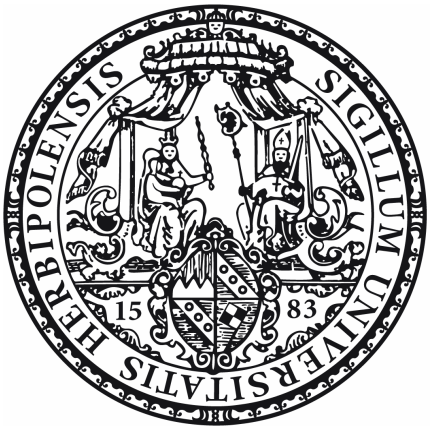
Towards the development of high affinity InhA and KasA inhibitors with activity against drug-resistant strains of *Mycobacterium tuberculosis*

Entwicklung von hoch-affinen InhA und KasA Inhibitoren gegen resistente Stämme von *Mycobacterium tuberculosis*



Doctoral thesis for a doctoral degree
at the Graduate School of Life Sciences,
Julius-Maximilians-Universität Würzburg,
Section Biomedicine

submitted by
Sylvia Rosie Luckner

from
Nürnberg

Würzburg 2009

Submitted on: ……..……..

Members of the *Promotionskomitee*:

Chairperson: Michael Sendtner

Primary Supervisor: Caroline Kisker

Supervisor (Second): Peter Tonge

Supervisor (Third): Thomas Müller

Date of Public Defence: ……

Date of Receipt of Certificates: ……

Summary

Mycobacterium tuberculosis is the causative agent of tuberculosis and responsible for more than eight million new infections and about two million deaths each year. Novel chemotherapeutics are urgently needed to treat the emerging threat of multi drug resistant and extensively drug resistant strains.

Cell wall biosynthesis is a widely used target for chemotherapeutic intervention in bacterial infections. In mycobacteria, the cell wall is comprised of mycolic acids, very long chain fatty acids that provide protection and allow the bacteria to persist in the human macrophage. The type II fatty acid biosynthesis pathway in *Mycobacterium tuberculosis* synthesizes fatty acids with a length of up to 56 carbon atoms that are the precursors of the critical mycobacterial cell wall components mycolic acids. KasA, the mycobacterial β-ketoacyl synthase and InhA, the mycobacterial enoyl reductase, are essential enzymes in the fatty acid biosynthesis pathway and validated drug targets.

In this work, KasA was expressed in *Mycobacterium smegmatis*, purified and co-crystallized in complex with the natural thiolactone antibiotic thiolactomycin (TLM). High-resolution crystal structures of KasA and the C171Q KasA variant, which mimics the acyl enzyme intermediate of the enzyme, were solved in absence and presence of bound TLM. The crystal structures reveal how the inhibitor is coordinated by the enzyme and thus specifically pinpoint towards possible modifications to increase the affinity of the compound and develop potent new drugs against tuberculosis. Comparisons between the TLM bound crystal structures explain the preferential binding of TLM to the acylated form of KasA. Furthermore, long polyethylene glycol molecules are bound to KasA that mimic a fatty acid substrate of approximately 40 carbon atoms length. These structures thus provide the first insights into the molecular mechanism of substrate recognition and reveal how a wax-like substance can be accommodated in a cytosolic environment.

InhA was purified and co-crystallized in complex with the slow, tight binding inhibitor 2-(o-tolyloxy)-5-hexylphenol (PT70). Two crystal structures of the ternary InhA-NAD^+-PT70 were solved and reveal how the inhibitor is bound to the substrate binding pocket. Both structures display an ordered substrate binding loop and corroborate the hypothesis that slow onset inhibition is coupled to loop ordering. Upon loop ordering, the active site entrance is more restricted and the inhibitor is kept inside more tightly. These studies provide additional information on the mechanistic imperatives for slow onset inhibition of enoyl ACP reductases.

Zusammenfassung

Mycobacterium tuberculosis, der Erreger der Tuberkulose ist für mehr als acht Millionen Neu-Infektionen und ungefähr zwei Millionen Todesfälle jedes Jahr verantwortlich. Besonders die Entwicklung von multiresistenten und extrem resistenten Stämmen macht die Entwicklung neuer Medikamente gegen Tuberkulose dringend erforderlich.

Die Zellwandbiosynthese ist ein validiertes Ziel für die Chemotherapie bei bakteriellen Infektionen. Bei Mycobakterien besteht die Zellwand zum Großteil aus Mykolsäuren, sehr langkettigen Fettsäuren, die den Bakterien Schutz bieten und ihnen ermöglichen, in Makrophagen zu überleben. Mycobakterien synthetisieren in der Fettsäurebiosynthese II (FAS-II) Fettsäuren bis zu einer Länge von 56 Kohlenstoffatomen, die Bestandteile der Mykolsäuren sind. KasA, die mycobakterielle β-ketoacyl Synthase und InhA, die mycobakterielle enoyl Reductase, sind essentielle Enzyme der FAS-II und geeignete Ziele für die Entwicklung neuer Antibiotika.

In dieser Arbeit wurde KasA in *Mycobacterium smegmatis* exprimiert und aufgereinigt. Das Protein wurde im Komplex mit dem natürlich vorkommenden Thiolacton-Antibiotikum Thiolactomycin (TLM) co-kristallisiert. Kristallstrukturen von KasA und der C171Q KasA Variante, die das acylierte Enzym-Intermediat darstellt, wurden als apo-Strukturen und im Komplex mit gebundenem TLM aufgeklärt. Die Kristallstrukturen zeigen, wie der Inhibitor an das Enzym gebunden ist und deuten darauf hin, wie das TLM Molekül verändert werden könnte, um seine Affinität für das Protein zu erhöhen und damit ein wirksames Medikament gegen Tuberkulose zu entwickeln. Vergleiche zwischen den TLM gebundenen Kristallstrukturen erklären, warum TLM bevorzugt an die acylierte Form des Enzyms bindet. Des Weiteren sind lange Polyethylenglykol-Moleküle an KasA gebunden, die ein Fettsäuresubstrat einer Länge von etwa 40 Kohlenstoff-Atomen nachahmen. Die Strukturen geben damit zum ersten Mal einen Einblick in den molekularen Mechanismus der Substrat-Erkennung und zeigen, wie eine wachsartige Substanz in einem cytosolischen Umfeld aufgenommen werden kann.

InhA wurde aufgereinigt und im Komplex mit dem „slow binding“ Inhibitor 2-(o-tolyloxy)-5-hexylphenol (PT70) co-kristallisiert. Zwei Kristallstrukturen des ternären InhA-NAD^+-PT70 Komplexes wurden gelöst und zeigen wie der Inhibitor in der Substratbindetasche gebunden ist. Beide Strukturen, weisen geordnete Substrat-Binde-Loops auf, die den Eingang zur „Active Site“ schließen und damit den gebundenen Inhibitor in der Tasche festhalten. Die Strukturen bestätigen damit die Hypothese, dass „Slow Binding Inhibition“ mit der Ordnung des Loops zusammenhängt. Diese Studien können als Basis für die Entwicklung weiterer „Slow Binding“ Inhibitoren verwendet werden.

Table of contents

1 Introduction

1.1 Tuberculosis

1.1.1 *Mycobacterium tuberculosis*

More than 70 species of mycobacteria exist but only two of them are major pathogens against humans: *Mycobacterium tuberculosis*, the causative agent of tuberculosis and *Mycobacterium leprae*, which causes leprosy and was discovered in 1874 by the Norwegian physician Gerhard Henrik Armauer Hansen.

Figure 1-1. Photograph of Robert Koch taken in 1883 after his discovery of *M. tuberculosis.* Taken from (Murray 2004).

Mycobacterium tuberculosis was first identified and described as the causative agent of tuberculosis by Robert Koch in 1882. At that time, however, tuberculosis was widely believed to be inherited because of its chronic nature and Koch's findings initially had opponents. In particular, Rudolf Virchow, a renowned pathologist, regarded tuberculosis as a result of a malfunction of the host cells rather than the result of a bacterial infection. This opposition urged Koch to develop the general concept of medical bacteriology, the famous Koch's postulates. Thereby he could clearly prove the infectious nature of *M. tuberculosis* and other infectious agents (Kaufmann 2005; Murray 2004). In 1905 he received the Nobel Prize in physiology or medicine for his work on tuberculosis.

Mycobacterium tuberculosis is a rod shaped, 2-4 µm long, non-motile bacillus that is classified as acid-fast and aerobic. The genome of *M. tuberculosis* is GC-rich (65%), consists of 4,411,529 base pairs and contains around 4,000 genes. As many as 8% of the genome is involved in lipid metabolism which indicates that lipids play an important role for the organism (Cole, Brosch et al. 1998). *M. tuberculosis* appears to be an obligate parasite and is not found as a free-living member of the genus, which means it is only found in host animals. The pathogen can infect a wide variety of animals but humans are the principal hosts. As the genomes of *M. tuberculosis* and *M. leprae* (3,268,203 base pairs encode only 1,604 proteins) are smaller than the genomes of members of the environmental mycobacteria (the *M. smegmatis* genome consists of 6,988,209 base pairs and 6716 genes), it is speculated that

the pathogens have lost the genes required for independent existence (Vissa and Brennan 2001).

M. tuberculosis belongs to the slow growing mycobacteria with a doubling time of about 20 h. The slow growth complicates diagnosis and makes long-term drug treatment necessary.

1.1.2 The disease

Tuberculosis is an airborne infection that spreads from person to person through sneezing and coughing. According to estimates of the World Health Organisation (WHO), two billion people, which is equivalent to one third of the world's population, are infected with tuberculosis. Tuberculosis causes more than 8 million new infections and about 2 million deaths each year and is for this reason the most effective human pathogen among all bacteria (WHO 2007).

In most people, the immune system can react to an infection and macrophages internalize the mycobacteria by phagocytosis. Usually, the bacteria are then transported to phagosomes that maturate and finally fuse with lysosomes, where enzymes destroy the bacteria. However, mycobacteria are able to escape lysosomal destruction and can survive within macrophages by arresting the phagosomes in an early maturation state and thus prevent fusion with lysosomes. In that way, *M. tuberculosis* is contained but not eradicated by the immune system and persists as a latent infection in the macrophages (Pieters 2008). The latent form of TB can persist the entire host's life while the individual is not infectious and shows nearly no symptoms. In 10% of the cases however, the bacteria reactivate, especially when the immune system is impaired like in cases of malnutrition, ageing or HIV infection. The active form of the disease is infectious and fatal in 50% of the cases, if left untreated, resulting in about 2 million deaths each year. In most cases TB affects the lungs (pulmonary TB) but it can also affect the bones, joints and kidneys, as well as the brain and leads to meningitis. The classical symptoms include loss of weight, loss of energy, poor appetite, fever, a productive cough, and night sweats (Kaufmann 2001). The definite diagnosis of tuberculosis relies on the demonstration of *M. tuberculosis* in a clinical sample like sputum or pus. An x-ray photograph of the chest that shows abnormalities or the tuberculin skin test, also called the Mantoux test, is used to confirm the diagnosis.

1.1.3 Current treatment against tuberculosis

Mycobacteria are naturally resistant against a broad range of antibiotics. This resistance is mainly due to the unusual waxy cell wall which is impermeable for most drugs (see chapter 1.1.4.). In addition, most mycobacteria possess an intrinsic β-lactamase. Consequently, the

first antibiotics in use, sulfonamides and penicillin, were ineffective against tuberculosis. The first effective drug against tuberculosis was streptomycin. It was isolated from *Streptomyces griseus* in 1943 in the laboratory of Selman Waksman and successfully administered to a tuberculosis patient in 1944 (Schatz, Bugie et al. 2005). In the years after the discovery of Streptomycin, several antibiotics were introduced as anti TB-agents: p-aminosalicylic acid in 1949, isoniazid in 1952, pyrazinamide in 1954, cycloserine in 1955, ethambutol in 1962 and rifampicin in 1963. Discovered more than 40 years ago, these drugs are still used today as the standard treatment for tuberculosis, which consists of isoniazid, rifampicin, pyrazinamid and ethambutol for two months followed by isoniazid and rifampicin for four months. The four drugs used for standard treatment are termed the first-line drugs for TB treatment. As it soon became clear that treatment with one drug alone (monotherapy) gives rise to drug resistant strains of *M. tuberculosis*, a combination of up to four drugs at a time is used to enhance the efficacy of the treatment and to prevent the rapid development of drug resistant strains.

However, drug resistant strains represent an increasing problem today. When the long, complex TB drug regimen is incomplete or inappropriate, for example when patients stop taking their medicines before the disease has been fully eradicated from their body or the access to the drugs is limited and patient take only one drug, the development of drug resistance is very likely. The drug resistant strains can then spread from person to person like drug susceptible bacteria. Strains of *M. tuberculosis* that are resistant against isoniazid and rifampicin, the most effective drugs against tuberculosis, are defined as multi drug resistant (MDR-TB). Every year, about 490,000 people get infected with MDR-TB. It is estimated that more than 5% of the world's TB cases are multi-drug resistant, 10% of those are extensively drug resistant (Zignol, Hosseini et al. 2006; Shah, Wright et al. 2007).

For treatment of patients with MDR-TB, a combination of second line drugs is used (second line drugs are listed in Figure 1-2). Compared to first-line drugs, second line drugs are less effective (p-aminosalicylic acid), show more side effects (cycloserines, fluoroquinolones) or are not available in many countries, probably because they are very expensive (linezolid). Furthermore, treatment with second-line drugs lengthens the treatment to up to 24 months and thereby increases the cost for the therapy dramatically.

Wherever second-line drugs to treat MDR-TB are being misused, MDR-TB strains can also get resistant to second line drugs and treatment is even more complicated (Toungoussova, Mariandyshev et al. 2005). In 2006 the WHO released the first data on extensively drug resistant strains (XDR-TB). These strains are resistant to any fluoroquinolones and at least to one of the injectable drugs kanamycin, capreomycin and amikacin, in addition to isoniazid

and rifampicin and occur in every part of the world, (Migliori, Loddenkemper et al. 2007). Patients infected with XDR-TB are virtually untreatable with current drugs (see Figure 1-2).

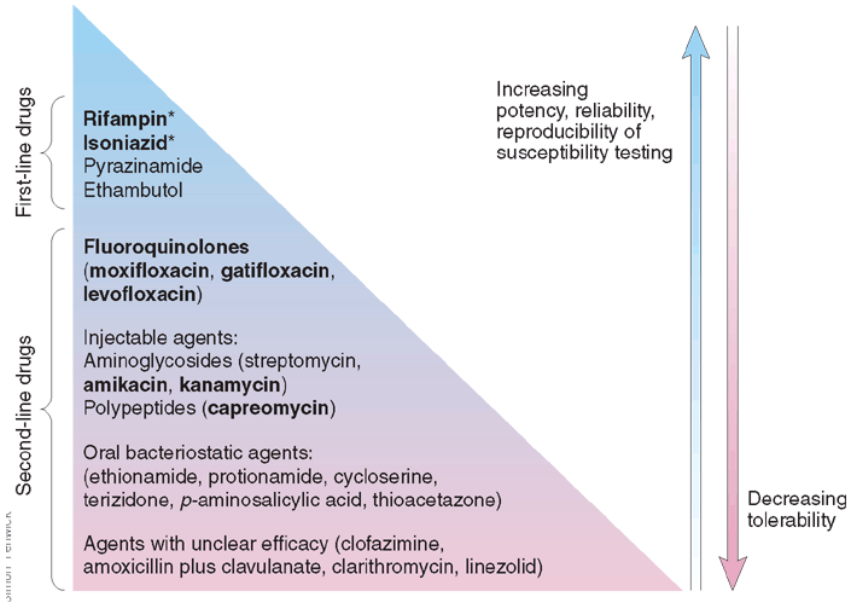


Figure 1-2. First-line drugs and second-line drugs against tuberculosis. MDR-TB strains are resistant to isoniazid and rifampicin (asterisks). XDR-TB strains are resistant to isoniazid, rifampicin plus fluoroquinolones and at least one of the injectable drugs capreomycon, kanamycin or amikacin (bold). Figure taken from (Dorman and Chaisson 2007).

1.1.4 Composition of the mycobacterial cell wall

The *myco* in mycobacteria means fungus or wax and refers to the waxy cell wall of mycobacteria. Indeed, mycobacteria have a quite unusual cell wall. It is lipid-rich and highly impermeable and thereby provides protection from many antibiotics and helps the bacteria to persist and proliferate in macrophages (Barry, Lee et al. 1998).

The cell wall of mycobacteria is composed of two segments. The lower segment is comprised of peptidoglycan that is covalently attached to arabinogalactan which in turn is attached to the mycolic acids that have one long meromycolate chain and one short α-chain. This entire structure is termed mycolyl arabinogalactan-peptidoglycan (mAGP) and forms the cell wall core (compare Figure 1-3). The upper part of the cell wall consists of free lipids that intercalate with the mycolic acids. These free lipids are for example phenolic glycolipids, phthiocerol dimycocerosates, cord factor or trehalose dimycolate, sulpholipids and phosphatidylinositol mannosides, most of which are specific for mycobacteria. While the free lipids play an important role as signaling molecules in the disease process, the cell wall core is

essential for the survival of the cell and the biosynthesis of its compounds display a good target for anti-mycobacterial drugs (Brennan 2003).

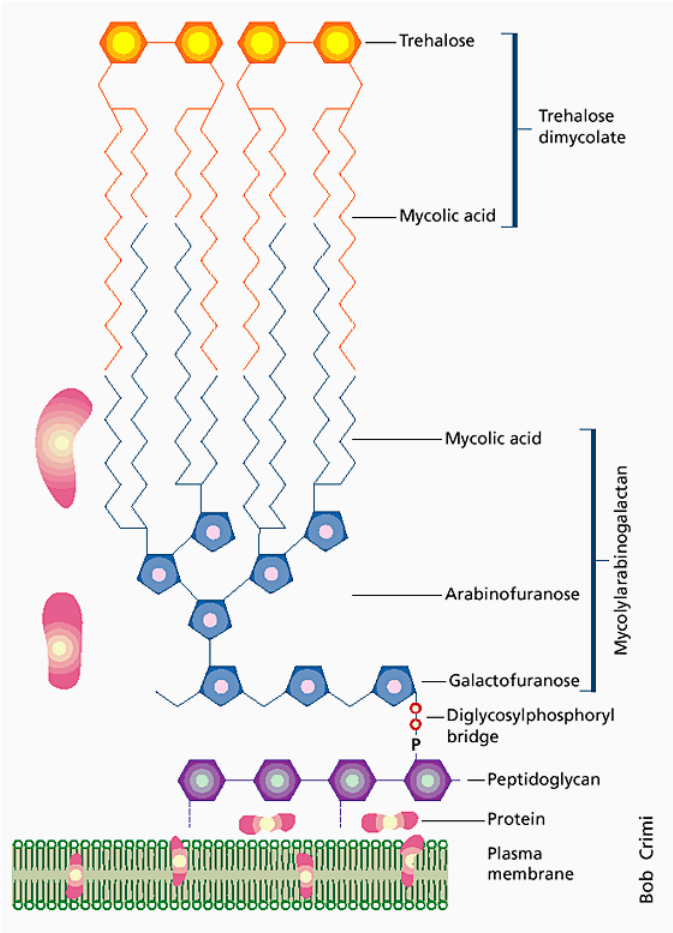


Figure 1-3. Schematic representation of the mycobacterial cell wall. The peptidoglycan is covalently attached to arabinogalactan, which in turn is covalently linked to mycolic acids. Also shown in yellow is trehalose dimycolate, the cord factor. The trehalose is esterified to two mycolic acids that intercalate with the mycolylarabinogalactan core. Figure was taken from (Tonge 2000).

Myoclic acids constitute up to 60% of the cell wall and are mainly responsible for the low permeability of the cell wall. Each molecule consists of a long α-alkyl chain and a shorter β-hydroxy chain and contains between 60 and 90 carbon atoms.

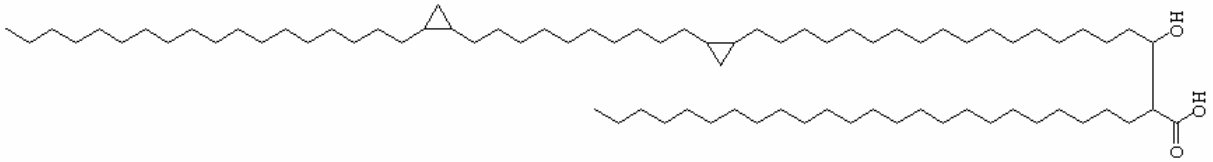


Figure 1-4. Example of an α-mycolic acid of *Mycobacterium tuberculosis*.

Most mycolic acids also contain various functional groups and the number of carbon atoms varies by species. Thus, the composition of the mycolic acids can be used to distinguish different mycobacteria species (Barry, Lee et al. 1998). Mycobacteria are related to members

of the genera *Nocardia*, *Rhodococcus* and *Corynebacterium* that contain similar but shorter chain length mycolic acids in their cell wall.

1.2 Development of novel antibiotics against tuberculosis

The development of new drugs against tuberculosis is a serious and inevitable challenge. Towards the containment of the bacteria, it is of particular importance to develop drugs that are able to shorten the treatment, combat drug resistant strains and are efficient against the dormant mycobacteria in the macrophages. Currently there are a few promising candidates for treatment of TB and MDR TB in clinical trials.

TMC-207 is patented by Johnson & Johnson and Tibotec is currently conducting the clinical trials on this compound. TMC-207 belongs to the diarylquinones and targets a subunit of ATP synthase. The compound is defined bactericidal and active against MDR-TB. *In vivo* studies showed that the compound has the potential to shorten treatment time (TB Alliance 2008).

PA-824 is a nitroimidazole and is the intellectual property of the non-profit organization TB Alliance and of Novartis. PA-824 seems to be a prodrug that acts via the generation of radicals and displays non-specific toxic effects in the mycobacterium. So far, it was shown to inhibit mycolic acid and protein biosynthesis. The compound is currently in Phase II clinical trials (TB Alliance 2008).

OPC-67683 is a highly potent compound developed by Otsuka pharmaceuticals. Like PA-824 it belongs to the nitroimidazole family and shows remarkable effects against MDR-TB. In addition it is also bactericidal against the bacteria in the anaerobic state and could therefore be used to treat both active and latent mycobacteria. OPC-67683 inhibits mycolic acid biosynthesis but the action seems to be different from that of isoniazid. OPC-67683 also requires metabolic activation by a mycobacterial enzyme. Rv3547 was reported as the enzyme that activates both PA-824 and OPC-67683, but the function of this enzyme is so far unknown. The drug is currently in Phase II clinical trials. In contrast to TMC-207, PA-824 and OPC-67683 are not active against *M. leprae*, because *M. leprae* lacks the Rv3547 enzyme required for the activation of the two compounds (Matsumoto, Hashizume et al. 2006).

Due to the limited research on tuberculosis for the last few decades, there is now a gap of new drugs to treat the resistant strains of *M. tuberculosis*. This problem becomes even more severe when it comes to clinical trials, since the gap of knowledge has increased over the years. For example, formal guidance on what exactly researchers need to measure and achieve in order to receive a licence for their new drug has yet to be published by the FDA (US Food and Drug Administration) (Sacchettini, Rubin et al. 2008).

1.2.1 Drug target fatty acid synthesis pathway II

To generate long chain fatty acids or mycolic acids, mycobacteria have, unlike other bacteria, two types of fatty acid synthesis pathways. The mammalian-like fatty acid synthesis pathway I (FAS-I) is a large multifunctional polypeptide complex and capable of de novo fatty acid synthesis. FAS-I produces fatty acids with a chain length of C_{16} and C_{26}. While the C_{26} fatty acids constitute the α-branch of the final mycolic acid, the C_{16} acyl-CoA products act as substrates for the fatty acid synthesis pathway II (FAS-II) where they are elongated up to C_{56} to build meromycolic acid. Polyketide synthase Pks13 catalyses the condensation of the α-branch and the meromycolic acid to produce mycolic acids.

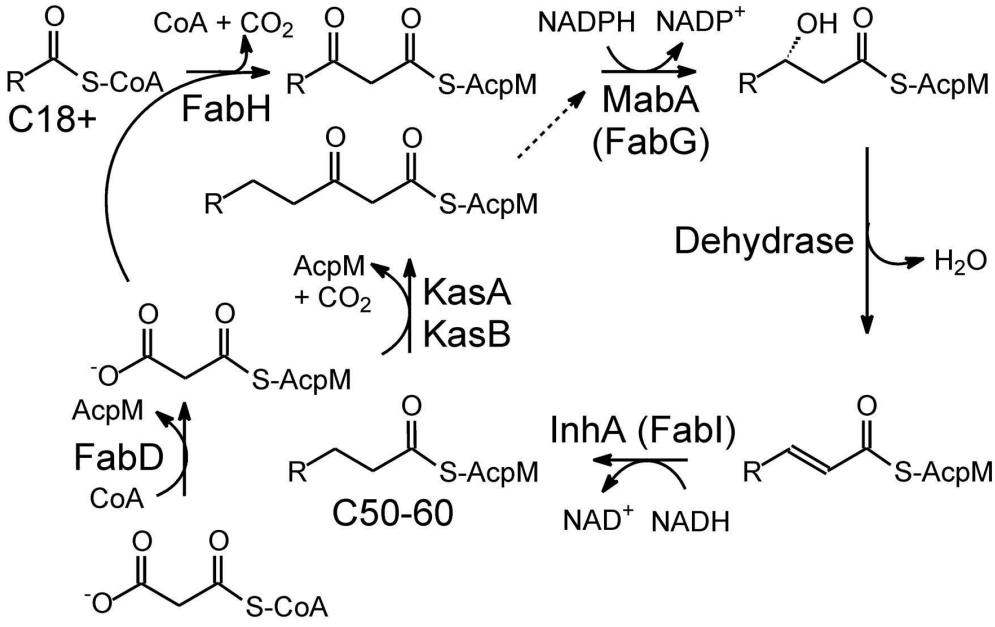


Figure 1-5. The fatty acid synthesis pathway II of *Mycobacterium tuberculosis*. Enzymes involved in the fatty acid elongation cycle are the condensing enzymes KasA and KasB, the keto-reductase MabA, an unidentified dehydratase, and the enoyl-reductase InhA. Also shown is the β-ketoacyl synthase III FabH that represents a link between FAS-I and FAS-II, and the transacylase FabD that synthesises the substrate malonyl-AcpM from malonyl –CoA.

The bacteria-like FAS-II pathway consists of discrete monofunctional enzymes that catalyze the individual steps of fatty acid synthesis (Figure 1-5). The acyl carrier protein AcpM transfers the growing fatty acid from one enzyme to the next in the cycle. The pathway is primed by the CoA-dependent β-ketoacyl-AcpM synthase FabH, which condenses the acyl-CoA with malonyl-AcpM to generate a β-ketoacyl-AcpM. The β-ketoacyl-AcpM is converted into a saturated enoyl-AcpM by the sequential actions of a β-ketoacyl-AcpM reductase (MabA), a dehydrase and a trans-2-enoyl-AcpM reductase (InhA). Subsequent rounds of elongation are initiated by either the KasA or KasB β-ketoacyl-AcpM synthases. KasA is

thought to be responsible for the early rounds of fatty acid elongation, while KasB may be involved in later steps of mycolic acid biosynthesis (Kremer, Douglas et al. 2000). The malonyl-CoA AcpM transacylase (FabD) is responsible for the synthesis of malonyl-AcpM.
Growing evidence points towards a direct interaction of FAS-II enzymes with each other in interconnected specialized complexes that are essential for mycobacterial survival (Kremer, Dover et al. 2003; Veyron-Churlet, Guerrini et al. 2004; Veyron-Churlet, Bigot et al. 2005). The molecular basis for these interactions, however, remains sketchy and, moreover, the capability of mycobacterial enzymes to interact with and efficiently transfer the extremely long hydrophobic fatty acids from one protein to the next within a cytosolic environment is not understood.
The enzymes of the FAS-II pathway of *Mycobacterium tuberculosis* represent attractive targets for the development of novel antibacterials. Inhibitors of FAS-II enzymes, with the first line antibiotic isoniazid that targets InhA as the most prominent example, impair the integrity of the cell wall and thereby act as bactericidal agents (Slayden, Lee et al. 1996). Mycolic acids are essential for the integrity of the cell wall and the mycobacteria can not survive by scavenging fatty acids from the host. Finally, despite the similar reactions, the difference of the bacterial FAS-II pathway to the mammalian FAS-I pathway is considerable so that a specific inhibition of the bacterial pathway is possible without affecting the human system.

1.2.2 Beta-ketoacyl synthase KasA

KasA is the mycobacterial β-ketoacyl synthase I and an important key enzyme within the FAS-II system. The active form of KasA is a homodimer that contains one active site per monomer. The residues of the catalytic triad are contained within the active site pocket and consist of one cysteine (Cys171) and two histidines (His311 and His345). KasA catalyzes the Claisen condensation between its two substrates malonyl-AcpM and acyl-AcpM *via* a ping pong mechanism and thereby adds two carbon atoms to the growing fatty acid chain (Figure 1-5). This reaction can be divided into three steps. In a first step, the acyl chain is covalently transferred to the active site cysteine resulting in an acylated KasA intermediate. The reduced AcpM is released. Subsequently, malonyl-AcpM binds to the active site and is decarboxylated. Thirdly, the condensation reaction occurs and the final product, β-ketoacyl-AcpM is released (Kremer, Dover et al. 2002; Bhatt, Molle et al. 2007).
KasB, the β-ketoacyl synthase II of *M. tuberculosis*, has a sequence identity of 67% to KasA. Despite the high similarity of the two enzymes, previous work has shown, that KasA and KasB have slightly separate functions. Genetic studies showed that KasA is essential for

mycobacterial cell growth, conditional depletion of the enzyme leads to cell lysis, while KasB was shown to be not essential (Bhatt, Kremer et al. 2005). Overexpression of KasA in *M. smegmatis* resulted in increased levels of shorter chain monounsaturated mycolic acids (C_{40}) compared to a strain overexpressing KasB, that led to longer chains of C_{54} (Slayden and Barry CE 2002). These results suggest that KasA is responsible for the early rounds of fatty acid elongation, whereas KasB elongates the chain to full length. However, the precise roles of the enzymes in mycolate biosynthesis are not completely understood. In addition, inhibition studies showed that KasA is more sensitive to thiolactomycin inhibition than KasB (Kremer, Douglas et al. 2000; Schaeffer, Agnihotri et al. 2001) and recent findings reveal that slow onset inhibition of TLM to the acylated form of KasA can not be observed for the acylated form of KasB (Machutta 2009). At the same time, the described characteristics make KasA a suitable target for the development of novel antibiotics without the risk that its function can be taken over by KasB or another enzyme within the organism.

