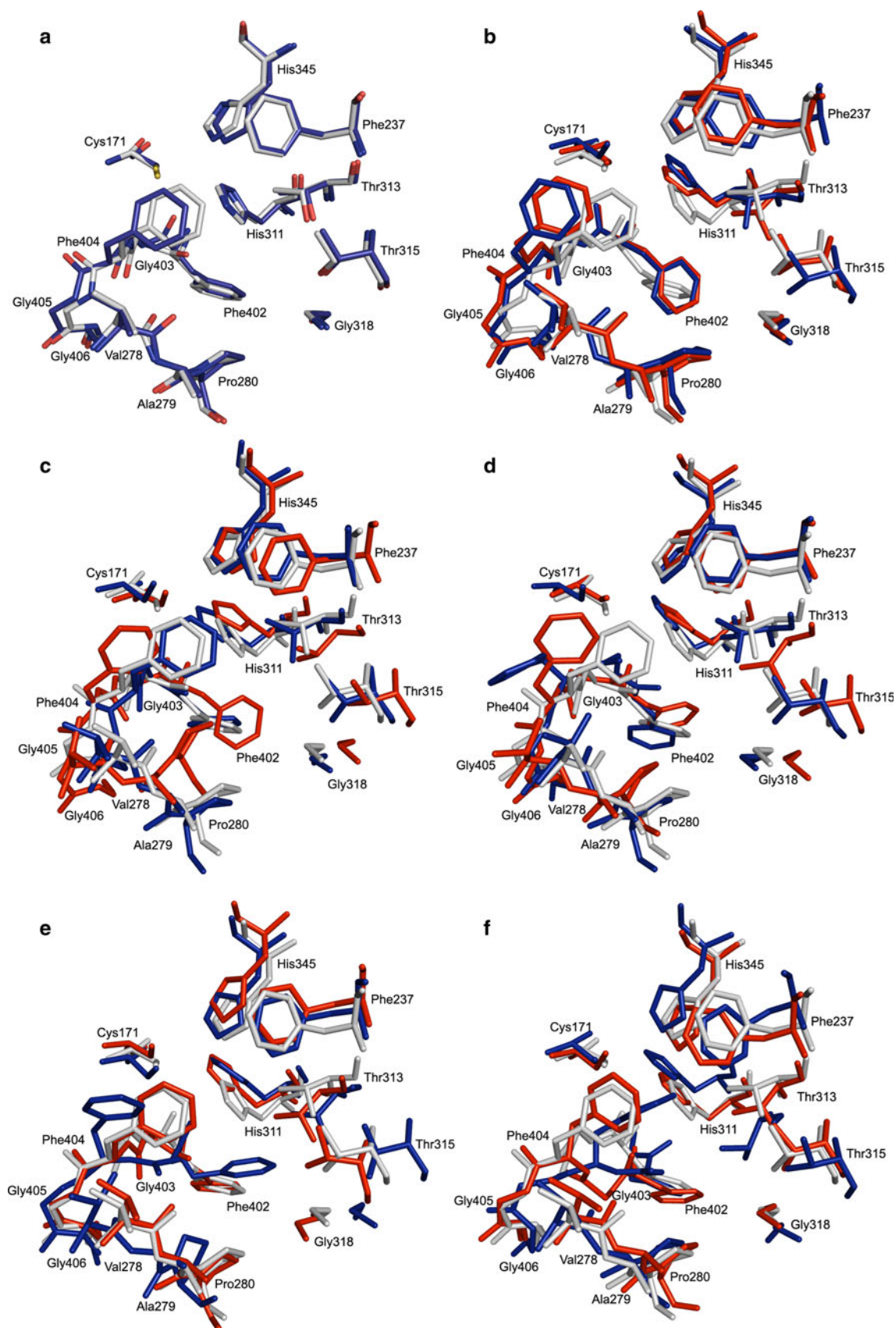
The more flexible half of the pocket is composed of Val278, Ala279, Pro280, Thr313, Thr315, Gly318, Phe402, Gly403, Phe404, Gly405, and Gly406. The highest conformational alterations are observed for Phe404 and its surrounding residues, especially Phe402 and Gly403. This is in line with the analysis of the superposed pockets of the apo and complex crystal structures (Fig. 5a): In the ligand-free state, Phe404 is shifted by ~ 0.8 Å into the binding site along the plane of the phenyl ring, while changes in the backbone torsions of Gly403 and Phe402 induce a twist of the carbonyl group of Phe402, whose oxygen, as a consequence, points into the pocket. Both orientations of this carbonyl functionality are revisited in the uncomplexed simulations, while in the complex simulation the more open orientation as observed in the X-ray structure is kept. Interestingly, in all simulations this open state mainly coincides with a different orientation of the Phe402 side chain as compared to the crystal structure. Here, the phenyl ring adopts a perpendicular pose relative to the X-ray orientation (cf. Fig. 5c, red sticks). This becomes most important in the complex simulation: In both pockets within the first 1.5 ns the χ_2 -torsion of Phe402 drops from initially -40° to -95° and stays there until the end of the simulation. This conformational alteration takes place due to a preceding twist of the Asn408 side chain that is located near Phe402. By a change of the χ_2 -torsion of Asn408 from -52° to -85° the side-chain amide group is brought close to the phenyl ring of Phe402 and thereby induces the above mentioned twist. However, side-chain conformations of Phe402 as observed in the X-ray structure are predominant when its carbonyl moiety adopts more closed, apo-like states (cf. Fig. 5c, blue sticks).

As mentioned above, in all simulations Phe404 is the most flexible residue of the pocket and shows the highest degree of structural changes. As a matter of fact, this residue is known to act as a gatekeeper, which, by a 60° -twist of the phenyl ring away from the active-site cysteine, provides direct access for long fatty acid substrates to enter the adjacent acyl-binding channel [18]. Therefore, the ability of Phe404 to adopt different conformations is hardly surprising. Open-state like conformations of Phe404 are also revisited in the uncomplexed simulations (cf. Fig. 5e, blue sticks). Additional conformational changes occur for the backbone atoms, which is not only due to changes of the backbone dihedrals of Phe404 itself, but also because of torsional alterations of Gly403 and Gly405. The resulting

Fig. 5 **a** Superimposed binding pockets of the apo (gray carbons) and complex crystal structure (blue carbons). **b–f** show a depiction of the ten distinct binding-pocket conformations chosen by means of hierarchical cluster analysis. Two representatives were extracted from the complex simulation (**b**) and four each from the apo (**c**, **d**) and the perturbed structure (**e**, **f**). MD-snapshots are colored red and blue, respectively. The original conformation of the crystal structure is shown in gray

movements of these backbone sections are most pronounced in the simulation of the perturbed structure, where Gly403 is shifted up to 3.7 Å into the pocket (cf. Fig. 5f, red sticks). As these conformations are based on a modified and, thus, to some extent artificial state of the protein, such large-scale shifts should not be overinterpreted. Nevertheless, significant changes of this backbone region are also found in the apo simulation, even if they are less pronounced than in the perturbed trajectory (cf. Fig. 5d, blue sticks). In the complex simulation, however, the original orientation of Phe404 is kept over a long time scale, supporting the observation of a stabilizing effect of the ligand on the binding-site residues (cf. Fig. 5b).

