Tomasz Manszewski

Structural studies of S-adenosyl-L-homocysteine hydrolase from symbiotic nitrogen-fixing bacteria

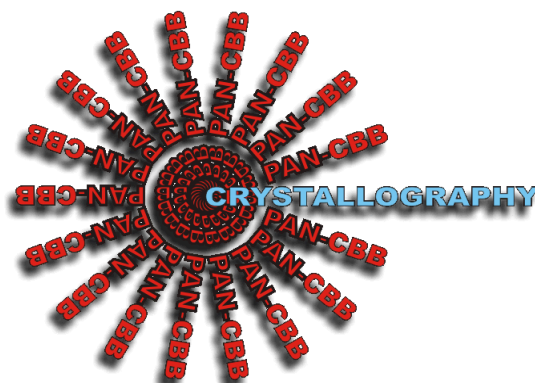*Thesis presented to the Scientific Council of the*

Institute of Bioorganic Chemistry

Polish Academy of Sciences in Poznan

as a Ph.D. dissertation

Poznań 2017

To my beloved wife

Joanna

The X-ray diffraction data were collected at

EMBL, DESY in Hamburg

and

HZB, BESSY in Berlin

Protein crystal structures deposited in the Protein Data Bank (PDB) as a result of this research:

4lvc   *Crystal structure of S-adenosyl-L-homocysteine hydrolase from Bradyrhizobium elkanii in complex with adenosine*

5m5k   *S-adenosyl-L-homocysteine hydrolase from Bradyrhizobium elkanii in complex with adenosine and cordycepin*

5m65   *Crystal structure of S-adenosyl-L-homocysteine hydrolase from Bradyrhizobium elkanii in complex with adenine*

5m66   *Crystal structure of S-adenosyl-L-homocysteine hydrolase from Bradyrhizobium elkanii in complex with adenosine*

5m67   *Crystal structure of S-adenosyl-L-homocysteine hydrolase from Bradyrhizobium elkanii in complex with adenine and 2'-deoxyadenosine*

THIS THESIS IS BASED ON THE FOLLOWING PUBLICATIONS:

1. **Manszewski T**, Singh K, Imiołczyk B & Jaskólski M (2013)
   Structural enzymology at the legume-microbe interface: S-adenosyl-L-homocysteine hydrolase of rhizobia. *Biotechologia* **91**, 38-39.

2. **Manszewski T**, Singh K, Imiołczyk B & Jaskólski M (2015)
   An enzyme captured in two conformational states: Crystal structure of S-adenosyl-L-homocysteine hydrolase from Bradyrhizobium elkanii. *Acta Cryst.* D**71**, 2422-2432.

3. **Manszewski T**, Szpotkowski K & Jaskólski M (2017)
   Crystallographic and SAXS studies on S-adenosyl-L-homocysteine hydrolase from Bradyrhizobium elkanii. *IUCrJ*, in press.

TABLE OF CONTENTS

1. ABSTRACT

The subject of the present structural investigations was S-adenosyl-L-homocysteine hydrolase (SAHase) from the symbiotic bacterium *Bradyrhizobium elkanii* (BeSAHase). The enzyme is crucial for sustaining the cellular methylation processes that depend on SAM (S-adenosyl-L-methionine). In humans, SAHase plays an important role in regulating homocysteine level in plasma, which - when elevated - is a risk factor in a number of diseases.

During this project, a high-performance protocol for enzyme overexpression, as well as procedures for ligand-free enzyme preparation in addition to optimal crystallization procedures, were developed. Five crystal structures of BeSAHase in complex with different ligands, such as adenosine, 2'-deoxyadenosine, 3'-deoxyadenosine and adenine, have been solved and refined.

During analysis of the first structure, which turned out to be a tetramer, adenosine molecules were found in the active sites of three subunits, despite the absence of adenosine during all of the purification and crystallization steps. Moreover, it turned out that the ligand-complexed molecules were in a different conformation, which was found to be closed, than the ligand-free subunit, which showed an open conformational state.

The second structure was determined from a protein sample incubated with an excess of 3'-deoxyadenosie. In this case, the desired ligand molecule was found only in one out of the four subunits, while the remaining three subunits were complexed, once again, with adenosine molecules.

The determination of the next three structures was preceded by an additional step in the purification protocol, which made it possible to prepare a ligand-free enzyme for incubation with the desired ligands. This is how the structures with adenosine, adenine and mixed adenine/2'-deoxyadenosine ligands were determined.

The crystal structures of complexes of BeSAHase with a number of different ligands became the subject of careful analysis and comparisons. During these studies, a new conformational state of SAHase, namely the semi-open state, was discovered. This state was observed for the protein complexed with 3'-deoxyadenosine. The presented studies also shed new light on the mechanisms of a molecular gate, which opens and shuts a channel leading to the active site. Also the role of monovalent cations (ammonium or sodium) in stabilizing the SAHase structure has been elucidated.

In summary, the work has revealed new structural aspects of S-adenosyl-L-homocysteine hydrolase showing in atomic detail the conformational transformations required for its enzymatic activity.

2. ABSTRACT IN POLISH

Przedmiotem badań strukturalnych była hydrolaza S-adenozylo-L-homocysteiny (SAHaza) z symbiotycznej bakterii *Bradyrhizobium elkanii* (BeSAHaza). Enzym ten reguluje komórkowe procesy metylacji zależne od SAM (S-adenozylo-L-metioniny). U człowieka SAHaza odgrywa ważną rolę w regulacji poziomu homocysteiny (Hcy) w osoczu, który - gdy jest podwyższony - jest przyczyną wielu chorób.

Podczas realizacji projektu opracowany został wysokowydajny protokół nadekspresji białka, jak również procedura otrzymywania enzymu w formie wolnej od jakichkolwiek ligandów oraz warunki krystalizacji. Rozwiązanych oraz udokładnionych zostało pięć struktur krystalicznych BeSAHazy w kompleksie z takimi ligandami jak adenozyna, 2'-deoksyadenozyna, 3'-deoksyadenozyna oraz adenina.

W trakcie analizy pierwszej ze struktur, która okazała się tetramerem, zauważono, że cząsteczki adenozyny znajdują się w miejscach aktywnych trzech podjednostek, pomimo że adenozyna nie była dodawana ani w trakcie oczyszczania, ani w trakcie krystalizacji. Co więcej, podjednostki, w których związane były cząsteczki liganda, miały inną konformację, która została określona jako zamknięta, niż podjednostka z wolnym miejscem aktywnym, której konformacja była otwarta.

Druga struktura została określona dla białka inkubowanego z nadmiarem 3'-deoksyadenozyny. W tym przypadku dodawany ligand zlokalizowany został tylko w jednej z czterech podjednostek, podczas gdy pozostałe podjednostki znów związały cząsteczki adenozyny w miejscach aktywnych.

Określenie pozostałych trzech struktur zostało poprzedzone wprowadzeniem dodatkowego etapu w protokole oczyszczania. Pozwoliło to na uzyskanie białka wolnego od jakichkolwiek ligandów, które mogło być użyte do inkubacji z pożądanymi ligandami. W ten sposób zostały określone i udokładnione struktury kompleksów z adenozyną, adeniną oraz kompleksu mieszanego z adeniną i 2'-deoksyadenozyną.

Wszystkie wyznaczone struktury krystaliczne BeSAHazy stały się podstawą szczegółowej analizy i wzajemnego porównania. W wyniku tych analiz opisana została nowa konformacja SAHazy, półotwarta, którą przyjął enzym w kompleksie z 3'-deoksyadenozyną. Przeprowadzone badania rzuciły również nowe światło na mechanizm działania "bramki molekularnej", która otwiera i zamyka kanał prowadzący do miejsca aktywnego. Wyjaśniona została również rola kationów (amonowego i sodowego) w stabilizacji struktury białka.

Podsumowując, odkryte zostały nowe aspekty strukturalne hydrolazy S-adenozylo-L-homocysteiny, dające wgląd na poziomie atomowym w zmiany, jakim ulega to białko w trakcie przemian konformacyjnych niezbędnych dla jego aktywności enzymatycznej.

3. AIMS OF THE THESIS

In this thesis I have presented the structural investigations on S-adenosyl-L-homocysteine hydrolase (SAHase) from *Bradyrhizobium elkanii* (BeSAHase).

The main goal of this work has been to determine the crystal structure of BeSAHase, the first crystal structure of SAHase from a nodulating bacterium, and to analyze the structural features of the protein.

On the way to solve the main goal of this work, I also had to solve many smaller problems. The most important one, among these minor issues, was to develop and optimize overexpression and purification protocols. These were important since, during the preparation of ligand-free protein, I was losing most of the enzyme because of poor yields. The next thing was to optimize the initial crystallization conditions to get good quality crystals for X-ray diffraction experiments.

To meet the main goal of this thesis, I also had to analyze five crystal structures of BeSAHase. Each structure determination was associated with the refinement of many thousands atomic parameters of atoms forming, in total, 18 subunits of the enzyme, as well as many ligand molecules. These five crystal structures of BeSAHase became the basis for the structural comparisons.

4. INTRODUCTION

Cellular methylation reactions of compounds such as nucleic acids, polysaccharides, lipids and proteins, most commonly use S-adenosyl-L-methionine (SAM) as the methyl group donor (Cantoni, 1975; Richards *et al*., 1978). During the methyl group transfer from SAM, S-adenosyl-L-homocysteine (SAH) is formed (Chiang & Cantoni, 1979; Cantoni & Chiang, 1980; Liu *et al*., 1992). SAH, in turn, is a strong inhibitor of SAM-dependant methyltransferases and needs to be removed from the reaction equilibrium. This is the moment when S-adenosyl-L-homocysteine hydrolase (SAHase) comes to the fore. SAHase splits SAH into adenosine (Ado) and homocysteine (Hcy), regulating the activity of SAM-dependant methyltranferases (Backlund *et al*., 1986) by maintaining the appropriate SAM:SAH ratio. Because SAHase also regulates the Hcy level, which when abnormally elevated seems to be an important independent risk factor for cardiovascular disease (Moller *et al*., 1997), coronary heart disease (Nygard *et al*., 1997), neural tube defects (Jakubowski, 2006) and as well as being a possible risk factor for amyloid diseases (White *et al*., 2001), the enzyme is an attractive target for antiparasitic, antivirial (Turner *et al*., 2000) and antitumor (Suarez & Chagoya de Sanchez, 1997) drugs.

SAHase was first identified in the late 1950s by De La Haba and Cantoni (1959). The structural knowledge of SAHases is based on the crystallographic models available in the Protein Data Bank (PDB; Berman *et al*., 2000), which include proteins from organisms such as: *H. sapiens* (PDB ID: 1a7a; Turner *et al*, 1998) *R. norvegicus* (PDB ID: 1b3r; Hu *et al*, 1999), *L. luteus* (PDB ID: 3ond, 3one, 3onf; Brzeziński *et al*., 2012), *M. tuberculosis* (PDB ID: 3ce6; Reddy *et al*., 2008) and *P. falciparum* (PDB ID: 1v8b; Tanaka *et al*., 2004).

SAHases are typically active as homotetramers with the exception of plant proteins, which are found to function as homodimers (Guranowski & Pawełkiewicz, 1977; Brzeziński *et al*., 2008). Each enzyme subunit, built of ~420 (archaeal and some bacterial) to ~480 (most eukaryotic and bacterial) amino acid residues, is divided into two main domains (Fig. 1), the substrate-binding domain and the cofactor-binding domain, and a small C-terminal domain which plays a role in dimerization (Manszewski *et al*., 2015). Moreover, two hinge regions, connecting the substrate-binding and cofactor-binding domains were distinguished by Manszewski *et al*. (2015). There is also

a specific insert in the substrate-binding domain of the eukaryotic and bacterial enzymes, making them slightly longer (Stępkowski *et al*., 2005). The enzymatic mechanism of SAHase involves a redox step (Palmer & Abeles, 1979; Fujioka & Takata, 1981) and, therefore, the protein needs NAD^+ as a cofactor, which is non-covalently bound in each protomer in a 1:1 ratio.

The two principal domains, the substrate-biding domain and the cofactor-binding domain, can oscillate between two major conformational forms. The first one, the so-called closed state, is achieved when the domains are close together and it is thought to be stabilized by a ligand molecule (a substrate, product or inhibitor) bound in the active site (Hu *et al*., 1999; Yin *et al*., 2000) and an alkali metal or ammonium cation coordinated in a metal-binding loop, overlapping with the second hinge region (Manszewski *et al.,* 2015), near the active site (Brzeziński *et al*., 2012; Manszewski *et al*., 2015). In the second state, the so-called open state, which is characteristic of ligand-free enzymes, these two domains are separated and form a substrate access channel leading to the active site. However, Zheng *et al*. (2015) observed that the ligand-free enzyme can also exist in the closed conformation.