1.2.3 Enoyl reductase InhA

InhA is the enoyl reductase of the fatty acid synthesis pathway II of *M. tuberculosis*. The enzyme catalyzes the reduction of the *trans* double bond between C2 and C3 of a fatty acid-AcpM in an NADH consuming reaction (Figure 1-5). The active form of InhA is a tetramer stabilized at one interface of a monomer by a four helix bundle and at the other interface by an extended β-sheet. Each monomer consists of seven β-strands that form one parallel beta sheet and eight α-helices. The β-strand forms together with six α-helices a Rossman fold in the lower part of the monomer to which the NADH (nicotinamide adenine dinucleotide) cofactor binds. In the upper part of the structure, a deep, hydrophobic pocket is formed by three α-helices, two of them form the flexible substrate binding loop. The substrate-binding loop of InhA is longer than that of other enoyl reductases and creates a deeper substrate-binding crevice, consistent with the ability of InhA to recognize longer fatty acid substrates (Dessen, Quemard et al. 1995).

InhA is an essential enzyme of FAS-II, its inhibition impairs the integrity of the cell wall and leads to cell lysis (Vilcheze, Morbidoni et al. 2000). Isoniazid (INH), one of the first line drugs against tuberculosis, targets InhA as well as ethionamide (ETA) that belongs to the second line drugs. Both compounds are prodrugs that have to be converted into the active forms by mycobacterial enzymes. INH is activated by KatG (Zhang, Heym et al. 1992), the mycobacterial catalase-peroxidase, and forms subsequently with NAD(H) the isonicotinoyl-NAD adduct (INH-NAD). The crystal structure of the complex between InhA and the INH-

NAD adduct shows that the adduct binds to the NADH binding site and thereby prevents the binding of the substrates NADH and the acyl chain (Vilcheze, Wang et al. 2006). In addition, the INH-NAD adduct was shown to be a slow, tight binding inhibitor of InhA with a half-life of the enzyme inhibitor complex of 43 min (Rawat, Whitty et al. 2003). The slow rate of the adduct dissociation has major impact on the *in vivo* activity of the drug.

The predominant mechanism of resistance of *M. tuberculosis* strains against isoniazid arises from mutations in the activating enzyme KatG rather than from mutations in InhA (Zhang, Heym et al. 1992). New compounds that directly target InhA and circumvent the activation step, are promising candidates for combating sensitive as well as drug resistant strains of *M. tuberculosis*.

1.2.4 Inhibitors of KasA and InhA

Thiolactomycin (TLM) is a natural product inhibitor isolated from *Nocardia* as well as *Streptomyces* species and was found by a Japanese company through screening experiments in the early 1980s (Sasaki, Oishi et al. 1982). TLM is a promising lead compound for the development of potent FAS-II inhibitors: it was shown to inhibit β-ketoacyl synthases with activity against gram positive and gram negative bacteria as well as mycobacteria (Slayden, Lee et al. 1996; Kremer, Douglas et al. 2000). TLM has favorable physicochemical properties, is effective in mouse infection models and it has been shown to inhibit the mycobacterial β-ketoacyl synthases KasA and KasB, with KasA being the most sensitive (Kremer, Douglas et al. 2000 ; Schaeffer, Agnihotri et al. 2001).

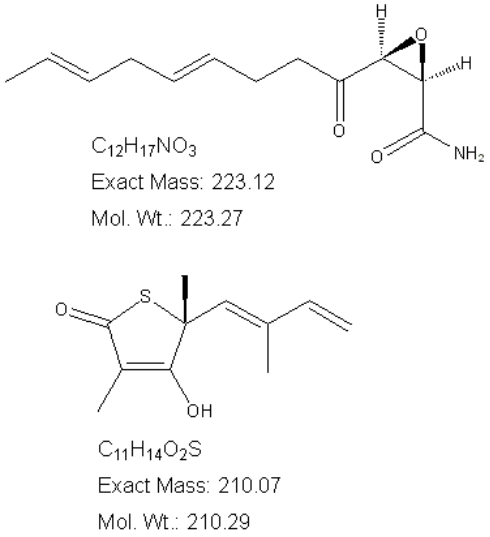


Figure 1-6. The structures of cerulenin (top) and thiolactomycin (bottom).

SAR studies showed that the isoprenoid moiety of TLM is very important for binding of the inhibitor to the enzyme. The removal of one or both of the double bonds decreased the affinity

dramatically (Kim, Zhang et al. 2006). Machutta et al. recently found by kinetic studies that TLM binds to both, the free enzyme and the acylated form of KasA with different affinities (Machutta 2009). The inhibitor preferentially binds to the acyl-enzyme intermediate and shows a slow binding step during the inhibition reaction, which plays a crucial role for the *in vivo* activity of the compound (Copeland, Pompliano et al. 2006).

Cerulenin is a natural compound from *Cephalosporium caerulens* and an irreversible inhibitor of β-ketoacyl synthases since it binds covalently to the active site cysteine of the enzymes. Due to the lack of specificity of the compound, it inhibits both, the bacterial FAS-II as well as the mammalian FAS-I complex, the development of cerulenin as a drug has not been widely pursued (Kauppinen, Siggaard-Andersen et al. 1988).

Triclosan is a synthetic molecule that is widely used as a biocide in toothpaste or antibacterial soaps. Triclosan was formerly thought to be a rather nonspecific biocide, but it was shown to be an inhibitor of the FAS-II of bacteria with the enoyl reductase as the primary target (McMurry, Oethinger et al. 1998). As triclosan directly targets the essential enoyl reductase of FAS-II with no need for activation, it is used as a lead compound for the development of inhibitors against several pathogens like *Staphylococcus aureus*, *Plasmodium falciparum* and *Mycobacterium tuberculosis*.

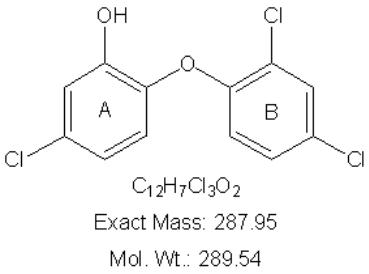


Figure 1-7. Structure of triclosan.

The mechanism of inhibition of triclosan with InhA has been studied in detail using enzyme kinetics and X-ray crystallography. Inhibition of InhA by triclosan is NAD^+ dependent and therefore triclosan is an uncompetitive inhibitor of InhA. The compound binds to the substrate-binding pocket on top of the NAD^+ molecule and stacking interactions are formed between the triclosan A-ring and the nicotinamide ring of the cofactor. A hydrogen-bonding network between triclosan, the NAD^+ cofactor and the active site residue Tyr158 further stabilizes the inhibitor. Another important feature of triclosan is that it evades the mycobacterial detoxification response (Tonge, Kisker et al. 2007).

Triclosan is a picomolar inhibitor of the enoyl reductase of *E. coli* (ecFabI) (Ward, Holdgate et al. 1999) but inhibits InhA of *M. tuberculosis* only in the sub-micromolar range (Parikh, Xiao et al. 2000). While the crystal structures of the complexes formed between Triclosan and

the enoyl reductases of *E. coli* and *M. tuberculosis* are very similar, one major difference is the substrate binding loop that is ordered in ecFabI and covers the substrate binding pocket (pdb code 1qg6, (Ward, Holdgate et al. 1999)) whereas it is disordered in InhA (pdb code 2b35, (Sullivan, Truglio et al. 2006)). As triclosan is a slow onset inhibitor of ecFabI and a rapid reversible inhibitor of InhA, it was speculated that ordering of the loop is an explanation for the slow step. In order to improve the triclosan molecule towards inhibition of InhA and to generate a slow onset inhibitor of InhA, triclosan derivatives were developed based on the structure of InhA with a bound C_{16} substrate (pdb code 1bvr, (Rozwarski, Vilcheze et al. 1999)). The best of these first generation diphenyl ethers (Figure 1-8) are nanomolar inhibitors of InhA with MIC_{99} values in the range of 1-2µg/mL. Still, these inhibitors are rapid reversible inhibitors of InhA and the crystal structures show in agreement with the hypothesis, that the substrate binding loop remains disordered (pdb codes 2b36 and 2b37, (Sullivan, Truglio et al. 2006)).

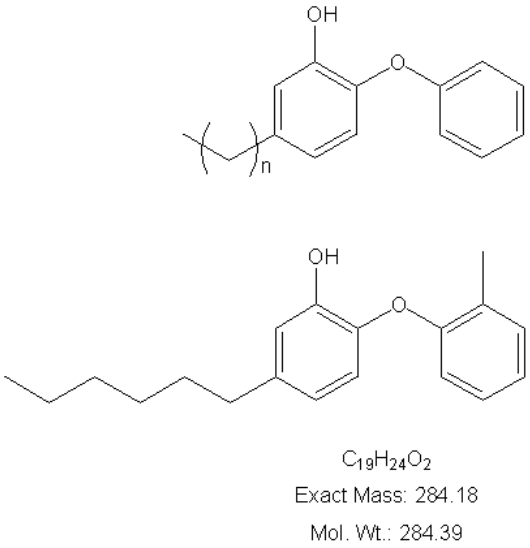


Figure 1-8. Structures of the 5-alkyl-diphenyl ethers (top) and the slow binding inhibitor 2-(o-tolyloxy)-5-hexylphenol (PT70, bottom).

Towards the design of a slow onset diphenyl ether, it was speculated that there must be an entropic penalty for loop ordering. Thus, reducing the conformational flexibility of the lead diphenyl ether might enable ordering of the active site loop which in turn results in slow onset inhibition. Introduction of a methyl group *ortho* to the diphenyl ether linkage resulted in a compound 2-(o-tolyloxy)-5-hexylphenol (PT70, Figure 1-8) that was shown to be a slow onset inhibitor of InhA with a K_i value of 0.5 nM.

1.3 Research objectives

In this work, the mycobacterial enzyme KasA was structurally characterized. As no structure of the mycobacterial KasA was available so far, structure based drug development efforts so far were based on the homologous enzyme of *E.coli* (ecFabB). In order to evaluate the interactions of TLM with KasA as a platform for rational inhibitor design, the crystal structures of KasA in the absence and presence of bound TLM were solved. Since TLM can bind to the apo form as well as to the acylated form of the enzyme, with preference for the acylated form, the crystal structures of an acyl enzyme mimic were determined. The replacement of the active site cysteine with a glutamine can mimic the carbonyl of the acyl moiety in the acyl-enzyme intermediate so that the mutant adopts an analogous conformation to that of the acyl-enzyme intermediate (Witkowski, Joshi et al. 1999). For structural studies, the C171Q mutant was used as a mimic of the acyl enzyme intermediate of KasA.

The crystal structures of the ternary InhA-NAD^+-PT70 complex were solved to further characterize the binding of the slow onset inhibitor PT70 to the key enzyme InhA. Of particular importance is the structural basis for the slow onset inhibition of InhA by PT70. These studies provide additional information on the mechanistic imperatives for slow onset inhibition of enoyl ACP reductases.

2 Material and methods

2.1 Material

2.1.1 Bacterial strains and plasmids

Mycobacterium smegmatis strain mc^2155, an efficient-plasmid-transformation (ept) mutant strain from the wild-type ATCC 607 (Snapper, Melton et al. 1990)

E. coli strain BL21(DE3)pLysS (Stratagene)

Mycobacterial expression vector pFPCA (Changsen, Franzblau et al. 2003)

E. coli expression vector pet15b (Novagen)

2.1.2 Media and antibiotics

Name	**Supplier**
7H10 solid medium for Mycobacteria	Middlebrook
7H9 medium for Mycobacteria	Middlebrook
Ampicillin	Carl Roth
Chloramphenicol	Carl Roth
Cycloheximide	AppliChem
Kanamycin	Carl Roth
LB medium	Carl Roth
LB agar	Carl Roth

2.1.3 Chemicals

If not noted differently, all chemicals and solutions were purchases from Carl Roth (Karlsruhe), Sigma Aldrich (Seelze), Fluka (Neu-Ulm), Hampton Research (Laguna Hills, USA), or Applichem (Darmstadt). All chemical were of analytical grade or better, the chemicals used for crystallization were of the highest available purity.

2.1.4 Relevant buffers and solutions

All buffers were adjusted to the respective pH value by adding either 32% NaOH or 37% HCl.

Table 2-1. Buffers and solutions for gel electrophoresis

Staining solution	Destaining solution	Running buffer (1x)	Loading buffer (1x)
50% Methanol	10%Methanol	192 mM Glycin	50 mM Tris pH 6.8
10% Acetic acid	5%Acetic acid	0.1% SDS	100 mM DTT
0.1% Coomassie brilliant blue		25 mM Tris	2% SDS
			0.1% Bromphenol blue
			10% Glycerol

Table 2-2. Buffers for KasA affinity chromatography

Binding buffer	Washing buffer	Elution buffer	Dilution buffer
20 mM CHES pH 9.5	20 mM CHES pH 9.5	20 mM CHES pH 9.5	20 mM CHES
500 mM NaCl	500 mM NaCl	500 mM NaCl	pH 9.5
5 mM imidazole	40 mM imidazole	1 M imidazole	

Table 2-3. Buffers for KasA anion exchange chromatography

Buffer A	Buffer B
20 mM Ches pH 9.5	20 mM Ches pH 9.5
50 mM NaCl	1 M NaCl

Table 2-4. Buffer for KasA size exclusion chromatography

Size exclusion buffer
20 mM Ches pH 9.5
500 mM NaCl

Table 2-5. Buffers for InhA affinity chromatography

Binding buffer	Washing buffer	Elution buffer
20 mM Tris pH 7.9	20 mM Tris pH 7.9	20 mM Tris pH 7.9
500 mM NaCl	500 mM NaCl	500 mM NaCl
5 mM imidazole	60 mM imidazole	500 mM imidazole

Table 2-6. Buffer for InhA size exclusion chromatography

Size exclusion buffer
20 mM Pipes pH 6.8
150 mM NaCl

2.1.5 Crystallization screens

Table 2-7. Crystallization Screens used in this work

Screen	Reference
Crystal Screen 1	Hampton Research
Crystal Screen 2	Hampton Research
Index Screen HT	Hampton Research
Nextal PEG suite	Qiagen
Wizard Screen I	Emerald Biosystems
Wizard Screen II	Emerald Biosystems

2.1.6 Equipment and instrumentation

Table 2-8. Important instruments used in this work

Name	Company
Äkta purifier 10	GE Healthcare
DynaPro Dynamic Light Scattering	Wyatt Technology Corporation
HoneyBee 963 crystallization robot	Zinsser Analytic
MicroMax 007 HF X-ray generator	Rigaku
NanoDrop ND 1000 Spectrophotometer	Peqlab
Raxis HTC X-ray detector	Rigaku
Real time PCR cycler (Thermofluor)	Stratagene
X-stream 2000 cryo system	Rigaku

2.1.7 Computer software and databases

Table 2-9. Software and databases used in this work

Software	Author / Reference
CastP	(Dundas, Ouyang et al. 2006)
Caver	(Petrek, Otyepka et al. 2006)
Ccp4 program suite	(CCP4 1994)
Coot	(Emsley and Cowtan 2004)
Crystal Twinning Server	(Yeates and Fam 1999)
D*Trek (part of Crystal Clear)	(Pflugrath 1999)
ExPASy Proteomics Server	http://www.expasy.ch/

MolProbity Server	(Davis, Leaver-Fay et al. 2007)
Mosflm	(Leslie 1992)
Phaser	(McCoy, Grosse-Kunstleve et al. 2007)
Phenix.refine	(Adams, Grosse-Kunstleve et al. 2002)
Phenix.xtriage	(Zwart, Afonine et al. 2008)
Pymol	(DeLano 2002)
Refmac	(Murshudov, Vagin et al. 1997)
Scala	(Kabsch 1988)
Stride Web Interface	(Frishman and Argos 1995)
TopMatch-web, protein structure comparison	(Sippl, Suhrer et al. 2008; Sippl and Wiederstein 2008)
XDS	(Kabsch 1993; Kursula 2004)

2.2 Methods

2.2.1 Transformation

Transformation is the directed uptake of DNA from the environment by competent acceptor cells. As only a few bacteria species are naturally competent, most other bacteria, including *E. coli* and *M. smegmatis*, have to be treated to make them competent. In case of *E. coli*, cells in the exponential growth phase are treated with ice-cold calcium chloride to make them chemically competent. Alternating heat shock and cold shock cycles are then used to transform the DNA into the bacteria. The plasmid pet15b which carries the gene for *Mycobacterium tuberculosis* InhA with an N-terminal His-tag and an ampicillin resistance marker was transformed into *E. coli* expression strain BL21(DE3)pLysS by heat shock transformation. 1 µL plasmid DNA was added to 100 µL competent cells and incubated for 10 min on ice. The solution was heat shocked for 90 sec at 42° C, cooled on ice and subsequently 1 mL LB liquid medium was added. The transformed cells were incubated by shaking at 37° C for 15 min before plating 100 µL on LB agar plates containing 100 µg/mL ampicillin and 34 µg/mL chloramphenicol. After incubation over night at 37° C positive transformants grew on the plate.

Transformation of mycobacteria can be difficult due to the thick and waxy cell wall. Nevertheless, it is possible, with the most successful method for transformation in *M. smegmatis* being electroporation. For this purpose, *M. smegmatis* cells in their exponential growth phase are washed several times with 10% ice-cold glycerol to make them

electrocompetent. By applying an electrical field to the cells, the cell membrane gets porous and the DNA is transformed into the bacteria. The plasmid pFPCA1, containing kanamycin resistance and the gene for KasA or the KasA C171Q mutant under the control of an acetamide inducible promoter, was transformed into *Mycobacterium smegmatis* strain mc^2155 by electroporation using a BioRad Micropulser.

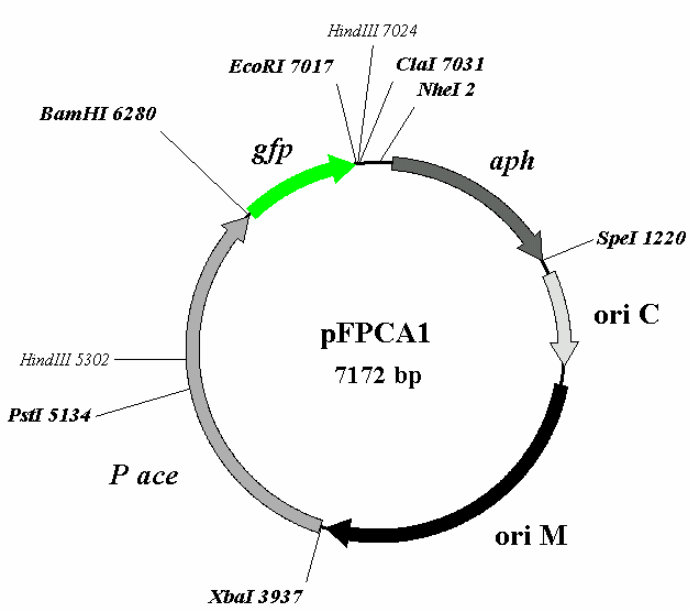


Figure 2-1. The pFPCA1 vector; origins of replication for both M. smegmatis (ori M) and E.coli (ori C), acetamidase promoter (P ace), kanamycin resistance gene aminoglycoside phosphotransferase (aph) and multiple cloning sites are illustrated. The GFP gene insert was excised by digestion with BamHI and EcoRI and the KasA gene was inserted. Figure obtained from Carl Machutta (personal communication).

1 µL of plasmid DNA was added to 200 µL competent *M. smegmatis* cells, and incubated for 10 min on ice. The mixture was electroporated and subsequently 600 µL 7H9 liquid media were added. The transformed cells were incubated by shaking at 37° C for 4h before plating them on 7H10 solid media containing 50 µg/mL kanamycin, 200 µg/mL ampicillin and 10 µg/mL cycloheximide. Mycobacteria are naturally resistant against ampicillin due to the highly impermeable cell wall and an intrinsic β-lactamase (Jarlier and Nikaido 1994). Ampicillin was used in the medium to prevent growth of *E. coli* and other fast growing bacteria while cycloheximide, an inhibitor of eukaryotic protein biosynthesis was used to prevent contamination with yeast and fungi. After 5 days of growth on the agar plate at 37° C colonies could be observed. As the plasmid carries a kanamycin resistance, mycobacteria containing the plasmid could be selected on the plates.

2.2.2 Protein expression

2.2.2.1 Expression of KasA in *M. smegmatis*

Past attempts to express *Mycobacterium tuberculosis* KasA in *E. coli* did not yield soluble protein, therefore KasA was expressed using *Mycobacterium smegmatis* as an expression host. The mycobacterial expression vector pFPCA1 contains an acetamidase promoter region, where the repressor molecule AmiA binds and prevents transcription of the target gene downstream of the promoter. In the presence of acetamide, AmiA is induced, dissociates from the promoter and transcription proceeds.

For the expression of KasA, a starter culture with 7H9 liquid medium supplemented with 0.5% glycerol and antibiotics was inoculated with selected individual colonies and grown by shaking at 37° C and 250 rpm for 3-4 days until a dense bacteria solution was achieved. 20 mL from the starter culture was used to inoculate 1.8 L of the main culture which was grown by shaking at 37° C and 200 rpm. At an OD_{600} of 0.6-0.8 protein expression was induced by adding 0.2% (w/v) solid acetamide. After an additional shaking of 36 h at 37° C, cells were harvested by two centrifugation steps each at 6000g for 20 min at 4° C. The obtained pellet was frozen at -80° C. Two centrifugation steps were necessary to finally pellet the cells.

2.2.2.2 Expression of InhA in *E. coli*

InhA was previously expressed in *E. coli* and yielded soluble protein. The expression vector pet15b contains a T7 RNA polymerase promoter and a Lac-operator region upstream of the target gene. The constitutively expressed Lac repressor LacI prevents the expression of T7 RNA polymerase in the genome of the *E. coli* (BL21) expression strains and the expression of the target gene. In the presence of IPTG LacI is released from the DNA and T7 RNA polymerase is produced and in turn transcribes the target gene.

For the expression of InhA, a starter culture with LB liquid medium supplemented with antibiotics was inoculated with selected colonies and grown by shaking at 37° C and 200 rpm over night. The starter culture was used to inoculate the main culture to obtain an initial optical density (OD_{600}) of 0.05. The main culture was grown initially at 37° C and shaken at 200 rpm to an OD_{600} of 0.5-0.6 at which protein expression was induced. IPTG was added to a final concentration of 0.5 mM and simultaneously the temperature was lowered to 30° C. After 4 h shaking at 30° C and 200 rpm, cells were harvested by centrifugation at 5000g for 15 min at 4° C. The obtained pellet was frozen at -80° C.

2.2.3 Chromatographic purification

The chromatographic purification of KasA and InhA was performed with an Äkta purifier 10 liquid chromatography system (GE Healthcare) equipped with a sample pump P960, a Monitor pH/C-900 for measurement of pH, conductivity, and temperature and a UV900 Monitor for UV-Vis detection of up to three wavelengths simultaneously. The fraction collector Frac-950 was connected to the Äkta system and was used for automated collection of the elution fractions. KasA was purified by nickel-affinity chromatography followed by anion-exchange and size-exclusion chromatography. Purification of InhA occurred in two steps, nickel-affinity followed by size-exclusion chromatography.

2.2.3.1 Nickel affinity chromatography

Affinity chromatography separates proteins on the basis of a reversible interaction between a protein and a specific ligand attached to a chromatographic matrix. The immobilized metal ion affinity (IMAC) uses the property of proteins to interact with metal ions. The most widely used form of IMAC is nickel affinity chromatography, where histidines interact with Ni^{2+} ions. Both, KasA and InhA were expressed with an N-terminal hexahistidine tag (His_6-tag) that binds to nickel chelate resin. The column used for affinity purification was a HisTrap HP 5 mL column (GE Healthcare) with a Ni Sepharose matrix.

For purification of KasA as well as InhA, the respective cell pellets were thawed, resuspended in binding buffer and lysed by sonication. The lysate was centrifuged for 1 h at 50000g and 4° C to separate cell fragments from the supernatant. The supernatant was directly loaded onto the nickel-affinity column with the Sample Pump P960. After the complete sample was loaded, two washing steps were applied before elution of the protein. First, the unbound proteins were washed out with eight column volumes of binding buffer. Subsequently, the unspecifically bound proteins were eluted with six column volumes of washing buffer that contains a slightly higher imidazole concentration (see Table 2-2 and Table 2-5 for exact buffer compositions). The target proteins were eluted *via* competitive displacement with an imidazole gradient from 40 mM to 1 M for KasA and 60 mM to 500 mM for InhA, in both cases over a range of 20 column volumes. The elution was collected in fractions of 3 mL. During the whole procedure the absorption at 280 nm was monitored so that the peak fractions could be defined. In combination with an SDS PAGE (described in chapter 2.2.4.1), the fractions containing the purified protein were identified and pooled for further purification. The KasA sample was diluted with 20 mM CHES buffer pH 9.5 (dilution buffer,

see Table 2-3 for exact buffer compositions) until the NaCl concentration was reduced to 50 mM to prepare the sample for the anion exchange chromatography.

The sample containing purified InhA was exchanged to size exclusion buffer (20 mM Pipes pH 6.8, 150mM NaCl) with PD10 desalting columns (GE Healthcare) to remove the imidazole and subsequently concentrated with a centrifugal filter device (Vivaspin, molecular weight cut off 5000 Dalton) to prepare it for enzymatic His-tag cleavage and size exclusion chromatography.

2.2.3.2 Cleavage of the His-tag

Before the final purification step of InhA by size exclusion chromatography, the His-tag was cleaved off by enzymatic digestion with thrombin. For that purpose biotinylated thrombin and thrombin cleavage buffer (Novagen, thrombin cleavage capture kit) was added to the concentrated protein solution and incubated over night at room temperature. The thrombin was removed after the cleavage with streptavidin agarose beads. The beads were added to the mixture, incubated for 1 h with gentle shaking at room temperature and removed by centrifugation. The supernatant was filtered with a 0.2 µm filter and reloaded on the nickel affinity column to separate uncut His-tagged protein from the digested protein. The flow through was collected while bound protein with the His-tag and the His-tags themselves were eluted with imidazole from the column and discarded. The successful cleavage was checked on an SDS gel.

2.2.3.3 Anion exchange chromatography

With ion exchange chromatography, proteins can be separated based on differences in their net charge. Charged molecules interact with an oppositely charged chromatography medium and are eluted by a gradient of competing counterions such as NaCl or other salts. When the pH of the buffer is above the pI (isoelectric point) of the protein then the protein is negatively charged and able to interact with the positively charged matrix of an anion exchange column.

KasA was purified in CHES buffer pH 9.5, while the calculated pI of KasA is 5.1. Thus, an anion exchange column was chosen as a second purification step of KasA using a MonoQ 10/100 column (GE Healthcare).

The diluted KasA sample was applied to the MonoQ with the Sample Pump P960. After loading was complete, the column was washed with six column volumes of buffer A (50 mM NaCl, 20 mM CHES pH 9.5) to wash out the unbound proteins before elution occurred with a NaCl gradient from 50 mM NaCl to 1M NaCl over 20 column volumes. The elution was collected in fractions of 1 mL. Fractions that contained the pure protein as judged by SDS

PAGE were combined and used for the final purification step, the size exclusion chromatography.

2.2.3.4 Size exclusion chromatography

Size exclusion chromatography is a method to separate proteins by their size and shape while they are in the native state. The matrix of the column consists of porous agarose and dextran beads with an average particle size of 34 µM. Large molecules pass the beads rapidly as they do not permeate into the pores, while small molecules diffuse into the pores of the beads and are delayed. Thus, large molecules elute first from the column followed by smaller molecules.
For the last polishing step, KasA and InhA were purified by size exclusion chromatography, using the HiLoad 26/60 Superdex 200 p.g. (GE Healthcare) column that has a volume of 320 mL and a separation range of 10 kDa to 600 kDa. To obtain a good resolution, the sample volume was always kept below 10 mL and the flow rate during elution was kept constant at 1.5 mL/min.
Purified KasA from the anion exchange column was already available in a concentrated form so that the sample was directly applied with a syringe to a 5 mL loop attached to the inlet valve of the Äkta. Isocratic elution with delayed fractionation was utilized and collected in fractions of 3.5 mL. The sample of InhA was applied in the same way: the concentrated protein after thrombin digestion was applied with a syringe to a 5 mL loop connected to the inlet valve of the Äkta. The isocratic elution with delayed fractionation was collected in fractions of 3.5 mL. The absorption reading at 280 nm in combination with SDS PAGE allowed the identification of the fractions containing the pure protein and the assessment of the grade of purity of the final samples. The elution profile gives additional information about the homogeneity and the oligomerization state of the protein.

2.2.4 Protein characterization

2.2.4.1 SDS polyacrylamid gelelectrophoresis

With SDS polyacrylamid gel electrophoresis (SDS PAGE) proteins are separated according to their molecular weight under denaturing conditions. SDS is a strong anionic detergent that binds to the main chain of the proteins so that the charge of the SDS-protein complexes is about proportional to the molecular weight of the protein. The protein sample is heated to break tertiary and quaternary structures. Additionally, a reducing agent is used to disrupt possible disulfide bonds. That way proteins can be separated in an electrical field, independently of their natural charge or conformation.

After each purification step, an SDS gel (Laemmli 1970) was run, to identify the fractions containing the protein and to analyze the grade of purity. For the estimation of the molecular weight of the different protein bands, the PageRuler prestained protein ladder (Fermentas) that shows bands between 10 and 170 kDa was loaded in one of the wells of the gel. Before loading on the gel, 10 µL of the protein sample was mixed with 3 µL 5x protein sample buffer, heated for 5 min at 95° C and the entire sample was loaded onto the gel. The run was performed for 45 min at 200 V. The gel was stained for 20 min in coomassie staining solution and the unbound stain was removed afterwards with destaining solution (for exact composition of the solutions see Table 2-1).

2.2.4.2 Dynamic light scattering

Dynamic light scattering (DLS) is used to determine the hydrodynamic sizes, polydispersities and aggregation effects of protein samples. DLS detects the laser light that is scattered by small particles in very short time periods and thus measures the rate of diffusion of the particles (dispersity). Since large molecules move slower in solution than small molecules, a correlation exists between the diffusion of the particles and the hydrodynamic radius or Stokes radius. Both mass and shape of the macromolecule define the hydrodynamic size.

Each protein batch was checked with the Dyna Pro DLS prior to crystallization to determine whether the protein is present in the form of a monomer or a higher oligomer and to detect the presence of aggregated protein. In general, protein at highest purity (after size exclusion chromatography) was used. The sample was adjusted to a concentration of 1 mg/mL and centrifuged at 25000g for at least 20 min right before usage. A polydispersity below 25% is generally desired for successful crystallization trials.

2.2.4.3 Thermofluor buffer screening

To find buffer conditions that stabilize the protein in solution a fluorescence based thermal shift assay was used. With this approach, it is also possible to identify additives, cofactors or inhibitors that enhance the stability of the protein, favour the crystallization process or help understanding the biology of the protein.