Val278, Ala279, and Pro280 are part of the entrance of the malonyl-binding pocket. Two alternate locations are deposited in the PDB file 2WGE for the backbone atoms of Val278 of the complex crystal structure. Location A perfectly fits the conformation observed in the uncomplexed state, whereas location B assumes a more open conformation by a twist of the ψ -torsion from -39° to 5° towards the outside of the pocket. In the setup of the complex simulation location B was chosen as starting orientation to provide somewhat more space for the isoprenoid chain of the ligand, and hardly any changes were observed for Val278 over the entire trajectory. However, both states are revisited in the simulation of the apo structure, which confirms that Val278 is able to adopt both alternate locations. Fig. 5c shows Val278 of the apo structure as gray sticks (which equals location A of the complex structure), while blue sticks depict an MD-snapshot of the uncomplexed state where Val278 adopts a B-like conformation. The backbone atoms of Ala279 are part of the pocket and positionally highly dependent on the orientation of Pro280. The flexibility of Pro280, in turn, is considerably susceptible to the conformation of the gatekeeper, Phe404: Closed conformations of the Phe404 side chain (as found in the wildtype crystal structures) retain the original locations of both Pro280 and Ala279 (cf. Fig. 5c, blue sticks). A movement of Pro280 into the binding pocket is only possible by a preceding shift of the Phe404 side chain out of the active site. The resulting positional change of Pro280 is nicely shown by the representative snapshot of the apo simulation in Fig. 5c (red sticks). A shift of mainly the Phe404 side chain out of the pocket creates sufficient space for Pro280 to protrude ~ 3 Å into the binding site. This



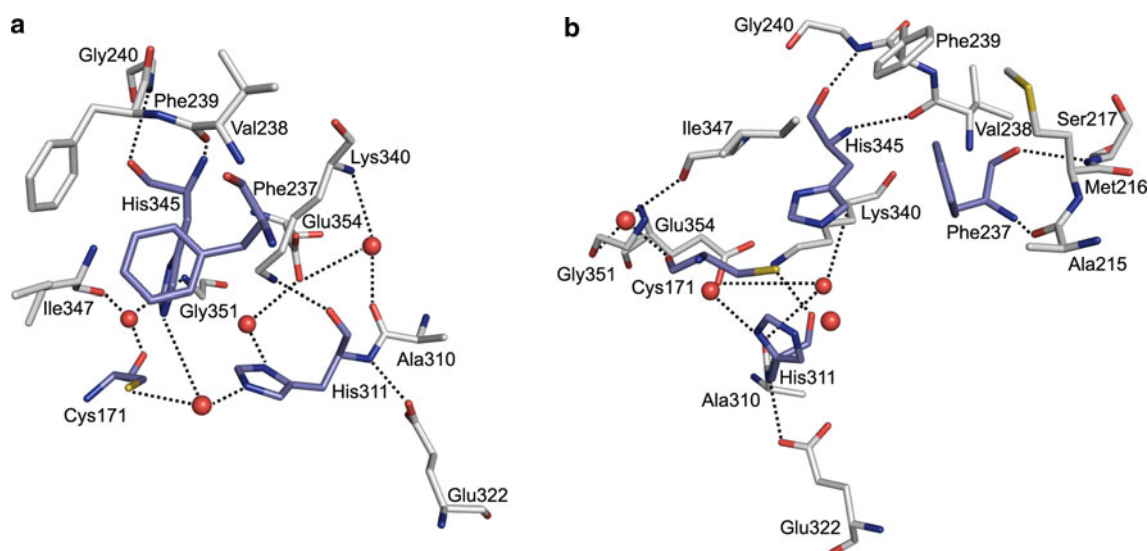


Fig. 6 Illustration of interaction patterns for Phe237 and the catalytic-triad residues, Cys171, His311, and His345 (blue sticks) which form a very stable part of the binding pocket. Shown are two different perspectives of the apo crystal structure (PDB code 2WGD). Stabilizing amino acids are illustrated with gray carbons. Water

also triggers a movement of Ala279, which follows by a shift of ~ 2 Å.

However, a conformational change of the Phe404 side chain does not necessarily result in a movement of Pro280. This is nicely illustrated in the representative snapshot of pocket B of the apo structure which is shown as blue sticks in Fig. 5d. Here, Phe404 assumes an open conformation that largely differs from the initial state, mostly due to a drastic change of both the φ -torsion of Phe404 from -53° to -123° and the χ_1 -dihedral from -64° to -165° . Despite the resulting widening of this part of the pocket, the position of Pro280 is hardly altered.

Thr315 and Gly318 are also directly located at the entrance of the malonyl-binding pocket. Both residues keep their original conformations in the majority of the simulation. Positional changes are mostly due to a shift of the backbone region which includes both Thr315 and Gly318. As a consequence, concerted motions are monitored for these two residues, resulting in an enlargement of the binding pocket. Movements involving a decrease in pocket size are not observed.

Like Thr315, Thr313 mainly retains its overall conformation throughout all simulations. However, in parts of the uncomplexed trajectories an alternative conformation is observed, as shown in Fig. 5c (red sticks). Here, the χ_1 -dihedral of Thr313 is twisted from 62° to -34° , such that the methyl group of the side chain is located at the starting position of the hydroxyl group. Furthermore, as opposed to Thr315 and Gly318, Thr313 is not located at the protein surface, and the close vicinity of the backbone atoms to other amino acids does not allow Thr313 to move

molecules are depicted as red spheres. Dashed lines show interactions that are observed over large parts of all three simulations. For reasons of clarity, interactions between the side chains of the catalytic triad and a water molecule are only depicted in a

out of the pocket. Instead, shifts along the backbone plane are observed, which result in larger distances to Phe237 and His345.

Dynamic behavior of the ligand structure

As mentioned above, a significantly lower mobility is observed for the binding-site residues when the ligand is present in the simulations. However, although KasA is a homodimer and identical preparation steps were performed for each monomeric subunit, the extent of this stabilization differs for the two binding sites of the complex simulation.

Pocket A of the complex structure displays the least structural diversity and mobility, presenting the lowest RMSD values and standard deviations (C^α : 0.82 ± 0.10 Å; all atoms: 1.41 ± 0.14 Å) of all simulations. The ligand remains tightly fixed within the pocket, as indicated by an RMSD of 0.81 ± 0.28 Å calculated for the ligand heavy atoms based on the C^α -fit of the binding-site residues. The overall conformation of the ligand itself is also well kept. The mean torsion angle around C^4 and C^5 (cf. Fig. 1 for atom naming) is $-179.1 \pm 8.0^\circ$ and thus very close to the conformation observed in the crystal structure (177.5°). With $157.4 \pm 15.9^\circ$, the largest deviation was measured for the torsion around C^6 and C^8 of the isoprenoid side chain (starting value: 178.7°). Also, the standard deviation of the average distance (2.84 Å) between the carbonyl oxygen of TLM and the ε -nitrogen of His345 is only 0.12 Å, which substantiates the importance of this hydrogen-bonding interaction to keep the ligand fixed in its position.

These findings are in clear contrast to pocket B of the complex simulation. Whereas comparison of both pockets reveals only moderately higher RMSD values for the binding-site residues of pocket B (C^2 : 1.11 ± 0.13 Å; all atoms: 1.78 ± 0.20 Å), they are about three times higher for the ligand (2.21 ± 0.87 Å vs. 0.81 ± 0.28 Å). The different behavior of TLM becomes immediately evident when plotting the RMSD as a function of the simulation time for both chains of the protein (Fig. 7). In pocket B, for the first 1.5 ns, the ligand stays close to the position observed in the crystal structure. The RMSD then steeply rises to 2.5 Å, which is caused by a turn of $\sim 80^\circ$ out of the plane of the thiolactone ring (Fig. 8b). From 5.2 ns on, this angle is decreased to $\sim 55^\circ$ and the ligand retains its position quite constantly until 12.4 ns (Fig. 8c). Interestingly, despite these changes, the original orientation within the pocket is kept up to this point of time. This is due to the hydrogen bond between the carbonyl oxygen of the ligand and the ϵ -nitrogen of His345 which is hitherto retained, as indicated by a virtually constant distance of 2.84 ± 0.12 Å between the two heavy atoms. However, at 12.4 ns, a sudden rise from 2.8 to 5 Å is observed, which illustrates the disruption of this hydrogen bond. This is immediately followed by a second steep increase of the ligand RMSD from 2.4 to 3.9 Å. At this point, TLM moves towards the entrance of the binding pocket alongside the thiolactone ring plane and stays there for the remaining 2.6 ns of the simulation. Interestingly, this position is stabilized exclusively by non-polar interactions, predominantly between the five-membered ring of TLM and the surrounding protein residues (Fig. 8d). The side chains of Phe404 and Phe237 vertically flank the thiolactone ring and thereby keep the position of this ligand moiety fixed. The adjacent methyl group (C^9 , cf. Fig. 1 for atom numbering) is nicely buried within a cavity formed by the side chains of Phe210

and Met213. In addition, the Pro280 ring stabilizes the region around the asymmetric carbon (C^4), including C^{10} and parts of the isoprenoid chain (mainly C^5 and C^6).