SAHase had already been studied in the Center for Biocrystallographic Research for several years, when this project started. The first single crystals of SAHase were grown by Brzeziński *et al*. (2008) for an enzyme from the legume plant *Lupinus luteus*, and their structure was solved at near-atomic resolution (Brzeziński *et al*., 2012) in our group. Structural studies of the same enzyme from the symbiotic partner - the nodulating bacteria *Bradyrhizobium elkanii* (BeSAHase) - were undertaken in this thesis to provide data for an orthologous enzyme for structural comparisons. This aspect is important in the light of the fact that the plant-bacteria symbiosis starts with the activation of the bacterial genes responsible for the biosynthesis of the nodulation (Nod) factor (NF). NodS, an enzyme involved in NF biosynthesis, is a SAM-dependant methyltransferase, whose activity is regulated through maintaining the appropriate SAM:SAH ratio, which is controlled by SAHase.

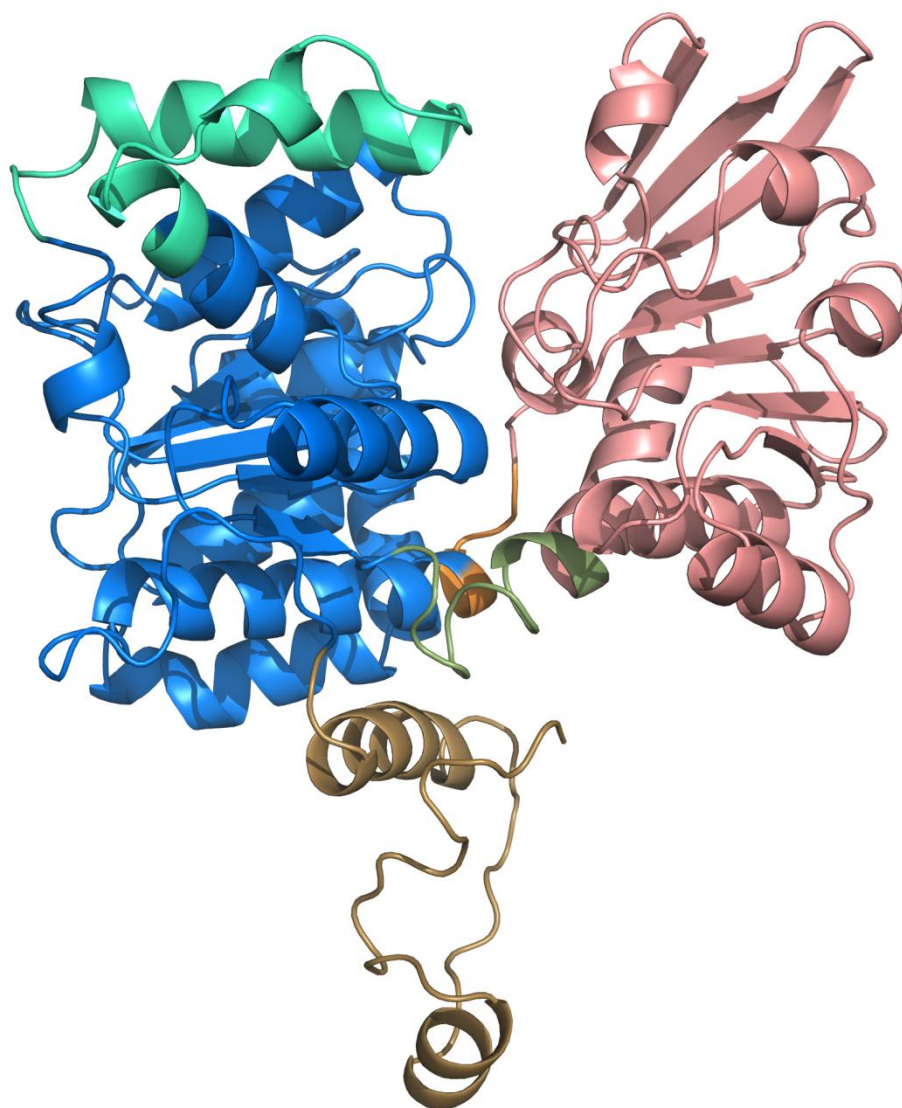Fig. 1. The structure of one subunit of SAHase from *B. elaknii*. The domains are color-coded as follows: blue, substrate-binding domain; salmon, cofactor-binding domain; sand, C-terminal domain. The hinge regions, corresponding to residues Asn222-Leu233 and His392-Val396, are shown in green and orange, respectively, and the insertion segment, located in the substrate-binding domain, is shown in cyan.

5. MATERIALS AND METHODS

5.1. Cloning and overexpression

The coding sequence of BeSAHase was amplified from *B. elkanii* cDNA library by PCR and cloned into pET151/D-TOPO plasmid, which adds a His_6 tag at the N-terminus of the expressed sequence.

Overexpression of the enzyme was carried out in BL21 Star (DE3) *E. coli* cells. The bacteria were cultured with shaking (230 rpm) at 310 K in LB media, containing 100 μg mL^{-1} ampicillin, until the OD_{600} reached 1.0. After that isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to the final concentration of 0.4 mM and the mixture was incubated for 4h at 310 K.

5.2. Purification

Bacterial cells, harvested after incubation, were resuspended in buffer *A* (50 mM Tris pH 8.0, 500 mM NaCl, 20 mM imidazole), disrupted by pulse sonication and centrifuged. The supernatant was loaded onto an Ni-NTA affinity column and the protein was eluted with buffer *B* (50 mM Tris pH 8.0, 500 mM NaCl, 300 mM imidazole). Next, the protein was dialyzed against buffer *A* and the His tag was removed by TEV protease. After dialysis, the protein was loaded onto desalting column equilibrated with buffer *C* (20 mM Tris pH 8.0, 50 mM NaCl) and then incubated on ice with 12-fold molar excess of NAD^+. As the last step, the whole mixture was loaded onto a gel-filtration column equilibrated with buffer *C*.

5.3. Ligand exchange procedure

To obtain ligand-free protein ready for incubation with desired ligands, one additional step in the purification protocol, a modified procedure of Yuan *et al*. (1993), was implemented before size-exclusion chromatography. Briefly, 5 mL of BeSAHase in buffer *A*, concentrated to 12 mg mL^{-1}, were mixed with 10 mL of saturated ammonium sulfate solution (pH 3.3) and stored on ice for 10 min. The mixture was centrifuged, the precipitate was dissolved in 5 mL of buffer *A*, and mixed again with ammonium sulfate solution as above. The pellet was again dissolved in 5 mL of buffer *A* and mixed with 10 mL of saturated ammonium sulfate solution of pH 7.0. Finally, the pellet was dissolved in 5 mL of buffer *A* and the protein concentration was adjusted to 8 mg mL^{-1}.

Ultimately, NAD$^+$ was added in 12-fold molar excess and the mixture was stored at 277 K for 12 h.

5.4. Crystallization and data collection

The crystallization trials were conducted by vapor diffusion in hanging or sitting drops at 292 K. Some samples were also crystallized at the High-Throughput Crystallization Facility at EMBL Hamburg (Müller-Dieckmann, 2006). Commercially available crystal screens (PEG/Ion, Hampton Research and Morpheus, Molecular Dimensions) as well as in-house prepared optimized conditions were used.

Low-temperature X-ray diffraction data were collected at BESSY, Berlin, beamlines 14.1 and 14.2, and at EMBL/DESY, Hamburg, Petra III beamlines P13 and P14.

5.5. Difficulties concerning overexpression and purification

During my PhD research I have encountered several difficulties. It may be of interest to discuss them briefly here.

The first difficulty was encountered when I was precipiting the protein with ammonium sulfate; originally, during this procedure the protein loss was nearly 90%. I had two possible ways out: (1) to modify the precipitation protocol to minimize the protein loss; or (2) to accept the poor yield and compensate for it by upscaling the protein overexpression. I modified the precipitation protocol by increasing the protein concentration from 6 to 12 mg mL^{-1} and by incubating the mixture on ice for 30 min instead of mixing it. This increased the yield of the purification procedure to nearly 70%. While working on the optimization of the overexpression protocol, I found out that the concentration of IPTG (up to 1 mM) was too high and that most of the protein was in insoluble form. Through a series of optimization steps, I was able to estimate that the best level of IPTG concentration was 0.4 mM. With those improvements to the overexpression and purification protocol, I was able to obtain more than 100 mg of pure protein from 1 L of bacterial culture.

Another problem was encountered during protein crystallization. When I obtained the crystals of the Ado/Cord complex, which diffracted to 1.84 Å resolution, I set up an optimization experiment starting from the original conditions (0.2 M sodium

acetate, 16% PEG 4000 and 0.1 M Tris pH 8.5). At first, I obtained many crystals, but their diffraction quality was very poor (weak diffraction of ~4 Å or no diffraction at all). I noticed, however, that the poor-quality crystals from the first crop, spontaneously dissolve after 1 or 2 weeks. After another 2 weeks small, well-shaped crystals of a new crop started to grow. Harvesting crystals from the second crop gave me the best specimens of BeSAHase in complex with 2'-deoxydenosine, which diffracted X-rays to 1.54 Å.

6. REVIEW OF THE PUBLICATIONS CONTAINED IN THIS THESIS

6.1. First structural insight into BeSAHase

The first X-ray diffraction dataset was collected for BeSAHase at BESSY, Berlin, using a single crystal grown by Kriti Singh, my predecessor on this project. The structure was solved at 1.74 Å resolution by molecular replacement, using chain A of SAHase from *B. melitensis* (PDB ID: 3n58; unpublished work) as a search model. It revealed a tetrameric protein in the asymmetric unit.

Difference electron density maps calculated after the first round of refinement showed that one NAD^+ molecule is bound in each subunit and that there is additional electron density present in the active sites of three (out of four) subunits. Careful analysis strongly suggested that this additional electron density corresponds to adenosine (Ado) molecules (Fig. 2), which was rather strange, because no adenosine had been added during any of the purification or crystallization steps. This observation suggests that the nucleoside was sequestered by the protein from the bacterial cell during overexpression.
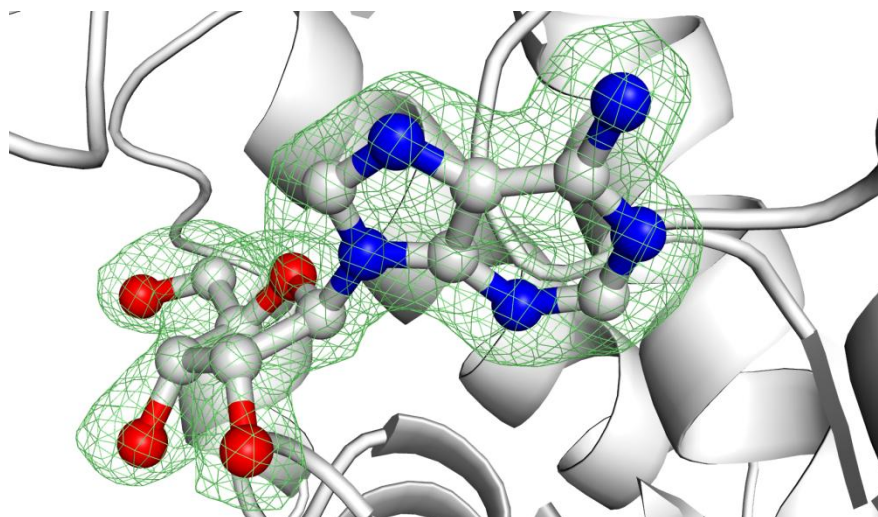
Fig. 2. Adenosine molecule in the active site of subunit A of the first BeSAHase structure (4lvc) in F_o-F_c OMIT electron-density map contoured at the 3σ level.

These first results, along with the information that crystals of other BeSAHase complexes had been obtained, were described in a short review paper (Manszewski *et al.*, 2013), which is the first part of this thesis.

During the next phase of the project, I carefully analyzed the first BeSAHase structure and the results were described in the second paper (Manszewski *et al.*, 2015)

published in *Acta Cryst. D.* The corresponding atomic coordinates and structure factors were deposited in PDB with accession number 4lvc.