During the first purification trials, it turned out that KasA was highly instable under the chosen conditions (20 mM Tris-HCl pH 7.4, 200 mM NaCl) and two chaperones were bound to the protein. Towards the crystallization of KasA the chaperones had to be removed, however, removal of the chaperones resulted in aggregation of KasA. A Thermofluor buffer screen was performed to identify a buffer condition suitable to stabilize KasA during the purification process and for subsequent crystallization trials. Each well of a 96-well plate was

filled with 5 µL protein solution (10 µM), 7.5 µL Sypro Orange (50x) and 12.5 µL of the relevant buffer solution (100 mM). The solutions were gently mixed by pipetting and heated in the real time PCR cycler in 1° C steps from 25° C to 95° C.

Upon melting or unfolding of the protein, hydrophobic parts of the protein are exposed. According to Ericsson and co-workers, the dye binds to the hydrophobic parts of the protein, which results in a significant increase in fluorescence emission that reaches a maximum and then starts to decrease. The decrease is probably due to precipitation of the complex formed between the fluorescent probe and denatured protein. The midpoint of temperature of the protein-unfolding transition is defined as the melting temperature T_m. An increased T_m indicates the stabilization effect of the buffer condition compared to the buffer condition of the control (Ericsson, Hallberg et al. 2006).

2.2.4.4 Preparation of the protein for crystallization

Towards crystallization of a protein, the protein solution has to be highly concentrated (typically 5-15 mg/mL). To adjust the concentration of KasA and InhA for the crystallization experiments, a centrifugal filter device (Vivaspin) with a molecular weight cut off of 10000 Dalton was used for either of the two proteins. With a known calculated extinction coefficient ε of a protein, the absorption at 280 nm (A_{280}) and the path length d of the cuvette, the protein concentration c can be calculated with Lambert Beer's-Law.

Formula 1. Lambert Beers Law to calculate the protein concentration

$$c = \frac{A_{280}}{\varepsilon \cdot d}$$

The extinction coefficient of a protein can be calculated by the number of tryptophanes, tyrosines and cysteines that are present in the protein.

Formula 2. Calculation of the extinction coefficient of a protein

$$\varepsilon = \left(n_{Trp} \cdot 5690 + n_{Tyr} \cdot 1280 + n_{Cys} \cdot 120\right)$$

KasA contains four tryptophanes, eight tyrosines and three cysteines and the calculated extinction coefficient is 33360 $M^{-1} \cdot cm^{-1}$. InhA contains four tryptophanes, six tyrosines and one cysteine, which results in a calculated extinction coefficient of 30560 $M^{-1} \cdot cm^{-1}$. The absorption of the protein solution at 280 nm was measured with the NanoDrop ND-1000 UV/Vis Spectrophotometer (Peqlab) to calculate the protein concentration between the purification steps and before freezing the concentrated protein.

For the purpose of co-crystallization, inhibitors and cofactors were added to the protein solution prior to the crystallization setups. NAD^+ was dissolved in water and pipetted to the

protein solution in a five- to ten-fold molar excess. Inhibitors that are not water soluble were dissolved in DMSO and added likewise to the protein solution in up to 200-fold excess. Alternatively, if the protein was sensitive to DMSO, the inhibitor was dissolved in isopropanol (as done for TLM) or added as a solid powder or oil (in case of Triclosan or PT70) to the protein solution. The protein was incubated with the inhibitors for 1-2 h on ice before the solution was centrifuged for 20 min at 25000g to remove aggregated protein or precipitated inhibitor following the setup of the vapor diffusion experiment.

2.2.5 Crystallization

2.2.5.1 Crystallization methods

In order to determine the 3-dimensional structure of a protein by x-ray diffraction, a well-ordered single protein crystal is needed. To grow a protein crystal, the highly pure, concentrated protein solution has to be brought in the state of supersaturation. Several techniques have been developed to achieve the supersaturation state with the vapor diffusion technique being the most frequently applied method.

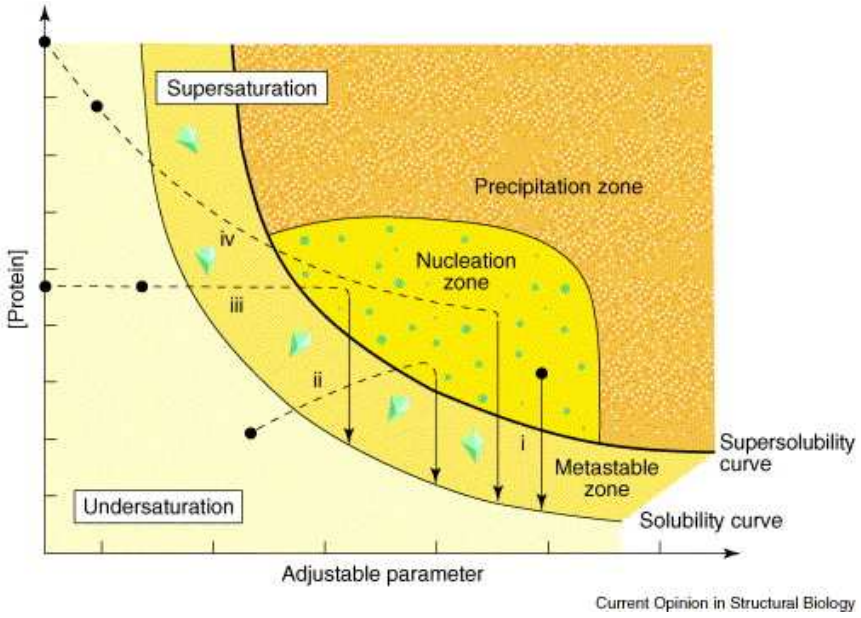


Figure 2-2. Crystallization phase diagram displaying how supersaturation is reached to trigger crystallization in different crystallization methods. Figure taken from (Chayen 2004).

For vapor diffusion (Figure 2-2, line ii, (Chayen 2004)) the protein solution is combined with crystallization solution containing an adequate precipitant (see chapter 2.2.5.2). In a sealed reservoir, the drop solution is equilibrated against a much larger reservoir of the crystallization solution. This technique can be performed by either hanging the drop from a cover slide towards the reservoir (*hanging drop*) or by placing the drop on a shelf that is

surrounded by the reservoir (*sitting drop*). In both cases, diffusion of H_2O from the drop to the reservoir slowly increases the precipitant and the protein concentration in the drop, the nucleation barrier is overcome and small crystal nuclei form. Upon formation of nuclei, the protein concentration in the drop is decreased and the condition in the drop is pushed towards the metastable zone where growth of crystals at the nucleation sites occurs. In this work the vapor diffusion method was used as the only method to obtain crystals. The sitting drop method was used for initial screening with the HoneyBee crystallization robot. The drop contained 0.2 µL protein solution and 0.2 µL crystallization solution and was equilibrated against 40 µL reservoir solution in a 96-well plate. The hanging drop method was used for the follow-up screens where the drop (mostly 1 µL protein solution pipetted on 1 µL crystallization solution) was equilibrated against 1 mL reservoir solution.

2.2.5.2 Crystallization screens

The precipitants of crystallization solutions are mainly inorganic salts, organic solvents or polyethylene glycols. Additionally, other factors like the protein concentration, the temperature and the pH of the solution have an influence on the development of protein crystals. As there is no obvious correlation between crystallization condition and protein structure or family so far, the conditions for every protein have to be identified by screening. Several strategies exist to test as many conditions as possible. The most frequently used screen for finding initial conditions is the *sparse matrix* screen, a screen that contains conditions that previously led to successful crystallization (Jancarik and Kim 1991). To optimize already known initial conditions, a *full factorial* screen is used in which every possible condition in a matrix of parameters is tested. The disadvantage of a full factorial screen is the large quantity of sample that is needed. When a limited amount of sample is available, a *random* screen can be used instead, in which a few parameters of the matrix are randomly chosen, or the *incomplete factorial* screen, in which the randomly chosen parameters are coordinated and distributed evenly in the matrix. Crystallization screens used in this work are listed in Table 2-7 and belong to the sparse matrix screens. Initial conditions identified by the sparse matrix screens were optimized by full factorial or incomplete factorial screens to obtain diffraction quality single crystals.

2.2.6 X-ray crystallography

2.2.6.1 Data collection

An x-ray beam that encounters a crystal lattice of atoms or molecules induces oscillation of the electrons in the lattice at the x-ray frequency. Upon returning to the unexcited state, the electrons emit radiation at the original energy and wavelength in a random direction. The scattered waves can be described as reflections at the lattice planes of the crystal. In most directions the scattered waves add up to zero (destructive interference). When the x-ray beam hits the lattice plane under the glancing angle θ and Braggs law (Formula 3) is satisfied, constructive interference takes place and reflections appear at the detector. The intensity and the position of the reflections are measured in the diffraction experiment. From the position of the reflections the Miller indices h, k, l can be determined (Blow 2003).

Formula 3. Bragg's law.

$$n \cdot \lambda = 2 \cdot d \cdot \sin\Theta$$

The electron density distribution ρ at every point in the unit cell (x,y,z) is calculated by the inverse Fourier transformation (Formula 4) over all structure factors F(hkl).

Formula 4. Fourier transform

$$\rho(x,y,z) = \frac{1}{V}\sum_{h}\sum_{k}\sum_{l} F(hkl)e^{[-2\pi \cdot i(hx+ky+lz)]}$$

The measured intensity (I(hkl)) of a reflection is proportional to the square of the structure factor amplitude |F(hkl)|. To calculate the structure factor F(hkl) from the intensity, the phase φ of the wave is needed (Formula 5). The phase information can not be determined during data collection and is lost. This fact is termed the *phase problem* of crystallography. Several methods exist to solve the phase problem. In this work molecular replacement was used (chapter 2.2.6.2).

Formula 5. Calculation of structure factors

$$I(hkl) = |F(hkl)|^2$$

$$F(hkl) = \left|F(hkl)e^{i\cdot\varphi(hkl)}\right|$$

For data collection, a single crystal was fished with a nylon loop (Hampton Research) out of the mother liquor, transferred into a cryo-protectant and frozen in liquid nitrogen. The cryo-protectant is used to prevent formation of ice crystals when the crystal is frozen and typically contains solvents like glycerol or polyethylene glycol together with the mother liquor of the protein crystal. As the KasA wild-type crystals grew in isopropanol, glycerol was chosen as a

cryo protectant. The crystals were directly transferred from their mother liquor to a cryo solution containing the crystallization solution and 30% glycerol. After one minute in the solution, the crystals were fished out with a cryo-loop and snap frozen in liquid nitrogen. The KasA C171Q crystals grew in PEG 3350, thus ethylene glycol was chosen as a cryo protectant. Experiments with different ethylene glycol concentrations showed that 25% ethylene glycol together with the crystallization solution was a successful cryo-solution. The crystals were directly transferred into the cryo-solution, fished out after 30 sec with a nylon loop and snap frozen in liquid nitrogen.

Crystals of the ternary InhA-NAD^+-PT70 complex grew in presence of 1% DMSO so that DMSO was used as a cryo protectant. The crystals were transferred to cryo solutions containing the crystallization solution and increasing amounts of DMSO in steps of 5% (v/v) to a final concentration of 25%. The crystals were soaked for 2 min in each well and subsequently frozen in liquid nitrogen.

For data collection, a cryo-loop containing one single crystal was mounted in the beam path and simultaneously cooled by a cryo stream of nitrogen at -180° C.

Diffraction data were collected *in-house* as well as at the synchrotron. The in-house x-ray source was a Rigaku MicroMax™-007HF generator with a rotating copper anode which produces X-rays with a fixed wavelength of 1.5418 Å. An image plate detector (Raxis HTC) was used to record the diffraction data. Image plate detectors store the energy of the X-ray photons that hit the detector until a laser beam reads the information and the plate is erased by visible light. The Raxis HTC detector contains three image plates which allows continuous data collection with one image plate being exposed while the other is read out and the third one is erased.

In a synchrotron electrons or positrons are accelerated and then injected into a storage ring where they circulate near the speed of light, guided by bending magnets. Synchrotron radiation is emitted when the charged particles go through the bending magnets, or through wigglers or undulators, which are additional magnetic devices inserted into the straight sections of the ring. The radiation is captured by beamlines and guided to experimental stations that contain instrumentation for diffraction experiments.

2.2.6.2 Molecular replacement

Molecular replacement is a method to solve the phase problem by using an already available structure with high homology to the unknown structure. This search model should have at least 30% sequence identity and has to be positioned correctly in the unit cell of the unknown

structure. This is done by calculating a Patterson function of the measured data of the unknown structure and of the structure of the search model. The advantage of the Patterson function is that no phase information is needed for calculation. In the rotation function the correct orientation of the search model is calculated by rotation of the self-Patterson vectors of the search model over the Patterson map of the unknown structure. The correct orientation of the search model is indicated by a maximum correlation of the intra-atomic vectors (*self-vectors*) of the two structures. The translation function is used to find the correct position of the oriented search model in the unit cell. Analogous to the rotation function, the translation function consists of superimposing the cross-Patterson vectors from the search model on the Patterson map of the unknown structure until a maximum of the inter-atomic vectors (*cross-vectors*) indicates the exact position of the search model. Finally, an initial electron density map can be calculated using the structure factor amplitudes of the measured data and the phases of the correctly positioned search model. As a six-dimensional search is computationally very intensive (three rotation variables and three translation variables), the six variables are determined in two steps with the rotation search first following the translation search.

With the known molecular weight of the protein (M_r), the volume of the unit cell (V_{EZ}) and the number of symmetry operators in the space group (z), the number of molecules (n) in the asymmetric unit can be estimated with the method of Matthews.

Formula 6. Calculation of Matthews coefficient

$$V_M = \frac{V_{EZ}}{M_r \cdot z \cdot n}$$

V_M is the packing parameter that describes crystal volume per protein mass. The values range between 1.8 $Å^3$/Da and 4.5 $Å^3$/Da with the maximum of the distribution at 2.4 $Å^3$/Da. With V_M the solvent content of the protein crystal can be calculated which is typically between 30% and 70% (Matthews 1968; Kantardjieff and Rupp 2003).

2.2.6.3 Model building and refinement

After structure solution, usually an incomplete model is obtained that contains many errors. In order to produce an accurate model, the structure is manually modified in real space by fitting the amino acids to the electron density and adding molecules like water, cofactors and inhibitors. For model building, the program Coot was used which is a molecular graphics application with the primary focus on crystallographic macromolecular model building and manipulation (Emsley and Cowtan 2004). During crystallographic refinement, the model

parameters like coordinates and B-factors are processed in the reciprocal space to improve the fit of observed and calculated structure-factor amplitudes. In order that the system is not under-determined, it is important that the number of parameters that define the structure do not exceed the number of observed reflections. This can especially be an issue when working with huge structures or at low resolution. Fortunately, it is possible to improve the data to parameter ratio by adding additional information to the measured intensities and the derived structure factor amplitudes. *Restraints* are conditions that apply within a certain standard deviation for the model. Information about the molecular geometry (bond lengths and bond angles), planarity of peptide bonds and phenyl groups and atom distances can be restrained. If more than one molecule is present in the asymmetric unit, they can be related by non-crystallographic symmetry (NCS) and the geometry can be restrained to be similar (NCS-restraints). Restraints increase the amount of data available for refinement. *Constraints* are conditions that apply exactly for the structure, reducing the number of independent parameters to be refined. The occupancy of one atom is usually constrained to be exactly 1.0. Displacement parameters can be constrained or the bond lengths and bond angles of hydrogen atoms.

Upon refinement, the phases become more accurate and an improved electron density can be calculated. The programs Refmac (Murshudov, Vagin et al. 1997) and Phenix.refine (Adams, Gopal et al. 2004) were used for maximum likelihood refinement or least squares twin refinement (see 2.2.6.4). Alternating rounds of model building and automated refinement were carried out to successively improve the atomic model. The success of refinement is evaluated by the R-factor that measures the agreement between the crystallographic model and the measured data. A decreasing R-factor indicates that the model has improved.

Formula 7. Calculation of the R-Factor

$$R = \frac{\sum_{hkl} \left\| F_{OBS} \right| - \left| F_{CALC} \right\|}{\sum_{hkl} \left| F_{OBS} \right|}$$

As the R-factor is related to data that is used during refinement, the obtained value might be biased towards the model. To prevent over-fitting of the data a more objective quality indicator, the free R-factor (R_{free}), was introduced (Brunger 1992). For calculation of the R_{free}, a fraction of 5-10% of the data are randomly chosen and not used for the refinement process. After each refinement cycle, the R_{free} is calculated from these data. If after refinement only the R-factor decreases and the R_{free} stays the same or even increases, the data might have been over fitted and the structure has to be carefully reanalyzed. In general, the R-factor is an

indicator of the precision of the model while the R_{free} displays more the accuracy of the model (Kleywegt and Jones 1997).

After interpretation of all features of the electron density map and when the R-factors indicate an appropriate value, refinement can be considered as complete. To evaluate the quality of the final model parameters like appropriate geometry, deviations of the bond lengths and angles from ideal values (Engh and Huber 1991) and unusual rotamers of amino acids are inspected. However, these parameters are often heavily restrained during refinement or closely monitored during rebuilding and are therefore not independent criteria of the model-quality. One independent criterion for model validation is the Ramachandran plot. In a Ramachandran plot, the dihedral angles φ and ψ of the amino acid residues in a protein structure are plotted against each other and shows the possible conformations of φ and ψ angles for a polypeptide. The diagram differentiates between the favoured regions, the additional allowed regions and the disallowed regions while exceptions apply only for proline and glycine (Ramachandran, Ramakrishnan et al. 1963). If an amino acid is found to lie in the disallowed region (outlier) there are two possibilities: either the conformation is indeed wrong or the residue actually has an unusual conformation. Amino acids with unusual conformations should have well defined electron density and are often found in important regions of the protein structure like the active site. In addition, the average coordinate error can be estimated, for example from a Luzzati plot. The MolProbity Server (Davis, Leaver-Fay et al. 2007) was used in this work to validate the final structures.

2.2.6.4 Merohedral twinning

Twinning is a severe crystal growth disorder, in which two distinct domains are related by a symmetry operator (*twin law*) that is part of the lattice symmetry but not part of the crystal point group.

Twinning in macromolecular crystals can be divided in two different categories: epitaxial twinning, where the twinning can be easily detected from the diffraction pattern, as the lattices are interpenetrating. Merohedral twinning can not be recognized from the diffraction pattern, the lattices from the distinct domains superimpose exactly in three dimensions so that each measured intensity results from the contribution of two separate twin-related reflections. The relative size of the smallest domain to the whole crystal is known as the *twin fraction* α. In case of merohedral twinning, the twin fraction can be nearly equal to 0.5 (perfect merohedral twinning) or it can be below 0.5 (partial merohedral twinning). A successful

structure determination is usually prevented in presence of twinning unless it is detected and either avoided or corrected (Yeates and Fam 1999).

The presence of merohedral twinning can be identified by carefully inspecting the intensity statistics in the output of the program Truncate (Padilla and Yeates 2003). As each reflection consists of two twin-related reflections there are less very weak and very strong reflections. This can be observed from the intensity distribution that differs from normal Wilson statistics. Furthermore, the cumulative intensity plot is distorted by twinning which can be seen as a sigmoidal curve. Finally, the theoretical values of the moments of E (normalized structure amplitude) are higher compared to untwinned data (between the theoretical value of untwinned data and perfectly twinned data). If twinning is detected, it is always recommended to avoid it by finding an untwinned crystal. However, it is still possible to solve a merohedrally twinned structure by molecular replacement and sometimes by experimental phasing if the twin fraction is not close to 0.5. In case of perfect merohedral twinning, the structure can only be solved by molecular replacement if the correct space group is known which can be difficult to identify as the two overlapping lattices simulate a higher symmetry than the true space group symmetry.

The datasets of both KasA C171Q structures were found to be twinned. Although the data were of good quality and the structures could be solved by molecular replacement, the R_{free} got stuck at high values (32% for the apo structure and 25% for the TLM-bound structure), during refinement with Refmac. Analyses of the data with phenix.xtriage revealed that the KasA C171Q TLM crystal was partially merohedrally twinned with a twin fraction of 0.29 and the KasA C171Q apo crystal had an even higher twin fraction of 0.46, which is almost a perfect twin. The twin law was determined by phenix.xtriage to be h,-h-k,-l for both datasets. If the twin fraction is estimated accurately and the twin law is known, phenix.refine can handle the refinement of merohedrally twinned data by least square twin refinement. Algebraic detwinning is performed for twin fractions below 40% and detwinning with proportionality rules for fractions above 40%.

3 Results

3.1 Purification of KasA

Both, wild-type KasA and the C171Q KasA variant were purified by the same steps. The results did not differ between the two variants, therefore only the results of wild-type KasA are described and apply accordingly for the KasA C171Q mutant.

3.1.1 Thermofluor buffer Screen

KasA was expressed in *M. smegmatis* to overcome solubility issues. Expression yielded soluble protein, but initially during the first purification trials, it turned out that under the chosen conditions (20 mM Tris-HCl pH 7.4, 200 mM NaCl) KasA was highly instable and two chaperones were bound to the protein. Towards the crystallization of KasA the chaperones had to be removed, however, removal of the chaperones resulted in aggregation of KasA. A Thermofluor buffer screen was performed to identify a buffer condition suitable to stabilize KasA during the purification process and for subsequent crystallization trials. In absence of a pure soluble protein solution, the aggregated protein was used for this analysis.

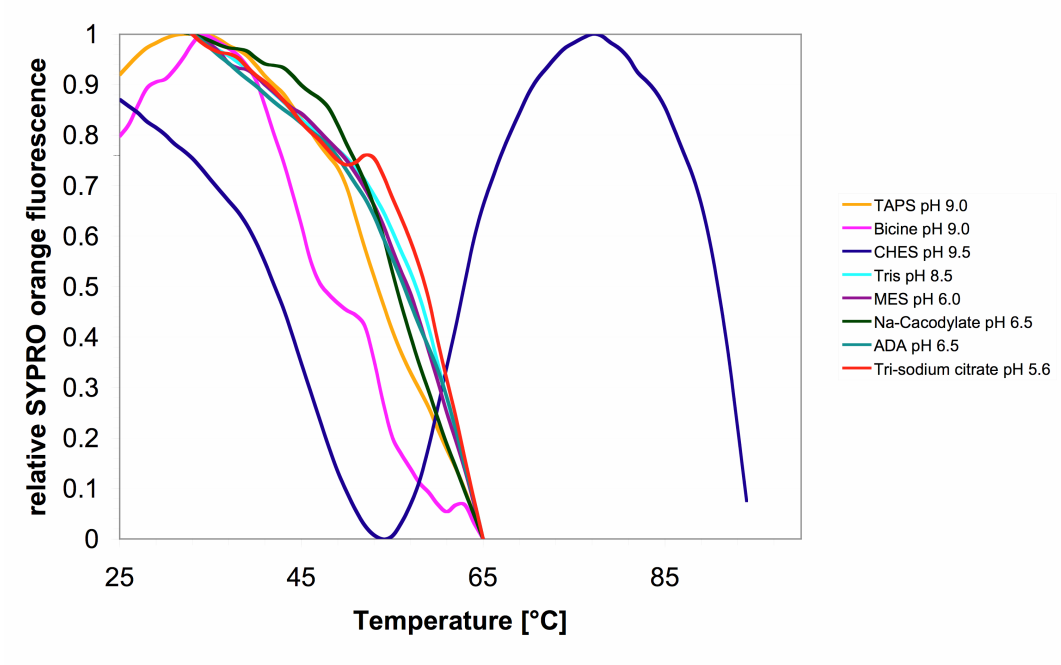


Figure 3-1. Normalized melting curves of KasA in the presence of different buffers. KasA in the presence of CHES pH 9.5 (blue) shows a typical melting curve with Tm = 62° C while the other curves indicate protein precipitation.

Each well of a 96-well plate was filled with 5 µL protein solution (10 µM), 7.5 µL Sypro Orange (50x) and 12.5 µL of the relevant buffer solution (100 mM). The solutions were gently mixed by pipetting and heated in the real time PCR cycler from 25° C to 95° C.
The protein showed a typical melting curve with only one of the tested buffers, namely CHES (N-Cyclohexyl-2-aminoethanesulfonic acid) at pH 9.5 with a Tm of 62° C. All other curves indicated precipitation of the protein which is in accordance with the fact that aggregated protein was used for the thermofluor experiments. All the more surprising is the ability of CHES buffer to solubilize KasA from the aggregated state. The following purification protocol further proved this ability. CHES buffer pH 9.5 was used throughout the purification and KasA remained stable and did not aggregate when the chaperones where removed.

3.1.2 Chromatographic purification

KasA was purified by a three step chromatography using affinity-, anion exchange- and size exclusion chromatography. Cell pellets were thawed, resuspended in binding buffer and lysed by sonication. The lysate was centrifuged (1 h, 50000g, 4° C) to separate cell fragments from the supernatant.

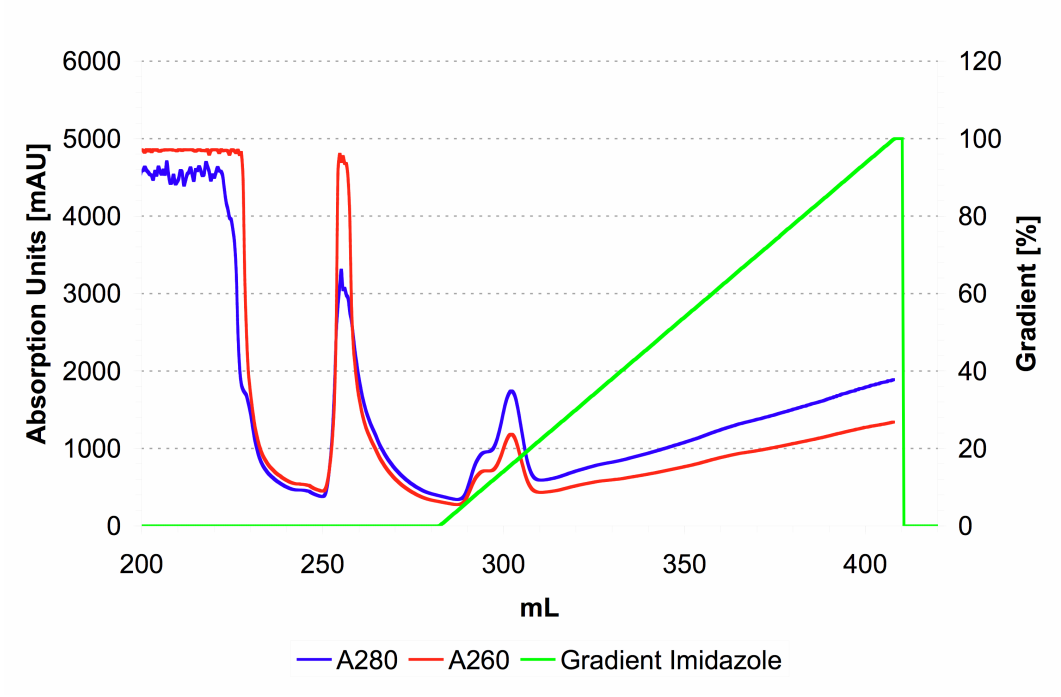


Figure 3-2. Purification of KasA by Nickel Affinity Chromatography. Blue: absorption at 280nm, red: absorption at 260nm, green: percentage of elution buffer, gradient. The peak in the first third of the gradient (green line) shows the characteristic shoulder of the KasA elution profile.

The supernatant was applied to a HisTrap HP column (GE Healthcare) using the sample pump P960 of the Äkta purifier system. After a washing step with buffer containing 40 mM imidazole, a gradient of 40 mM - 1 M imidazole was used for elution. KasA was eluted from

the column in the first third of the gradient at about 200 mM imidazole (16% elution buffer) with a shoulder in front of the main peak (Figure 3-2).

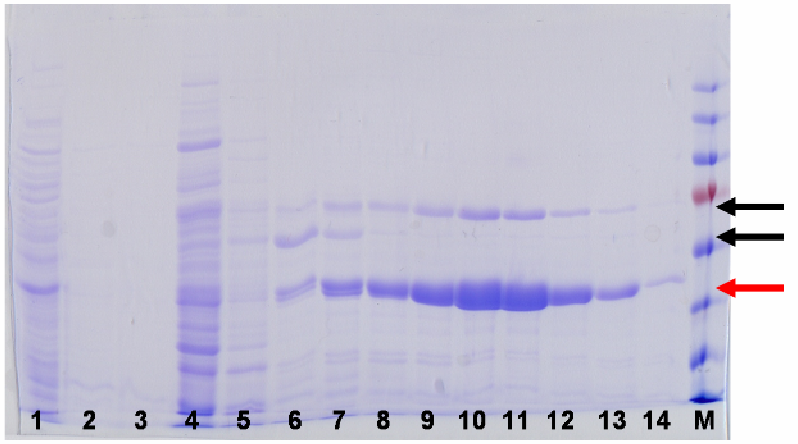


Figure 3-3. 10% SDS Gel showing the fractions of the affinity chromatography. Lanes 1-3 are the flow-through fractions followed by the washing step fractions (4-5), while lane 6-14 belong to the elution fractions. M indicates the protein ladder. The red arrow points to the position of the KasA protein while the two black arrows indicate the positions of the two chaperones that were co-purified with KasA.

Samples of different peak fractions were loaded on a 10% SDS gel which shows that the shoulder (lanes 6-8) contains KasA with two chaperones and the main peak (lanes 9-14) the KasA protein with only one chaperone (Figure 3-3).

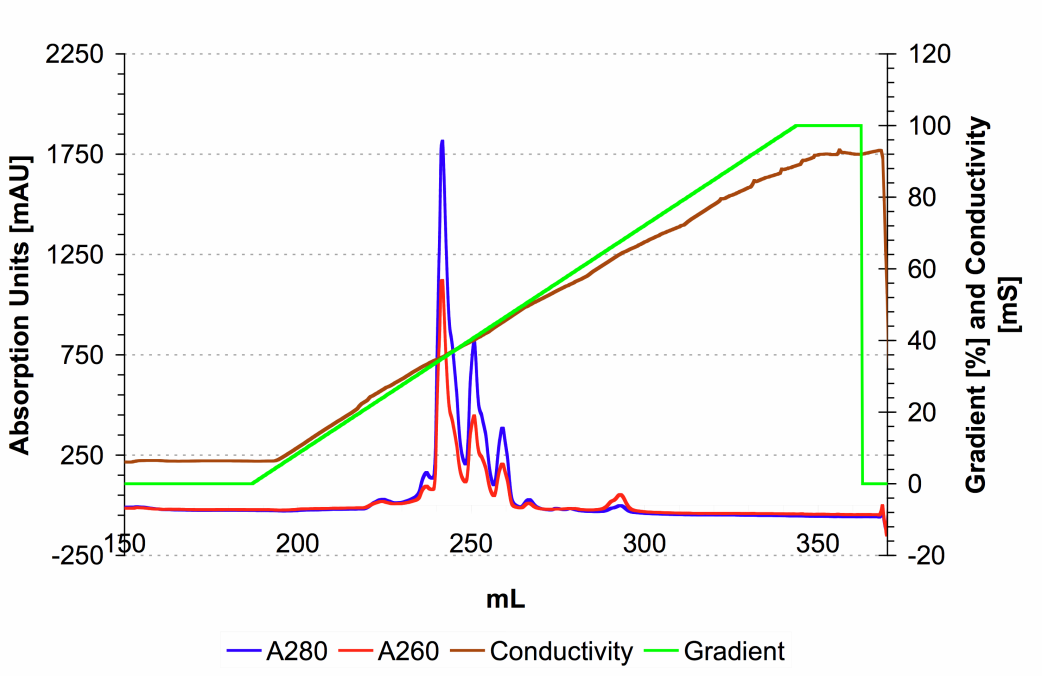


Figure 3-4. Purification of KasA by Anion Exchange Chromatography. Blue: absorption at 280 nm, red: absorption at 260 nm, brown: conductivity (mS), green: percentage of elution buffer, gradient. The elution profile shows three major peaks. The first and largest peak contained KasA without chaperones.