What is the reason for the drastic changes of the ligand in pocket B while its original orientation is persistently kept in pocket A? Comparing the per-residue RMSD values of the binding-site residues reveals a substantially higher value for Phe404 in pocket B (2.91 ± 0.58 Å) than in pocket A (1.88 ± 0.51 Å), indicating that structural changes must have occurred that did not take place in pocket A. When monitoring the dihedrals of Phe404 for both pockets, the by far largest—and also most important—discrepancy is observed for the χ_1 -torsion of the side chain. Within pocket A, the starting value of -52.9° is precisely kept over the entire simulation, as indicated by a mean torsion angle of $-54.5 \pm 11.0^\circ$ (cf. Fig. 8a), while in pocket B it is $-23.3 \pm 53.9^\circ$, which immediately discloses a drastically larger degree of structural change and mobility (cf. Fig. 8b).

Interestingly, the origin of this observation is not found within the immediate proximity of Phe404, but located at a distance of 9.5 Å within the interface of the KasA dimer. Calculation of occupancy rates for water molecules shows that in (the structurally stable) pocket A, for 73% of the time a water molecule is found between the backbone oxygens of Ala209 of chain A and Ser138 of chain B. Distance measurements combined with inspection of the interaction geometries over the trajectory clearly prove that a hydrogen-bonding network is present, involving a water molecule as ‘dual donor’ and the two backbone oxygens mentioned above as acceptors (Fig. 8a). In total, seven alternating water molecules take part in this interaction motif, where three of them—summed over the 15 ns trajectory—cover a total time of 11.7 ns (cf. Fig. 9). This interaction engenders a stabilizing effect to an α -helix

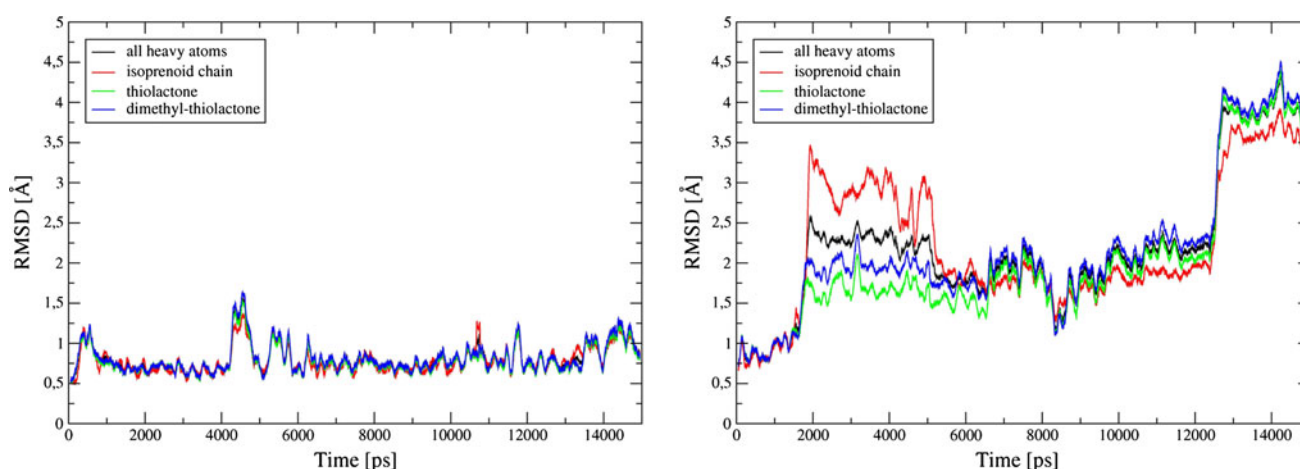


Fig. 7 RMSD with respect to the starting structure for TLM bound to pocket A (*left*) and pocket B (*right*) of the complex simulation as a function of simulation time

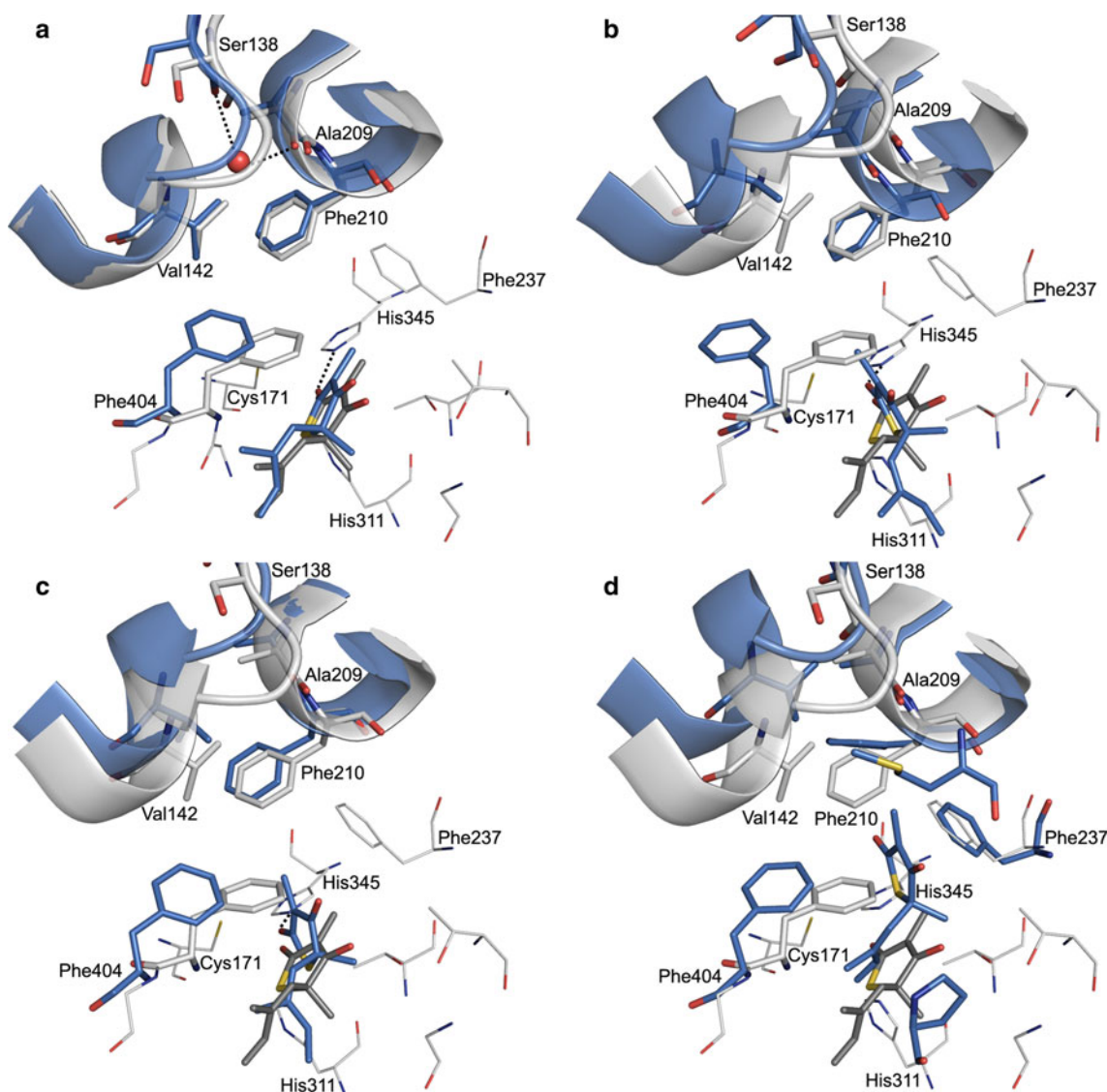


Fig. 8 a Representative snapshot for pocket A of the complex simulation at 5 ns simulation time. Hydrogen bonds (*dashed lines*) between a water molecule, Ser138, and Ala209 stabilize the α -helix including Val142 which, in turn, stabilizes the Phe404 side chain and