6.2. Adenosine molecules and their influence on BeSAHase conformation

In the 4lvc crystal structure, Ado molecules were found in three out of four subunits of the BeSAHase tetramer. Since the conformational state of the subunit is expected to be affected by the presence of the ligand (Hu *et al.*, 1999; Yin *et al.*, 2000), it was clear to me that the Ado-free subunit should have a different conformation. A comparison of the conformation of the protein chains by means of root-mean-square (rms) deviations calculated for the Cα atom positions showed that the protomers with the Ado molecules bound in the active site adopt the same conformation, with rms deviations ranging from 0.12 to 0.21 Å, while the subunit with the empty active site differs from the others by 2.26-2.32 Å. In addition, I was able to delineate two amino acid residues, His342-Phe343, whose side chains form a characteristic molecular gate (MG). The presence of the molecular gate had been already noted earlier by Reddy *et al.* (2008) in SAHase from *M. tuberculosis* (MtSAHase) and by Brzeziński *et al.* (2012) in SAHase from *L. luteus* (LlSAHase). It is of note that in MtSAHase only one amino acid residue - His363 (corresponding to His342 in BeSAHase and His350 in LlSAHase) - was considered to form the molecular gate, while in LlSAHase and in BeSAHse one additional residue was also implicated in this mechanism. This
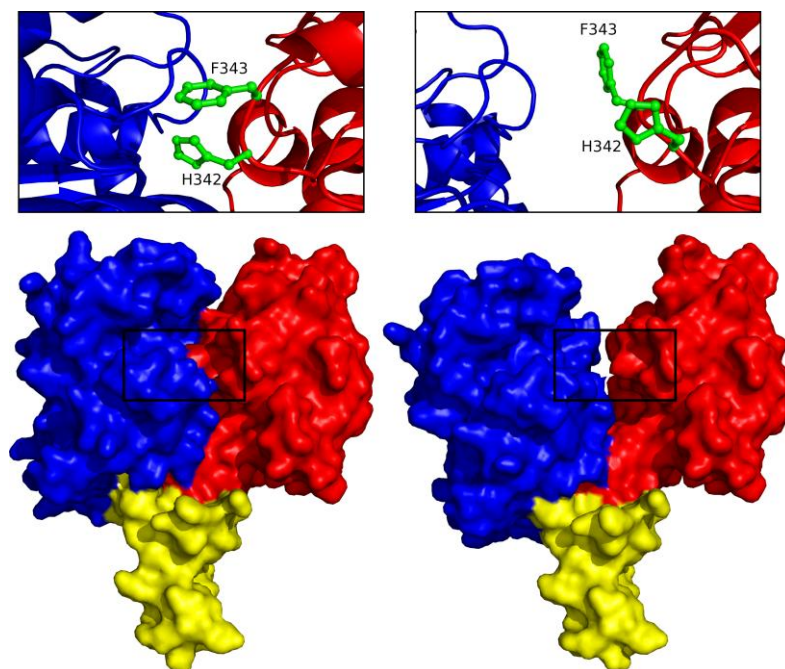
Fig. 3. BeSAHase subunits from the 4lvc structure in the closed (left) and open (right) conformation. The domains are color-coded as follows: blue, substrate-binding domain; red, cofactor-binding domain; yellow, C-terminal dimerization domain. The side chains of His342-Phe343, forming the molecular gate, are highlighted in the enlargements.

residue is, respectively, Phe343 in BeSAHase and Phe351 in LlSAHase. The torsion angles of the residues forming the MG element suggested that the open/shut state of the molecular gate is generally connected with the open/closed state of the protein (Fig. 3), so that the ligand bound in the active site is additionally blocked by the side chains of the MG residues.

Thanks to the fact that two different conformational states were observed in one structure, I was able to investigate how this transition affects the values of the torsion angles at the joints between the substrate-binding and cofactor-binding domains. Since the cofactor-binding domain is, in some sense, an insert in the substrate-binging domain, I was looking for two hinge regions. Careful analysis of the main-chain torsion angles allowed me to distinguish these two regions, which corresponded to amino acid residues Asn222-Leu233 and His392-Val396 (green and orange, respectively, in Fig. 1). In these regions the torsion angles differ by more than 10° between the subunits in the closed and open states.
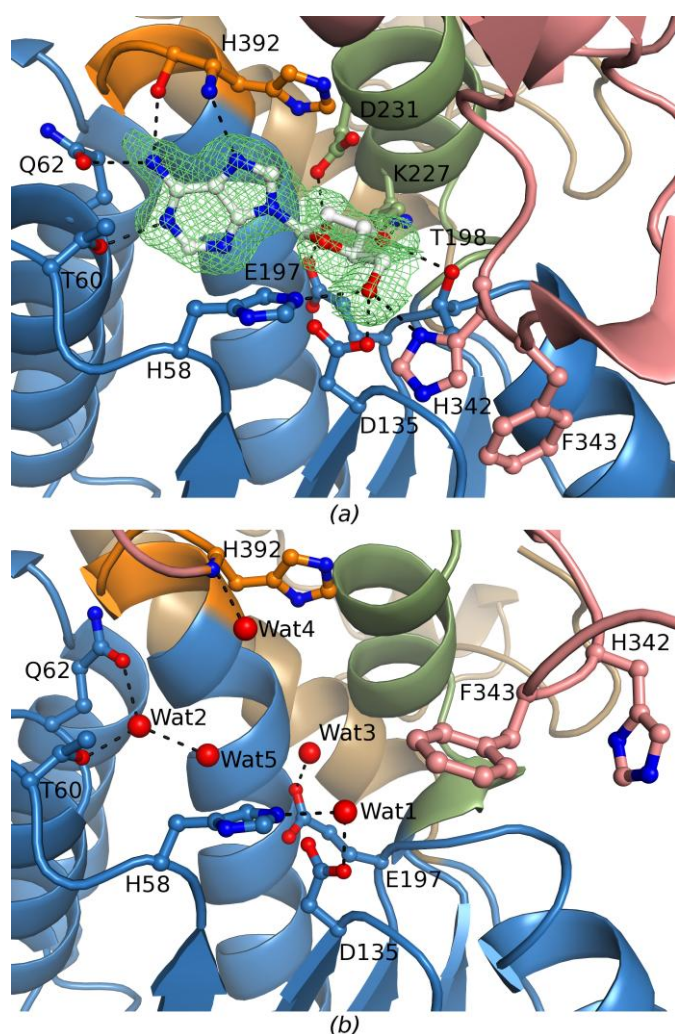
Fig. 4. Comparison of the active site in subunit A (closed form) (a) and subunit D (open form) (b) from structure 4lvc. The Ado molecule in (a) is shown in F_o-F_c OMIT electron-density map contoured at the 3σ level. Water molecules are represented as red spheres and dashed lines represent hydrogen bonds. The domains are color-coded as follows: blue, substrate-binding domain; salmon, cofactor-binding domain; sand, C-terminal domain.

6.3. The active site in Ado-complexed and ligand-free subunits of BeSAHase

The structure 4lvc also gave me the opportunity to compare the active sites of BeSAHase when a ligand is bound in the active site and when the active site is empty (Fig. 4). Although the active sites of the

complexed and ligand-free proteins cannot be compared directly because the corresponding residues are too far apart, the general hydrogen bond network in the open subunit is preserved thanks to the presence of five water molecules, four of which mimic the positions of the heteroatoms of the Ado molecule. In addition, it is noteworthy that all of the Ado heteroatoms, with the exception of N3, take part in hydrogen bonding with both the main chain and side chains of the protein.

6.4. Ammonium ions and their role in the stabilization of protein conformation

The next structural result from the investigation of the first structure, which was described in Manszewski *et al*. (2015), concerns the role of the ammonium ions in the stabilization of BeSAHase in the closed state. These NH_4^+ ions, which originate from the crystallization buffer, where their concentration was relatively high (0.2 M), were found only in the adenosine complexed subunits. Carbonyl atoms of two amino acid residues (Met390 and His392), from the Ala389-Pro393 loop as well as three water molecules create five hydrogen bonds with the ammonium ion. One of the NH donors must, therefore, form a bifurcated hydrogen bond. On the other hand, in the ligand-free subunit, no cationic species were found in the cation-binding loop; only a poorly bound water molecule is situated near the carbonyl group of Met390. Since the cation-binding loop partially overlaps the second hinge region (His392-Val396) and there is no cation bound in the subunit in the open conformation, the cation's role is evidently to stabilize the closed conformation of the BeSAHase subunit. The identity of the ammonium cations was confirmed by the calcium bond-valence sum method (Müller *et al*., 2003).

6.5. BeSAHase complexes with 3'-deoxydenosine, adenine, adenosine and 2'-deoxyadenosine

The third paper (Manszewski *et al*., 2017), accepted for publication in *IUCrJ*, reports structural investigations of four complexes of BeSAHase with different ligands: 3'-deoxyadenosie (cordycepin; Cord; resolution 1.84 Å), adenine (Ade; resolution 1.95 Å), adenosine (resolution 1.95 Å) and 2'-deoxyadnosine (2'-dAdo; resolution 1.54 Å).

The first structure, which is a tetramer with three subunits complexed with Ado and one with 3'-deoxyadenosine (cordycepin), was solved using data from a crystal obtained by cocrystallization of a protein sample with an 8-fold molar excess of

cordycepin. Since the desired ligand was bound only in one out of four subunits, it is clear that BeSAHase has higher affinity for adenosine and sequesters it from the bacterial cells during overexpression. This result made it clear that a ligand-free protein is required for the crystallization experiments. For this reason, the ligand exchange procedure has been introduced. This allowed me to determine the next three crystal structures. In two of them, only the desired ligand was present (complexes with Ade or Ado) and in one of them two subunits were complexed with 2'-dAdo and the other two with Ade, which is a product of 2'-dAdo hydrolysis (Abeles *et al.*, 1980, 1982). The atomic coordinates and the structure factors were deposited in PDB with the following accession codes: 5m5k (Ado/Cord complex), 5m65 (Ade), 5m66 (Ado) and 5m67 (2'-dAdo/Ade).

All but one of the complexes crystallized as a tetramer in the asymmetric unit. The only exception is the complex with Ade, which crystallized with two subunits in the asymmetric unit, from which the tetrameric state is created by crystallographic symmetry (twofold axis along z). There was a well defined electron density corresponding to NAD^+ at full occupancy in each subunit.

6.6. Protein conformation influenced by the ligand type bound in the active site

Having at hand crystal structures of the enzyme with four different ligands, I was able to investigate whether there is a relationship between the ligand type that is bound in the active site, and the protein conformation. Apart from the two previously determined conformational states, the closed state and the open state (Manszewski *et al.*, 2015), I found in the new set of structures a new, intermediate state. This conformation, termed the semi-open state, was observed in the BeSAHase complex with cordycepin. Comparison of the semi-open subunit with the closed (subunit A from 4lvc, complexed with adenosine) and open (subunit D from 4lvc) subunits (Fig. 5) by means of rms deviations for Cα atomic positions, shows that the differences amount to 1.71 and 1.29 Å, respectively. It is worth noting that SAHase from *L. luteus* in complex with Cord (PDB ID: 3onf; Brzeziński *et al.*, 2012) adopts the closed conformation. This shows that the enzyme can adopt not only two ultimate end conformations, but that some intermediate states are also possible, and that, even with the same ligand in the active site, SAHases can adopt different conformational states.
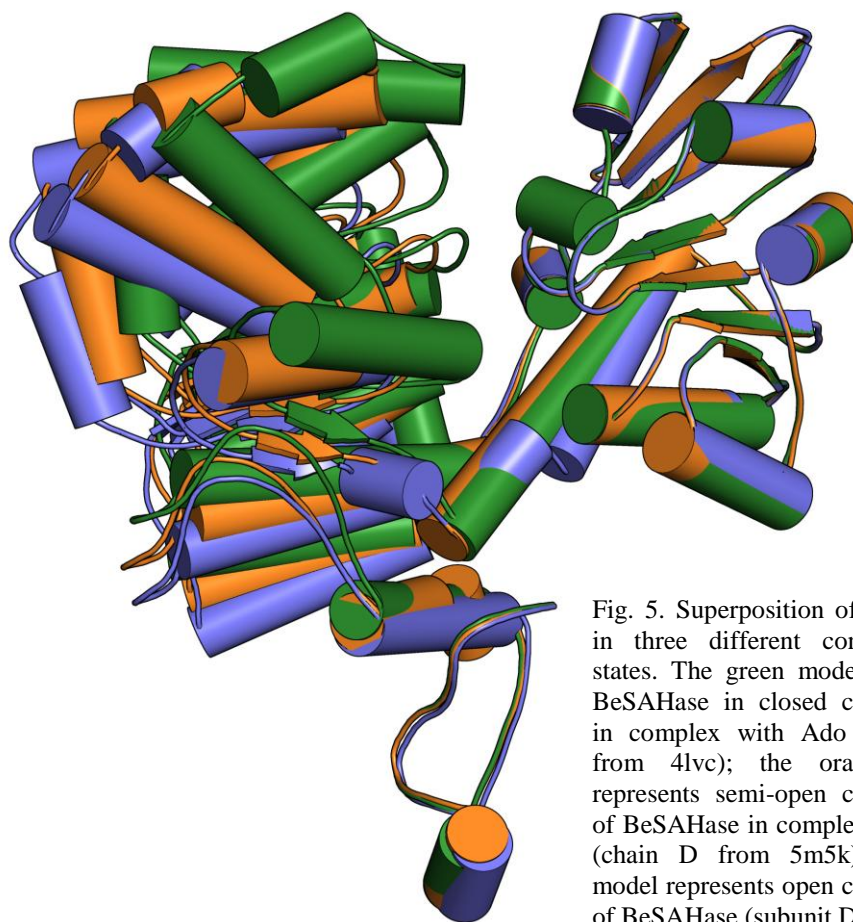
Fig. 5. Superposition of BeSAHase in three different conformational states. The green model represents BeSAHase in closed conformation in complex with Ado (subunit A from 4lvc); the orange model represents semi-open conformation of BeSAHase in complex with Cord (chain D from 5m5k); the blue model represents open conformation of BeSAHase (subunit D from 4lvc).

6.7. The influence of ligand type on the architecture of the active site

Since I had five different structures with four different ligands, I could not only check how the protein conformation was affected by the ligand type, but also how the architecture of the active site and the conformation of the molecular gate residues changed depending on the ligand.