Fractions of the main peak were pooled and diluted 1:10 with dilution buffer containing 20 mM CHES pH 9.5 to reduce the NaCl concentration of the protein solution to 50 mM. The diluted protein solution was applied to a MonoQ 10/100 anion exchange column (GE Healthcare) and eluted with a gradient of 50 mM to 1 M NaCl (Figure 3-4). Three peaks eluting closely together were observed, at 420 mM NaCl (37% elution buffer), 460mM NaCl (41% elution buffer) and 510 mM NaCl (46% elution buffer) concentration.

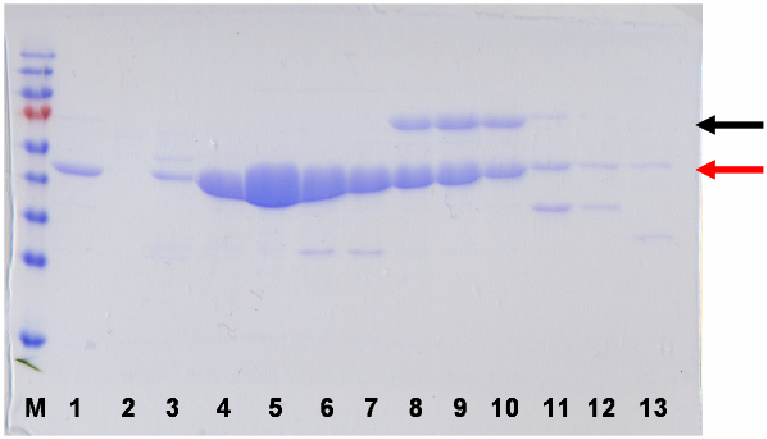


Figure 3-5. 12% SDS Gel showing the elution fractions of the anion exchange chromatography. M indicates the protein ladder, lane 1 is the diluted sample after Ni-NTA chromatography, lane 2 is the flow-through fraction, lanes 3-13 correspond to the elution fractions: peak one lanes 4-7, peak two lanes 8-10 and peak three 11-13. The red arrow points to the position of KasA while the black arrow depicts the position of one of the chaperones.

Samples of the elution fractions were examined on a 12% SDS gel (Figure 3-5), which shows that KasA without chaperones eluted in the first peak (lanes 4-7), while KasA with one chaperone still present eluted in the second peak (lane 8-10). The content of the third peak is shown in lanes 11-13.

The fractions of the first peak were pooled and applied to a Superdex 26/60 200 p.g. size exclusion column (GE Healthcare). The elution profile (Figure 3-6) shows one major peak at an elution volume (205 mL) which corresponds to a KasA dimer.

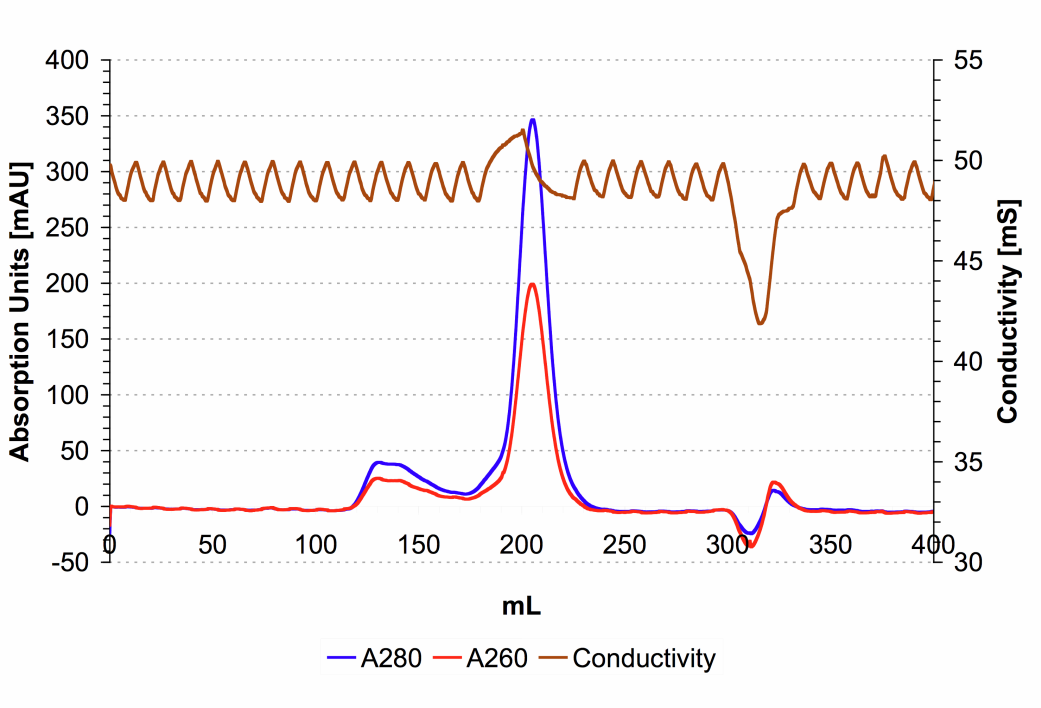


Figure 3-6. Purification of KasA by Size Exclusion Chromatography. Blue: absorption at 280 nm, red: absorption at 260 nm, brown: conductivity (mS). The elution volume of the main peak (205 mL) corresponds to a KasA dimer.

The 12% SDS gel in Figure 3-7 shows that the broad non-symmetric peak at approximately 135 mL also contains KasA and was considered as aggregated protein (lanes 2 and 3). The main peak contains the pure KasA (lanes 4-10). Lanes 4 and 5 correspond to a tiny shoulder that contains additional protein with lower molecular weight, as can be seen from the SDS gel. These fractions were discarded and fractions corresponding to lanes 6-10 were pooled.

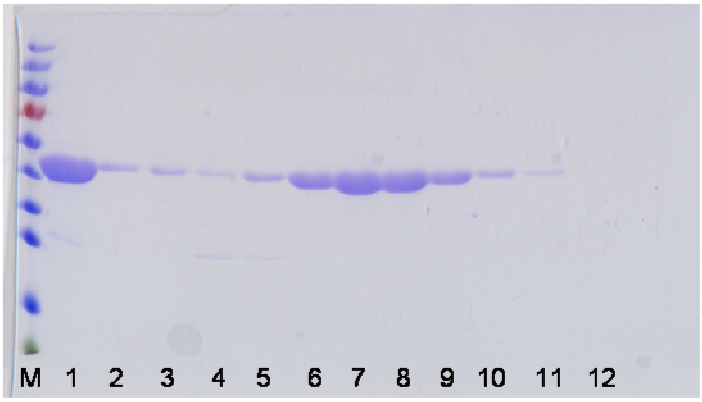


Figure 3-7. 12% SDS Gel showing the pure protein after the size exclusion column. Lane 1 contains the pooled sample after the anion exchange column, lanes 2 and 3 the aggregated KasA, lanes 4 and 5 the content of the tiny shoulder of the main peak and lanes 6-10 the fractions of the main peak.

Wild-type KasA and the C171Q KasA mutant were concentrated with a centrifugal filter device to final concentrations of 4.3 mg/mL (100 µM) and 8.5 mg/mL (200 µM), respectively and aliquoted in 0.5 mL Eppendorf tubes. Aliquots containing 31 µL of the concentrated protein solution were snap-frozen in liquid nitrogen and stored at −80° C.

The homogeneity of the protein was checked with DLS after the purification procedure before freezing and after thawing a frozen sample. DLS experiments were carried out at 20° C with the KasA wild-type protein and indicated that the protein is a dimer is solution with an apparent molecular mass of around 95 kDa (the dimer has a calculated molecular weight of 86.6 kDa). The solution has a low polydispersity below 14% and freezing does not harm the homogeneity of the protein.

3.2 Crystallization of KasA

3.2.1 Crystallization of wild-type KasA

KasA was crystallized by hanging drop vapor diffusion, mixing equal volumes of protein solution and crystallization solution. The crystallization solution that yielded diffraction quality crystals contained 10% (v/v) isopropanol, 0.2 M NaCl and 0.1 M Hepes pH 7.5. First structural analyses revealed that the active site cysteine was oxidized (see chapter 3.3.1.). To prevent oxidation in subsequent setups the reducing agent TCEP (*tris*(2-carboxyethyl) phosphine) was included in the crystallization condition. With 10 mM TCEP the crystals still grew applying the above mentioned crystallization conditions and the resulting crystal structure showed a reduced cysteine. The crystals appeared after one night and grew to their full size within two days to an approximate size of 200 x 100 x 80 μm^3. For a reproducible result it was necessary to trap an air bubble in the hanging drop, where the crystals grew around.

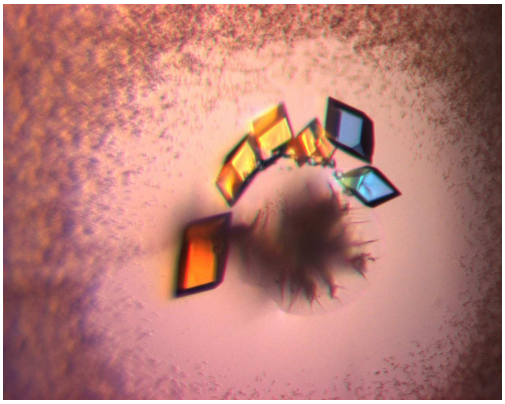


Figure 3-8 Crystals of KasA growing around a bubble in the drop.

For the formation of the wild-type KasA inhibitor complex, thiolactomycin (TLM, Sigma) was dissolved in isopropanol and added to the concentrated protein solution in a 200 fold molar excess. After 2 h incubation on ice, the KasA-TLM solution was centrifuged (25000g, 4° C, 20 min) to remove aggregates and undissolved TLM chunks. The KasA-TLM complex could be crystallized utilizing the same conditions as for the apo protein.

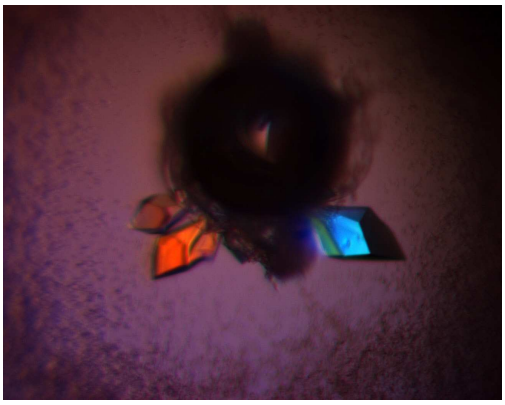


Figure 3-9 Crystals of the binary KasA-TLM complex

Before cryo-cooling in liquid nitrogen, the KasA wild-type crystals were transferred for two minutes into a cryo-protectant solution that contained the respective precipitant and 30% glycerol. The cryosolution of the KasA-TLM co-crystals additionally contained an excess of TLM and the crystals were soaked for up to 15 min in the solution before freezing.

3.2.2 Crystallization of the C171Q KasA variant

For the formation of the C171Q KasA inhibitor complexes, thiolactomycin (TLM, Sigma) was dissolved in isopropanol and added to the concentrated protein solution in a 10 fold molar excess. The crystallization solution of the C171Q KasA crystals contained 20% PEG 3350 and 0.2 M potassium formate. Diffraction quality crystals grew within four days to their final size of 60 x 50 x 50 μm^3 for the C171Q KasA apo crystals and 500 x 100 x 100 μm^3 for the C171Q KasA TLM bound crystals (compare Figure 3-10 and Figure 3-11).

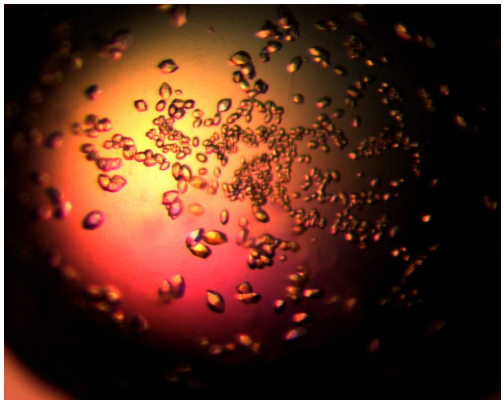


Figure 3-10 Crystals of the KasA C171Q variant grown in PEG

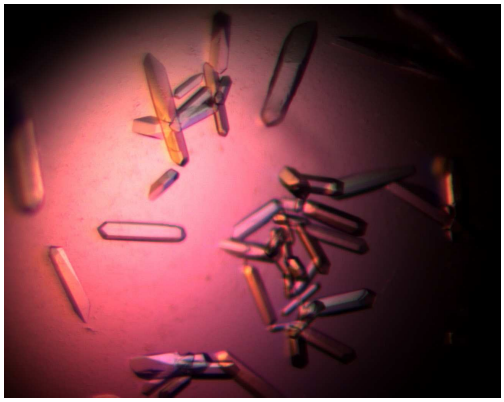


Figure 3-11 Crystals of the binary KasA C171Q-TLM complex grown in PEG have a different shape compared to the KasA C171Q apo crystals

Prior to cryo-cooling in liquid nitrogen, the crystals were transferred for a few seconds into a cryoprotectant solution that contained the respective precipitant and 25% ethylene glycol.

3.3 Structures of KasA

3.3.1 Structure solution and refinement of wild-type KasA

Diffraction data of apo wild-type KasA were collected at a Rigaku MicroMax™-007HF generator with an Raxis HTC Detector. Datasets were processed using D*Trek implemented in the CrystalClear software (Rigaku). The data set of the binary wild-type KasA TLM complex was collected at beam line ID14 of the European Synchrotron Radiation Facility, Grenoble, and was indexed and integrated using Mosflm and scaled with Scala. Wild-type KasA crystals belong to space group $P3_121$ and contain one monomer in the asymmetric unit. The probability of a certain number of molecules in the asymmetric unit was determined with the Matthews coefficient. The most plausible value of 2.9 $Å^3$/Da was obtained for one monomer in the asymmetric unit with a respective solvent content of 58.2% (KasA wild-type apo structure). The most plausible value of 3.0 $Å^3$/Da for the KasA wild-type TLM bound structure was obtained for one monomer in the asymmetric unit which reflects a solvent content of 59.2%.

All structures were solved by molecular replacement using Phaser. As a search model for the wild-type KasA apo structure, one monomer of the *M. tuberculosis* KasB structure (pdb code 2gp6 (Sridharan, Wang et al. 2007)) was used. For all subsequent structures, the coordinates of the refined wild-type KasA apo structure were used as a search model. Data collection statistics are given in Table 3-1.

Table 3-1 Data collection and refinement statistics for the KasA wild-type structures.

	KasA wild-type apo	KasA wild-type TLM
Data collection		
Space group	$P3_121$	$P3_121$
Cell dimensions		
a, *b*, *c* (Å)	77.53, 77.53, 146.80	77.78, 77.78, 149.52
α, β, γ (°)	90, 90, 120	90, 90, 120
Resolution (Å)	39.55 – 2.01 (2.08 – 2.01)*	50.06 – 1.80 (1.90 – 1.80)
R_{merge}	0.079 (0.338)	0.057 (0.315)
$I / \sigma I$	9.0 (3.0)	19.1 (7.5)
Completeness (%)	99.7 (100)	99.0 (93.2)
Redundancy	4.4 (4.3)	6.7 (6.5)
Refinement		
Resolution (Å)	2.01	1.80
No. reflections	32,867	46,350
R_{work} / R_{free} (%)	18.0 / 21.4	16.7 / 20.1
No. atoms		
Protein	3,065	3,096
TLM	-	14
PEG	-	-
Water	181	281
B-factors		
Protein	40.4	21.1
TLM	-	26.0
PEG	-	-
Water	43.7	30.7
Rms deviations		
Bond lengths (Å)	0.012	0.013
Bond angles (°)	1.296	1.421

*values in parenthesis refer to the highest resolution shell

Model building and refinement of the wild-type KasA structures was carried out using alternating rounds of Coot for manual model building and Refmac5 for maximum likelihood refinement. The TLM molecule and water molecules were built into the electron density using Coot.

At this point it became clear that the active site cysteine (Cys171) of the first structures of KasA was oxidized. The Fo-Fc difference density clearly showed an extension at the side chain (Figure 3-12, panel A) and a sulfinic acid fitted the density very well (Figure 3-12, panel B). The oxidation of the cysteine was considered an artifact that was produced during

purification or crystallization. Therefore a reducing agent was included into the crystallization setups and subsequently a structure with a reduced cysteine was obtained as shown in Figure 3-12, panel C.

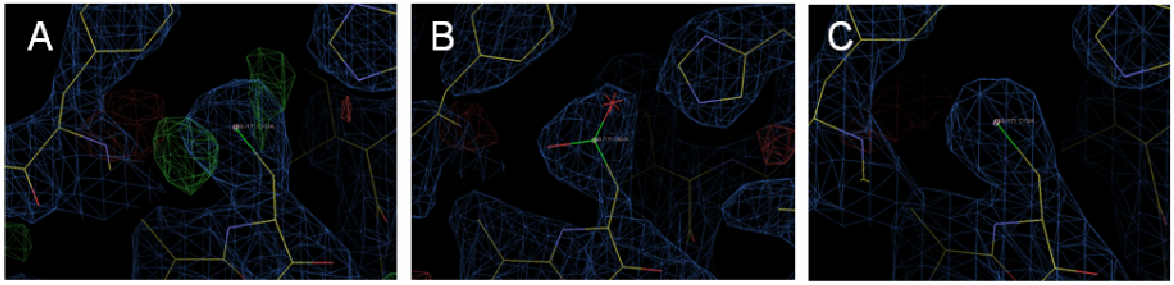


Figure 3-12 Oxidized cysteine 171 in the KasA structures. (A) Fo-Fc difference density (green) at +3 σ. (B) Sulfinic acid fits the density. (C) Reduced cysteine after crystallization with TCEP.

3.3.2 Validation of wild-type KasA structures

Both KasA wild-type structures contain one monomer in the asymmetric unit that forms a dimer with a symmetry related molecule. The homodimeric assembly of the protein in the crystal structures is consistent with the knowledge that KasA forms a homodimer in solution.

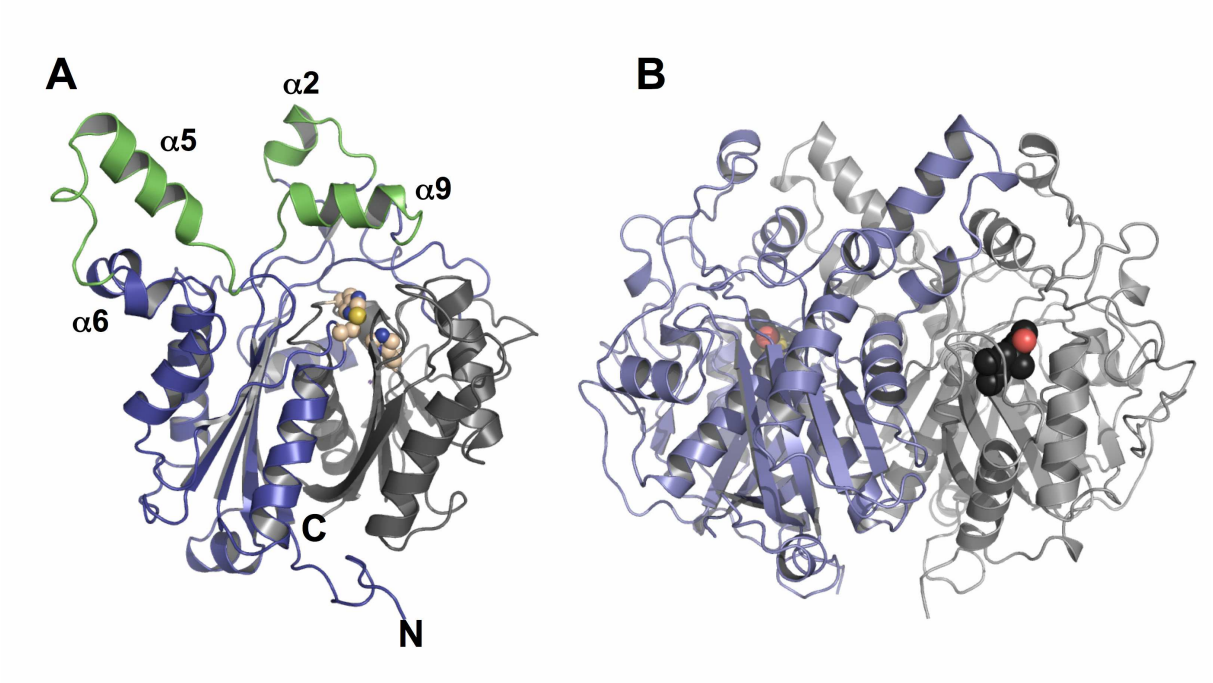


Figure 3-13. Overall structure of wild-type KasA. (A) One monomer of KasA is shown and contains the characteristic thiolase fold. Gray: C-terminal core domain, blue: N-terminal core domain, green: cap. The residues of the catalytic triad (Cys171, His311, His345) are shown in space filling representation in light brown. (B) Dimer of KasA showing the thiolactomycin molecule (black) bound to the active site.

Each monomer is comprised of a core domain and a cap. The core domain can be divided into two halves with similar topology, a mixed five-stranded β-sheet covered on each face by

α–helices. Residues 2-259 form the N-terminal half and residues 260-416 the C-terminal half of the protein. The two core domains form a five-layered αβαβα structure, characteristic of the thiolase superfamily. Residues of the catalytic triad (Cys171, His311 and His345) are located in the core domain. While the catalytically essential cysteine that becomes covalently modified during the reaction is located in the N-terminal domain and lies at the N-terminus of an α-helix, all other catalytic residues are contained within the C-terminal core domain. The more flexible cap consists of helices α2, α5 and α9 which form together with α5' of the second monomer the acyl-binding channel. The secondary structure elements were assigned to the amino acid sequence with Stride (Figure 3-14).

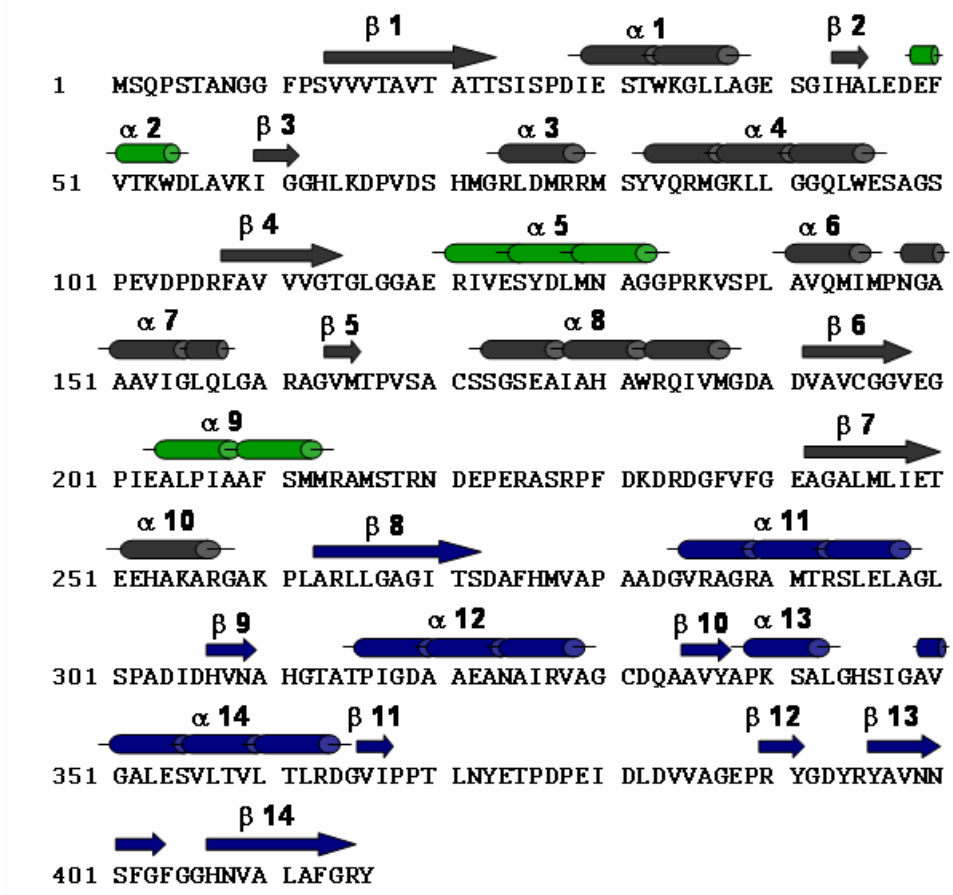


Figure 3-14. Secondary structure elements were assigned by stride for the apo structure of wild-type KasA. Arrows show β-strands, tubes stand for α-helices. The colours are according to the domains described in the text and coloured like in Figure 3-14A. Black: N-terminal core domain, blue: C-terminal core domain and green: helices forming the cap.

The quality of the structures was analyzed by the MolProbity Server. The KasA wild-type apo structure was solved at a resolution of 2.01 Å and a total of 415 amino acid residues are present of which 6 residues have alternate conformations. The structure was refined to an R-factor of 18.0% and an R_{free} of 21.4%. Two residues (Leu128 and Leu140) display an

unusual rotamer. However, these residues are part of a very flexible loop with weak electron density due to high B-factors. The analysis of the φ-ψ dihedral angles in the Ramachandran plot shows that 97.4% of all residues are in the favored region, 99.8% in the allowed region and one residue is an outlier (Figure 3-15). The outlier is Ile347, a residue close to the active site residue His345. The electron density of Ile347 is well defined so that it is conceivable that this residue has to be in this unusual conformation to properly form the active site.

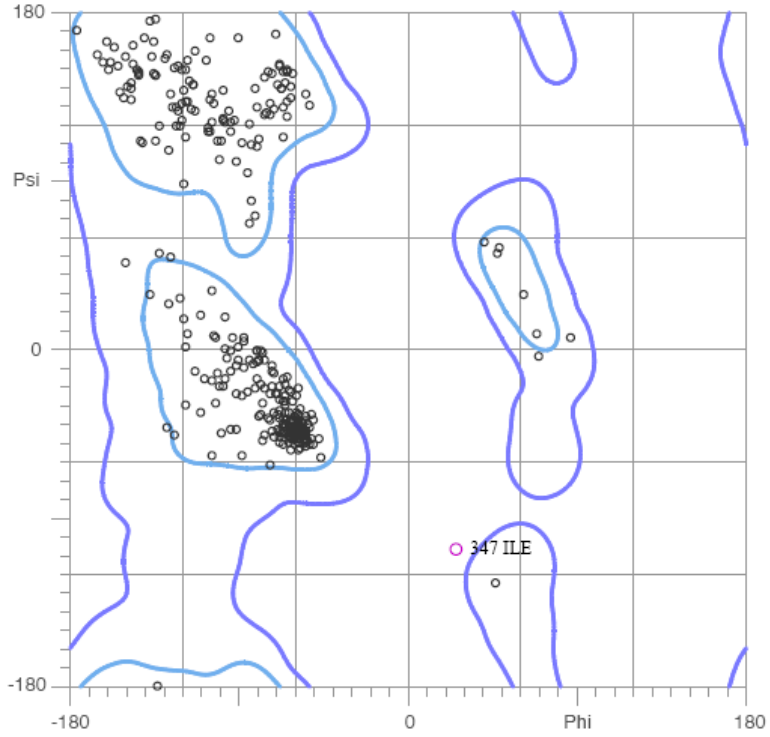


Figure 3-15. Ramachandran plot of the KasA wild-type apo structure. Light blue areas delineate favoured regions, dark blue areas are additional allowed areas and pink circles indicate outliers.

The wild-type KasA TLM bound structure was solved at a resolution of 1.80 Å. The asymmetric unit contains one protein chain with a total of 415 amino acid residues and one thiolactomycin molecule. 13 amino acids display alternate conformations. The structure was refined to an R-factor of 16.7% and an R_{free} of 20.1%. The Ramachandran plot shows that 97.9% of the residues are in the favored region and 100% are in the allowed region. Four residues are rotamer outliers, of which three (Leu116, Ile122, Leu128) are part of the flexible region and one is the active site His345. As thiolactomycin is bound to the active site and forms hydrogen bonds to the catalytic histidines, the unusual rotamer of His345 may result from the interaction with the inhibitor.

The rms deviation between the two structures using LSQ superpose utilizing all main chain atoms is 0.34 Å. This low value suggests that no major conformational changes occur upon TLM binding.

To further evaluate the quality of the structures, the B-factors were analyzed. The mean B-factor of the KasA wild-type apo and TLM bound structure is 40.4 Å^2 and 21.1 Å^2, respectively (Table 3-1). A graph of the B-factors of the main chain atoms plotted against the residue number shows several small peaks that indicate flexible loops at the protein surface and one prominent peak that is present in both structures points to a flexible region (residues 115 to 147) in the structure (Figure 3-16, left panel). The average B-factor of this region is much higher than the average B-factor of the entire structure (60 Å^2 for residues 115 to 147 of the KasA wild-type apo structure and 40 Å^2 for the same region of the KasA wild-type TLM bound structure). The B-factor distribution is further illustrated in the right panel of Figure 3-16 in which the main chain of one monomer is color coded according to the B-factors. The core of the protein displays relatively low B-factors, while some surface exposed loops have higher B-factors. The highest B-factors can be found in the flexible region (residues 115-147) that is involved in dimer formation and forms part of the acyl binding channel. The putative function of the flexible region is discussed in chapter 4.3.