TLM. The remaining figures show snapshots of pocket B at **b** 3 ns, **c** 7 ns, and **d** 13 ns, respectively. MD-snapshots are colored in *blue*, the crystal structure in *gray*

(residues 140–145), which shows low average RMSD values of 0.85 ± 0.22 Å for the C^α -atoms and 1.13 ± 0.21 Å for all atoms, respectively. In turn, the side chain of Val142 that is part of the helix stays close to Phe404 over the entire simulation, which limits the conformational scope of the gatekeeper. Due to this restriction, a stabilizing effect of the Phe404 side chain on the thiolactone moiety of TLM is maintained. Combined with the strong hydrogen bond towards His345, it is impossible for the ligand to leave its original position.

As for pocket B, however, no occupancy is seen for a water molecule between Ala209 and Ser138. The resulting lack of stabilization is expressed by an increase of the α -helix RMSD to 2.26 ± 0.65 Å for the C^α -trace and

2.38 ± 0.62 Å for all atoms, respectively. Accordingly, Val142 moves out of the binding site and, thus, provides a gain of mobility for the Phe404 side chain (vide supra for comparison of torsion angles). The movement of Val142—and consequently of the entire α -helix—away from the binding site is nicely shown when monitoring the distance between the C^α -atoms of Val142 and Phe402 of the binding pocket (with Phe402 hardly altering its position over the trajectory, presenting a per-atom RMSD of 0.54 ± 0.20 Å), which from 0.5 ns on increases from ~ 15 up to ~ 20 Å. As a consequence, from 0.9 ns on, the side chain of Phe404 turns away from the inside of the pocket. The resulting enlargement of the binding site remains until the end of the simulation and reaches a maximum between 1.3 and 6.6 ns,

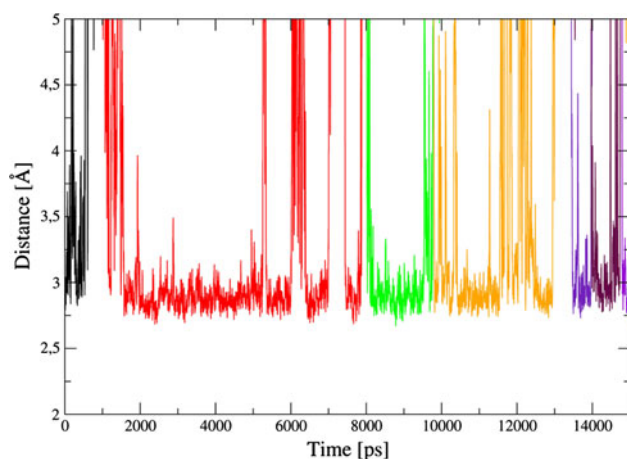


Fig. 9 Distance measurements for the seven water molecules involved successively in the hydrogen-bonding network which stabilizes pocket A of the complex simulation. For reasons of clarity, only the distances towards the oxygen of Ala209 (chain A) are shown. Similar plots are obtained for Ser138 (chain B)

when the χ_1 -torsion is increased from initially -52.9° to $\sim 45^\circ$, as shown in Fig. 8b. The mutual dependencies of the ligand, Phe404, and His345 are illustrated in Fig. 10. The motion of TLM in pocket B (shown by the change of RMSD) is in perfect agreement with the change of distances towards both the side chain of Phe404 and the ϵ -nitrogen of His345. The first rise of the ligand RMSD at 1.5 ns is tightly coupled to an increase of the distance between the ring systems of TLM and Phe404, while the second slope at 12.4 ns is associated with the rapid increase of the distance to His345, which corresponds to the disruption of this hydrogen bond.

Taken together, the simulation of the complex structure reveals that the presence of a water molecule acting as a dual hydrogen bond donor between Ala209 and Ser138 is of pivotal functional relevance. This water molecule may be thought of as a molecular switch toggling the flexibility of (primarily) the Phe404 side chain and, thus, the ability of the gatekeeper to adopt conformations that enable access from the active site to the acyl-binding channel and vice versa. Interestingly, also in the complex X-ray structure this position is occupied by a water molecule which, however, was not included in the MD-setup. This proves that the high occupancy rates observed in this region for pocket A are not just artifacts resulting from the simulation. Furthermore, the different behavior of TLM within the two binding sites may illustrate that this ligand is only a moderate inhibitor of KasA, presenting an IC_{50} value of $242 \pm 60 \mu M$ and a K_d of $225 \pm 9 \mu M$, respectively [35]. The complex simulation clearly shows that the fixation of TLM within the binding pocket is highly susceptible to conformational alterations of the Phe404 side chain. Changes of mainly its χ_1 -torsion angle result in a widening

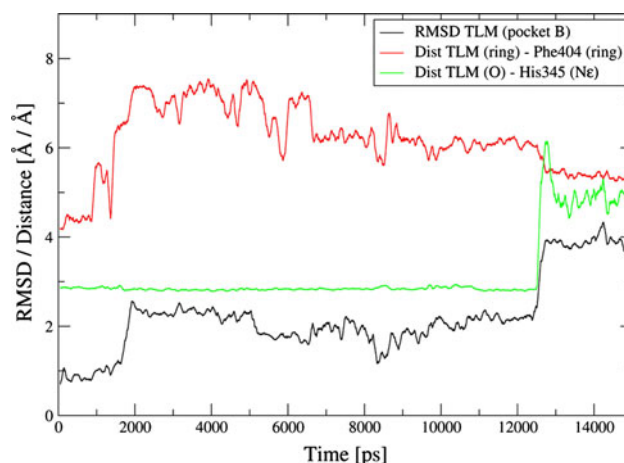


Fig. 10 Relationship between ligand RMSD and the distances of TLM towards Phe404 and His345, respectively

of the binding pocket. This additional space, in turn, offsets the tight fixation of TLM and is, thus, detrimental to the binding mode of the ligand.

As a consequence, it is obvious to assume that ligands which a priori stabilize the conformation of an enlarged pocket and, accordingly, limit the variability of Phe404 might show a more robust binding behavior as compared to TLM. Such augmented pockets are especially found in the uncomplexed simulations (cf. previous section). Hence, the binding-site representatives that were extracted from the unliganded trajectories (Fig. 5c–f) present a useful basis for the structure-based development of ligands that are likely to stabilize an enlarged pocket. The question remains open, however, whether this eventually results in an overall gain in affinity and hence an increase of inhibitory potency. Nevertheless, these findings depict useful information for ligand design approaches and should be considered in future virtual screening studies.

Implications of binding-site water molecules for drug design

Handling of water in structure-based drug design is a difficult task. With a sufficiently high resolution of the X-ray diffraction pattern, it is, in principle, possible to depict the positions of ordered water molecules in a crystal structure. However, to date only a small number of structures has been resolved with resolutions of $\sim 0.8 \text{ \AA}$ or lower, permitting an assignment of hydrogen atoms and, consequently, providing instant information on possible hydrogen-bonding networks [36]. Furthermore, when it comes to the design of new ligands for the binding site of a given protein, it is not straightforward to estimate out of a single crystal structure whether a certain water molecule located in the pocket should be regarded as virtually fixed part of the protein (due to tight interactions) or rather as

displaceable by a ligand moiety [37]. It is also possible that ‘conserved’ water positions displayed in crystal structures are artifacts or at least only of transitory nature under physiological conditions and, hence, do not need to be further taken into account. MD-simulations may help to clarify such issues, as it is possible to follow water movements and hereby detect conserved—or at least preferential—positions for water molecules within the protein [36].