6.7.1. Adenosine

In the case of the Ado molecules (Fig. 6a), which were present in two complexes, the interactions between the nucleosides and the protein atoms were the same and were similar to those described in Manszewski *et al*. (2015; PDB ID: 4lvc). There was, however, one exception which consisted in a different conformation of the side chain in His342 which forms a part of the MG. In the BeSAHase complexes with Ado described in Manszewski *et al*. (2017) this side chain is shifted far away from the active site and therefore cannot interact with the Ado O5' atom, as it does in the 4lvc structure.

6.7.2. Cordycepin

A more complicated situation occurs in the subunit that is complexed with Cord (Fig. 6b). Due to the conformational changes in this subunit, which adopts the semi-open conformation, as well as to the absence of the ribose O3' atom, the hydrogen-bond network around the cordycepin ligand is significantly different. Also, the cordycepin torsion angle γ, defined by the O5'-C5'-C4'-C3' atoms, is rotated by ~100° with respect to the angle in the Ado molecules, making it impossible for the O5' atom to participate in the hydrogen bonding interactions.

6.7.3. Adenine

An interesting situation was found in the active site of the subunits where adenine is bound (Fig. 6c). This ligand was found in two structures, one obtained by cocrystallization with Ade (i) and one obtained by cocrystallization with 2'-dAdo (ii). In this structure the architecture of the active site is preserved thanks to the four water molecules that mimic the positions of the ribose oxygen atoms. Therefore, I was able to
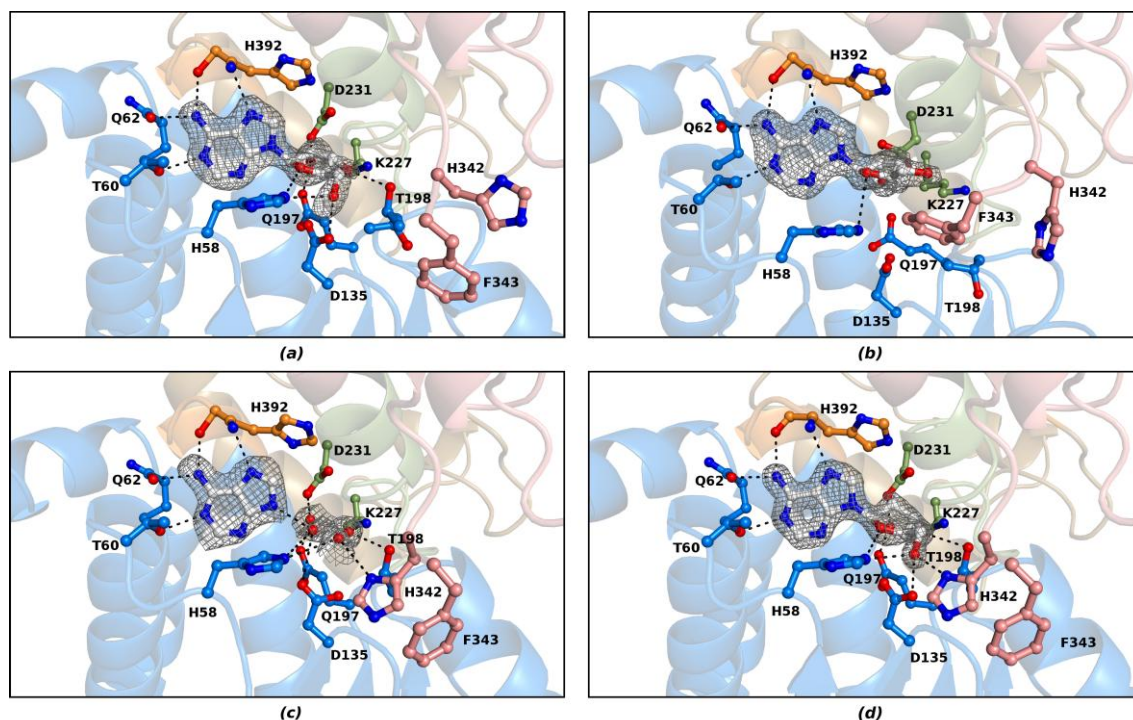
Fig. 6. Comparison of BeSAHase active sites. (a) Ado-complexed subunit A from the Ado/Cord complex; (b) Cord-complexed subunit D from the Ado/Cord complex; (c) Ade-complexed subunit A from the Ade complex; (d) 2'-dAdo-complexed subunit C from the Ade/2'-dAdo complex. Since the mode of binding of all Ado molecules in all respective complexes is identical, only the ligand from the Ado/Cord structure is shown. Similarly, as the mode of binding of all Ade ligands is identical, only Ade from the Ade complex is shown. The ligands and water molecules are shown in F_o-F_c OMIT electron-density (calculated without the contribution of the nucleosides, Ade and water atoms to F_c). The maps are contoured at 4σ, 3σ, 3σ and 4σ, respectively.

analyze how well the water molecules that complement the ligand in the active site are able to mimic the positions of the ribose oxygen atoms of Ado. In the Ade-complexed subunits from (i), four water molecules mimic the positions of the ribose O2', O3', O4' and O5' atoms at positions that are, respectively, 0.23, 0.26, 0.77, 1.08 Å (subunit A) and 0.32, 0.20, 1.01, 1.11 Å (subunit B) away from the template Ado ribose atoms. In the Ade complex (ii) these values are 0.65, 0.39, 0.45, 0.21 Å and 0.69, 0.29, 0.34, 0.39 Å, respectively for the Ade molecules bound in subunit A and subunit B. The hydrogen bonds created by these water molecules are the same in each case, although their distances vary from ~0.4 to 0.6 Å. This is because the positions of the equivalent water molecules differ by roughly 0.55, 0.33, 0.59 and 0.87 Å.

6.7.4. 2'-Deoxyadenosine

When 2'-deoxyadenosine is bound in the active site (Fig. 6d), the hydrogen-bond network between the nucleoside and protein atoms is almost the same as in Ade complexed subunits. This is because the 2'-dAdo O atoms are superimposable with the water molecules from the Ade complexes within 0.19 Å (O3'), 0.47 Å (O4') and 0.26 Å (O5').

6.8 Molecular gate state vs. ligand type bound in the active site

Availability of a set of structures of different complexes allowed me to compare the conformation of the molecular gate residues, His342-Phe343. In the 4lvc structure (Manszewski *et al*., 2015) the conformation of the MG residues was clearly correlated with the conformation of the whole subunit; that is to say, when the protein was in the closed conformation, the MG was shut, and when the protein was in the open conformation, the gate was open. That preliminary observation would seem to suggest that there should be a simple correlation between the conformational state of the MG residues and the ligand type, but the final conclusions were quite different.

Ultimately, it turned out that there is no simple correlation between the conformation of the MG residues, the ligand type, and the protein conformation. To illustrate this point, in the Ado-bound subunits from the Ado/Cord complex (Fig. 7a) and in chains A and D of the Ado complex (Fig. 7b), the MG element was found to be in the open state, i.e. in a different state than in the Ado-bound closed subunits of the 4lvc structure (Manszewski *et al*., 2015). Furthermore, even when the same ligand is present

in the active site, a variety of different MG conformations are possible. This is seen in subunits B and C of the Ado complex (Fig. 7c), where the side chain of His342 remains in the open state while the side chain of Phe343 assumes a conformation that is compatible with the shut state of the MG. In the Cord-complexed subunit (Fig. 7d), the His342 side chain adopts exactly the same conformation as it does in the open subunit of BeSAHase (chain D of 4lvc), but the side chain of Phe343, while remaining in the open conformation, is rotated by ~180° around the Cα-Cβ bond. The MG residues of all the Ade- and 2'-dAdo-complexed subunits (Fig. 7e) are in the closed conformation.
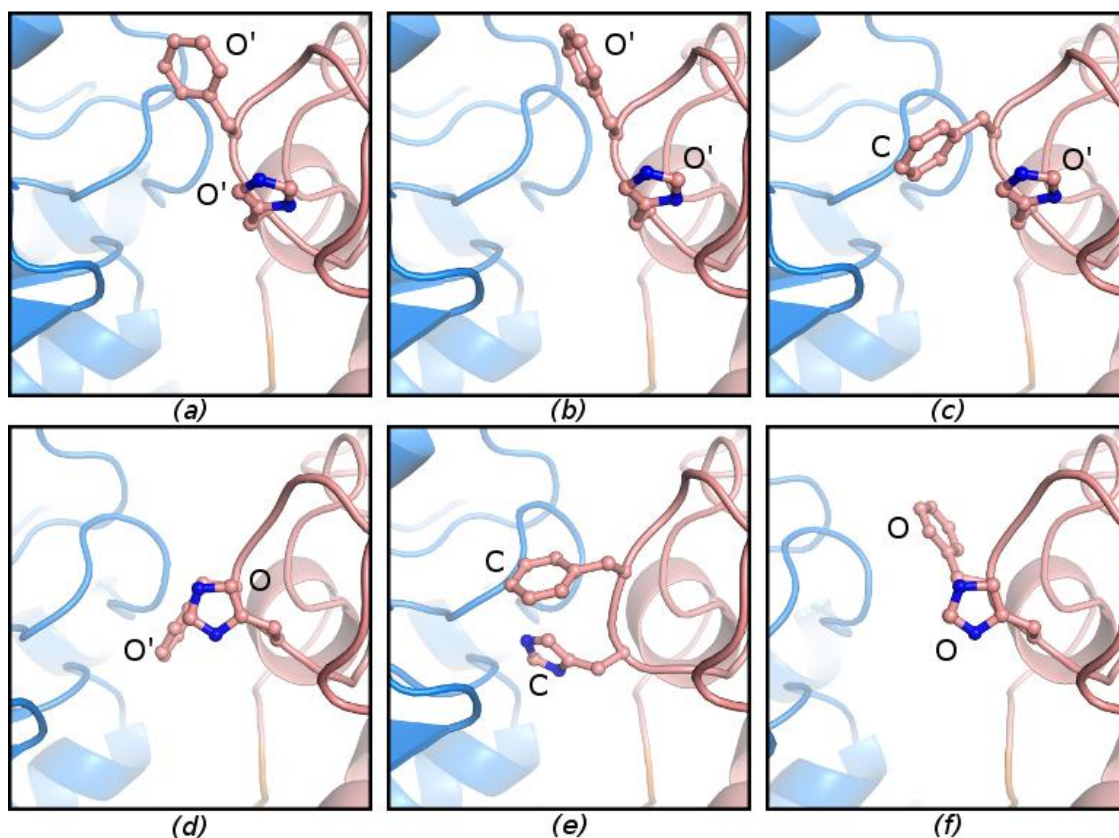
Fig. 7. Comparison of side chain conformations of the MG residues in: (a) Ado-complexed subunit of the Ado/Cord complex; (b) subunit A of the Ado complex; (c) subunit B of the Ado complex; (d) Cord-complexed subunit of the Ado/Cord complex; (e) subunit A of the Ade complex; (f) open subunit (D) of the PDB model 4lvc. The side chains of the molecular gate residues His342 and Phe343 are shown in ball-and-stick representation and annotated as C (closed state), O (open state) or O' (variant open state).

6.9. Sodium cations bound to BeSAHase

The set of structures of BeSAHase also allows a discussion of the role of the monovalent cations bound near the active site. In the 4lvc structure (Manszewski *et al.*, 2015) an ammonium cation was found in the Ado-complexed subunits in the closed conformation. I postulated then that the presence of the cation stabilizes the closed

conformation of the protein. In all of the structures presented in Manszewski *et al.* (2017) a sodium cation was located in the cation-binding loop (Fig. 8a), with one exception. As expected, the Cord-complexed subunit, which adopts a semi-open conformation, has no sodium cation bound in the cation-binding loop. Since ammonium or sodium cations were found only is the fully closed subunits, I can conclude that the interactions (coordination or hydrogen bonds) between the monovalent ions and the amino acid residues from the metal-binding loop (which overlaps with the second hinge region) play a role in stabilizing the closed conformation.

I also found an interesting sodium coordination site in the highest resolution (1.54 Å) structure of the 2'-dAdo/Ade complex. An additional sodium ion was found right in the center of this BeSAHase tetramer (Fig. 8b). This sodium cation is coordinated by six water molecules, four of which are hydrogen-bonded with the Gln275 Oε1 and Nε2 atoms from all four of the subunits. The coordination of this central cation does not seem to play any significant role in the stabilization of the quaternary structure of BeSAHase, since the interaction interfaces between the subunits are very large, ~2900 Å² between subunits forming a tight dimmer and ~1600 Å² between juxtaposed subunits.

The identification of all the sodium cations was confirmed by the *B* factors and by coordination geometry, and was additionally validated using

Fig. 8. Coordination of sodium cations (purple spheres). (a) The Na⁺ cation bound in subunit A of the Ade/2'-dAdo complex. (b) The Na⁺ cation bound in the core of the Ade/2'-dAdo complex. Hydrogen bonds are represented by red dash lines and coordination bonds are represented by black dash lines. In both panels, the F_o-F_c OMIT electron density map was calculated without the contribution of Na⁺ to F_c and contoured at 5σ. The domains are color-coded as in Fig. 1. Water molecules are represented as red spheres.

the *CheckMyMetal* server (Zheng *et al*., 2014).