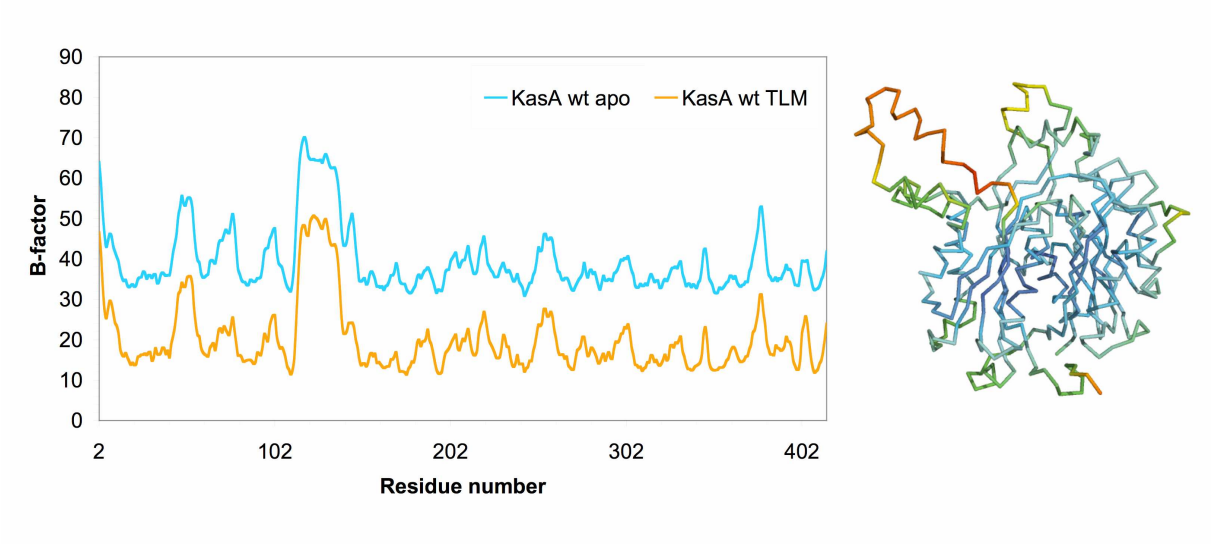


Figure 3-16. Illustration of the B-factors of the KasA wild-type structures. (Left) Plot of the B-factors of the main chain atoms against the residues of KasA wild-type (blue) and TLM bound KasA wild-type (orange). The major peak delineates the flexible region (residues 115-147). (Right) monomer of KasA wild type colored according to its B-factors. The color scale is shown in ten steps from blue (low B-factor) to green to red (high B-factor).

3.3.3 Structure solution and refinement of C171Q KasA

Diffraction data of the C171Q mutant were collected on a Rigaku MicroMax™-007HF generator with an Raxis HTC Detector and processed using D*Trek implemented in the CrystalClear software (Rigaku). The dataset of the binary C171Q KasA TLM complex was collected at the protein structure factory beamline BL14.2 at BESSYII, Berlin, indexed and

integrated using Mosflm (Leslie 1992) and scaled with Scala (CCP4 1994). The C171Q KasA variant crystallized in space group $P3_1$ with eight monomers in the asymmetric unit. The most plausible Matthews coefficient of 2.8 Å3/Da was determined for 8 monomers in the asymmetric unit corresponding to a solvent content of 56.5% for the C171Q KasA apo structure. Eight monomers in the asymmetric unit of the C171Q KasA TLM bound structures resulted also in the most plausible Matthews coefficient of 2.8 Å3/Da with 56.3% solvent content.

Table 3-2. Data collection and refinement statistics for the KasA C171Q variants.

	KasAC171Q apo	KasAC171Q TLM
Data collection		
Space group	$P3_1$	$P3_1$
Cell dimensions		
a, *b*, *c* (Å)	151.50, 151.50, 147.83	151.27, 151.27, 147.95
α, β, γ (°)	90, 90, 120	90, 90, 120
Resolution (Å)	33.71 – 2.15 (2.27 – 2.15)*	46.18 – 2.00 (2.11 – 2.00)
R_{merge}	0.080 (0.479)	0.069 (0.348)
$I / \sigma I$	9.6 (1.6)	14.4 (3.8)
Completeness (%)	97.9 (96.2)	100 (100)
Redundancy	1.7 (1.6)	3.9 (3.8)
Refinement		
Resolution (Å)	2.15	2.00
No. reflections	280,329	511,630
R_{work} / R_{free} (%)	16.5 / 21.2	13.6 / 17.8
No. atoms		
Protein	24,274	24,349
TLM	-	112
PEG	98	273
Water	35	870
B-factors		
Protein	55.0	29.1
TLM	-	29.8
PEG	57.7	45.4
Water	47.5	29.9
Rms deviations		
Bond lengths (Å)	0.005	0.011
Bond angles (°)	0.917	1.532
Twin fraction	0.459	0.287

*values in parenthesis refer to the highest resolution shell

The structures were solved by molecular replacement using Phaser (McCoy, Grosse-Kunstleve et al. 2007) with the KasA wild-type apo structure as a search model. Data collection statistics are given in Table 3-2.

3.3.4 Merohedral twinning

Due to problems during refinement of the C171Q KasA structures with Refmac, a twinning test with Phenix.xtriage (Zwart, Grosse-Kunstleve et al. 2008) was performed that indicated that the C171Q KasA TLM dataset was merohedrally twinned to 29% and the C171Q KasA dataset to 46%. Inspection of the cumulative intensity distribution of the truncate output confirmed the twinning. The sigmoid curve of the observed acentrical reflections in Figure 3-17 clearly indicates merohedral twinning.

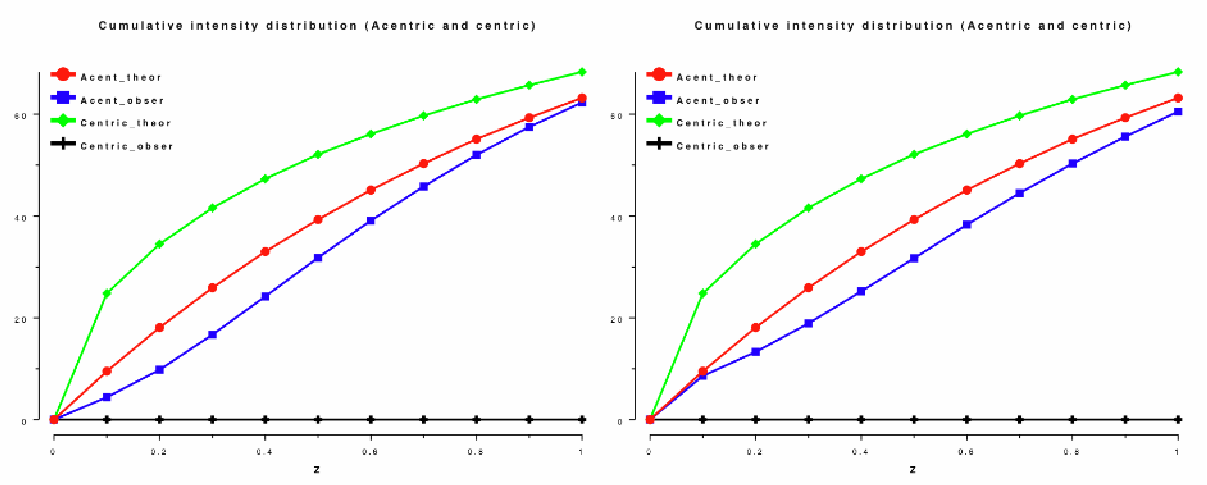


Figure 3-17. Cumulative intensity plots of KasA C171Q TLM (left) and KasA C171Q apo (right). Red, theoretical distribution of acentric reflections; blue, observed distribution of acentric reflections; green, theoretical distribution of centric reflections; black, observed distribution of centric reflections. The sigmoidal blue curve indicates merohedral twinning.

Further refinement was carried out using Phenix.refine, taking the twin law (h,-h –k,-l) into account and refining the twin fraction. Upon twin refinement, the density improved and the R_{free} decreased significantly. The improved electron density permitted further model building in Coot and allowed the unambiguous modeling of TLM and water molecules into the electron density maps. In addition, several polyethylene glycol (PEG) chains originating from the crystallization condition could be identified in the density and were built into the electron density using Coot. To avoid bias of the R_{free} due to the presence of 8-fold NCS and twinning, five percent of all reflections that were omitted during refinement to calculate the R_{free} were selected in 20 thin resolution shells.

3.3.5 Validation of the KasA C171Q variant structures

The KasA C171Q structures contain each eight monomers in the asymmetric unit that form four dimers. The overall fold of the protein is the same as already described for the KasA wild-type structures (see chapter 3.3.2). Additionally, polyethylene glycol (PEG) chains that originate from the crystallization condition are bound to every monomer. The PEG chains are bound to the acyl channels of the structure and are all oriented in the same way so that it can be assumed that the PEG chains mimic a long chain fatty acid. Thus the binding mode of a natural substrate can be identified.

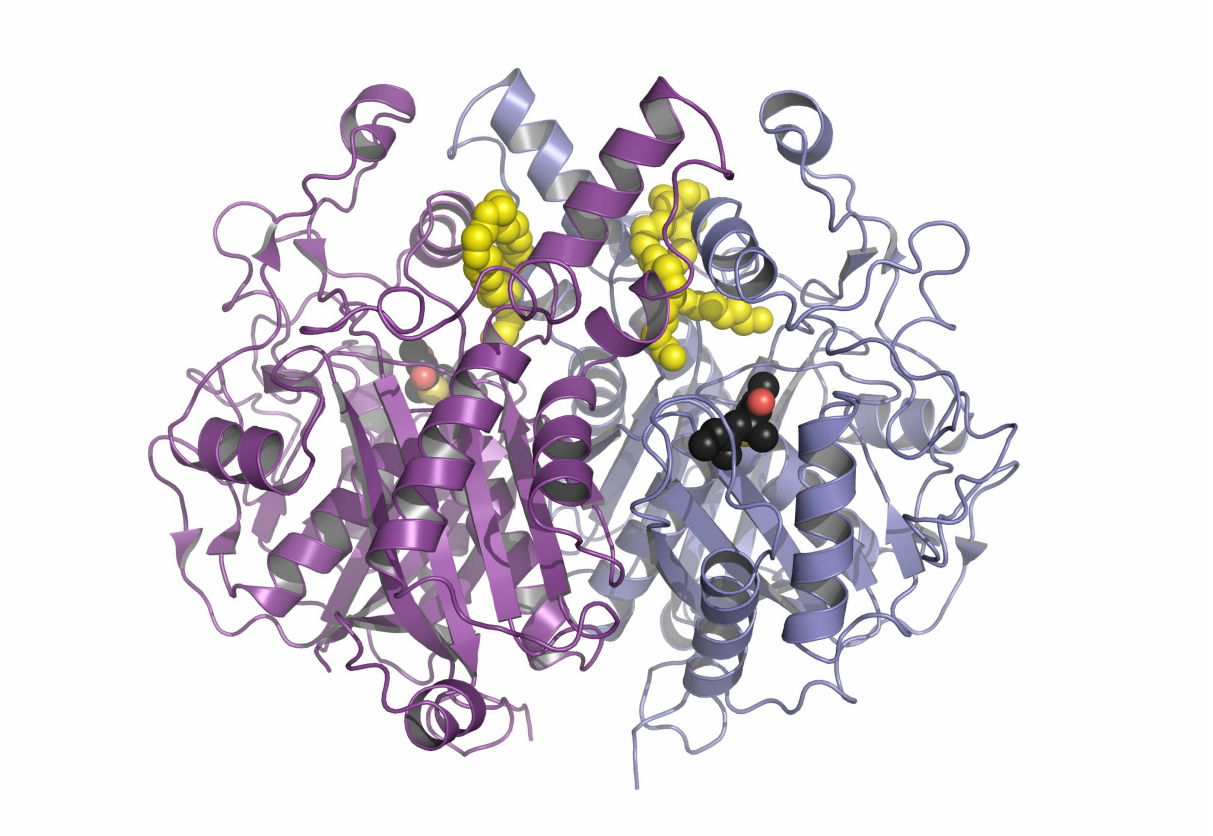


Figure 3-18. Dimer of the KasA C171Q TLM structure. The bound TLM molecules are shown in black space filling representation. The polyethylene glycol chains that are bound to the acyl channel are depicted in yellow space filling representation.

The C171Q KasA apo structure was solved at 2.15 Å resolution, contains 8 polypeptide chains in the asymmetric unit and a total of 3320 amino acid residues of which 3 protein residues have alternate conformations. The structure was refined to an R-factor of 16.6% and an R_{free} of 21.1%. The Ramachandran diagram shows that 95.7% of the residues are in the favored region, 99.6% in the allowed region and 13 residues are outliers. A few outliers are part of flexible loops on the protein surface (Asp383 chain E, Asp55 chain F and Asp381 chain F) while the majority of outliers are glycines that are part of the active site (Gly312 of chains B-H, Ala 314 chain F) or participate in the formation of the substrate binding channel

(Gly240 of chain A and F). As the C171Q variant of KasA is a mimic of the acylated form of KasA, it might well be that the unusual conformation of the above mentioned residues results from the changes in the active site and the presence of the PEG chains in the acyl channel.

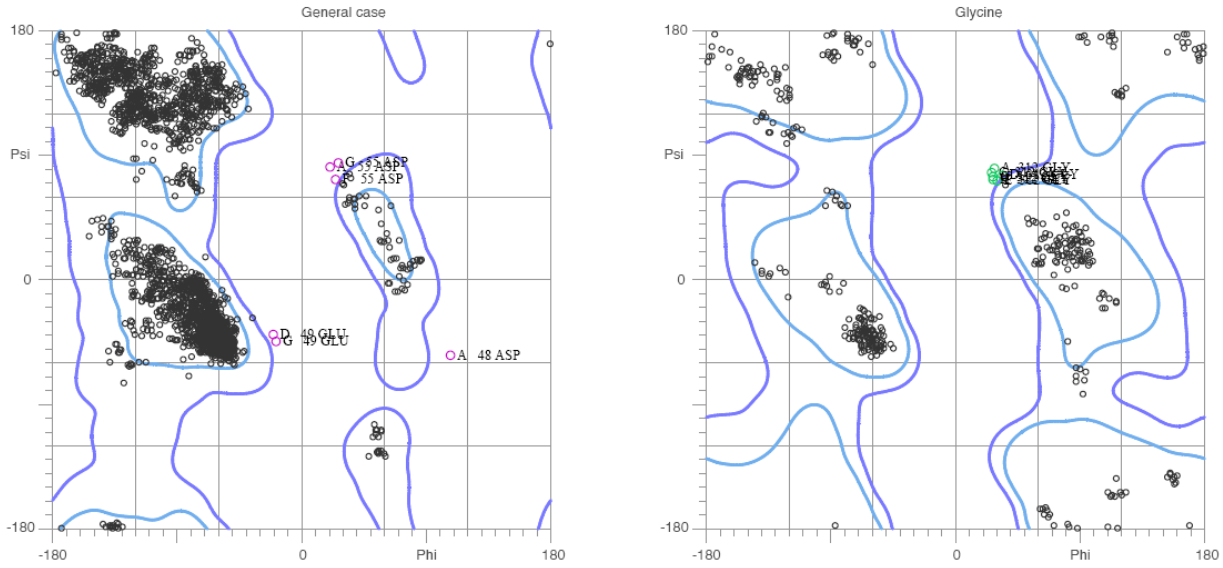


Figure 3-19. Ramachandran plot of the TLM bound C171Q KasA variant. Light blue areas delineate favored regions, dark blue areas are allowed regions. The left panel shows the plot for all amino acids except for glycines and prolines. Pink circles indicate the outliers. The right panel shows the plot for glycines only, green circles indicate the outliers.

The C171Q KasA TLM bound structure was solved at a resolution of 2.00 Å and also contains 8 chains in the asymmetric unit (chains A-H). Each monomer contains one thiolactomycin molecule bound to the active site. In total there are 3320 amino acid residues of which 18 protein residues display alternate conformations. The structure was refined to an R-factor of 13.7% and an R_{free} of 17.9%. Analysis of the φ-ψ dihedral angels with the Ramachandran plot shows that 96.7% of the residues are in the favored region, 99.6% in the allowed region and the structure contains 12 Ramachandran outliers, as calculated with the MolProbity server. Similar to the C171Q KasA apo structure, a few of the outliers belong to flexible loop regions (Asp48 Asp55 chain A, Asp55 chain F, Glu 49 and Asp 55 chain G) and their conformation can not be defined more clearly. All other outliers are glycine 312 (of chains A, C, D, E, G and H) which is the neighboring residue to the catalytically active His311 of the active site. The unusual conformation might be due to rearrangements in the active site upon TLM binding.

The structural difference between the apo structure and the TLM bound structure is very small (rms deviation of 0.37 Å) which shows that no major rearrangements occur upon binding of TLM to the protein. Superposition of the main chain atoms of the KasA wild type structure

and one monomer of the KasA C171Q structure, however, results in an rms deviation of 0.80 Å.

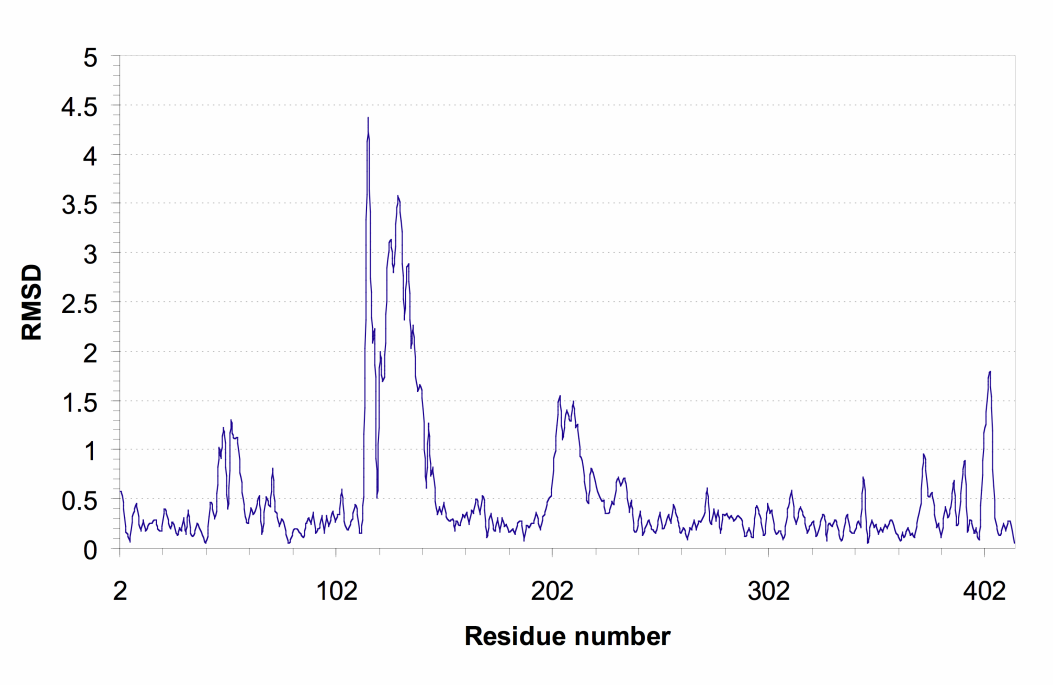


Figure 3-20. Structural differences between the KasA wild-type structure and the KasA C171Q structure. The graph shows the rms deviation between wild type KasA and monomer A of the KasA C171Q structure. Residues corresponding to the main peaks in the graph correspond to helices α2, α5 and α9 of the cap domain.

While the difference between the cores of the structures is still very small, the cap regions differ significantly from each other. A graph of the rms deviation between the two structures shows three main peaks that belong to residues 48-60, 115-147 and 203-223 (Figure 3-20).

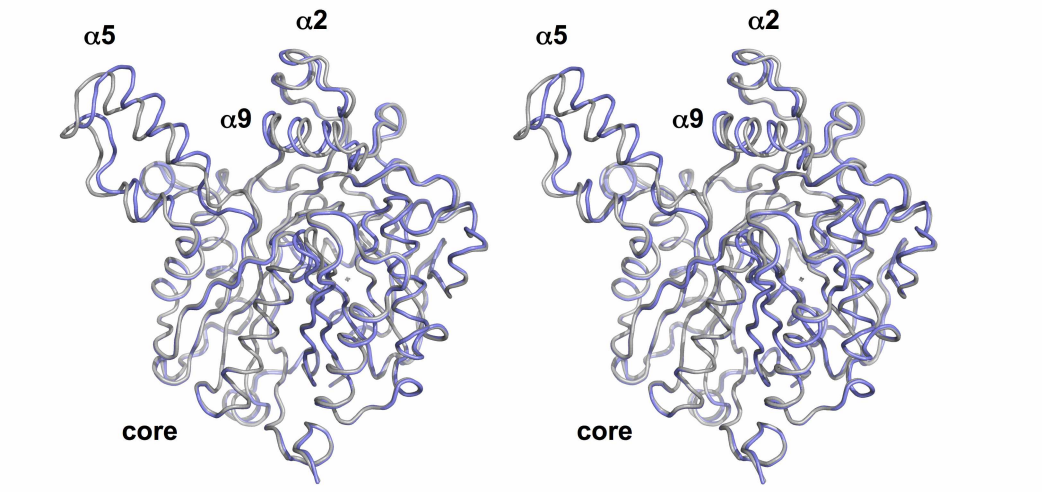


Figure 3-21. Stereo view of the superposition of wild-type KasA (gray) and C171Q KasA (blue). It is clearly visible that the cores of the two protein structures are highly similar and the three helices from the cap (α2, α5 and α9) differ most between the two structures.

These residues correspond to helices α2, α5 and α9, respectively. A superposition of two monomers of wild-type KasA and C171Q KasA is shown in Figure 3-21 and clearly shows that the main structural difference is located in the cap region. As large parts of the cap region are part of the acyl binding channel, the reason for the different conformations is probably the cysteine to glutamine mutation that mimics the acylated form of the enzyme and therefore triggers these conformational changes. Substrate binding is further discussed in chapter 4.2 and 4.3.

The average B-factor of the KasA C171Q apo structure is 55.0 $Å^2$ and 29.1 $Å^2$ of the KasA C171Q TLM structure. A graph of the B-factors of the main chain atoms (Figure 3-22) shows several small peaks that indicate flexible loops at the protein surface. In contrast to the KasA wild type structures, the graph of the KasA C171Q variant structures does not show any large areas that exhibit major deviations from the mean value. Also shown is the main chain of one monomer of the KasA C171Q TLM bound structure which is color coded according to the B-factors. It can be clearly seen, that the core of the protein displays low B-factors and the surface exposed loops B-factors with higher values. The large flexible area (residues 115-147) described for the KasA wild type structures displays average B-factors in the KasA C171Q structures.

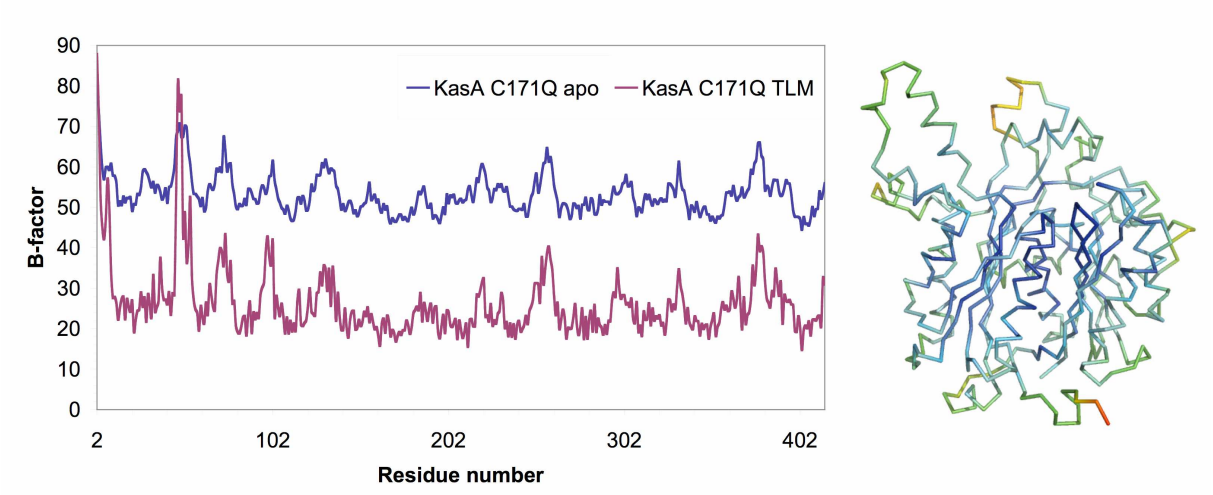


Figure 3-22. Illustration of the B-factors of the KasA C171Q variant structures. (Left) Plot of the B-factors of the main chain atoms against the residues of KasA C171Q apo (blue) and TLM bound KasA C171Q (plum). (Right) monomer of the KasA C171Q variant colored according to the B-factors. The color scale is depicted in ten steps from blue (low B-factor) to green to red (high B-factors).

3.4 Chromatographic purification of InhA

The bacterial pellet was resuspended in 20 mM Tris-HCl buffer, pH 7.9, containing 500 mM NaCl and 5 mM imidazole and lysed by sonication. The lysate was centrifuged (1 h, 50000g, 4° C) to separate cell fragments from the supernatant. The supernatant was applied to a HisTrap HP column (GE Healthcare) using the sample pump P960 of the Äkta purifier system. After a washing step with buffer containing 60 mM imidazole, a gradient of 60 mM to 500 mM imidazole was used for elution. The chromatogram in Figure 3-23 shows that InhA elutes in the middle of the gradient at about 300 mM imidazole (50% elution buffer) with a broad peak.

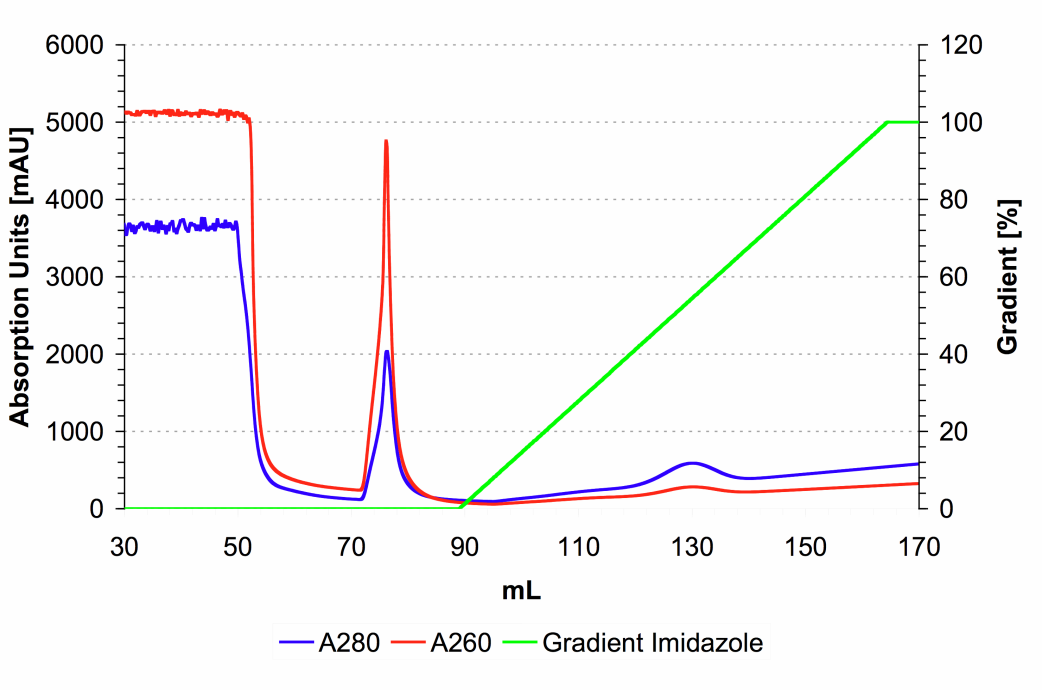


Figure 3-23 Purification of InhA by Nickel Affinity Chromatography. Blue: Absorption at 280nm, red: Absorption at 260nm, green: percentage of buffer C, gradient. InhA elutes in the middle of the gradient with a broad peak at approximately 130 mL.

To check the grade of purity, samples of the flow-through, the wash step and the elution were loaded on a 15% SDS gel (Figure 3-24). Lane 1 shows that the over expression worked very well as the major band corresponds to the molecular weight of InhA. The protein is not present in the flow through (lane 2 and 3), indicating that the protein binds tightly to the affinity matrix. In the washing step a small amount of InhA is eluted (lane 4 and 5), while the major peak during the gradient is almost pure InhA (lane 6-14).

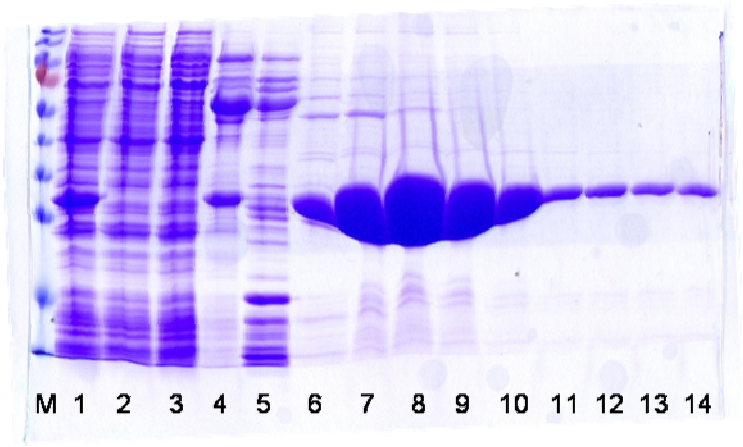


Figure 3-24. 15% SDS gel showing fractions of the affinity chromatography of InhA. M indicates the protein ladder, lane 1 is a sample of the supernatant before affinity purification, lanes 1-5 correspond to the flowthrough (lanes 2 and 3) and washing steps (lanes 4 and 5), lanes 6-14 correspond to the elution fractions.

The fractions corresponding to lanes 6-14 on the gel were pooled and exchanged to the size exclusion buffer utilizing PD10 desalting columns (GE Healthcare) and concentrated to a volume of 5 mL with a centrifugal filter device. Before loading the sample on a size exclusion column (HiLoad 26/60 Superdex 200 prep grade, GE Healthcare), the His tag was cleaved off by thrombin digestion over night.