The crystal structures of KasA used in this study show well-solvated binding sites for both the complexed and the uncomplexed state. Twelve water molecules are located within the immediate pocket of the apo structure, whereas only six are observed when TLM is bound to the active site (Fig. 11); the positions of these remaining water molecules are virtually identical with the locations observed in the apo state. All these crystallographic water molecules were included in the respective MD-setup (cf. Materials and Methods for details), presenting the starting point and reference for further analysis. Water occupancies were calculated for the binding pockets of all simulations using the VolMap plugin of VMD [32] to identify positions that are preferentially frequented by one or several water molecules.

Evaluation of the resulting water maps reveals two areas with significantly increased occupancies compared to the remaining part of the binding pocket (cf. Fig. 11). The first

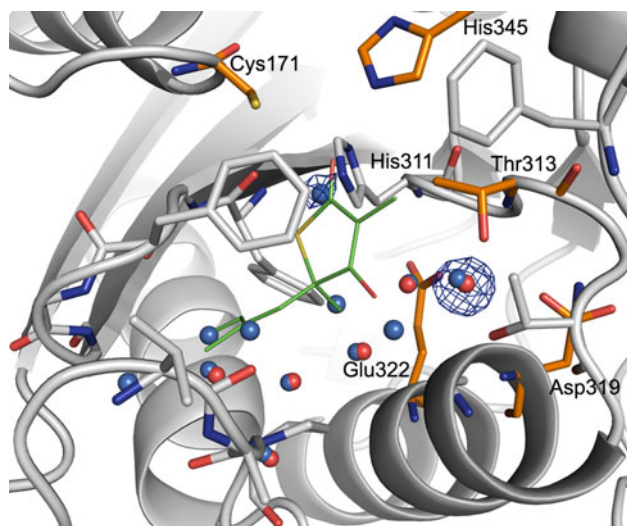


Fig. 11 Depiction of two positions within the immediate KasA binding site that are highly frequented by water molecules during the simulation, using pocket A of the apo structure as an example. Water occupancy maps are shown as blue meshes at 66% occupancy. Residues that mainly fix the water molecules are colored orange. Blue and red spheres depict the binding site water molecules from the crystal structures of the apo and the aligned complex structure, respectively. The ligand is shown in green carbons to illustrate the positional match of the thioester carbonyl oxygen as hydrogen-bonding acceptor and the respective water occupancy map

position is located near the residues of the catalytic triad, showing occupancy rates of 77%/61% for the pockets of the apo structure and 66%/85% for the perturbed simulation, respectively. Interestingly, the highest rate was calculated for the perturbed simulation, although this part of the pocket was artificially ‘evacuated’ by extracting the ligand before starting the simulation. This underlines the importance as a preferred spot of interaction with the protein. Key interactions related to this position are already experimentally proven: In the apo crystal structure, this position is also occupied by a water molecule, whereas in the complex it is the carbonyl oxygen of the ligand which establishes a hydrogen bond only towards the ϵ -nitrogen of His345. However, visual inspection of several snapshots shows water molecules in an orientation which permits their lone pairs to address both His345 and the thiol group of Cys171 simultaneously. Such an establishment of a dual hydrogen bond is not observed for TLM. The thiolactone ring would have to turn $\sim 30^\circ$ out of its plane to move the carbonyl lone pairs in a geometrically favorable position for this interaction motif, which, in turn, is not possible, as this would result in a steric clash with the side-chain atoms of Thr313. Generally speaking, addressing Cys171 in addition to His345 with the same ligand atom does not require further enthalpically unfavorable desolvation of this part of the pocket. It is, hence, likely that the establishment of a geometrically well-defined dual hydrogen bond instead of addressing only one residue of the catalytic triad will result in an increase of binding affinity. However, a correct orientation of the lone pairs of the ligand atom is a necessary precondition. Therefore, as an advice for drug-design approaches, putative binders should be able to fulfill this criterion without having to adopt energetically unfavorable conformations.

A second spot that is highly frequented by water molecules is located near Thr313, as well as Asp319 and Glu322, which are both located directly adjacent to the immediate binding pocket of TLM. The calculated occupancy rates are even higher than for the position near the catalytic triad, with 93%/97% for the complex, 94%/75% for the apo, and 73%/87% for the perturbed simulation, respectively. These values are hardly surprising, as inspection of this pocket moiety reveals ideal conditions for interaction with a water molecule: Both the torsion angles and the intermolecular distances of Thr313, Asp319, and Glu322 are favorable to saturate a hydrogen-bonding acceptor as well as two donors from a single atom. Accordingly, water molecules are most commonly found in a very similar orientation over all three simulations, where the hydrogens point towards the acid functionalities of Asp319 and Glu322, while one lone pair is directed towards the backbone amino group of Thr313. The conformations of the interacting protein residues are stabilized

by this water molecule, which becomes most evident when measuring the torsion angles of the carboxy groups of the Asp319 and Glu322 side chains. Taken together, it is reasonable to consider this specific position for drug design. However, it is hardly possible to give advice whether (a) this water molecule should be replaced by an appropriate functional group of a ligand or (b) rather be regarded as part of the protein and, thus, addressed at the free lone pair with a donor functionality of the molecule. At least, some hint is provided that option b) is plausible, as a second water molecule is regularly found (41 to 85%) to address this lone pair. Nevertheless, one should consider both options in virtual screening instead of concentrating on just one of them.

Furthermore, we analyzed the hydration of a loop comprising the binding-pocket residues Val278, Ala279, and Pro280, plus Asp273 to Met277. This part is supposed to be stabilized by two water molecules that are present in all KasA crystal structures [18]. However, our studies reveal that these water molecules are not conserved during the simulations, but move away from this part of the protein at early stages. Furthermore, occupancy rates are either very low or even not present in this area, indicating that a stabilization by several alternating water molecules—as found for pocket A of the complex simulation (cf. Fig. 9)—can also be ruled out here. It is, hence, comprehensible why residues 278–280 are part of the more flexible half of the binding site, as already presented above (cf. Fig. 5c–f).

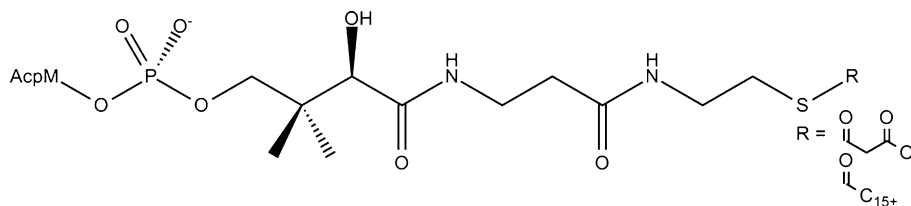
Substrate binding to KasA

KasA catalyzes the elongation of fatty acids by two carbon atoms per cycle via a ping-pong mechanism [35], resulting in an extension from C_{16} up to C_{56} . Both substrates—a fatty acid chain and malonic acid, which allocates the two carbon atoms for extension—form a thioester bond with the thiol moiety of a phosphopantetheine molecule which, in turn, acts as a linker to the mycobacterial acyl carrier protein (AcpM) (Fig. 12) [38]. At first, acyl-AcpM enters the malonyl-binding pocket and the thioester group is cleaved by Cys171 of the catalytic triad to form an acyl-enzyme intermediate. After diffusion of the deacylated phosphopantetheinyl-AcpM out of the pocket, malonyl-AcpM enters and, after decarboxylation, is attached to the

acyl chain by cleaving the thioester bond of the intermediate, thereby elongating the fatty acid by two carbons. The catalytic cycle is completed by release of the elongated acyl-AcpM product.