6.10. Cofactor molecules

For their enzymatic activity, SAHases require one molecule of nicotinamide adenine dinucleotide in its oxidized form (NAD$^+$) bound near the active site of each subunit. The cofactor molecules are hydrogen-bonded by residues of both the cofactor-binding and substrate-binding domains. Moreover, the side chains of residues Lys467 and Tyr471 from the C-terminal domain of an adjacent subunit are also involved in cofactor binding, highlighting the role of this domain in the dimerization process. This way, the cofactor-binding site is created
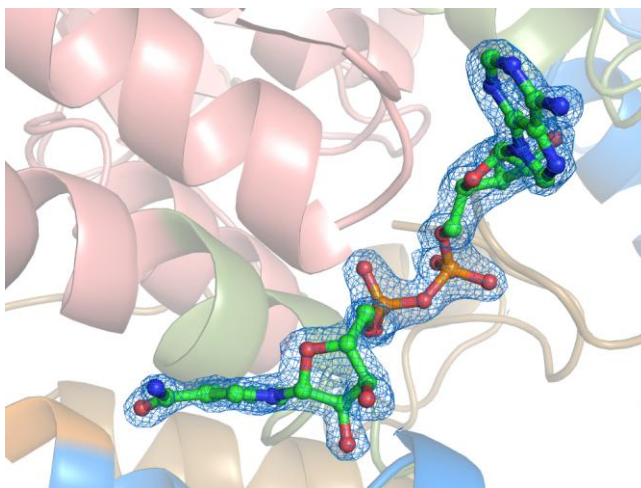
Fig. 9. The NAD$^+$ cofactor molecule bound in subunit A of the Ade/2'-dAdo complex (5m67), shown in *2F$_o$-F$_c$* electron density contoured at 1σ.

with the participation of both protomers forming the tight dimer. It has been also shown by Ault-Riché *et al.* (1994) that mutation of Lys426 in human SAHase (corresponding to Lys467 in BeSAHase) to alanine or glutamate results in inactive, monomeric protein without the cofactor.

In all five BeSAHase structures presented in this work, the cofactor molecules were clearly defined in the electron density maps and refined at full occupancy. To support the conclusion that the cofactor is in its oxidized (NAD$^+$) state, I analyzed the puckering of the nicotinamide moiety. Meijers *et al.* (2001) showed that the addition of the hydride ion to the nicotinamide ring causes a loss of its aromatic character and its deformation into a boat conformation, which is characteristic of the reduced (NADH) form. In the present studies, I refined the cofactor molecules in all structures without planarity restraints. In each case the refinement was stable and converged with a flat nicotinamide ring (Fig. 9).

7. CONCLUSIONS

All three papers presented as this dissertation are connected through the common theme of BeSAHase structure. The results, together with the existing information, provide a unique opportunity to draw general conclusions about this group of enzymes. It is quite clear now that SAHases can adopt several different conformations, including an intermediate state between the fully open and fully closed forms. On the other hand, more research is needed to fully elucidate the exact mechanism of the molecular gate. The data, that I have presented, show a puzzling situation, namely that the MG residues can adopt different conformations even with the same ligand molecule bound in the active site. There is also no apparent correlation between the closed/open conformation of the protein subunit and the shut/open state of the molecular gate. What is clear though, is the role of the sodium or ammonium cations in the stabilization of the protein structure. These ions, which are bound near the active site within one of the interdomain hinge regions, are present only when the protein subunit assumes the closed conformation. For subunits in the open or semi-open conformation, the cation-binding loop remains empty. Intriguingly, the monovalent cation can exert its stabilizing effect either via coordination bonds (Na^+) or hydrogen bonds (NH_4^+).

8. REFERENCES

Abeles RH, Tashjian AH & Fish S (1980) The mechanism of inactivation of S-adenosylhomocysteinase by 2'-deoxyadenosine. *Biochem Biophys Res Commun* **95**, 612-617.

Abeles RH, Fish S & Lapinskas B (1982) S-Adenosylhomocysteinase: mechanism of inactivation by 2'-deoxyadenosine and interaction with other nucleosides. *Biochemistry* **21**, 5557-5562.

Ault-Riché DB, Yuan CS & Borchardt RT (1994) A single mutation at lysine 426 of human placental S-adenosylhomocysteine hydrolase inactivates the enzyme. *J Biol Chem* **269**, 31472-31478.

Backlund PS, Carotti Jr. D & Cantoni GL (1986) Effects of the S-adenosylhomocysteine hydrolase inhibitors 3-deazaadenosine and 3-deazaaristeromycin on RNA methylation and synthesis. *Eur J Biochem* **160**, 245-251.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN & Bourne PE (2000) The Protein Data Bank *Nucelic Acids Res* **28**, 235-242.

Brzeziński K, Bujacz G & Jaskólski M (2008) Purification, crystallization and preliminary crystallographic studies of plant S-adenosyl-L-homocysteine hydrolase (Lupinus luteus). *Acta Cryst.* F**64**, 671-673.

Brzeziński K, Dauter Z & Jaskólski M (2012) High-resolution structures of complexes of plant S-adenosyl-L-homocysteine hydrolase (Lupinus luteus). *Acta Cryst.* D**68**, 218-231.

Cantoni GL (1975) Biological methylation: selected aspects. *Annu Rev Biochem* **44**, 435-451.

Cantoni GL. & Chiang PK (1980) The role of S-adenosylhomocysteine and S-adenosylhomocyteine hydrolase in the control of biological methylation. In*: Natural Sulfur Compounds* (Cavallini D, Gaull GR & Zappia V, eds), pp. 67-80. Plenum Press. New York.

Chiang PK & Cantoni GL (1979) Perturbation of biochemical transmethylations by 3-deazaadenosine in vivo. *Biochem Pharmacol* **28**, 1897-1902.

De La Haba G & Cantoni GL (1959) The enzymatic synthesis of S-adenosyl-L-homocysteine from adenosine and homocysteine. *J Biol Chem* **234**, 603-608.

Fujioka M & Takata Y (1981) S-Adenosylhomocysteine hydrolase from rat liver. Purification and some properties. *J Biol Chem,* **256,** 1631-1635.

Guranowski A & Pawełkiewicz J (1977) Adenosylhomocysteinase from yellow lupin seeds. Purification and properties. *Eur J Biochem* **80**, 517-523.

Hu Y, Komoto J, Huang Y, Gomi T, Ogawa H, Takata Y, Fujioka M & Takusagawa F (1999) Crystal structure of S-adenosylhomocysteine hydrolase from rat liver. *Biochemistry* **38**, 8323-8333.

Jakubowski H (2006) Pathophysiological consequences of homocysteine excess. *J Nutr* **136**, 1741S-1749S.

Manszewski T, Singh K, Imiołczyk B & Jaskólski M (2013) Structural enzymology at the legume-microbe interface: S-adenosyl-L-homocysteine hydrolase of rhizobia. *Biotechologia* **91**, 38-39.

Manszewski T, Singh K, Imiołczyk B & Jaskólski M (2015) An enzyme captured in two conformational states: Crystal structure of S-adenosyl-l-homocysteine hydrolase from Bradyrhizobium elkanii. *Acta Cryst.* D**71**, 2422-2432.

Manszewski T, Szpotkowski K & Jaskólski M (2017) Crystallographic and SAXS studies on S-adenosyl-L-homocysteine hydrolase from Bradyrhizobium elkanii. *IUCrJ*, in press.

Meijers R, Morris RJ, Adolph WH, Merli A, Lamzin VS & Cedergren-Zeppezauer ES (2001) On the enzymatic activation of NADH. *J Biol Chem* **276**, 9316-9321.

Moller JM, Nielsen GL, Ekelund S, Schmidt EB & Dyerberg J (1997) Homocysteine in Greenland Inuits. *Thromb Res* **86**, 333-335.

Müller-Dieckmann J (2006) The open-access high-throughput crystallization facility at EMBL Hamburg. *Acta Cryst.* D**62**, 1446-1452.

Müller P, Köpke S & Sheldrick GM (2003) Is the bond-valence method able to identify metal atoms in protein structures? *Acta Cryst.* D**59**, 32-37.

Nygard O, Nordrehaug JE, Refsum H, Ueland PM, Farstad M & Vollset SE (1997) Plasma homocysteine levels and mortality in patients with coronary artery disease. *N Engl J Med* **337**, 230-237.

Palmer JL & Abeles RH (1979) The mechanism of action of S-adenosylhomocysteinase. *J Biol Chem* **254**, 1217-1226.

Reddy MC, Kuppan G, Shetty ND, Owen JL, Ioerger TR & Sacchettini JC (2008) Crystal structures of Mycobacterium tuberculosis S-adenosyl-L-homocysteine hydrolase in ternary complex with substrate and inhibitors. *Protein Sci* **17**, 2134-2144.

Richards HH, Chiang PK & Cantoni GL (1978) Adenosylhomocysteine hydrolase. Crystallization of the purified enzyme and its properties. *J Biol Chem* **253**, 4476-4480.

Stępkowski T, Brzeziński K, Legocki AB, Jaskólski M & Bena G (2005) Bayesian phylogenetic analysis reveals two-domain topology of S-adenosylhomocysteine hydrolase protein sequences. *Mol Phylogenet Evol* **34**, 15-28.

Suarez J & Chagoya de Sanchez V (1997) Inhibition of S-adenosyl-L-homocysteine hydrolase by adrenaline in isolated guinea-pig papillary muscles. *Int J Biochem Cell Biol* **29**, 1279-1284.

Tanaka N, Nakanishi M, Kusakabe Y, Shiraiwa K, Yabe S, Ito Y, Kitade Y & Nakamura KT (2004) Crystal structure of S-adenosyl-L-homocysteine hydrolase from the human malaria parasite Plasmodium falciparum. *J Mol Biol* **343**, 1007-1017.

Turner M, Yang X, Yin D & Borchardt RT (2000) Structure and function of S-adenosylhomocysteine hydrolase. *Cell Biochem Biophys* **33**, 101-125.

Turner MA, Yuan CS, Borchardt RT, Hershfield MS, Smith GD & Howell PL (1998) Structure determination of selenomethionyl S-adenosylhomocysteine hydrolase using data at a single wavelength. *Nature Struct Biol* **5**, 369-376.

White AR, Huang X, Jobling MF, Barrow CJ, Beyreuther K, Masters CL, Bush AI & Cappai R (2001) Homocysteine potentiates copper- and amyloid beta peptide-mediated toxicity in primary neuronal cultures: possible risk factors in the Alzheimer's-type neurodegenerative pathways. *J Neurochem* **76**, 1509-1520.

Hu Y, Komoto J, Huang Y, Gomi T, Ogawa H, Takata Y, Fujioka M & Takusagawa F (1999) Crystal structure of S-adenosylhomocysteine hydrolase from rat liver. *Biochemistry* **38**, 8323-8333.

Yin D, Yang X, Hu Y, Kuczera K, Schowen RL, Borchardt RT & Squier TC (2000) Substrate binding stabilizes S-adenosylhomocysteine hydrolase in a closed conformation. *Biochemistry* **39**, 9811-9818.

Yuan CS, Yeh J, Liu S & Borchardt RT (1993) Mechanism of inactivation of S-adenosylhomocysteine hydrolase by (Z)-4',5'-didehydro-5'-deoxy-5'-fluoroadenosine. *J Biol Chem* **268**, 17030-17037.

Zheng H, Chordia MD, Cooper DR, Chruszcz M, Müller P, Sheldrick GM & Minor W (2014) Validation of metal-binding sites in macromolecular structures with the CheckMyMetal web server. *Nature Protoc.* **9**, 156-170.

Zheng Y, Chen CC, Ko TP, Xiao X, Yang Y, Huang CH, Qian G, Shao W & Guo RT (2015) Crystal structures of S-adenosylhomocysteine hydrolase from the thermophilic bacterium Thermotoga maritima. *J Struct Biol* **190**, 135-142.

(37 references in total)

COPIES OF THE

PUBLICATIONS

INCLUDED

IN THIS THESIS

1

Manszewski T, Singh K, Imiołczyk B, Jaskólski M. (2013) Structural enzymology at the legume-microbe interface: S-adenosyl-L-homocysteine hydrolase of rhizobia. *Biotechologia* **91**, 38-39.

Structural enzymology at the legume-microbe interface: S-adenosyl-L-homocysteine hydrolase of rhizobia

TOMASZ MANSZEWSKI [1], KRITI SINGH [1], BARBARA IMIOLCZYK [1], MARIUSZ JASKOLSKI [1,2*]

[1] Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznań, Poland

[2] Faculty of Chemistry, Adam Mickiewicz University, Poznań, Poland

* Corresponding author: mariuszj@amu.edu.pl

Abstract

S-adenosyl-L-methionine (SAM) is the most common substrate used in biological methylation reactions. During methyl group transfer, S-adenosyl-L-homocysteine (SAH) is formed as a byproduct. Since SAH is a strong inhibitor of the ubiquitous SAM dependent methylation reactions, removal of SAH by enzymatic hydrolysis serves as an important mechanism in the control of cellular methylation processes.