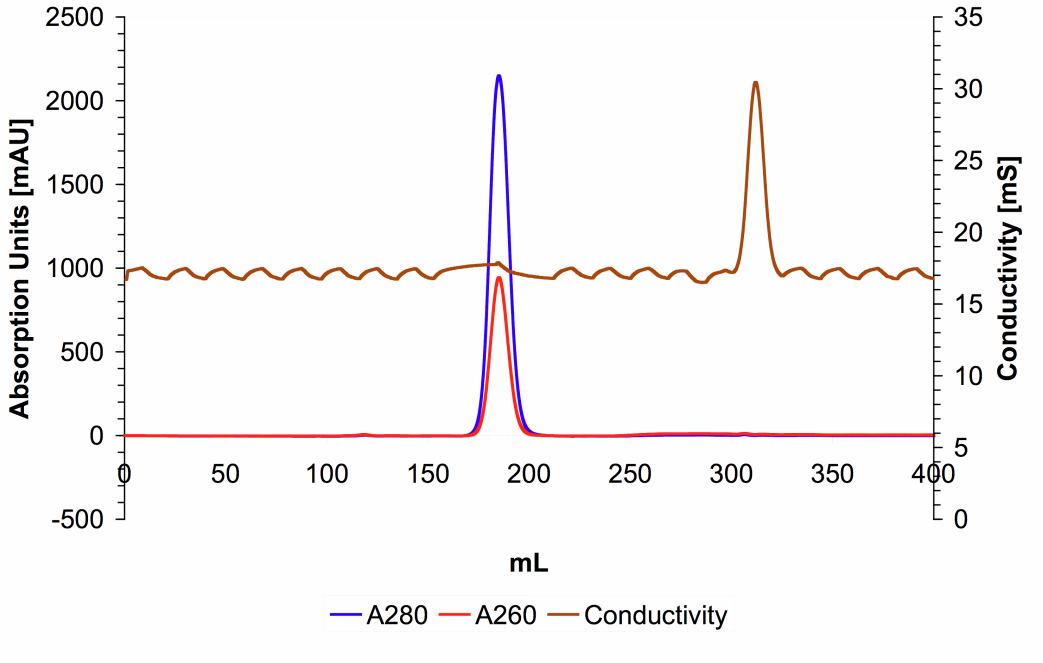


Figure 3-25. Purification of InhA by size exclusion chromatography. Blue: absorption at 280nm, red: absorption at 260 nm, brown: conductivity in mS depicting the salt concentration.

The elution profile shows a homogenous peak with the elution volume corresponding to a tetramer of InhA.

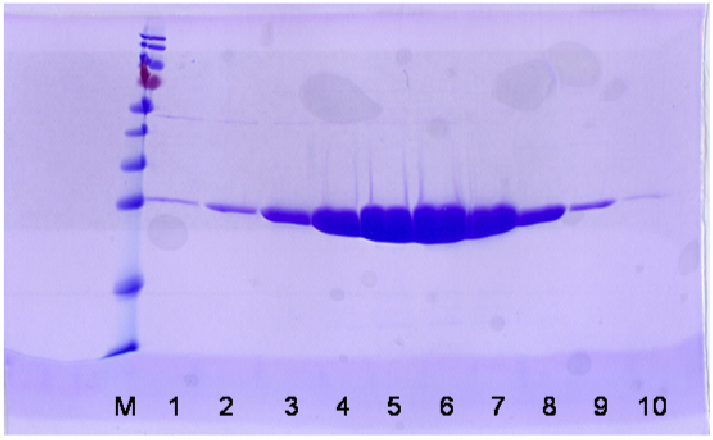


Figure 3-26. 15% SDS gel showing the elution fractions of the size exclusion chromatography of InhA. M indicates the protein ladder, lanes 1-10 correspond to the elution fractions.

The 15% SDS gel in Figure 3-26 shows that the peak contains pure InhA. Fractions containing the pure protein were pooled and concentrated with a centrifugal filter device to 10 mg/mL (330 µM). The protein was directly dropped into liquid nitrogen in fractions of 30 µL and stored as small globules at -80° C. Due to the very symmetric peak of the size exclusion chromatography that already indicated a very homogenous protein solution, no further measurements were taken by DLS.

3.5 Crystallization of the ternary InhA complex

For the formation of the ternary complex formed between the protein, its cofactor NAD^+ and the inhibitor PT70 (InhA-NAD^+-PT70) the protein solution (10 mg/mL InhA, 20mM Pipes pH 6.8, 150mM NaCl) was combined with NAD^+ and PT70 at a molar ratio of 1:5:200. After 2 h incubation on ice, the solution was centrifuged at 25000g and 4° C for 20 min and the supernatant was used for crystallization by the hanging drop vapor diffusion technique. Equal volumes of protein solution and crystallization solution (12-16% (w/v) PEG 4000, 1% DMSO, 100 mM ADA, pH 6.8, 100-250 mM ammonium acetate) were mixed and equilibrated against a reservoir containing the crystallization solution. Crystals of the complex grew within 4 days at 22° C to a maximum size of 900 x 100 x 100 μm^3. The crystals were transferred to cryo solutions containing NAD^+ and PT70 in excess and increasing amounts of DMSO in steps of 5% (v/v) to a final concentration of 25%, soaked in each well for 2 min and subsequently frozen in liquid nitrogen.

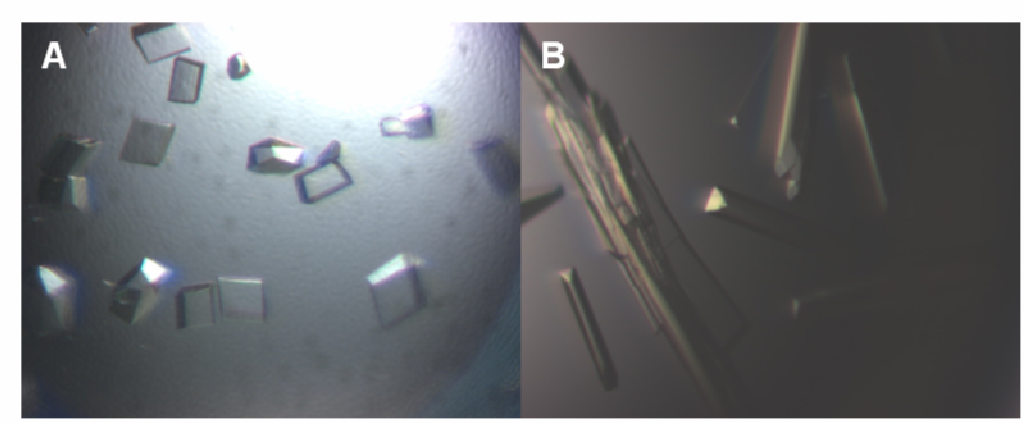


Figure 3-27. Crystals of InhA. (A) Crystals of the binary InhA-NAD$^+$ complex grown in the same conditions as the crystals of the ternary InhA-NAD$^+$-PT70 complex (B) have a different shape.

3.6 Structures of InhA

3.6.1 Structure solution and refinement

Two diffraction data sets were collected at a Rigaku MicroMax™-007HF generator with a Raxis HTC Detector. The first crystal diffracted to 2.1 Å, the second crystal to 1.8 Å. Datasets one and two were indexed and processed using mosflm and xds, respectively. Although the crystals grew in the same drop, they belong to different space groups: $C222_1$ for dataset one and $P2_1$ for dataset two. The most plausible Matthews coefficient of 2.9 Å^3/Da was calculated for two monomers in the asymmetric unit for dataset one with a solvent content of 57.3%. For dataset two the most plausible Matthews coefficient of 2.8 Å^3/Da was determined for four monomers in the asymmetric unit which corresponds to 56.1% solvent content. Data collection statistics are provided in Table 3-3. The two structures were solved by molecular replacement using Phaser. As a search model for the first dataset, a monomer was derived from the protein data bank entry 2b37 (Sullivan, Truglio et al. 2006), omitting the NAD$^+$ and the inhibitor. Molecular replacement yielded two molecules in the crystallographic asymmetric unit that form half of a tetramer. Model building and refinement was carried out using alternating rounds of Coot for manual model building and Refmac5 for maximum likelihood refinement. Five percent of all reflections were omitted throughout the refinement for the calculation of the R_{free}. In the initial refinement rounds, the density improved significantly and allowed the unambiguous modeling of the NAD$^+$ molecule and the inhibitor PT70 into the electron density using Coot. For molecular replacement of the second dataset the coordinates of one monomer of the refined first structure were used as a search model, omitting only the inhibitor. Molecular replacement yielded one tetramer in the

crystallographic asymmetric unit. Refinement and model building was carried out as described above. The refinement statistics for the two datasets are listed in Table 3-3.

Table 3-3. Data collection and refinement statistics for the two datasets of the InhA-NAD^+-PT70 complex.

	InhA NAD^+ PT70 Crystal One	InhA NAD^+ PT70 Crystal Two
Data collection		
Space group	$C222_1$	$P2_1$
Cell dimensions		
a, *b*, *c* (Å)	89.80, 157.51, 91.23	88.48, 90.27, 89.56
α, β, γ (°)	90, 90, 90	90, 118.76, 90
Resolution (Å)	36.16-2.10 (2.21-2.10)*	78.57-1.81 (1.90-1.80)
R_{merge}	0.080 (0.300)	0.101 (0.438)
$I / \sigma I$	16.0 (2.5)	6.1 (1.7)
Completeness (%)	98.1 (89.7)	92.5 (77.8)
Redundancy	4.4 (2.9)	1.6 (1.4)
Refinement		
Resolution (Å)	2.10	1.81
No. reflections	38141	112861
R_{work} / R_{free} (%)	17.2 / 21.6	16.8 / 20.3
No. atoms		
Protein	4063	8033
PT70	42	84
NAD^+	88	176
Water	143	682
B-factors		
Protein	26.2	10.4
PT70	26.9	7.3
NAD^+	19.7	3.9
Water	29.9	16.0
Rms deviations		
Bond lengths (Å)	0.013	0.015
Bond angles (°)	1.465	1.534

*values in parenthesis are for highest resolution shell

3.6.2 Validation

Two crystal structures of InhA with bound NAD^+ and PT70 were solved in different space groups. Structure one belongs to space group $C222_1$ and contains two monomers in the asymmetric unit that form the functional tetramer together with two symmetry related InhA molecules. Structure two belongs to space group $P2_1$, a space group that has not been reported

for InhA so far. The structure contains four chains in the asymmetric unit that form one tetramer. One monomer is composed of 8 α-helices and 7 β-strands. The β-strands form one parallel β-sheet that generates together with 6 α-helices a Rossman-fold in the lower part of the structure to which the NAD^+ molecule is bound.

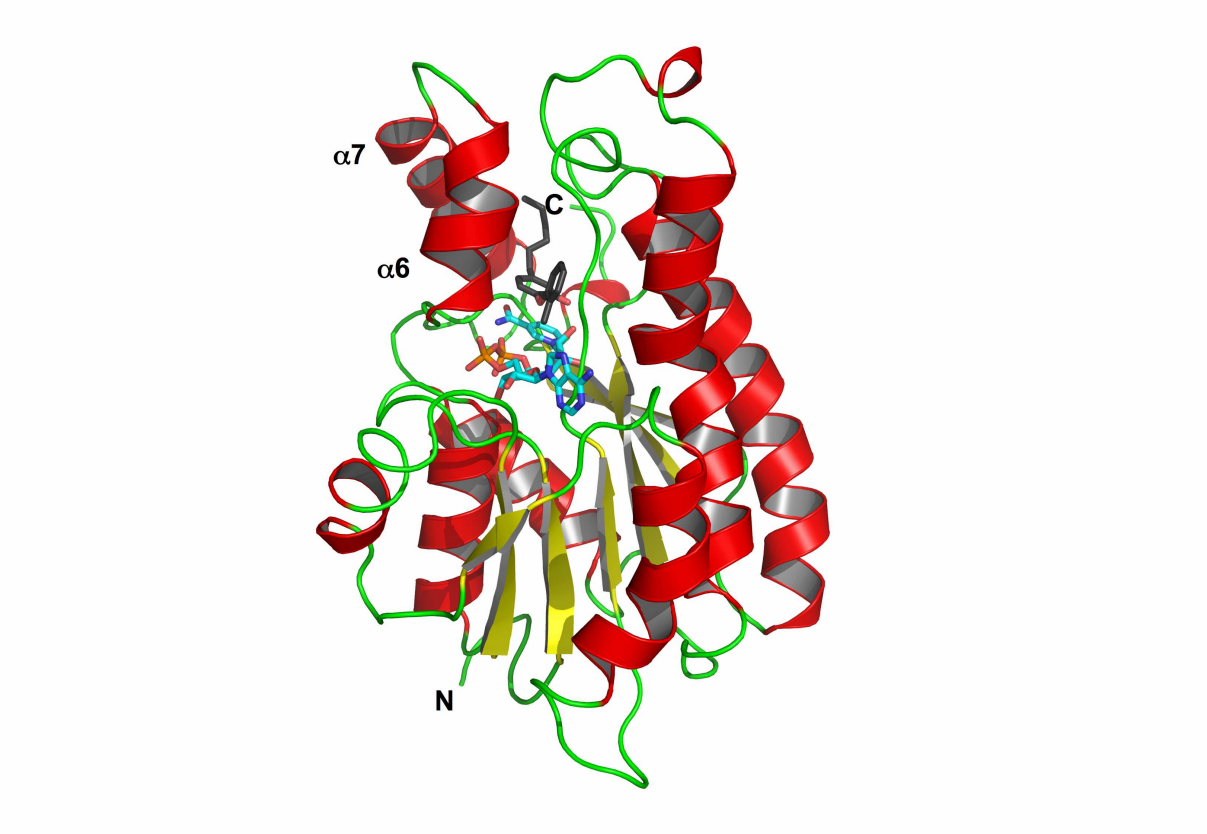


Figure 3-28. One monomer of the ternary structure of InhA-NAD^+-PT70. InhA is color coded according to the secondary structure elements (sheets yellow, helices red and loops green). The bound NAD^+ molecule is colored in cyan, the inhibitor PT70 in the substrate binding pocket is black. Helices α6 and α7 that form the substrate binding loop are labelled.

In the upper part of the structure two α helices (helix α6 and α7) form the so called substrate binding loop that composes one part of the substrate binding pocket. The inhibitor PT70 is bound to this substrate binding pocket.

The rms deviation between different monomers of the two structures is below 0.63 Å, and it can be assumed that despite the different space groups the two structures are very similar.

The Ramachandran plot shows that the structures are of good quality. 96.9% of all residues in structure one are in the favored regions and 100% are in the allowed regions. Structure two has 96.4% of all residues in the favored regions and 99.9% in the allowed regions; one residue (Asp42 of chain E) is an outlier that lies on a flexible loop at the surface of the protein.

The average B-factor of structure one is 26.2 $Å^2$ and 10.4 $Å^2$ for structure two. A graph of the B-factor distribution for all residues shows that the B-factors of most residues is in the

average range, with the exception of one monomer which has increased B-factors in the range of residues 206-220, helix α7 of the substrate binding loop. The increased B-factors of the substrate binding loop can only be found in one of the two monomers, the other monomer is stabilized by contacts to symmetry related molecules. In structure two the same phenomenon is visible; here two of four substrate binding loops display increased B-factors and the loops of the other monomers are stabilized by contacts to symmetry related molecules.

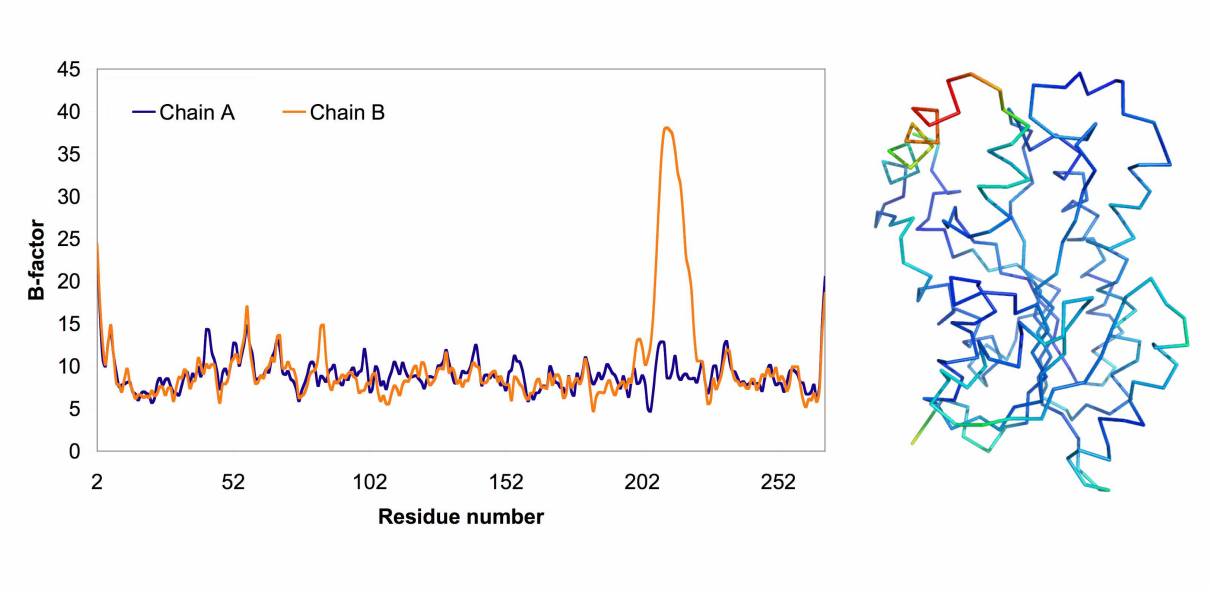


Figure 3-29. (left) The graph shows the average B-factors of the main chain atoms of InhA structure two. One area, helix α7 of the substrate binding loop (residues 206-220) displays increased B-factors in one monomer. (right) A monomer of the InhA structure two is colored according to the B-factors. Red indicates high B-values and blue low values.

4 Discussion

4.1 Thiolactomycin binding to KasA

In order to evaluate the interactions of TLM with KasA as a platform for rational inhibitor design and to explore the molecular basis for the preferential binding of TLM to acyl-KasA, the X-ray structures of wild-type KasA and the acyl enzyme mimic C171Q, both unliganded (apo) and with bound TLM were determined at resolutions ranging from 1.8 to 2.2 Å.

4.1.1 The structures of wild-type KasA

To obtain a complex structure of wild-type KasA with bound TLM, it was necessary to crystallize the protein with the inhibitor in a 1:200 molar ratio. Albeit this high ratio, KasA is not fully occupied by TLM, consistent with the biochemical observation that TLM binds only weakly to the free enzyme. Nevertheless, the conformation of the TLM molecule is clearly defined in the structure and allows an unambiguous assignment of its protein interactions. Detailed interactions between the KasA atoms and the atoms of the TLM molecule are listed in Table 4-1.

Table 4-1. Interactions between atoms of the KasA wild-type protein and atoms of TLM.

KasAwt residue / atom	TLM atom	Distance (Å)
Hydrogen bond		
His 311 / NE2	O1	2.9
His 345 / NE2	O1	2.7
Hydrophobic interaction		
Phe 237 / CD2	C9	3.7
Phe 237 / CE2	C9	3.9
Val 278 / C	C11	3.1
Pro 280 / CD	C5, C6	3.7, 3.7
Pro 280 / CB	C7	3.8
Pro 280 / CA	C8	4.0
Pro 280 / CG	C10	3.7
Thr 313 / CG2	C1, C9	3.8, 3.7
Thr 313 / CB	C9	3.9
Phe 402 / CB	C7, C8	4.0, 4.1
Gly 403 / C	C7	3.8
Phe 404 / CE1	C3	3.3
Phe 404 / CD1	C5	3.5
Phe 404 / CZ	C2, C9	3.2, 3.2
Phe 404 / CA	C11	3.7
Gly 406 / CA	C8	3.6
Gly 406 / C	C8	4.2

TLM binds to the malonyl binding pocket of KasA. The methyl groups C-9 and C-10 are positioned in two hydrophobic pockets formed by Pro280, Gly318 and Phe402, Phe237, respectively. The intercalation of the isoprenoid tail into the space between two peptide bonds, namely Ala279-Pro280 and Gly403-Phe404, further stabilizes the interaction.

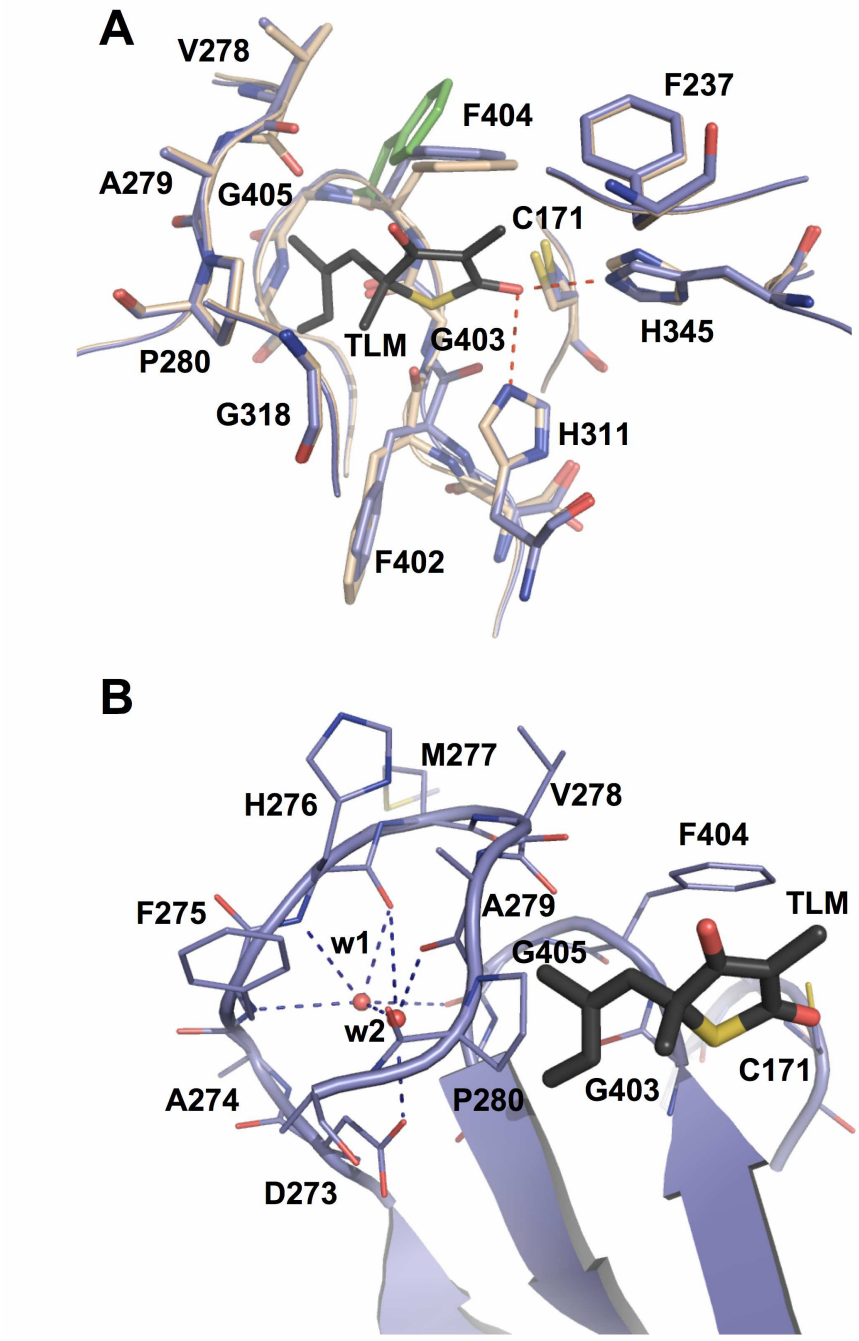


Figure 4-1. The active site of wild-type KasA. (A) Superposition of the TLM bound structure (blue) with the apo structure (light brown). The TLM molecule is depicted in black all-bonds representation; the red dotted lines indicate hydrogen bonds. Shown in green is the position of the phenylalanine in the *E. coli* FabB structure. (B) The isoprene side chain of TLM (black) is sandwiched between two peptide bonds formed by Ala279-Pro280 and Gly403-Phe404. Two water molecules are shown in the lipophilic pocket (red spheres), blue dotted lines indicate hydrogen bonds.

Two strong hydrogen bonds are formed between the O-1 oxygen and the nitrogens of the active site histidines, His311 and His345. The isoprenoid moiety of TLM points towards an extended lipophilic pocket where two important water molecules are present in all structures and stabilize the loop from Asp273 to Pro280 (Figure 4-1).

Upon binding of TLM to the active site, Phe404 and the associated loop comprising residues Phe402-Gly406 are shifted by 0.9 Å to avoid clashes with the inhibitor. Additionally, the main chain oxygens of Phe402 and Val278 are rotated away from the TLM molecule. Binding of TLM to the active site in *M. tuberculosis* KasA can be compared to the related *E. coli* FabB (ecFabB) protein in the presence of TLM (PDB code 1fj4 (Price, Choi et al. 2001); Figure 4-1) Unexpectedly, the isoprenoid tail of TLM and, most strikingly, the active site residue Phe404 adopt different conformations. Phe404 acts as a gatekeeper to the adjacent acyl channel and assumes a closed conformation in both the free and TLM-bound KasA wild-type structures. This feature has not been observed in the TLM bound and apo structures of ecFabB, where the corresponding residue Phe392 is always observed in an open conformation leading to an almost perpendicular orientation of its aromatic ring relative to the TLM thiolactone ring. In contrast, the aromatic ring of Phe404 in KasA is oriented parallel to the thiolactone ring and TLM binding does not appear to alter the position of Phe404 in KasA which remains in the 'closed' conformation in both wild-type structures.

4.1.2 The structures of C171Q KasA

For the C171Q KasA TLM complex, co-crystallization with a protein to inhibitor ratio of 1:10 led to a structure that was fully occupied by TLM. TLM forms the same hydrophobic interactions as described for the wild-type KasA TLM complex and, likewise, the hydrogen bonds with the active site histidines are unchanged (Figure 4-2 and Table 4-2).

Table 4-2. Interactions between atoms of the KasA C171Q protein and atoms of TLM.

KasA C171Q residue / atom	TLM atom	Distance (Å)
Hydrogen bond		
His 311 / NE2	O1	3.0
His 345 / NE2	O1	2.8
Hydrophilic interaction		
Phe 237 / CD2	C9	3.9
Phe 237 / CE2	C9	4.1
Val 278 / C	C7	3.5
Pro 280 / CD	C6	3.8
Pro 280 / CA	C8	3.9
Pro 280 / CD	C10	4.0
Pro 280 / CG	C10	4.0

Thr 313 / CG2	C1, C2, C9	3.7, 3.8, 4.2
Thr 313 / CB	C9	4.5
Gly 318 / CA	C10	4.1
Phe 402 / CB	C11	3.6
Phe 404 / CE1	C3, C5, C9	3.5, 3.6, 3.9
Phe 404 / CD1	C6, C7	4.0, 3.9
Phe 404 / CZ	C9	4.0
Gly 406 / CA	C8	3.8

One fundamental difference is, however, the orientation of the Phe404 side chain in the C171Q KasA structure. Phe404 adopts an open conformation that is also observed in the apo C171Q KasA structure. The side chain oxygen of the mutated Gln171 generates a hydrogen bond to the amide nitrogen of Phe404, thereby shifting the phenylalanine side chain by 60° and by 3.3 Å out of the active site relative to the position of this residue in the KasA wild-type structure.

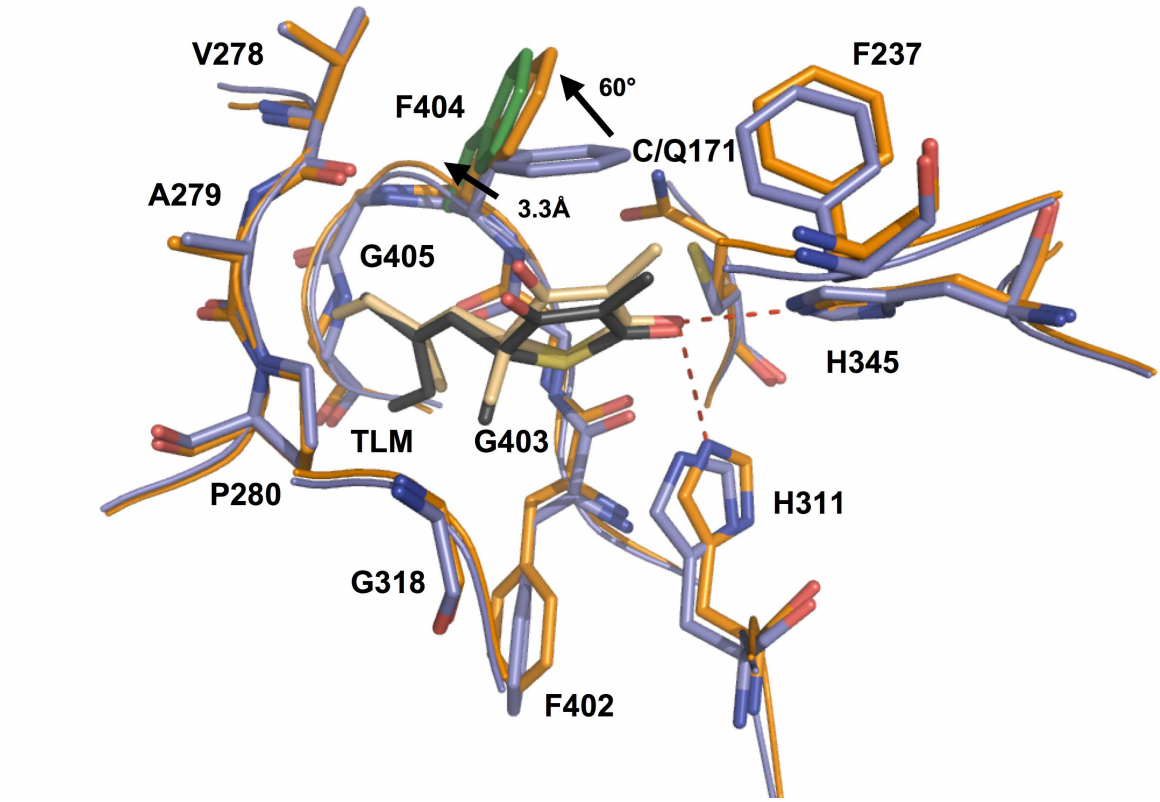


Figure 4-2. The structure of C171Q KasA with bound TLM (colored orange with TLM in light orange) has been superimposed with the wild-type KasA TLM structure (blue with black TLM). Also shown in green is the position of the phenylalanine in the C163Q variant of *E. coli* FabF.

In this open conformation the Phe404 aromatic ring forms an energetically favorable edge-to-face interaction with the thiolactone ring of TLM (Hunter, Singh et al. 1991). The very flexible loop containing residues Phe402, Gly403, Gly405 and Gly406 which immediately

precede and follow Phe404 is shifted out of the binding pocket by about 1.5 Å, while Phe210 and helix α9 are moved away from the cavity by ~2 Å. Thereby the solvent accessible volume of the active site pocket is increased from 97 Å^3 in the wild-type to 150 Å^3 in the mutant structure as calculated with the program CASTp. Possibly due to additional space in the cavity of the KasA C171Q mutant, the isoprenoid tail of TLM in the KasA C171Q structure adopts a different conformation to that observed in the binary wild-type KasA TLM structure. Calculations of the pathways leading from the inside of the pocket to the outside solvent with CAVER showed that the entrance to the active site is more restricted in the wild-type enzyme (2 Å radius) than in the mutant (2.5 Å radius), suggesting that the binding site is more accessible for TLM, and presumably also the malonyl group, when the enzyme is acylated (Figure 4-3).