When Luckner et al. [18] solved the crystal structures of KasA, they also revealed for the first time how long-chain fatty acids are accommodated in the acyl-binding channel, which is formed by residues 115–147 of the monomeric subunits and provides direct access to the malonyl-binding pocket: For crystallization of the KasA C171Q mutant, polyethylene glycol (PEG) was used as part of the precipitant solution and, serendipitously, a long PEG chain crystallized within the acyl channel, mimicking a fatty acid of $\sim C_{40}$ length (cf. Fig. 13a). As the residues of the acyl channel show increased temperature factors ($\sim 60 \text{ \AA}^2$) as compared to the whole protein (40 \AA^2), Luckner et al. assumed that the helices of the respective monomer move in a scissor-like motion, thereby permitting the fatty-acid chains to access the long acyl channel. In fact, for all three simulations increased temperature factors are also found in this area. They are 58 \AA^2 for the complex, 46 \AA^2 for the apo, and 86 \AA^2 for the perturbed structure, as compared to 28, 28, and 35 \AA^2 for the whole protein. However, distance measurements between the acyl-channel helices indicate that no large-scale movements of the entire secondary structure take place in this region; the increased temperature factors are rather due to conformational alterations of the side chains. Nevertheless, a small loop of the acyl channel comprising residues 115–119 additionally shows increased backbone flexibility. Fig. 13b illustrates that this part is also shifted in the C171Q mutant complex structure, thereby creating space for the long PEG chain to bind to the acyl channel. In this context, Leu116 plays a crucial role. The positional movement of its side chain leads to an opening of the acyl channel, thus providing a direct link to the protein surface. Therefore, residues 115–119 and, in particular, the Leu116 side chain may be regarded as a second gatekeeper which determines the open and closed conformation of the acyl channel to the outside of the enzyme. Accordingly, open conformations of both gatekeeping units—as present in the mutant structures—are needed: On one hand, a twist of the Phe404 side chain is necessary, thereby allowing the acyl substrate to access the binding channel with its long fatty-acid chain through the malonyl-binding pocket and to appropriately place the

Fig. 12 The KasA substrates—malonic acids and fatty acids—are bound to a phosphopantetheine group which acts as a linker to the mycobacterial acyl-carrier protein (AcpM)



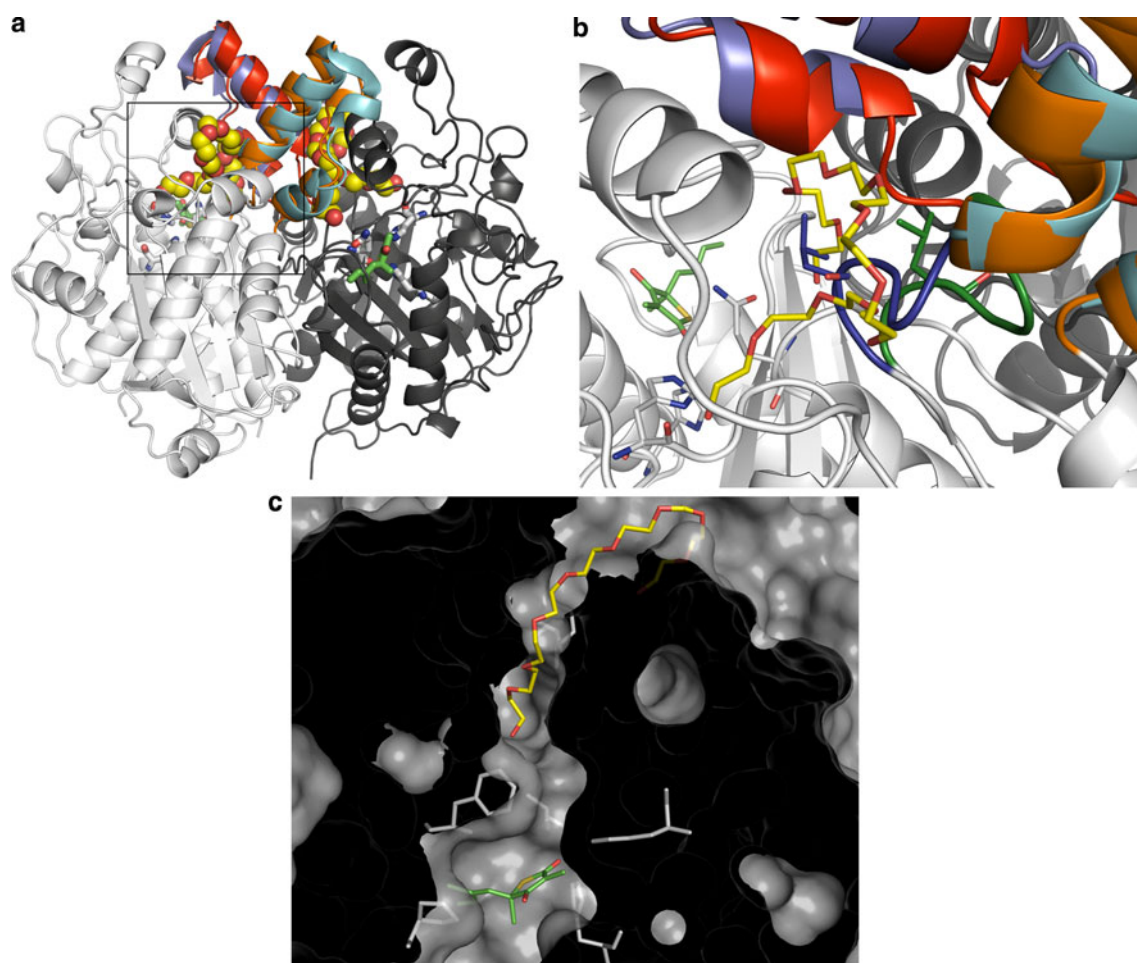


Fig. 13 **a** Crystal structure of the complexed KasA C171Q mutant dimer (PDB code 2WGG). Parts forming the acyl channel (residues 115–147 of each monomer) are colored *orange* and *red*. Ligand and catalytic triad residues are shown as *green* and *gray sticks*. The PEG chains are depicted as *yellow* and *red spheres*. Additionally, the acyl-channel segments of the wildtype complex structure (PDB code 2WGE) are superimposed and colored *cyan* and *purple*. **b** Image section of **a** (*black rectangle*) showing the shift of a channel segment (residues 115–119) upon binding of a PEG molecule (*yellow* and *red sticks* representation) to the acyl channel of pocket B. The loop is