Introduction

S-adenosyl-L-homocysteine hydrolase (SAHase) catalyzes the hydrolysis of SAH to adenosine and L-homocysteine in a reversible reaction (de la Haba et al., 1959). The equilibrium of the hydrolysis reaction is shifted far in the direction of the substrate synthesis but under physiological conditions the products are quickly removed by other enzymatic processes (Richards et al., 1978). The enzymatic process catalyzed by SAHases is quite complex and involves a redox step. This is why the enzyme requires NAD^+ as a cofactor for its enzymatic competence.

Crystal structures of several SAHases are known, including enzymes from mammals, protozoans and bacteria. The enzymes are typically homotetrameric, with each subunit comprised of two principal domains: a substrate binding domain and a cofactor (NAD^+) binding domain. Phylogenetic and biochemical analyses indicate that some groups of SAHases may be significantly different. In particular, plant SAHases have a unique sequence signature and are believed to function as dimers (Guranowski et. al., 1977; Brzezinski et al., 2008).

In this project, we focus on the crystal structure of SAHases from legume plants, such as yellow lupine (Brzezinski et al., 2012), and from their nitrogen-fixing bacterial symbionts (rhizobia), looking for possible interdependencies during the formation and existence of the symbiotic association. Here, we present the results of our crystallographic analysis of recombinant SAHase
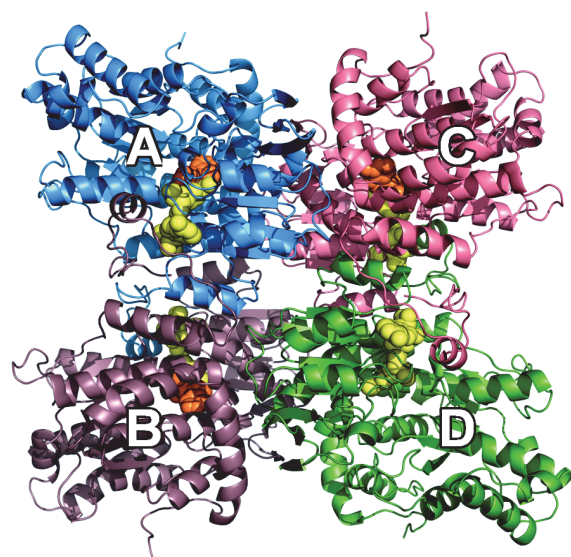
Fig. 1. A ribbon diagram of the BeSAHase tetramer, with subunits labeled A, B, C, D. NAD^+ and adenosine molecules are in space-filling representation

from *Bradyrhizobium elkanii* (BeSAHase), which infects soybean. It is the first structure of this enzyme from a bacterium capable of assimilating atmospheric nitrogen.

BeSAHase in complex with adenosine molecules

Single crystals of BeSAHase were grown by vapor diffusion in hanging drops using protein preparations preincubated with the NAD^+ cofactor. They diffracted synchrotron radiation to 1.7 Å. The structure was solved by molecular replacement, using as a model the struc-
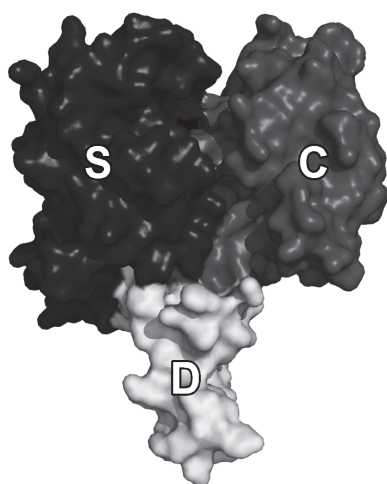
Fig. 2a. A subunit of BeSAHase in closed conformation
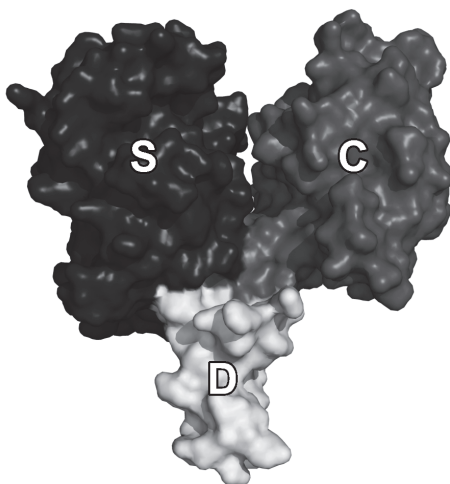
Fig. 2b. A subunit of BeSAHase in open conformation. The domains, shown in shades of gray, are labeled S (substrate-binding domain), C (cofactor binding domain), and D (C-terminal dimerization domain)

ture of SAHase from *B. melitensis*, and refined to R/R$_{\text{free}}$ of 0.1496/0.1757. The crystals are orthorhombic, space group $P2_12_12$, with four protein molecules in the asymmetric unit (473 amino acid residues each) forming a tetrameric enzyme (Fig. 1). There is one cofactor molecule bound in each subunit, as required by the enzymatic mechanism. In addition, a molecule of adenosine (Ado) is present in three subunits. In consequence, the homotetrameric enzyme is found in a mixed conformational state: while the Ado-binding subunits are in a closed conformation, the fourth subunit is open (Fig. 2). The structure gives us a unique glimpse of SAHase during a conformational transformation required for its enzymatic cycle.

BeSAHase in complexes with other ligand molecules

The protein was also studied after preincubation with 3-deoxyadenosine (cordycepin). The crystals have the same $P2_12_12$ symmetry and diffracted X-rays to a 1.84 Å resolution. The cordycepin molecule was found in only one out of the four subunits. The remaining three subunits are occupied with adenosine ligands.

To obtain BeSAHase crystals with uniform ligand occupation, we elaborated a new purification protocol that includes an extra step designed to wash out all of the bound ligand molecules first, according to a modified procedure of Yuan et al. (1993). Following this procedure, a complex with 2′-deoxyadenosine was crystallized and its crystal structure was solved to 1.54 Å resolution, also in space group $P2_12_12$. The structure is currently being refined.

Other biophysical studies and the future

The protein and its complexes have also been characterized by several biophysical methods, including Dynamic Light Scattering (DLS), Small-Angle X-ray Scattering (SAXS), Isothermal Titration Calorimetry (ITC), Mass Spectroscopy (MS) and Thermofluor analysis. The results are currently being analyzed. Our crystallographic observations of this tetrameric protein in complexes with different ligands and in different binding modes are used to map the geometrical transformations connected with the complicated chemistry catalysed by this enzyme.

Acknowledgments

References

Brzezinski K., Bujacz G., Jaskolski M. (2008) Acta Cryst. F64: 671-673.

Brzezinski K., Dauter Z., Jaskolski M. (2012) Acta Cryst. D68: 218-231.

De La Haba G., Cantoni G. L. (1959) J. Biol. Chem. 234: 603-608.

Guranowski A., Pawełkiewicz J. (1977) Eur. J. Biochem. 80: 517-523.

Richards H.H., Chiang P.K., Cantoni G.L. (1978) J. Biol. Chem. 253: 4476-4480.

Yuan C.S., Yeh J., Liu S., Borchardt R.T. (1993) J. Biol. Chem. 268: 17030-17037.

2

Manszewski T, Singh K, Imiołczyk B, Jaskólski M. (2015)

An enzyme captured in two conformational states: Crystal structure of S-adenosyl-l-homocysteine hydrolase from Bradyrhizobium elkanii. *Acta Cryst.* **D71**, 2422-2432.

CrossMark

An enzyme captured in two conformational states: crystal structure of S-adenosyl-L-homocysteine hydrolase from Bradyrhizobium elkanii

Tomasz Manszewski,[a] Kriti Singh,[a] Barbara Imiolczyk[a] and Mariusz Jaskolski[a,b]*

[a]Center for Biocrystallographic Research, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan, Poland, and [b]Department of Crystallography, Faculty of Chemistry, A. Mickiewicz University, Poznan, Poland. *Correspondence e-mail: mariuszj@amu.edu.pl

S-Adenosyl-L-homocysteine hydrolase (SAHase) is involved in the enzymatic regulation of S-adenosyl-L-methionine (SAM)-dependent methylation reactions. After methyl-group transfer from SAM, S-adenosyl-L-homocysteine (SAH) is formed as a byproduct, which in turn is hydrolyzed to adenosine (Ado) and homocysteine (Hcy) by SAHase. The crystal structure of BeSAHase, an SAHase from *Bradyrhizobium elkanii*, which is a nitrogen-fixing bacterial symbiont of legume plants, was determined at 1.7 Å resolution, showing the domain organization (substrate-binding domain, NAD$^+$ cofactor-binding domain and dimerization domain) of the subunits. The protein crystallized in its biologically relevant tetrameric form, with three subunits in a closed conformation enforced by complex formation with the Ado product of the enzymatic reaction. The fourth subunit is ligand-free and has an open conformation. The BeSAHase structure therefore provides a unique snapshot of the domain movement of the enzyme induced by the binding of its natural ligands.

1. Introduction

S-Adenosyl-L-methionine (SAM) is used in the vast majority of enzymatic methylation reactions because of its methyl-group transfer potential (Cantoni, 1975; Richards *et al.*, 1978). SAM acts as the methyl-donor substrate in the methylation of a large variety of biological molecules, ranging from macromolecules such as proteins, nucleic acids and polysaccharides to small molecules such as catecholamines, histamines and norepinephrine (Keller & Borchardt, 1988). Upon methyl-group transfer, SAM is converted to S-adenosyl-L-homocysteine (SAH). This byproduct of methylation is a strong inhibitor of SAM-dependent methyltransferases (Chiang & Cantoni, 1979; Cantoni & Chiang, 1980; Liu *et al.*, 1992) and must be enzymatically decomposed to adenosine (Ado) and homocysteine (Hcy) by S-adenosyl-L-homocysteine hydrolase (SAHase; De la Haba & Cantoni, 1959). Thus, by maintaining the appropriate SAM:SAH ratio, SAHase serves as an important regulator of cellular SAM-dependent methyltransferases. Plant SAHases have also been implicated in cytokinin binding (Mitsui *et al.*, 1993), but in view of the recent structural studies of SAHase from yellow lupin (Brzezinski *et al.*, 2012) this hypothesis is probably not true, as the N6 substituent of cytokinins would generate steric clashes with the adenine-binding site of the protein

The equilibrium of the SAH hydrolysis reaction lies far in the direction of substrate synthesis; however, under physiological conditions the products, Ado and Hcy, are quickly removed (Richards *et al.*, 1978). Ado is converted to inosine or
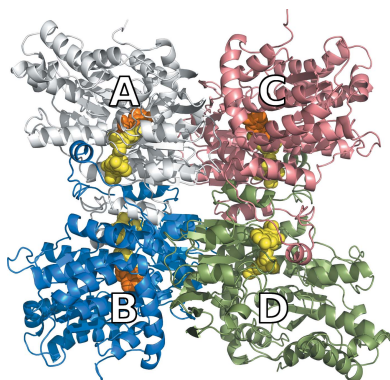
adenosine triphosphate by adenosine deaminase (ADA) or adenosine kinase (AK), respectively. It has been shown that accumulation of Ado causes inhibition of SAHase, which in turn leads to combined immune deficiency (Kredich & Martin, 1977; Hershfield & Kredich, 1978; Hershfield, 1979; Hershfield *et al.*, 1979). Hcy is used in remethylation and transsulfuration reactions in an approximately 1:1 ratio (Robinson *et al.*, 1994). Remethylation leads to methionine regeneration *via* two different pathways. One of them is irreversible, takes place in the liver and kidneys and involves betaine (McKeever *et al.*, 1991), while the more important one requires a methyl-group transfer to Hcy from 5-methyltetrahydrofolate by 5-methyl-tetrahydrofolate homocysteine methyltransferase in the folate cycle. It has been shown that an elevated plasma level of Hcy is a risk factor in coronary heart disease (Nygård *et al.*, 1997).

SAHases are usually active as tetramers, with the only exception being plant SAHases, which function as homodimers (Guranowski & Pawełkiewicz, 1977; Brzezinski *et al.*, 2008). The SAHase subunit is comprised of ~420 (for archaeal and some bacterial enzymes) or ~480 (for most eukaryotic and bacterial enzymes) residues and has a molecular weight of about 45–55 kDa. This difference is owing to the presence of an ~40-residue insert in the latter group of SAHases (Stępkowski *et al.*, 2005). In each subunit there is one tightly but noncovalently bound molecule of the NAD^+ cofactor, which is required for enzymatic activity (Palmer & Abeles, 1979; Fujioka & Takata, 1981). Each subunit consists of two main domains, the substrate-binding domain and the cofactor-binding domain, and a small C-terminal dimerization domain.

SAHases can exist in two different conformations described as open and closed. The difference between these two states is in the relative position of the substrate-binding and cofactor-binding domains. In the open conformation, which is characteristic of the substrate/product-free enzyme, the domains are swung apart and the active site is accessible. The opposite situation occurs when the active site is occupied by a ligand molecule: the main domains come close together, adopting the closed conformation.