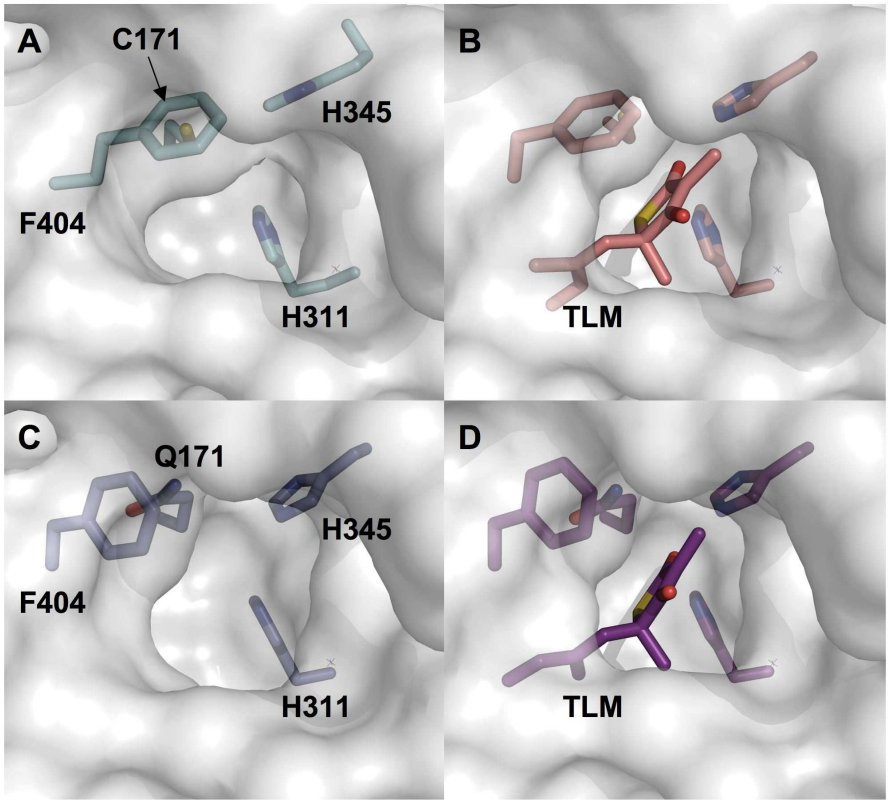


Figure 4-3. Widening of the malonyl binding pocket in the C171Q KasA structure. (A) Surface representation of the binding pocket of the wild-type KasA apo structure. The active site residues are shown in all-bonds representation. (B) Binding pocket of wild-type KasA with bound TLM. (C) Binding pocket of the C171Q KasA variant in the apo form and (D) Binding pocket of the TLM-bound KasA C171Q variant.

These results support the contention that replacement of the active site cysteine with a glutamine mimics structural changes caused by acyl-enzyme formation. Similar observations are reported for the *E. coli* FabF enzyme where Phe400 adopts a closed conformation in the apo enzyme but an open conformation in the acyl-enzyme and the C163Q mutant structures

(pdb codes 2gfw, 2gfv, 2gfy and 2gfx (Wang, Soisson et al. 2006)). In FabF as well as in KasA this structural rearrangement has great impact on the binding efficiency of inhibitors. Binding of TLM to C171Q KasA perturbs the protein structure only very slightly and just the main chain oxygen of Val278 undergoes a significant movement upon TLM binding to avoid steric clashes with the inhibitor. These structural differences suggest that the binding pocket is already preformed to accept the TLM molecule, providing a molecular explanation for the differential binding behavior of TLM to KasA as observed by Machutta and coworkers (Machutta 2009). In this work, direct binding experiments also revealed that addition of TLM to the acyl-enzyme form of KasA resulted in a much slower decrease in the fluorescence signal than the instantaneous change observed for the free enzyme, which suggested that there is a slow-onset component to the interaction of TLM with acyl-KasA. However, no structural evidence could be found to explain this phenomenon.

4.2 Substrate mimic binding to KasA

The acyl channel in KasA is clearly defined in two of the structures due to the presence of a bound PEG molecule, which mimics an acyl chain of about 40 carbon atoms in length.

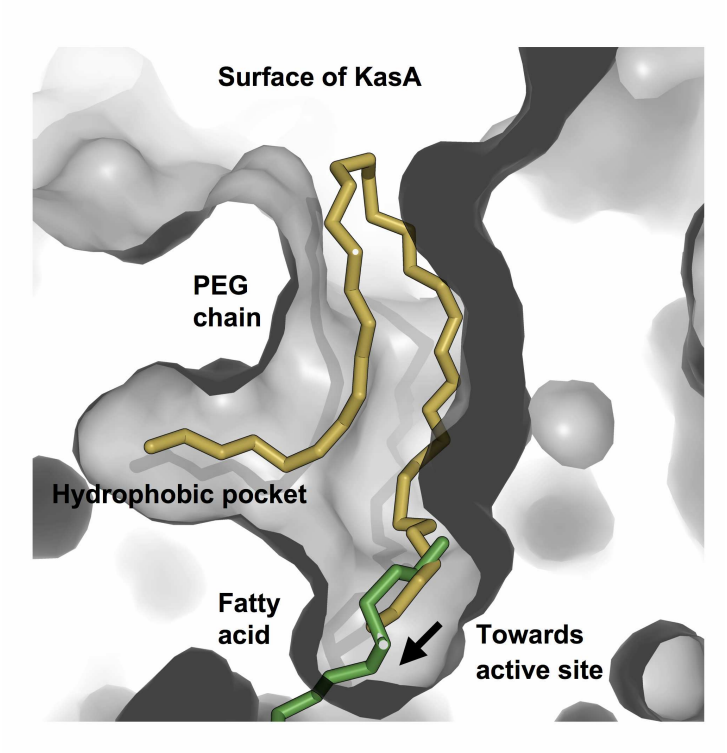


Figure 4-4. Surface representation of the acyl binding channel of KasA with the bound PEG molecule. For clarity, the front part of the protein is removed to allow a view into the hydrophobic channel. The superposition of the acyl chain bound *E. coli* FabB structure (pdb code 1f91) with C171Q KasA clearly shows that the fatty acid of FabB (green) merges with the PEG molecule (yellow) bound to the KasA structure.

These structures thus provide for the first time a view into the hydrophobic acyl binding cavity, which is lined with numerous hydrophobic amino acids and perfectly accommodates the growing fatty acid chain (Figure 4-5 A).

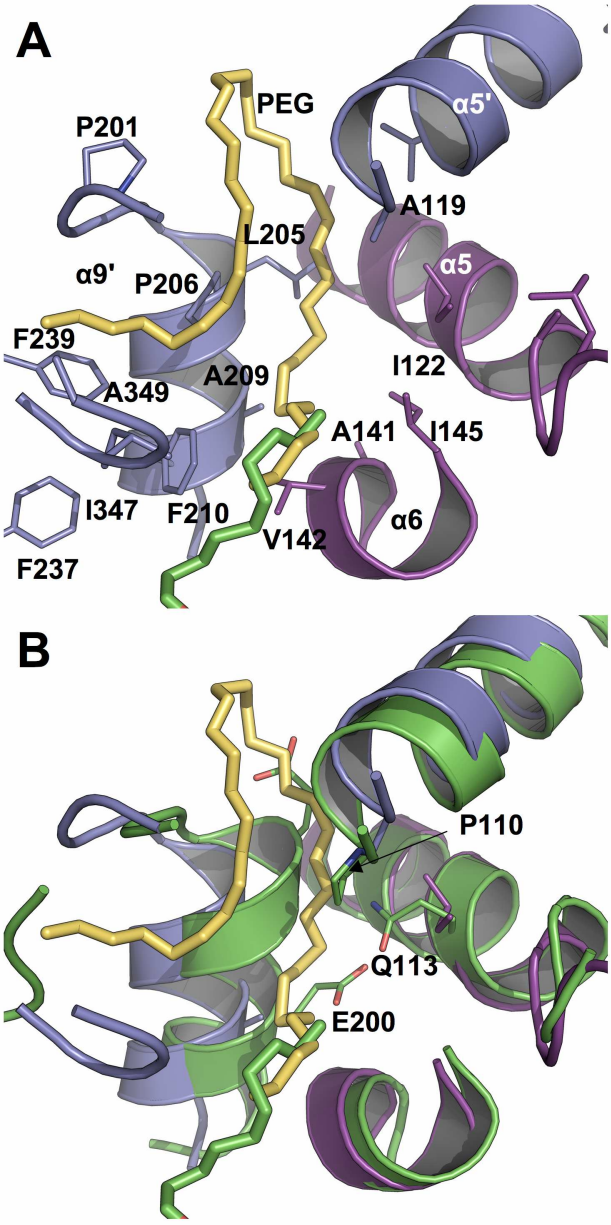


Figure 4-5. (A) The acyl channel lies in the interface of the KasA monomers (blue and purple). The hydrophobic residues in KasA that line the channel are shown in all-bonds representation. (B) Superposition of C171Q KasA with the *E. coli* FabB structure (pdb code 1f91). Shown is only the part of the acyl binding channel. The helices and the polar residues in FabB truncate the acyl binding channel. Large polar residues in the FabB structure (Glu200 and Gln113) together with Pro110 correspond to small hydrophobic residues in KasA (Ala209, Ile122 and Ala119).

A superposition of KasA main chain atoms of a dimer with the *E. coli* FabB structure (PDB code 1f91) in complex with a short fatty acid clearly displays the difference between proteins capable of binding short versus long fatty acid chains. Initially the two fatty acids superimpose but then the PEG molecule extends through the channel to the surface of the protein, bends back into the channel and terminates in a hydrophobic pocket. In the *E. coli* FabB structure, side-chain and backbone atoms from a helix comprising residues Trp195-Gly205 significantly truncate the acyl binding channel. This helix is reoriented in KasA and thereby permits binding of the longer acyl chain substrates. Additionally, long polar side chains point into the channel in *E. coli* FabB, namely Glu200 and Gln113, as well as Pro110, thereby blocking further access. In KasA short hydrophobic side chains Ala209, Ile122 and Ala119 facilitate fatty acid binding (Figure 4-5 B).

4.3 Proposed substrate binding mechanism

The acyl channel of KasA is directly accessible through two openings, the malonyl binding pocket and the opening of the acyl channel at the surface of the protein. However, migration of such a long fatty acid chain through the malonyl binding channel past the hydrophilic and charged active site residues appears to be energetically and sterically unfavorable. Equally unfavorable is the migration of the pantetheine group through the hydrophobic environment of the long acyl channel to the active site cysteine. Both entries would afford a threading mechanism of a long fatty acid which would be too time-consuming for efficient transfer between the different enzymes involved in the individual steps of fatty acid synthesis. A separate mechanism of substrate binding and product release therefore seems plausible. Comparisons of the wild-type KasA structures with the C171Q KasA structures indicate that a polypeptide segment in KasA containing helices α5 and α6 (residues 115-147) shows increased B-factors relative to the average of the whole protein in the wild-type KasA structures. The B-factors in the respective residues of the mutant structures remain at the same level as the average. These findings suggest that this loop remains very flexible in the absence of a bound acyl substrate and becomes ordered upon binding of the fatty acid to KasA. It is conceivable that residues 115-147 move in a concerted way in a scissor like motion in the dimer, thereby providing direct access to the long acyl channel (Figure 4-6). Most likely, the opening of the channel is further guided by the interaction with the acyl carrier protein. These findings are further supported by Sachdeva et al., who characterized the binding mode of fatty acids to *M. tuberculosis* FabH and proposed an open conformational state of FabH where the acyl channel is accessible for acyl substrates or inhibitors (Sachdeva, Musayev et al. 2008).

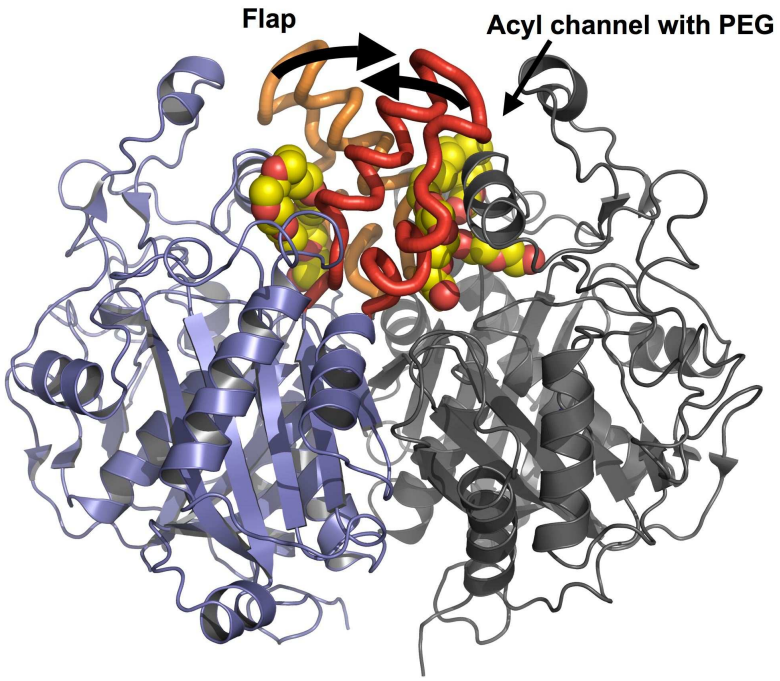


Figure 4-6. Proposed substrate binding mechanism of KasA. A dimer of the KasA C171Q variant is displayed in blue and gray cartoons. The hypothetical gate segments (residues 115-147) are shown as ribbons in red and orange. Arrows indicate the movement of the gates that is necessary to open the acyl channel. PEG chains that occupy the acyl channel are shown as yellow and red spheres.

4.4 Comparison of KasA and KasB

KasB is the β-ketoacyl synthase II of *M. tuberculosis*. The sequence identity between KasA and KasB is 67% and the structures of the two enzymes are highly similar, the rms deviation between the two enzymes after superposition using secondary structure matching is only 1.09 Å. Despite the high similarity of the two enzymes, previous work has shown that KasA and KasB have slightly separate functions (compare chapter 1.2.2 for details). A careful comparison between the KasA structure and the structure of KasB (pdb code 2gp6 (Sridharan, Wang et al. 2007)) could give some hints that would explain the observed functional differences in substrate chain length preference and TLM affinity. The two structures were superimposed by the TopMatch-web server and according to this superposition, 391 out of 415 residue pairs are structurally equivalent resulting in a relative similarity of the two structures of 94%. Six areas show structural differences. Two of the areas (159 and 331-332, see Figure 4-7) are surface exposed loops and considered as not relevant for substrate specificity or inhibitor binding. Three areas (49-56, 65-70 and 117-120) are in the region of

the acyl channel. However, the structural differences are so minor, that the difference in substrate specificity still can not be explained.

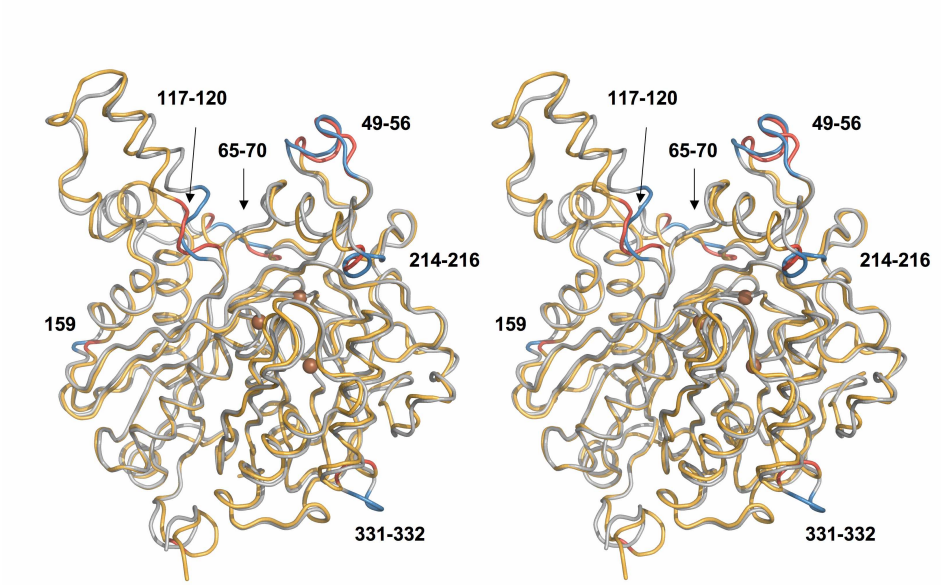


Figure 4-7. Stereo view of the superposition of KasA and KasB monomers (both wild-type apo forms) shows the regions of structural differences. KasA is shown in orange ribbons with the areas containing structural differences in red and KasB is depicted in gray ribbons with blue areas of difference. For orientation, the Cα atoms of the active site residues C171, H311 and H345 of KasA are shown as brown spheres.

Finally, area 214-216 that is close to the active site entrance was found to be different in KasA and KasB. It was already speculated in the work about KasB based on a model of KasA that the active site entrance is more restricted in KasB due to this loop that has an insertion of one amino acid in KasB (Sridharan, Wang et al. 2007). The comparison of the two structures seems to confirm this hypothesis. The loop is shorter in KasA resulting in an increased accessibility of the active site. That could account for the higher affinity of TLM for KasA in comparison to KasB. In conclusion, the structures of KasA and KasB are highly similar and it is not possible to explain the different substrate specificities on the basis of the two apo structures. To further investigate the substrate specificities of the two enzymes, a structure of the KasB acyl enzyme mimic (C170Q KasB) might provide more insights into the substrate binding mode of KasB.

4.5 Crystal structures of InhA

The crystal structures of the ternary InhA-NAD^+-PT70 complex were solved to further characterize the binding of the inhibitor PT70 to the key enzyme InhA. Of particular importance is the structural basis for the slow onset inhibition of InhA by PT70.

4.5.1 PT70 binding to InhA

The slow onset inhibitor PT70 binds to the substrate binding site with the phenyl-rings of the inhibitor oriented almost 90° to each other. Hydrogen bonds are formed between the inhibitor hydroxyl group and Tyr158 as well as the 2'-hydroxyl group of NAD^+ and a hydrogen bonding network is formed to Lys165 (Figure 4-8).

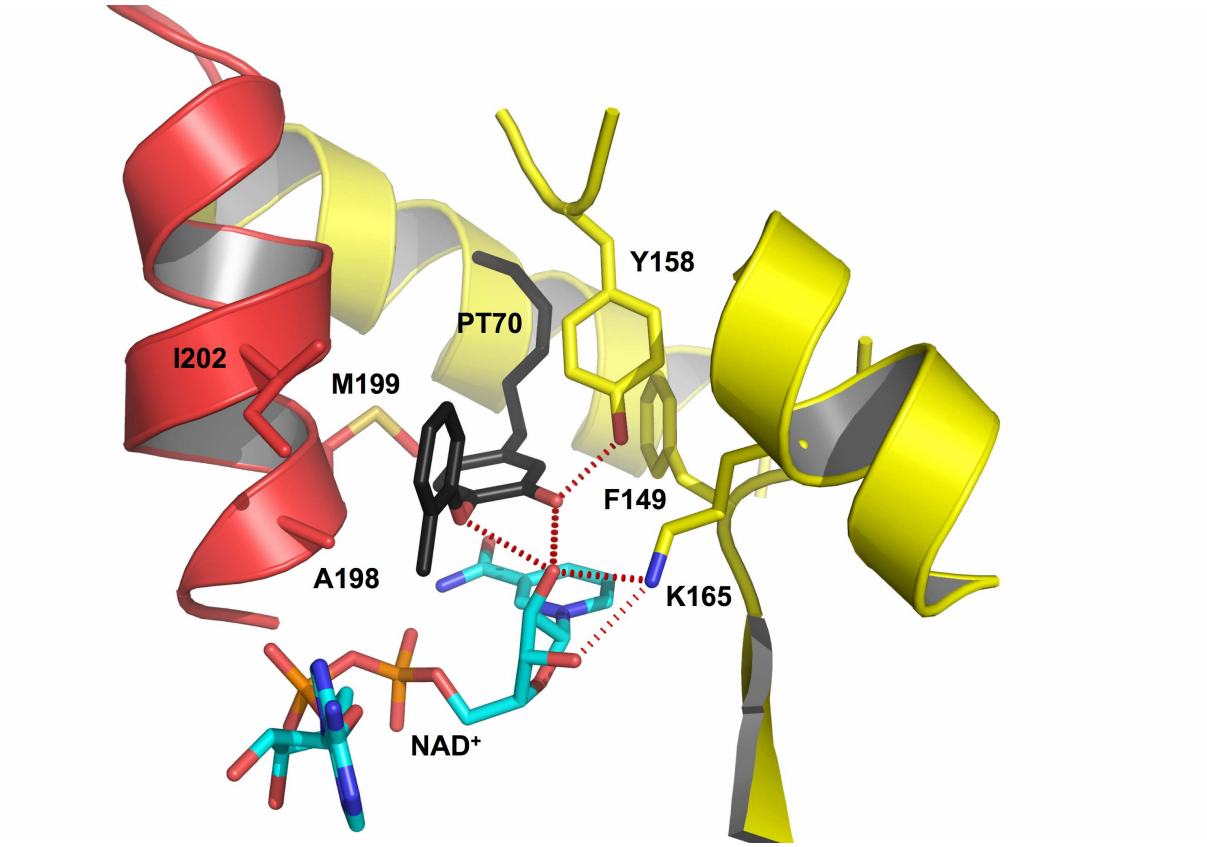


Figure 4-8. Close-up of the binding pocket of InhA with bound NAD^+ and PT70. Hydrogen bonds between PT70 (black) and Tyr158 (yellow) as well as the NAD^+ molecule (cyan) are indicated as red dotted lines. The important hydrophobic residues Ala198, Met199 and Ile202 of the substrate binding loop are shown as all bonds representation in red.

A π-π stacking interaction between the B-ring of PT70 and the nicotinamide ring of NAD^+ further stabilizes the inhibitor in the pocket. The alkyl chain of PT70 extends into the hydrophobic environment of the substrate binding cavity and forms hydrophobic interactions with residues Phe149 and Tyr158. Most importantly, however, is the substrate binding loop

(residues 195-210) of the two structures reported here, which is ordered and thus allows an analysis of the interactions of the loop residues with the inhibitor. Hydrophobic interactions are formed between the phenyl rings of the inhibitor and the amino acids Ala198, Ile202 and Met199 of the substrate binding loop. Ala198 forms a hydrophobic interaction with the B-ring *ortho* methyl group at a distance of 3.4 Å, Ile202 with the B-ring at a distance of 3.8 Å and Met199 interacts with the A-ring at a distance of 3.8 Å (Figure 4-8 and Table 4-3).

Table 4-3. Interactions between atoms of the InhA protein or the cofactor NAD^+ and the atoms of the PT70 inhibitor.

InhA or cofactor residue /atom	PT70 atom	Distance (Å)
Hydrogen bond		
Tyr 158 / OH	O17	2.6
NAD / NBJ	O17	3.4
NAD / OBH	O17, O7	2.5, 3.2
Hydrophobic interaction		
Gly 96 / C	C12, C13, C14	3.6, 4.0, 3.7
Gly 96 / CA	C14	3.8
Phe 97 / CA	C11, C12	4.1, 3.6
Phe 97 / C	C11, C12	4.0, 4.0
Met 103 / CE	C10	3.9
Phe 149 / CG	C1	4.0
Phe 149 / CZ	C16, C18	3.9, 3.7
Phe 149 / CE1	C19	4.0
Ala 157 / CB	C20,C21	4.0, 3.6
Tyr 158 / CE1	C6, C1,C9	3.6, 3.4, 4.0
Tyr 158 / CZ	C6, C1	3.9, 3.8
Tyr 158 / CD1	C17, C19	4.0, 4.0
Met 161 / CG	C9, C10, C11	4.0, 3.6, 4.0
Met 161 / CE	C11, C12, C13	3.8, 3.7, 4.0
Ala 198 / CB	C8, C13, C14	3.7, 3.6, 3.4
Met 199 / CE	C16, C3	3.9, 3.8
Met 199 / CB	C3	4.0
Met 199 / CA	C4	3.9
Ile 202 / CG2	C10, C11	4.0, 3.8
Ile 202 / CD1	C11, C12	4.0, 3.8
Val 203 / CG1	C20, C21	3.6, 3.6
Val 203 / CG2	C3, C4, C9	4.0, 4.0, 3.9
Leu 218 / CD2	C18	3.7

4.5.1 Slow tight binding inhibition

The substrate binding loop is disordered in InhA structures in the presence of triclosan as well as in the structures with the improved 5PP and 8PP triclosan derivatives which, despite their

improved IC_{50} values, remain to be rapid reversible inhibitors (Sullivan, Truglio et al. 2006). In contrast, triclosan is a slow onset inhibitor of *E. coli* FabI and the corresponding crystal structure shows that the substrate binding loop is ordered (Ward, Holdgate et al. 1999). The substrate binding loop is also ordered in the crystal structure of InhA inhibited by the slow onset INH-NAD adduct (Rawat, Whitty et al. 2003; Vilcheze, Wang et al. 2006).

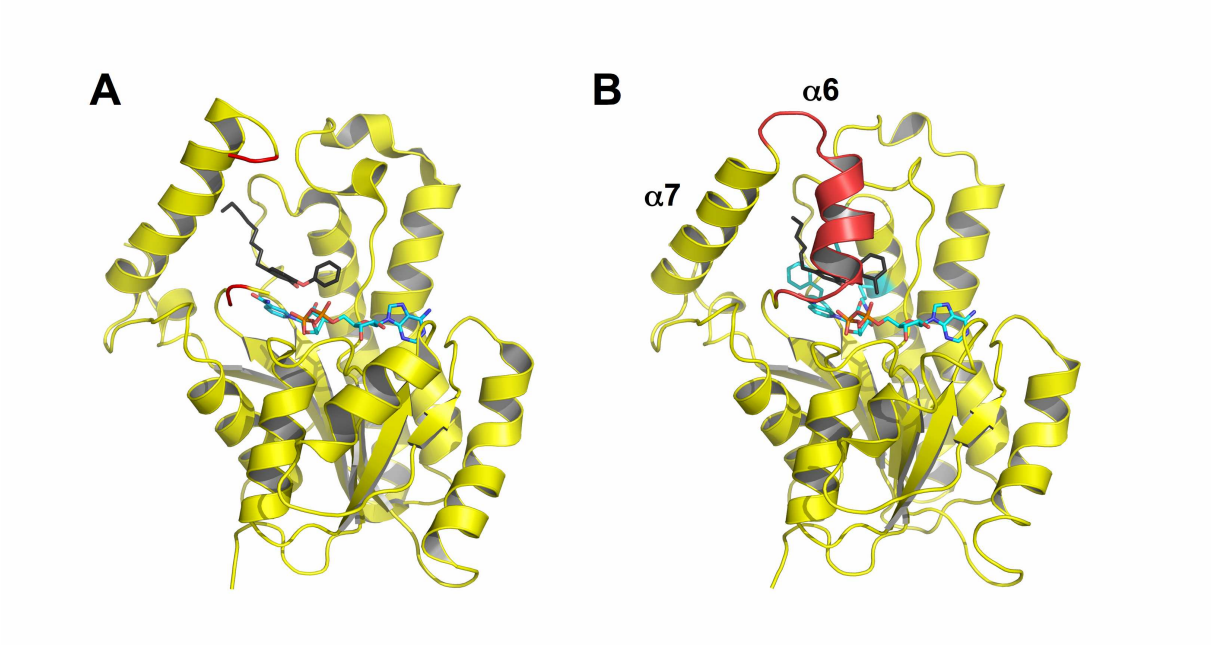


Figure 4-9. Loop ordering upon PT70 binding. (A) Monomer of InhA inhibited by 8PP (pdb code 2b37). The inhibitor is colored in black all-bonds representation, the NAD^+ cofactor is shown in cyan. Helix α6 of the substrate binding loop is disordered. The loop ends are shown in red. (B) Monomer of InhA inhibited by the slow onset inhibitor PT70, colors are chosen as in (A). Helix α6 is ordered and covers the active site where PT70 is bound

These observations indicate that compounds with the ability to cause loop ordering are slow, tight binding inhibitors of InhA (Sullivan, Truglio et al. 2006). Loop ordering with respect to INH-NAD binding remains elusive because no direct contacts between the adduct and the substrate binding loop can be observed in the crystal structures (pdb code 2idz (Dias, Vasconcelos et al. 2007) and 2nv6 (Vilcheze, Wang et al. 2006)). However, the *E. coli* FabI structure with bound triclosan clearly reveals that the inhibitor forms hydrophobic interactions with the loop-amino acids Ala196 and Ala197, which correspond to Ala198 and Met199 in *M. tuberculosis* InhA. Additionally, Rozwarski and coworkers solved the crystal structure of InhA with a bound C16 fatty acyl substrate (Rozwarski, Vilcheze et al. 1999), and could find that hydrophobic residues of the loop are important for proper substrate binding into the cavity. Interestingly, the last few carbon atoms of the fatty acid interact with the hydrophobic amino acids Ala198, Met199 and Ile202. A fatty acid shorter than 16 carbons might not be

accommodated correctly by the enzyme due to missing interactions with Ala198, Met199 and Ile202 which is in accordance with the finding that InhA and the whole FASII pathway of mycobacteria prefer fatty acids longer than 16 carbon atoms.

These observations strongly suggest that the interactions with the three amino acids (Ala198, Met199, Ile202) are important determinants for loop ordering. An inhibitor, like PT70, that is able to directly interact with these residues leads to a defined loop structure. The ordered substrate binding loop covers the entrance to the binding pocket and thereby locks the inhibitor into the cavity and increases its residence time. It is conceivable that the conformational change of the loop poses the slow step observed in the binding studies.

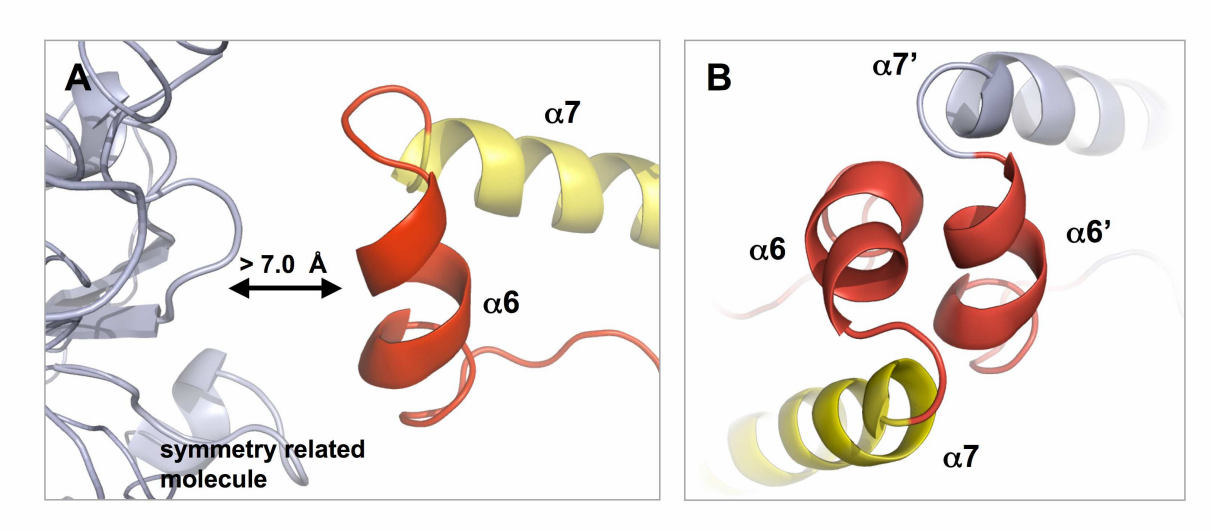


Figure 4-10. Crystal contacts of different monomers to symmetry related molecules (blue) in the region of the substrate binding loop (red and yellow). (A) No contacts are generated between the ordered loop region (red cartoon) and the closest symmetry related molecule (blue cartoon). (B) The substrate binding loop (red) of another monomer in the tetramer generates close contacts to the same residues of the closest symmetry related molecule.