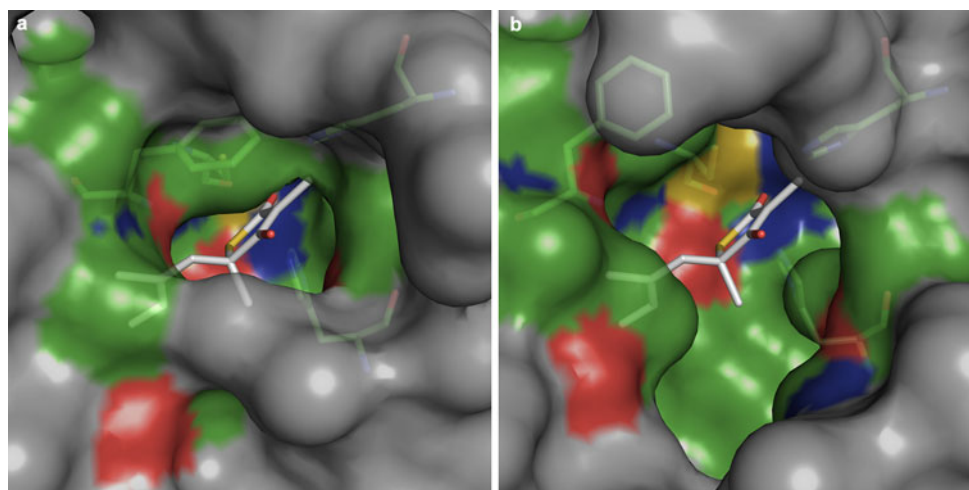
colored *blue* for the wildtype and *green* for the mutant structure. In addition, Leu116 as an important ‘second gatekeeper’ is shown in sticks representation in *blue* and *green* as well. Remaining coloring is according to **a**. **c** Depiction of the same representative MD-snapshot of the apo structure as used in Fig. 5d (*red sticks*), showing a completely open state of the acyl channel. PEG chain (*yellow* and *red sticks*) and ligand (*green* carbons) are taken from the aligned mutant structure. Catalytic-triad residues, as well as Leu116, Pro280, and Phe404 are shown as *gray sticks*

thioester group within the active site to enable the actual enzymatic reaction. On the other hand, a shift of residues 115–119 is needed to accommodate the fatty-acid chain that can protrude through the acyl channel to the protein surface and then bend back into the channel. Indeed, such open conformations are found in the simulations, as shown in Fig. 13c by a representative snapshot of the apo structure. Distance measurements, however, show that mutant-like conformational states of residues 115–119 and Phe404 accounting for this dual opening of the acyl channel are rather occasionally found in the simulations: in 0.7%/0.2% (pocket A/B) of the apo-structure simulation snapshots, never in the perturbed simulation, and in 0.2%/4.5% (pocket A/B) of the complex-structure

simulation snapshots. The higher occurrence frequency (4.5%) in pocket B of the complex simulation is related to the enlarged distance between Val142 and Phe404 and the resulting gain of flexibility of the gatekeeper in this pocket (vide supra).

Furthermore, the same representative snapshot as shown in Fig. 13c reveals a significantly augmented opening of the binding pocket in relation to the crystal structure (cf. Fig. 14), facilitating substrate access to both the active site and the acyl-binding channel via the pocket. This finding supports the assumption that, besides malonyl-AcpM, also acyl-AcpM uses the opening of the binding pocket to access the enzyme. The acyl chain should enter the pocket, with the tail of the carbon chain first, and protrude into the

Fig. 14 The entrance of the malonyl-binding pocket of **a** the crystal structure and **b** the same MD-snapshot as used in Fig. 13c of the wildtype apo structure. The ligand of the superimposed complex structure (gray carbons) is shown to better illustrate the substantial enlargement of the entrance. Residues of the catalytic triad and Phe404 are depicted as sticks with green carbons



acyl channel until the thioester bond towards the phosphopantetheine group is located near the residues of the catalytic triad. The phosphopantetheine moiety, in turn, acts as a spacer to the adjacent AcpM, enabling the thioester group to be transferred deeply enough into the binding site for subsequent cleavage. After formation of the acyl-enzyme intermediate and diffusion of phosphopantetheinyl-AcpM out of the binding site, malonyl-AcpM is able to enter the pocket and elongate the fatty acid as described above. After the elongation process, the acyl-AcpM product should then be released out of the binding site.

Taken together, the simulations provide insights into the dynamic behavior of the acyl-binding channel and first hints how substrate binding and product release may take place. To further probe the above hypotheses, it is conceivable to make use of enhanced sampling techniques. By means of steered molecular dynamics (SMD), for example, one could study the stepwise unbinding process of the PEG chain from the KasA mutant structures. In this context, SMD simulations would help to make a reliable guess of the exit path of the PEG chain and allow to monitor the unbinding process within reasonable computational timescales.

Conclusions

This study presents a comprehensive and detailed analysis of the dynamics of KasA. We performed three 15 ns MD simulations of the homodimer and, hence, sampled the conformational space of the binding pocket over a period of 90 ns in total. By means of 2D-RMSD calculations and hierarchical cluster analysis, we extracted binding-site representatives as the basis for further investigations.

Our study reveals a dual nature of the binding pocket in terms of flexibility. The residues of the catalytic triad, Cys171, His311, and His345, plus Phe237, form the more

rigid part and are stabilized by highly conserved water molecules, intramolecular interactions, and the carbonyl oxygen of the ligand. A more flexible behavior is observed for the remaining residues, with Phe404 presenting the largest conformational alterations, which is in line with its known role as a gatekeeper between the malonyl-binding pocket and the acyl-binding channel. Transitions between open and closed conformations of this gatekeeper are revisited in the uncomplexed simulations. Besides, closed states of the Phe404 side chain reveal a stabilizing effect on the ligand. However, due to the high flexibility of Phe404 the stability of the closed conformation is limited, which may account for the weak binding affinity of TLM to wildtype KasA. In this context, a water molecule between Ala209 and Ser138 was found to be of functional relevance, acting as a molecular switch toggling the flexibility of the Phe404 side chain and, thereby, the fixation of TLM in the binding pocket. Furthermore, two positions within the active site show high occupancy rates for a water molecule—one near Thr313, Asp319, and Glu322, and another near the residues of the catalytic triad.

Based on these findings, it may be useful to consider the following suggestions for the design of new KasA inhibitors:

- To reduce susceptibility to conformational alterations of the Phe404 side chain, the shape of the ligand should promote the formation of an enlarged binding site by stabilizing the Phe404 side chain in the conformation turned out of the pocket.
- Ligands should present a functionality permitting the establishment of a geometrically well-defined dual hydrogen bond towards the thiol group of Cys171 and the ϵ -NH moiety of His345.
- The preferred spot for a water molecule near Thr313, Asp319, and Glu322 should be occupied either by an appropriate functional group of the ligand (e.g. a

hydroxyl group) or by a water molecule that is regarded and addressed as integral part of the active site.

In addition, the dynamics of the acyl-binding channel were analyzed. In accordance with the crystal structure, this region shows increased temperature factors which, however, are rather due to conformational side-chain alterations than to large-scale movements of the secondary structure. The conformation of residues 115–119, including especially the side chain of Leu116, are further found to be crucial for accommodation of the long-chain fatty-acid substrates and may therefore be regarded as a second gatekeeper regulating the opening between the acyl channel and the outside of the enzyme.

Acknowledgments Financial support by the DFG (SFB630) is gratefully acknowledged.