The mechanism of the reaction catalyzed by SAHase, described by Palmer & Abeles (1976, 1979), can be briefly summarized as follows. Firstly, SAH is oxidized to a 3′-keto derivative with NAD^+ as the oxidant. The 4′ proton is then abstracted by the N^ζ atom of a lysine located in the active site and the Hcy moiety at the 5′ position is eliminated, leading to 3′-keto-4′,5′-dehydroadenosine. The last step is Michael addition of a water molecule and reduction of the 3′-keto derivative to adenosine with simultaneous conversion of NADH to NAD^+.

The crystal structures of several SAHases are known, including mammalian [*Homo sapiens* (Turner *et al.*, 1998, PDB entry 1a7a; Yang *et al.*, 2003, PDB entry 1li4), *Rattus norvegicus* (Hu *et al.*, 1999, PDB entry 1b3r; Huang *et al.*, 2002, PDB entry 1k0u; Takata *et al.*, 2002, PDB entry 1ky5)], plant [*Lupinus luteus* (Brzezinski *et al.*, 2012, PDB entries 3ond, 3one and 3onf)], protozoan [*Plasmodium falciparum* (Tanaka *et al.*, 2004, PDB entry 1v8b), *Trypanosoma brucei* (Structural Genomics Consortium, unpublished work, PDB entry 3h9u)]

and bacterial [*Mycobacterium tuberculosis* (Reddy *et al.*, 2008, PDB entry 3ce6), *Burkholderia pseudomallei* (Seattle Structural Genomics Center for Infectious Disease, unpublished work, PDB entry 3d64), *Brucella melitensis* (unpublished work, PDB entry 3n58)] proteins. Here, we present the first crystal structure of SAHase from a nodulating bacterium, *Bradyrhizobium elkanii* (BeSAHase), in complex with the reaction product adenosine (Ado). Nodulating bacteria colonize the root nodules of legume plants in a symbiotic association that leads to atmospheric nitrogen fixation. Together with the structure of the enzyme from a legume plant (LlSAHase; Brzezinski *et al.*, 2012), the structure of BeSAHase also provides the first example of a pair of orthologous enzymes from symbiotic partners for structural comparison.

The plant–bacteria symbiosis, leading to the formation of nodules and allowing atmospheric nitrogen fixation by plants, starts with the activation of bacterial genes responsible for the biosynthesis of Nod factor (NF). One of the NF modification enzymes, NodS, is a SAM-dependent methyltransferase and its activity is controlled by a SAHase, which maintains the appropriate SAM:SAH ratio.

2. Materials and methods

2.1. Cloning, overexpression and purification

The coding sequence of SAHase was amplified from a *B. elkanii* cDNA library by PCR and cloned into the pET151/D-TOPO plasmid, which adds a His_6 tag at the N-terminus of the expressed sequence. The vector was used to transform the BL21 Star (DE3) strain of *E. coli*. 6 l of LB medium containing 100 µg ml^{-1} ampicillin was inoculated with 150 ml of overnight culture and grown at 310 K to an OD_{600} of 1.0. The culture was induced with IPTG at a final concentration of 0.4 m*M*. The cells were harvested after 4 h of induction by centrifugation, the supernatant was removed and the cell pellet was resuspended in buffer *A* consisting of 50 m*M* Tris pH 8.0, 500 m*M* NaCl, 20 m*M* imidazole. The cells were disrupted by pulse sonication on ice and centrifuged. The supernatant was loaded onto an Ni–NTA affinity column and the protein was eluted with 15 ml buffer *B* (50 m*M* Tris pH 8.0, 500 m*M* NaCl, 300 m*M* imidazole). The protein concentration was estimated to be 8 mg ml^{-1} according to Bradford (1976). SDS–PAGE analysis confirmed that the molecular weight of the expressed protein was ~52 kDa. Following elution, the protein was dialyzed against buffer *A*. 100 µg of TEV protease per milligram of protein was added to remove the His tag. After 20 h incubation, the protein solution was again loaded onto the nickel-affinity column to remove the His-tag debris and His-tagged TEV protease. The purified protein was loaded onto a desalting column equilibrated with buffer *C* (20 m*M* Tris pH 8.0, 50 m*M* NaCl). To ensure the presence of the cofactor in the uniformly oxidized state, NAD^+ was added (as the sodium salt) in a 12-fold molar excess and the mixture was incubated for 3 h on ice. Finally, the protein solution was loaded onto a gel-filtration column equilibrated with buffer *C*.

Table 1
Data-collection and structure-refinement statistics.

Values in parentheses are for the last resolution shell.

Data collection	
Beamline	14.1, BESSY
Wavelength (Å)	0.918
Temperature (K)	100
Crystal system	Orthorhombic
Space group	$P2_12_12$
Unit-cell parameters (Å)	$a = 107.7$, $b = 176.5$, $c = 104.3$
Mosaicity (°)	0.23
Resolution range (Å)	47.85–1.74 (1.79–1.74)
Total reflections	1278402
Unique reflections	200549
Multiplicity	6.4 (4.2)
Completeness (%)	98.7 (97.0)
$\langle I/\sigma(I)\rangle$	12.34 (1.95)
R_{merge}†	0.086 (0.694)
Refinement	
No. of working/test reflections	199538/1011
R/R_{free}‡	0.148/0.172
No. of protein atoms	14468
No. of Ado atoms	57
No. of NAD$^+$ atoms	176
No. of water atoms	1704
No. of NH$_4^+$ ions	3
No. of acetate ions	3
No. of glycerol molecules	15
$\langle B\rangle$ (Å^2)	
Protein	27.1
Ado	29.3
NAD$^+$	21.3
Solvent	39.4
R.m.s.d. from ideality for bonds (Å)	0.018
Ramachandran statistics	
Favoured (%)	97.2
Allowed (%)	2.8
PDB code	4lvc

† $R_{merge} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the average intensity of reflection hkl. ‡ $R = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively. R_{free} is calculated analogously for the test reflections, which were randomly selected and excluded from refinement.

Fractions containing BeSAHase were collected, concentrated to 12 mg ml^{-1} and used for crystallization.

2.2. Crystallization

Single crystals suitable for diffraction experiments were obtained in condition No. 30 of the PEG/Ion screen (Hampton Research), consisting of 0.2 M ammonium acetate, 20%(w/v) polyethylene glycol 3350 pH 7.1. Crystallization was conducted by vapour diffusion in hanging drops at 292 K.

2.3. Data collection, structure solution and refinement

Low-temperature X-ray diffraction data extending to 1.74 Å resolution were collected on beamline 14.1 at BESSY at a wavelength of 0.918 Å. 20%(v/v) glycerol was used as a cryoprotectant. The crystals belonged to the orthorhombic system with space group $P2_12_12$. The diffraction data were processed and scaled with *XDS* (Kabsch, 2010) as summarized in Table 1.

The structure was solved by molecular replacement using *Phaser* (McCoy *et al.*, 2007). The atomic coordinates of chain *A* of SAHase from *B. melitensis* (PDB entry 3n58) served as the search model. The search procedure found four copies (labelled *A*, *B*, *C* and *D*) of the model chain arranged into a tetrameric assembly, which was built with the correct sequence using the *ARP/wARP* server (Langer *et al.*, 2008). The model was refined in *REFMAC*5 (Murshudov *et al.*, 2011) with maximum-likelihood targets. Three groups of TLS parameters (Painter & Merritt, 2006) were used per subunit. *Coot* (Emsley & Cowtan, 2004) was used for manual model rebuilding between rounds of refinement. A test refinement without planarity restraints for the nicotinamide moiety of the NAD$^+$ molecule was calculated to probe the oxidation state of the cofactor. The final structure-refinement statistics are shown in Table 1.

The polypeptide chain of BeSAHase is comprised of 473 amino-acid residues. The construct used in this work has six extra residues (GIDPFT–) at the N-terminal end that were introduced by the cloning vector. However, in the crystal structure these extra residues plus the first five genuine residues of the protein (–MNAKP–) were not modelled in the electron density because of disorder. Difference Fourier maps clearly indicated the presence of one NAD$^+$ cofactor molecule in each subunit. 1704 water molecules, 15 glycerol molecules, three acetate ions and three ammonium ions (the latter found only in the subunits complexed with Ado) were also modelled in the electron-density maps. During the analysis of difference Fourier maps, clear electron density corresponding to three adenosine (Ado) molecules was detected in subunits *A*, *B* and *C*.

The stereochemical restraints for the protein chains were as defined in *REFMAC*, while the restraint libraries for the Ado and NAD$^+$ ligands were built in *elBOW* (Moriarty *et al.*, 2009) as implemented in *PHENIX* (Adams *et al.*, 2010).

The *PyMOL* molecular-graphics system (DeLano, 2002) was used to prepare all of the figures presented in this paper.

3. Results and discussion

3.1. Overall structure of BeSAHase

The enzyme crystallizes as a 222 pseudosymmetric homo-tetramer (Fig. 1), with each subunit composed of 473 amino-acid residues, of which residues 6–473 are well ordered and could be modelled in the electron density. Each protomer consists of two large domains, the substrate-binding domain and the cofactor-binding domain, which are separated by a deep crevice forming the substrate-access channel to the active site, and a small C-terminal oligomerization domain.

The substrate-binding domain, built from amino-acid residues Gly6–Val221 and Met397–Val426, has an α/β-fold. The central parallel β-sheet is built from seven β-strands with the following topology: −1x, −1x, +3x, +1x, +1x, +1x. The cofactor-binding domain is comprised of residues Tyr234–Gly391 and has an unusual Rossmann fold, with the central mixed β-sheet composed of eight β-strands with topology +1x, +1x, −3x, −1x, −3x, +1, +1 as in previously reported SAHase structures (Brzezinski *et al.*, 2012). These two main domains are joined by two hinge regions built from amino-acid residues

Asn222–Leu233 and His392–Val396. The C-terminal domain, formed by residues Leu427–Tyr473, has a helix–loop–helix–loop fold. This domain is located away from the catalytic region of its own subunit (*e.g.* *A*) and is inserted in a mutual fashion into the cofactor-binding domain of an adjacent subunit (*B*). The mutual swapping of the C-terminal domains between pairs of subunits leads to their discernible tight dimerization. The tetrameric enzyme molecule must therefore be considered as a dimer of intimate dimers.

The centre of the tetramer is formed by the cofactor-binding domains, while the substrate-binding domains are located outside the core of the enzyme, where they are much more mobile than the rigid central part.

3.2. The adenosine molecules

Upon inspection of the difference Fourier maps, clear electron density corresponding to three adenosine molecules was identified (Fig. 2*a*) in subunits *A*, *B* and *C*, which are in a closed conformation (see §3.4), even though the nucleoside was not added at any stage of protein purification or crystallization. The identification of the unknown molecule as Ado was straightforward, as the location, conformation and protein interactions of this ligand in the substrate-binding domain were known to us from our previous studies of the LlSAHase–Ado complex (Brzezinski *et al.*, 2012).

During the refinement and electron-density modelling, the occupancy of the three Ado molecules was adjusted to 0.7. The incomplete saturation of the Ado-binding sites is explained by the fact the ligand was sequestered from the

limited pool available in the bacterial cell during the high-level overexpression of the recombinant protein.

The binding mode of the Ado molecules is similar to those observed in previously reported structures containing Ado or its analogues (Turner *et al.*, 1998; Tanaka *et al.*, 2004; Reddy *et al.*, 2008). The Ado-binding site is situated in the crevice between the substrate-binding and cofactor-binding domains. The Ado molecules interact with amino-acid residues from both the substrate-binding and cofactor-binding domain, and are stabilized in the active site by a number of hydrogen bonds and hydrophobic interactions. The hydrogen bonds are formed with the participation of both the main chain and side chains of the protein (Table 2) using all of the heteroatoms of

(*a*)

(*b*)

Figure 2
Comparison of the active site in subunit *A* (closed form) (*a*) and subunit *D* (open form) (*b*). The Ado molecule in (*a*) is shown in an $F_o - F_c$ OMIT electron-density map (calculated without the contribution of the Ado atoms to F_c) contoured at the 3σ level (corresponding to 0.17 e Å^{-3}). Water molecules in (*b*) are represented as red spheres. Dashed lines represent the hydrogen bonds listed in Table 2. The different domains of the BeSAHase subunit are highlighted by colours: blue, substrate-binding domain; salmon, cofactor-binding domain; sand, C-terminal domain. Green and orange colours represent, respectively, the first and second hinge regions between the substrate-binding and cofactor-binding domains.