Freundlich and coworkers recently reported InhA structures inhibited by triclosan derivatives (Freundlich, Wang et al. 2009) that have an ordered substrate binding loop in two of the four structures. However, it was not mentioned if the inhibitors show slow binding inhibition. The possibility that loop ordering is due to intermolecular contacts in space group $I4_122$ can not be excluded. In case of the PT70 structures, crystal contacts can be ruled out to be responsible for loop ordering, rather the hydrophobic interactions of the residues with the inhibitor itself lead to the formation of the stable structure. Close inspection of the solvent areas between the molecules in the crystal packing demonstrate, that the closest contact between two of the loops within the tetramer and a symmetry related molecule is at least ~5 Å and no salt bridges or hydrogen bonds exist between the molecules that directly stabilize the loop (Figure 4-10 A).

4.5.2 Comparison to other triclosan derivatives

A lot of progress has been made to optimize triclosan as an inhibitor of InhA. Structures are available for InhA with bound triclosan, with the improved triclosan derivative 8PP (5-octyl-2-phenoxyphenol) of Sullivan and coworkers (Sullivan, Truglio et al. 2006) and with the triclosan derivative JPL (5-(cyclohexa-1,5-dien-1-ylmethyl)-2-(2,4-dichlorophenoxy)phenol) of Freundlich and coworkers (Freundlich, Wang et al. 2009). A superposition of these inhibitors was performed to identify differences of the binding modes (Figure 4-11). While the A-ring of the inhibitors Triclosan, 8PP, JPL and PT70 are always orientated in the same way in the active site of InhA, the B-rings differs significantly from each other. These differences offer an explanation for the different affinities and binding behaviors towards InhA. The PT70 methyl group forms van-der-Waals contacts to the NAD^+ cofactor, resulting in a 1 Å shift of the B-ring upward relative to its position in 8PP (Figure 4-11B).

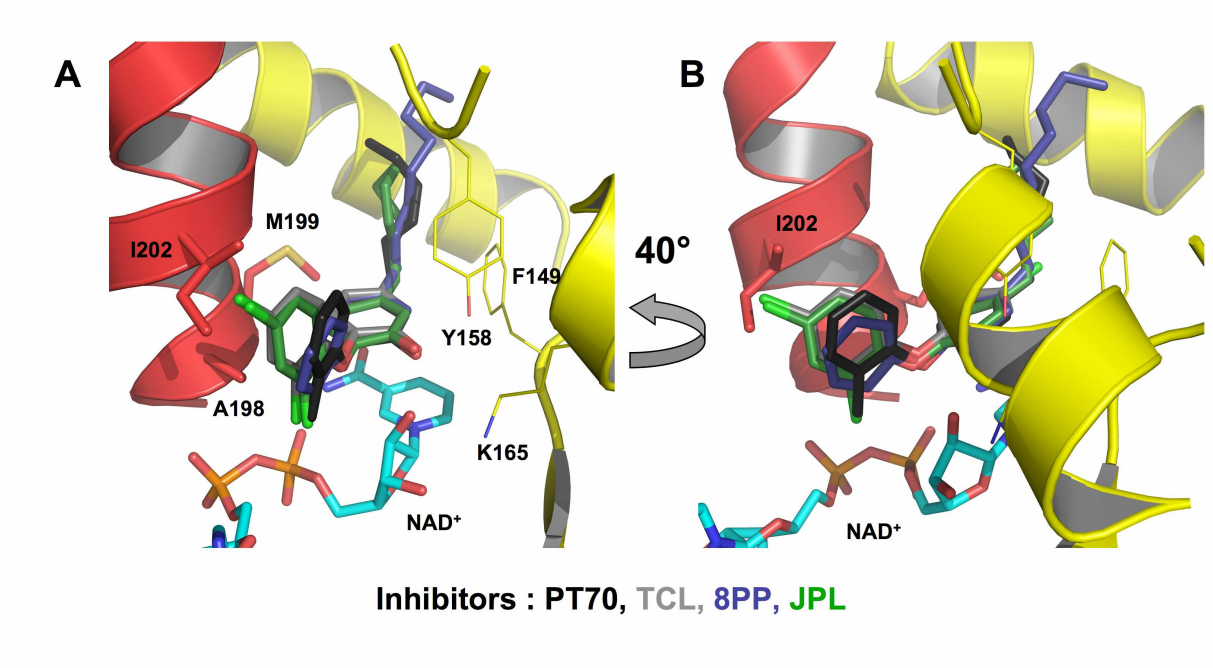


Figure 4-11.Overlay of different triclosan based inhibitors in the binding pocket of InhA (PT70 black, triclosan gray, 8PP blue and JPL green). The different orientations of the B-rings of the inhibitors are clearly visible. (A) Triclosan (TCL, gray) and JPL (green) are tilted towards the substrate binding loop (red cartoons) compared to the inhibitors 8PP and PT70. (B) PT70 (black) is shifted upwards by 1 Å relative to the position of 8PP (blue) and more comparable to the position of TCL (gray) and JPL (green).

In this orientation, the hydrophobic interactions of PT70 to the important hydrophobic residues Ala198, Met199 and Ile202 are facilitated. Additionally, the reduced flexibility of the phenyl-rings enhances the stability of the described interactions. In the Triclosan bound and JPL bound InhA structures it can be readily identified that the B-ring is tilted relative to the B-ring of PT70 by ~25° towards the substrate binding loop and thereby interferes sterically

with the loop residues, pushing the loop away from the substrate cavity instead of keeping it in place (Figure 4-11). All these observations provide an explanation why the new characteristics of PT70 facilitate the hydrophobic interactions to the important loop residues of InhA and thus make it a slow onset inhibitor.

4.6 Outlook for future drug design

4.6.1 Inhibitors of KasA

The β-ketoacyl synthase KasA in the *M. tuberculosis* fatty acid biosynthesis pathway is a validated but unexploited target for the development of novel TB chemotherapeutics. In the present work, the interaction of KasA with the natural inhibitor TLM was characterized. The crystal structures provide direct evidence why the inhibitor preferentially binds to the acylated form of KasA rather than to the apo form. It was shown that the main structural change upon KasA acylation involves a rotation of the gatekeeper residue Phe404 from a closed to an open conformation based on the presumption that replacement of the catalytic cysteine with a glutamine causes similar structural changes to KasA that occur upon acylation. These structural changes appear to be driven by the formation of hydrogen bonds between the oxyanion hole and the carbonyl groups of either the glutamine side chain or the acyl-enzyme thioester. Formation of these hydrogen bonds causes Phe404 to rotate, resulting in an increase in the size of the malonyl binding pocket and a widening of the entrance to the pocket that facilitates TLM binding. These rearrangements in KasA combined with the energetically favorable edge-to-face interaction between Phe404 and TLM provide an explanation for the increase in affinity of the inhibitor for the acylated enzyme.

Previous studies (Kim, Zhang et al. 2006) aimed towards the modification of the isoprenoid moiety of TLM due to the observation that the isoprene side chain is pointing towards a lipophilic pocket, which could be filled and thereby increase the affinity of the lead compound. Although this pocket is slightly larger in KasA compared to homologous enzymes of *E. coli*, the KasA structures clearly reveal, that this pocket can not accommodate a longer hydrophobic chain, since two important water molecules are present in all structures and stabilize the loop from Asp273 to Pro280. It was also attempted to remove the double bonds in the isoprene side chain, which again yielded no improvement of the lead compound. In the KasA structures, the isoprene side chain is sandwiched between two peptide bonds. This arrangement only allows the position of a planar group whereas non-planar groups would sterically interfere with either of the two walls formed by residues 279-280 and 403-404.

Further work to improve the thiolactomycin lead structure will have to focus on modifications of the thiolactone ring rather than the isoprene moiety.

Based on the crystal structures, another approach towards improvement of the lead compound TLM can be envisioned. The bound PEG molecules in the crystal structures directly point towards the possibility to combine the TLM molecule with a polyethylene glycol molecule of defined length. This so-called PEGylation is a common process to modify drugs or therapeutic molecules, mostly, to improve solubility and decrease immunogenicity. Additionally, PEGylation can increase the stability of the drug and the retention time of the conjugates in blood, thereby allowing a reduced dosing frequency and reduced cost (Veronese and Pasut 2005). In case of drug development against tuberculosis, where current treatment can take longer than 12 months, this improvement is very desirable. A PEG chain with a length of 40-50 atoms attached to TLM could also have the advantage of achieving increased specificity, targeting only the mycobacterial FAS-II enzymes without affecting the human system that is not able to accommodate such long substrates. Further studies will have to prove if a combined TLM-PEG molecule is able to bind to KasA and trigger the conformational change of the gatekeeper residue Phe404 to induce the favorable interaction between the TLM thiolactone ring and the phenylalanine side chain.

The crystal structures presented here are the first structures of *Mycobacterium tuberculosis* KasA and provide a good starting point for the modeling of novel inhibitors that specifically target this essential enzyme. Especially the crystal structures of KasA in complex with thiolactomycin provide important information about the composition of the binding pocket and reveal essential protein ligand interactions and the mechanism of inhibition. With this information, new lead compounds are identified by Christine Topf in the laboratories of Prof. Sotriffer and Prof. Holzgrabe (University of Würzburg) by virtual screening. A pharmacophore model based on the TLM binding mode was generated which was then used for virtual screening of databases of commercially available compounds. The identified compounds were docked into the KasA binding pocket to inspect the predicted binding modes. In addition, the chemical accessibility of possible derivatisations was checked and finally selected substances with high chemical diversity were synthesized.

To further characterize the activity of the identified compounds *in vitro*, a fluorescence assay will be used. The fluorescence assay is based on the intrinsic tryptophan fluorescence of the protein as developed for KasA (Machutta 2009). Briefly, the protein is excited at a wavelength of 290 nm and the emission is recorded at 340 nm. Upon binding of a compound

to the protein, the fluorescence is quenched which correlates to the formation of the protein-inhibitor complex. After plotting the fluorescence change against the inhibitor concentration, the dissociation constant K_d can be calculated and the different compounds can be ranked accordingly. Compounds with the potency to bind to KasA will be used for future co-crystallization experiments.

4.6.2 Inhibitors of InhA

Although the triclosan derivative PT70 shows improved *in vitro* activity compared to its predecessors 8PP or 5PP (Sullivan, Truglio et al. 2006) the compound is even more hydrophobic with a calculated logP (ClogP) value of 6.5. Drug-like molecules are expected to have ClogP values less than 5.0 according to Lipinski's rule of five (Table 4-4, (Lipinski, Lombardo et al. 2001)). Towards the improvement of the bioavailability of PT70 several structural modifications are conceivable. In collaboration with the laboratory of Prof. Peter Tonge (SUNY, Stony Brook) some modifications are currently being tested. The replacement of the alkyl chain of PT70 with a polyethoxy chain would result in a significant decrease of ClogP. As already described for the improvement of the TLM molecule PEGylation is a common modification of drugs and might result in improvement of the *in vivo* properties (chapter 4.6.1).

Table 4-4. Lipinski's rule of five to evaluate the drug likeliness of a compound

Lipinski's rule of five
Not more than 5 hydrogen bond donors
Not more than 10 hydrogen bond acceptors
A molecular weight under 500 daltons
An octanol-water partition coefficient logP of less than 5

A second possibility would be the attachment of piperazine or morpholine groups as substituents to the B-ring of the diphenyl ether and finally it will be tested if the attachment of pyridine or pyrazine groups to the diphenyl ether B-ring or the replacement of the B-ring with saturated and unsaturated nitrogen-containing rings will result in increased bioavailability without sacrificing the drug-efficacy.

To search for novel InhA inhibitors, a virtual screening approach is pursued by Constanze Waltenberger in the laboratories of Prof. Sotriffer and Prof. Schirmeister. Three pharmacophore models were defined based on the information of the target and the ligand

which was in this case the InhA structure complexed with 1-cyclohexyl-N-(3,5-dichlorophenyl)-5-oxopyrrolidine-3-carboxamide (pdb code 2h7m (He, Alian et al. 2006)). The pharmacophore models were used to screen the drug-like subset of the ZINC database and the vendor database distributed by CCG. After hierarchical filtering of the hits and visual inspection, the most suitable candidates were selected for docking experiments. The most promising hits after docking were chosen to be analyzed with respect to their inhibitory activity utilizing an enzyme activity assay. The assay is based on the decrease of the absorption value upon reduction of the cofactor NADH.

Appendix

Affidavit

I hereby declare that my thesis entitled “Towards the development of high affinity InhA and KasA inhibitors with activity against drug-resistant strains of *Mycobacterium tuberculosis*” is the result of my own work. I did not receive any help or support from commercial consultants. All sources and materials applied are listed and specified in the thesis.

Furthermore, I verify that this thesis has not yet been submitted as part of another examination process neither in identical nor in similar form.

Würzburg, ……

Date Signature

References

Adams, P. D., K. Gopal, et al. (2004). "Recent developments in the PHENIX software for automated crystallographic structure determination." Journal of Synchrotron Radiation **11**(1): 53-55.

Adams, P. D., R. W. Grosse-Kunstleve, et al. (2002). PHENIX: building new software for automated crystallographic structure determination. Acta Crystallographica Section D. **58:** 1948-1954.

Barry, C. E., 3rd, R. E. Lee, et al. (1998). "Mycolic acids: structure, biosynthesis and physiological functions." Prog Lipid Res **37**(2-3): 143-79.

Bhatt, A., L. Kremer, et al. (2005). "Conditional Depletion of KasA, a Key Enzyme of Mycolic Acid Biosynthesis, Leads to Mycobacterial Cell Lysis." J. Bacteriol. **187**(22): 7596-7606.

Bhatt, A., V. Molle, et al. (2007). "The Mycobacterium tuberculosis FAS-II condensing enzymes: their role in mycolic acid biosynthesis, acid-fastness, pathogenesis and in future drug development." Mol Microbiol **64**(6): 1442-54.

Blow, D. (2003). Outline of Crystallography for Biologists. New York, Oxford University Press.

Brennan, P. J. (2003). "Structure, function, and biogenesis of the cell wall of Mycobacterium tuberculosis." Tuberculosis (Edinb) **83**(1-3): 91-7.

Brunger, A. T. (1992). "Free R value: a novel statistical quantity for assessing the accuracy of crystal structures." Nature **355**(6359): 472-5.

CCP4 (1994). "The CCP4 suite: programs for protein crystallography." Acta Crystallographica Section D **50**(5): 760-763.

Changsen, C., S. G. Franzblau, et al. (2003). "Improved green fluorescent protein reporter gene-based microplate screening for antituberculosis compounds by utilizing an acetamidase promoter." Antimicrob Agents Chemother **47**(12): 3682-7.

Chayen, N. E. (2004). "Turning protein crystallisation from an art into a science." Curr Opin Struct Biol **14**(5): 577-83.

Cole, S. T., R. Brosch, et al. (1998). "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence." Nature **393**(6685): 537-44.

Copeland, R. A., D. L. Pompliano, et al. (2006). "Drug-target residence time and its implications for lead optimization." Nat. Rev. Drug Discov. **5**(9): 730-9.

Davis, I. W., A. Leaver-Fay, et al. (2007). "MolProbity: all-atom contacts and structure validation for proteins and nucleic acids." Nucleic Acids Res **35**(Web Server issue): W375-83.

DeLano, W. L. (2002). "The PyMOL Molecular Graphics System." from http://www.pymol.org.

Dessen, A., A. Quemard, et al. (1995). "Crystal structure and function of the isoniazid target of Mycobacterium tuberculosis." Science **267**(5204): 1638-41.

Dias, M. V., I. B. Vasconcelos, et al. (2007). "Crystallographic studies on the binding of isonicotinyl-NAD adduct to wild-type and isoniazid resistant 2-trans-enoyl-ACP (CoA) reductase from Mycobacterium tuberculosis." J Struct Biol **159**(3): 369-80.

Dorman, S. E. and R. E. Chaisson (2007). "From magic bullets back to the magic mountain: the rise of extensively drug-resistant tuberculosis." Nat Med **13**(3): 295-8.

Dundas, J., Z. Ouyang, et al. (2006). "CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues." Nucleic Acids Res **34**(Web Server issue): W116-8.

Emsley, P. and K. Cowtan (2004). "Coot: model-building tools for molecular graphics." Acta Crystallogr D Biol Crystallogr **60**(Pt 12 Pt 1): 2126-32.

Engh, R. A. and R. Huber (1991). Accurate bond and angle parameters for X-ray protein structure refinement. Acta Crystallographica Section A. **47:** 392-400.

Ericsson, U. B., B. M. Hallberg, et al. (2006). "Thermofluor-based high-throughput stability optimization of proteins for structural studies." Anal Biochem **357**(2): 289-98.

Freundlich, J. S., F. Wang, et al. (2009). "Triclosan derivatives: towards potent inhibitors of drug-sensitive and drug-resistant Mycobacterium tuberculosis." ChemMedChem **4**(2): 241-8.

Frishman, D. and P. Argos (1995). "Knowledge-based protein secondary structure assignment." Proteins **23**(4): 566-79.

He, X., A. Alian, et al. (2006). "Pyrrolidine Carboxamides as a Novel Class of Inhibitors of Enoyl Acyl Carrier Protein Reductase from Mycobacterium tuberculosis." Journal of Medicinal Chemistry **49**(21): 6308-6323.

Hunter, C. A., J. Singh, et al. (1991). "Pi-pi interactions: the geometry and energetics of phenylalanine-phenylalanine interactions in proteins." J. Mol. Biol. **218**(4): 837-46.

Jancarik, J. and S. H. Kim (1991). Sparse matrix sampling: a screening method for crystallization of proteins. Journal of Applied Crystallography. **24:** 409-411.

Jarlier, V. and H. Nikaido (1994). "Mycobacterial cell wall: structure and role in natural resistance to antibiotics." FEMS Microbiol Lett **123**(1-2): 11-8.

Kabsch, W. (1988). "Automatic indexing of rotation diffraction patterns." Journal of Applied Crystallography **21**(1): 67-72.

Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. Journal of Applied Crystallography. **26:** 795-800.

Kantardjieff, K. A. and B. Rupp (2003). "Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals." Protein Sci **12**(9): 1865-71.

Kaufmann, S. H. (2001). "How can immunology contribute to the control of tuberculosis?" Nat Rev Immunol **1**(1): 20-30.

Kaufmann, S. H. (2005). "Robert Koch, the Nobel Prize, and the ongoing threat of tuberculosis." N Engl J Med **353**(23): 2423-6.

Kauppinen, S., M. Siggaard-Andersen, et al. (1988). "beta-Ketoacyl-ACP synthase I of Escherichia coli: nucleotide sequence of the fabB gene and identification of the cerulenin binding residue." Carlsberg Res Commun **53**(6): 357-70.

Kim, P., Y. M. Zhang, et al. (2006). "Structure-activity relationships at the 5-position of thiolactomycin: An intact (5R)-isoprene unit is required for activity against the condensing enzymes from Mycobacterium tuberculosis and Escherichia coli." J. Med. Chem. **49**(1): 159-171.

Kleywegt, G. J. and T. A. Jones (1997). "Model building and refinement practice." Methods Enzymol **277**: 208-30.

Kremer, L., J. D. Douglas, et al. (2000). "Thiolactomycin and related analogues as novel anti-mycobacterial agents targeting KasA and KasB condensing enzymes in Mycobacterium tuberculosis." J. Biol. Chem. **275**(22): 16857-64.

Kremer, L., L. G. Dover, et al. (2002). "Mycolic acid biosynthesis and enzymic characterization of the beta-ketoacyl-ACP synthase A-condensing enzyme from Mycobacterium tuberculosis." Biochem J **364**(Pt 2): 423-30.

Kremer, L., L. G. Dover, et al. (2003). "Inhibition of InhA activity, but not KasA activity, induces formation of a KasA-containing complex in mycobacteria." J Biol Chem **278**(23): 20547-54.

Kursula, P. (2004). XDSi: a graphical interface for the data processing program XDS. Journal of Applied Crystallography. **37:** 347-348.

Laemmli, U. K. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." Nature **227**(5259): 680-5.

Leslie, A. G. W. (1992). Recent changes to the MOSFLM package for processing film and image plate data Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography.

Lipinski, C. A., F. Lombardo, et al. (2001). "Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings." Advanced Drug Delivery Reviews **46**(1-3): 3-26.

Machutta, C. A., Reddy, B.G., Luckner, S.R., Kapilashrami, K., Ruzsicska, B., Simmerling, C., Kisker, C., Tonge, P.J. (2009). Slow-Onset Inhibition of KasA by Thiolactomycin: Implications for Drug Design and Development. *Unpublished manuscript*.

Matsumoto, M., H. Hashizume, et al. (2006). "OPC-67683, a nitro-dihydro-imidazooxazole derivative with promising action against tuberculosis in vitro and in mice." PLoS Med **3**(11): e466.

Matthews, B. W. (1968). "Solvent content of protein crystals." J Mol Biol **33**(2): 491-7.

McCoy, A. J., R. W. Grosse-Kunstleve, et al. (2007). Phaser crystallographic software. Journal of Applied Crystallography. **40:** 658-674.

McMurry, L. M., M. Oethinger, et al. (1998). "Triclosan targets lipid synthesis." Nature **394**(6693): 531-2.

Migliori, G. B., R. Loddenkemper, et al. (2007). "125 years after Robert Koch's discovery of the tubercle bacillus: the new XDR-TB threat. Is "science" enough to tackle the epidemic?" Eur Respir J **29**(3): 423-7.

Murray, J. F. (2004). "Mycobacterium tuberculosis and the cause of consumption: from discovery to fact." Am J Respir Crit Care Med **169**(10): 1086-8.

Murshudov, G. N., A. A. Vagin, et al. (1997). "Refinement of macromolecular structures by the maximum-likelihood method." Acta Crystallogr D Biol Crystallogr **53**(Pt 3): 240-55.

Padilla, J. E. and T. O. Yeates (2003). "A statistic for local intensity differences: robustness to anisotropy and pseudo-centering and utility for detecting twinning." Acta Crystallogr D Biol Crystallogr **59**(Pt 7): 1124-30.

Parikh, S. L., G. Xiao, et al. (2000). "Inhibition of InhA, the enoyl reductase from Mycobacterium tuberculosis, by triclosan and isoniazid." Biochemistry **39**(26): 7645-50.

Petrek, M., M. Otyepka, et al. (2006). "CAVER: a new tool to explore routes from protein clefts, pockets and cavities." BMC Bioinformatics **7**: 316.

Pflugrath, J. W. (1999). "The finer things in X-ray diffraction data collection." Acta Crystallogr D Biol Crystallogr **55**(Pt 10): 1718-25.

Pieters, J. (2008). "Mycobacterium tuberculosis and the macrophage: maintaining a balance." Cell Host Microbe **3**(6): 399-407.

Price, A. C., K. H. Choi, et al. (2001). "Inhibition of beta-ketoacyl-acyl carrier protein synthases by thiolactomycin and cerulenin. Structure and mechanism." J. Biol. Chem. **276**(9): 6551-9.

Ramachandran, G. N., C. Ramakrishnan, et al. (1963). "Stereochemistry of polypeptide chain configurations." J Mol Biol **7**: 95-9.

Rawat, R., A. Whitty, et al. (2003). "The Isoniazid-NAD Adduct is a Slow, Tight-Binding Inhibitor of InhA, the Mycobacterium Tuberculosis Enoyl Reductase; Adduct Affinity and Drug Resistance." Proc. Nat. Acad. Sci. U.S.A. **100**: 13881-13886.

Rozwarski, D. A., C. Vilcheze, et al. (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): 15582-9.

Sacchettini, J. C., E. J. Rubin, et al. (2008). "Drugs versus bugs: in pursuit of the persistent predator Mycobacterium tuberculosis." Nat Rev Microbiol **6**(1): 41-52.

Sachdeva, S., F. N. Musayev, et al. (2008). "Separate entrance and exit portals for ligand traffic in Mycobacterium tuberculosis FabH." Chem Biol **15**(4): 402-12.

Sasaki, H., H. Oishi, et al. (1982). "Thiolactomycin, a new antibiotic. II. Structure elucidation." J Antibiot (Tokyo) **35**(4): 396-400.

Schaeffer, M. L., G. Agnihotri, et al. (2001). "Purification and biochemical characterization of the Mycobacterium tuberculosis beta-ketoacyl-acyl carrier protein synthases KasA and KasB." J. Biol. Chem. **276**(50): 47029-37.

Schatz, A., E. Bugie, et al. (2005). "Streptomycin, a substance exhibiting antibiotic activity against gram-positive and gram-negative bacteria. 1944." Clin Orthop Relat Res(437): 3-6.

Shah, N. S., A. Wright, et al. (2007). "Worldwide emergence of extensively drug-resistant tuberculosis." Emerg Infect Dis **13**(3): 380-7.

Sippl, M. J., S. J. Suhrer, et al. (2008). "A discrete view on fold space." Bioinformatics **24**(6): 870-1.

Sippl, M. J. and M. Wiederstein (2008). "A note on difficult structure alignment problems." Bioinformatics **24**(3): 426-7.

Slayden, R. A. and r. Barry CE (2002). "The role of KasA and KasB in the biosynthesis of meromycolic acids and isoniazid resistance in *Mycobacterium tuberculosis*." Tuberculosis **81**(1): 1-12.

Slayden, R. A., R. E. Lee, et al. (1996). "Antimycobacterial action of thiolactomycin: an inhibitor of fatty acid and mycolic acid synthesis." Antimicrob. Agents Chemother. **40**(12): 2813-9.

Snapper, S. B., R. E. Melton, et al. (1990). "Isolation and characterization of efficient plasmid transformation mutants of Mycobacterium smegmatis." Mol Microbiol **4**(11): 1911-9.

Sridharan, S., L. Wang, et al. (2007). "X-ray crystal structure of Mycobacterium tuberculosis beta-ketoacyl acyl carrier protein synthase II (mtKasB)." J. Mol. Biol. **366**(2): 469-80.

Sullivan, T. J., J. J. Truglio, et al. (2006). "High affinity InhA inhibitors with activity against drug-resistant strains of Mycobacterium tuberculosis." ACS Chem Biol **1**(1): 43-53.

TB Alliance, (2008). "Pa-824." Tuberculosis (Edinb) **88**(2): 134-6.

TB Alliance, (2008). "Tmc-207." Tuberculosis (Edinb) **88**(2): 168-9.

Tonge, P. J. (2000). "Another brick in the wall." Nat Struct Biol **7**(2): 94-6.

Tonge, P. J., C. Kisker, et al. (2007). "Development of modern InhA inhibitors to combat drug resistant strains of Mycobacterium tuberculosis." Curr Top Med Chem **7**(5): 489-98.

Toungoussova, O. S., A. O. Mariandyshev, et al. (2005). "Resistance of multidrug-resistant strains of Mycobacterium tuberculosis from the Archangel oblast, Russia, to second-line anti-tuberculosis drugs." Eur J Clin Microbiol Infect Dis **24**(3): 202-6.

Veronese, F. M. and G. Pasut (2005). "PEGylation, successful approach to drug delivery." Drug Discovery Today **10**(21): 1451-1458.

Veyron-Churlet, R., S. Bigot, et al. (2005). "The biosynthesis of mycolic acids in Mycobacterium tuberculosis relies on multiple specialized elongation complexes interconnected by specific protein-protein interactions." J Mol Biol **353**(4): 847-58.

Veyron-Churlet, R., O. Guerrini, et al. (2004). "Protein-protein interactions within the Fatty Acid Synthase-II system of Mycobacterium tuberculosis are essential for mycobacterial viability." Mol Microbiol **54**(5): 1161-72.

Vilcheze, C., H. R. Morbidoni, et al. (2000). "Inactivation of the inhA-encoded fatty acid synthase II (FASII) enoyl-acyl carrier protein reductase induces accumulation of the FASI end products and cell lysis of Mycobacterium smegmatis." J Bacteriol **182**(14): 4059-67.

Vilcheze, C., F. Wang, et al. (2006). "Transfer of a point mutation in Mycobacterium tuberculosis inhA resolves the target of isoniazid." Nat Med **12**(9): 1027-9.

Vissa, V. D. and P. J. Brennan (2001). "The genome of Mycobacterium leprae: a minimal mycobacterial gene set." Genome Biol **2**(8): REVIEWS1023.

Wang, J., S. M. Soisson, et al. (2006). "Platensimycin is a selective FabF inhibitor with potent antibiotic properties." Nature **441**(7091): 358-61.

Ward, W. H., G. A. Holdgate, et al. (1999). "Kinetic and structural characteristics of the inhibition of enoyl (acyl carrier protein) reductase by triclosan." Biochemistry **38**(38): 12514-25.

WHO (2007). "World Health Organization Global Tuberculosis Control: surveillance, planning, financing." WHO Report www.who.int/tb/publications/global_report/2007/pdf/full.pdf.

Witkowski, A., A. K. Joshi, et al. (1999). "Conversion of a beta-ketoacyl synthase to a malonyl decarboxylase by replacement of the active-site cysteine with glutamine." Biochemistry **38**(36): 11643-50.

Yeates, T. O. and B. C. Fam (1999). "Protein crystals and their evil twins." Structure **7**(2): R25-9.

Zhang, Y., B. Heym, et al. (1992). "The catalase-peroxidase gene and isoniazid resistance of Mycobacterium tuberculosis." Nature **358**(6387): 591-3.

Zignol, M., M. S. Hosseini, et al. (2006). "Global incidence of multidrug-resistant tuberculosis." J Infect Dis **194**(4): 479-85.

Zwart, P. H., P. V. Afonine, et al. (2008). "Automated structure solution with the PHENIX suite." Methods Mol Biol **426**: 419-35.

Zwart, P. H., R. W. Grosse-Kunstleve, et al. (2008). Surprises and pitfalls arising from (pseudo)symmetry. Acta Crystallographica Section D. **64:** 99-107.