References

- World Health Organization (2009) Global tuberculosis control: epidemiology, strategy, financing: WHO report 2009
- World Health Organization (2009) Factsheet
- Dye C, Williams BG (2010) The population dynamics and control of tuberculosis. *Science* 328(5980):856–861. doi:10.1126/science.1185449
- Koul A, Arnoult E, Lounis N, Guillemont J, Andries K (2011) The challenge of new drug discovery for tuberculosis. *Nature* 469(7331):483–490. doi:10.1038/nature09657
- TB Alliance (2011). <http://www.tballiance.org>
- Glickman MS, Jacobs WR Jr (2001) Microbial pathogenesis of *Mycobacterium tuberculosis*: dawn of a discipline. *Cell* 104(4):477–485
- He X, Alian A, Stroud R, Ortiz de Montellano PR (2006) Pyrrolidine carboxamides as a novel class of inhibitors of enoyl acyl carrier protein reductase from *Mycobacterium tuberculosis*. *J Med Chem* 49(21):6308–6323. doi:10.1021/jm060715y
- Rozwarski DA, Vilcheze C, Sugantino M, Bittman R, Sacchettini JC (1999) Crystal structure of the *Mycobacterium tuberculosis* enoyl-ACP reductase, InhA, in complex with NAD⁺ and a C16 fatty acyl substrate. *J Biol Chem* 274(22):15582–15589
- Vilcheze C, Jacobs WR Jr (2007) The mechanism of isoniazid killing: clarity through the scope of genetics. *Annu Rev Microbiol* 61:35–50. doi:10.1146/annurev.micro.61.111606.122346
- Kuo MR, Morbidoni HR, Alland D, Sneddon SF, Gourlie BB, Staveski MM, Leonard M, Gregory JS, Janjigian AD, Yee C, Musser JM, Kreiswirth B, Iwamoto H, Perozzo R, Jacobs WR Jr, Sacchettini JC, Fidock DA (2003) Targeting tuberculosis and malaria through inhibition of Enoyl reductase: compound activity and structural data. *J Biol Chem* 278(23):20851–20859. doi:10.1074/jbc.M211968200
- He X, Alian A, Ortiz de Montellano PR (2007) Inhibition of the *Mycobacterium tuberculosis* enoyl acyl carrier protein reductase InhA by arylamides. *Bioorg Med Chem* 15(21):6649–6658. doi:10.1016/j.bmc.2007.08.013
- Ghiladi RA, Medzihiradzsky KF, Rusnak FM, Ortiz de Montellano PR (2005) Correlation between isoniazid resistance and superoxide reactivity in *Mycobacterium tuberculosis* KatG. *J Am Chem Soc* 127(38):13428–13442. doi:10.1021/ja054366t
- Bhatt A, Kremer L, Dai AZ, Sacchettini JC, Jacobs WR Jr (2005) Conditional depletion of KasA, a key enzyme of mycolic acid biosynthesis, leads to mycobacterial cell lysis. *J Bacteriol* 187(22):7596–7606. doi:10.1128/JB.187.22.7596-7606.2005
- Kremer L, Douglas JD, Baulard AR, Morehouse C, Guy MR, Alland D, Dover LG, Lakey JH, Jacobs WR Jr, Brennan PJ, Minnikin DE, Besra GS (2000) Thiolactomycin and related analogues as novel anti-mycobacterial agents targeting KasA and KasB condensing enzymes in *Mycobacterium tuberculosis*. *J Biol Chem* 275(22):16857–16864. doi:10.1074/jbc.M000569200
- Schaeffer ML, Agnihotri G, Volker C, Kallender H, Brennan PJ, Lonsdale JT (2001) Purification and biochemical characterization of the *Mycobacterium tuberculosis* beta-ketoacyl-acyl carrier protein synthases KasA and KasB. *J Biol Chem* 276(50):47029–47037. doi:10.1074/jbc.M108903200
- Price AC, Choi KH, Heath RJ, Li Z, White SW, Rock CO (2001) Inhibition of beta-ketoacyl-acyl carrier protein synthases by thiolactomycin and cerulenin. Structure and mechanism. *J Biol Chem* 276 (9):6551–6559. doi:10.1074/jbc.M007101200
- Wang J, Soisson SM, Young K, Shoop W, Kodali S, Galsgocci A, Painter R, Parthasarathy G, Tang YS, Cummings R, Ha S, Dorso K, Motyl M, Jayasuriya H, Ondeyka J, Herath K, Zhang C, Hernandez L, Allocco J, Basilio A, Tormo JR, Genilloud O, Vicente F, Pelaez F, Colwell L, Lee SH, Michael B, Felcetto T, Gill C, Silver LL, Hermes JD, Bartizal K, Barrett J, Schmatz D, Becker JW, Cully D, Singh SB (2006) Platensimycin is a selective FabF inhibitor with potent antibiotic properties. *Nature* 441(7091):358–361. doi:10.1038/nature04784
- Luckner SR, Machutta CA, Tonge PJ, Kisker C (2009) Crystal structures of *Mycobacterium tuberculosis* KasA show mode of action within cell wall biosynthesis and its inhibition by thiolactomycin. *Structure* 17(7):1004–1013. doi:10.1016/j.str.2009.04.012
- Kremer L, Dover LG, Carrere S, Nampoothiri KM, Lesjean S, Brown AK, Brennan PJ, Minnikin DE, Loch C, Besra GS (2002) Mycolic acid biosynthesis and enzymic characterization of the beta-ketoacyl-ACP synthase A-condensing enzyme from *Mycobacterium tuberculosis*. *Biochem J* 364(Pt 2):423–430. doi:10.1042/BJ20011628
- The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC
- Case DA, TA, Cheatham TE, III, Simmerling CL, Wang J, Duke RE, Luo R, Crowley M, Walker RC, Zhang W, Merz KM, Wang B, Hayik S, Roitberg A, Seabra G, Kolossvary I, Wong KF, Paesani F, Vanicek J, Wu X, Brozell SR, Steinbrecher T, Gohlke H, Yang L, Tan C, Mongan J, Hornak V, Cui G, Mathews DH, Seetin MG, Sagui C, Babin V, Kollman PA (2008) AMBER 10, University of California, San Francisco
- Lee W, Luckner SR, Kisker C, Tonge PJ, Engels B (2011) Elucidation of the protonation states of the catalytic residues in mtKasA: implications for inhibitor design. *Biochemistry* 50(25):5743–5756. doi:10.1021/bi200006t
- Bayly CI, Cieplak P, Cornell WD, Kollman PA (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J Phys Chem* 97:10269–10280
- Frisch MJT, GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JRJ, Montgomery A Jr., Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Cliford S, Cioslowski

- J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2004) Gaussian, Inc., Wallingford CT
25. Wang J, Wang W, Kollman PA, Case DA (2006) Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model* 25(2):247–260. doi:[10.1016/j.jmgm.2005.12.005](https://doi.org/10.1016/j.jmgm.2005.12.005)
26. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general amber force field. *J Comput Chem* 25(9):1157–1174. doi:[10.1002/jcc.20035](https://doi.org/10.1002/jcc.20035)
27. Still WCT A, Hawley RC, Hendrickson T (1990) Semianalytical treatment of solvation for molecular mechanics and dynamics. *J Am Chem Soc* 112:6127–6129
28. Srinivasan J, Trevathan MW, Beroza P, Case DA (1999) Application of a pairwise generalized born model to proteins and nucleic acids: inclusion of salt effects. *Theor Chem Acc* 101(6): 426–434
29. Jorgensen WLC J, Madura J, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935
30. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical-integration of cartesian equations of motion of a system with constraints—molecular-dynamics of N-alkanes. *J Comput Phys* 23(3): 327–341
31. Darden T, York D, Pedersen L (1993) Particle mesh Ewald—an N.Log(N) method for Ewald sums in large systems. *J Chem Phys* 98 (12): 10089–10092
32. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graph* 14 (1):33–38
33. SYBYL-X 1.0, Tripos International, 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA (2010)
34. Luque I, Freire E (2000) Structural stability of binding sites: consequences for binding affinity and allosteric effects. *Proteins Suppl* 4:63–71
35. Machutta CA, Bommineni GR, Luckner SR, Kapilashrami K, Ruzsicska B, Simmerling C, Kisker C, Tonge PJ (2010) Slow onset inhibition of bacterial beta-ketoacyl-acyl carrier protein synthases by thiolactomycin. *J Biol Chem* 285(9):6161–6169. doi:[10.1074/jbc.M109.077909](https://doi.org/10.1074/jbc.M109.077909)
36. Kirchmair J, Spitzer GM, Liedl KR (2011) Consideration of water and solvation effects in virtual screening. In: Sotriffer C (ed) *Virtual screening. Principles, challenges, and practical guidelines*, vol 48, 1st edn. Wiley-VCH, Weinheim, pp 263–289
37. Ladbury JE (1996) Just add water! The effect of water on the specificity of protein-ligand binding sites and its potential application to drug design. *Chem Biol* 3(12):973–980
38. Wong HC, Liu G, Zhang YM, Rock CO, Zheng J (2002) The solution structure of acyl carrier protein from *Mycobacterium tuberculosis*. *J Biol Chem* 277(18):15874–15880. doi:[10.1074/jbc.M112300200](https://doi.org/10.1074/jbc.M112300200)