Figure 1
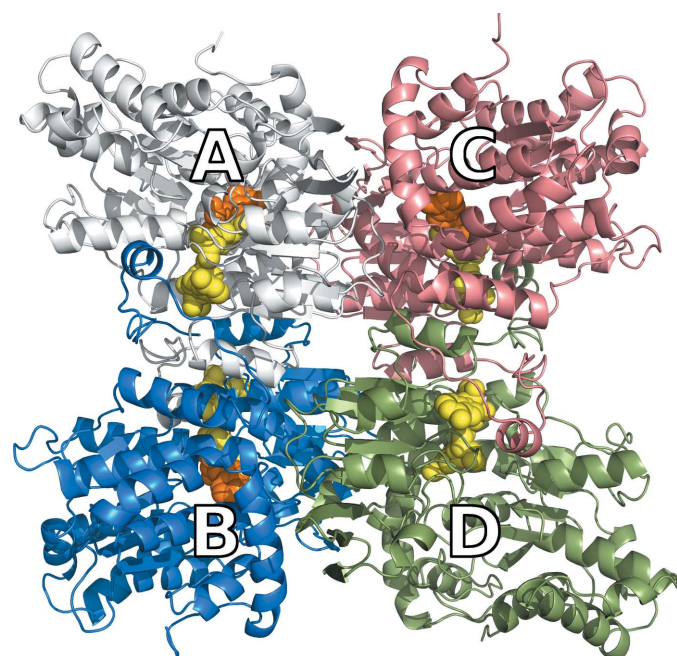A ribbon diagram of the BeSAHase tetramer, with the subunits labelled *A*, *B*, *C* and *D*. The NAD$^+$ (yellow) and adenosine (orange) molecules are shown in space-filling representation. The pseudodyad relating the subunits of the intimate dimers (*AB* and *CD*) is horizontal in this view. The remaining axes of the 222 pseudosymmetric tetramer are vertical and perpendicular to the paper.
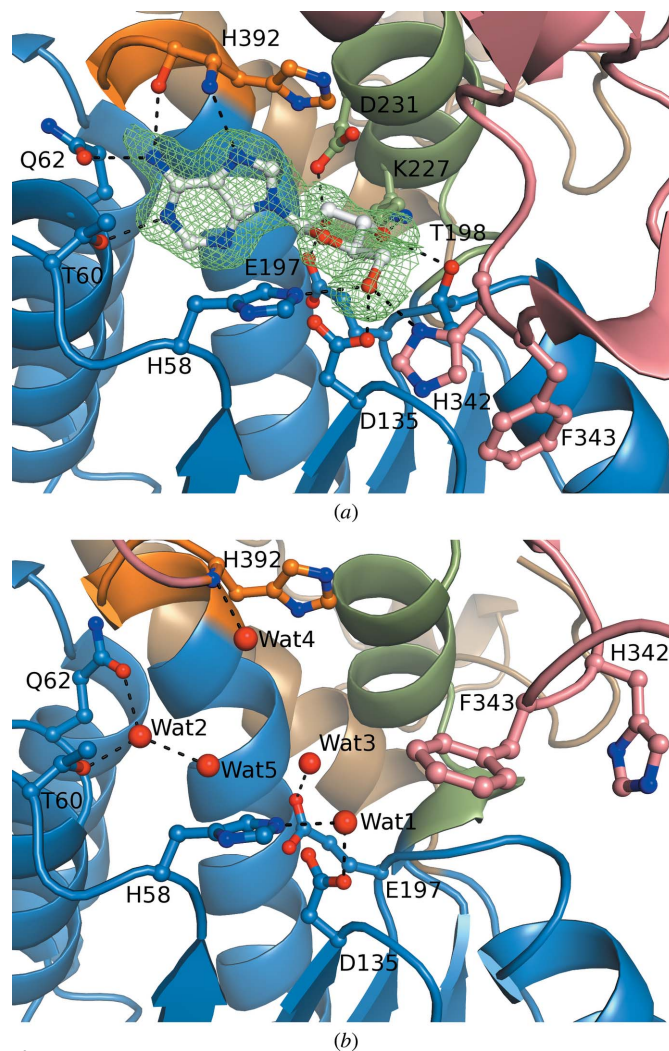
Table 2
Hydrogen bonds between the Ado molecules bound in subunits *A*/*B*/*C* and protein atoms.

For comparison, the contacts of the corresponding residues from subunit *D* are also shown.

Residue No. and atom	Ado atom	Distance (Å)	Interactions in subunit *D*
His58 N$^{\varepsilon2}$	O4′	3.17/3.08/3.22	No interaction
His58 N$^{\varepsilon2}$	O5′	2.86/2.76/2.86	Water(1)
Thr60 O$^{\gamma1}$	N1	2.76/2.74/2.73	Water(2)
Gln62 O$^{\varepsilon1}$	N6	2.82/3.01/2.95	Water(2)
Asp135 O$^{\delta1}$	O5′	2.90/2.82/2.87	Water(1)
Glu197 O$^{\varepsilon2}$	O2′	2.65/2.80/2.59	Water(3)
Thr198 O$^{\gamma1}$	O3′	3.21/2.98/3.05	No interaction
Lys227 N$^{\zeta}$	O3′	2.69/2.82/2.79	No interaction
Asp231 O$^{\delta1}$	O2′	2.72/2.51/2.70	No interaction
His342 N$^{\delta1}$	O5′	2.73/2.81/2.68	No interaction
His392 N	N7	2.86/2.88/2.99	Water(4)
His392 O	N6	3.07/3.09/3.01	No interaction

the Ado molecules except N3. The purine N6 atom is a donor of two hydrogen bonds to the His392 O atom, which is also involved in the coordination of the ammonium ion, and the Gln62 O$^{\varepsilon1}$ atom. These hydrogen bonds indicate that the adenine moiety is in the amino rather than the imino form. The ribose ring has a strained O4′-*endo* conformation, stabilized by two hydrogen bonds between the 3′-hydroxyl and the Thr198 O$^{\gamma1}$ and Lys227 N$^{\zeta}$ atoms.

A comparison with the Ado molecules bound to SAHase from *L. luteus* (Brzezinski *et al.*, 2012) reveals a high level of similarity of the conformational parameters, such as the *anti* orientation of the adenine ring around the glycosidic bond and the relative orientation of the O5′ atom against the ribose moiety, described by the C3′−C4′−C5′−O5′ torsion angle (γ), which is *trans*. Full conformational details of the Ado nucleosides, including the pseudorotation parameters of the ribofuranose rings, are listed in Table 3.

3.3. The adenosine-binding site in the ligand-free subunit

The active site of the ligand-free subunit *D* cannot be overlaid directly on those of subunits *A*, *B* and *C* in the closed conformation because the corresponding residues are too far apart. In general, the area occupied by Ado in subunit *A* is filled with five water molecules in subunit *D*. Four of these solvent molecules mimic the positions of Ado heteroatoms and therefore create a similar pattern of hydrogen bonds to the same amino-acid residues (Fig. 2*b*).

The position of the O5′ atom of the Ado molecule is mimicked in subunit *D* by a water molecule [Water(1) in Table 2] hydrogen-bonded to the His58 N$^{\varepsilon2}$ and Asp135 O$^{\delta1}$ atoms, which in subunit *A* interact with O5′. In the active sites occupied by Ado, there is also a hydrogen bond between O5′ and His342 N$^{\delta1}$, but His342, which is an element of the 'molecular gate' (see §3.5), in subunit *D* does not interact with Water(1) as it changes conformation completely to open the access to the empty active site (see below). The second water molecule [Water(2)] occupies the position corresponding to N1 of Ado and likewise creates a hydrogen bond to the O$^{\gamma1}$ atom of Thr60 and additionally to Gln62 O$^{\varepsilon1}$. This solvent

Table 3
Conformation of the Ado molecules found in subunits *A*, *B* and *C*.

The amplitude (τ_{m}) and phase angle (*P*) of pseudorotation were calculated by the method of Jaskólski (1984).

	Adenosine (*A*)	Adenosine (*B*)	Adenosine (*C*)
Glycosidic bond	*Anti*	*Anti*	*Anti*
Angle (°)†			
χ	−117.1	−106.7	−119.8
γ	−177.9	−179.1	−167.7
ν_0	−40.0	−28.1	−42.7
ν_1	18.5	11.2	19.0
ν_2	7.8	7.8	9.5
ν_3	−31.2	−24.3	−34.9
ν_4	45.8	33.9	49.8
P	79.8 (10)	75.8 (11)	78.4 (9)
τ_{m}	45.0 (8)	32.9 (7)	48.8 (8)
Sugar pucker	O4′-*endo* ($^0T_4$)	O4′-*endo* ($^0T_4$)	O4′-*endo* ($^0T_4$)

† The torsion angles defining the nucleoside conformation are as follows: χ, C4−N9−C1′−O4′; γ, O5′−C5′−C4′−C3′; ν_0, C4′−O4′−C1′−C2′; ν_1, O4′−C1′−C2′−C3′ *etc.*

Table 4
R.m.s. deviations (Å) for C$^{\alpha}$ atoms of superposed subunits of BeSAHase.

Values were calculated with *ALIGN* (Cohen, 1997).

Subunit	*A*	*B*	*C*
B	0.20		
C	0.12	0.21	
D	2.32	2.28	2.26

molecule also interacts *via* a hydrogen bond with another water molecule [Water(5)], which occupies the position of the N3 Ado atom, although the N3 atom itself does not form any hydrogen bonds to the enzyme (see above). The position of O2′ is occupied in subunit *D* by a third water molecule [Water(3)] which is hydrogen-bonded to the Glu197 O$^{\varepsilon1}$ atom. Water(4) does not mimic any Ado atom, but is involved in hydrogen bonding to the His392 N atom and in this way completes the active-site architecture of subunit *D*. In addition, the side chain of Gln62 is shifted into the active site in the ligand-free subunit, placing its O$^{\varepsilon1}$ atom in the position of the Ado N6 atom in the Ado complex.

3.4. Ligand-induced conformational change

Ado molecules were found in subunits *A*, *B* and *C*, but not in subunit *D*. Hu *et al.* (1999) showed that ligand binding in the active site of SAHase switches the subunit conformation from open to closed. In agreement with this observation, the subunits of the BeSAHase homotetramer are found in two conformational states: while the Ado-binding subunits *A*, *B* and *C* are in the closed conformation, the *D* subunit is open. The differences in terms of C$^{\alpha}$ r.m.s. deviations are shown in Table 4. Despite the difference in the overall conformation of the complexed and ligand-free subunits, the folds of the individual domains are the same, as illustrated by the small r.m.s.d. values calculated for C$^{\alpha}$ superpositions of subunits *A* and *D*, which are 0.32, 0.15 and 0.39 Å for the substrate-binding, cofactor-binding and dimerization domains, respectively. This indicates that there should be a region between the
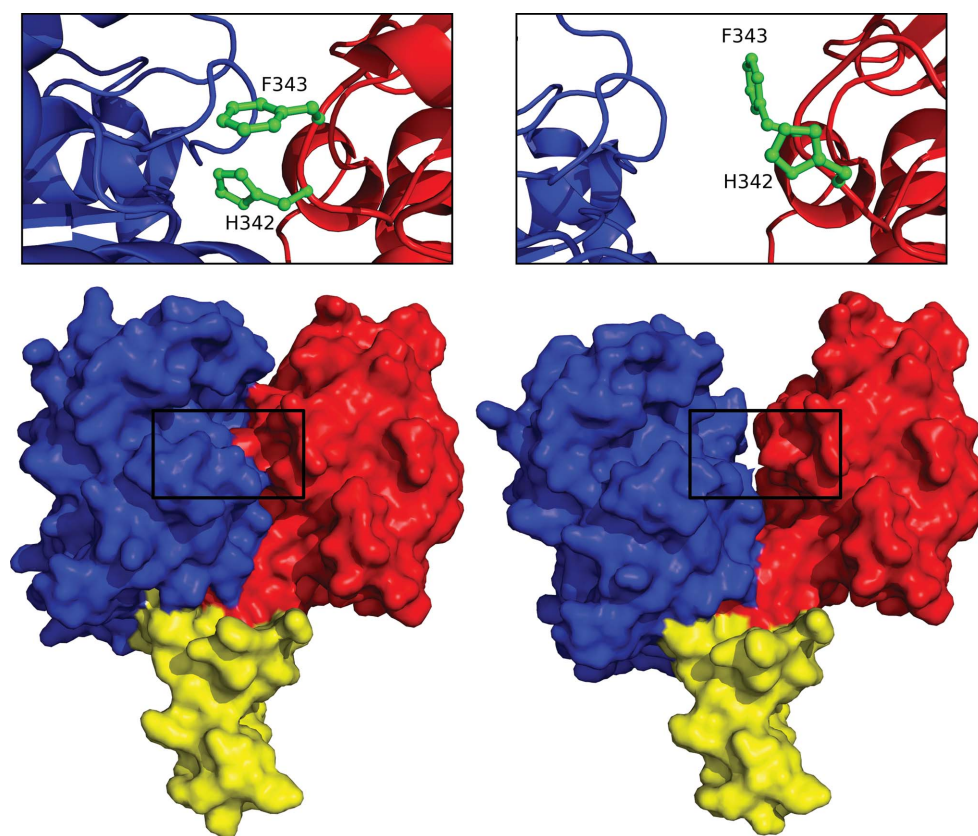
Table 5
Main-chain torsion angles (°) in subunits A (closed form) and D (open form) that undergo the most pronounced change ($|\Delta| > 10°$) upon domain movement.

Because of the similarity of subunits A, B and C, only values for subunit A are given. $|\Delta|$ changes in excess of 20° are shown in bold. The two hinge regions are separated by a line.

Residue No.	Torsion angle	Subunit A	Subunit D	$|\Delta|$
222	ω	178.7	−169.6	11.7
223	ψ	5.0	−11.2	16.2
223	ω	**161.5**	**−174.1**	**24.4**
224	φ	−58.2	−69.7	11.5
224	ψ	139.6	155.1	15.6
226	ψ	**1.8**	**−36.8**	**38.6**
226	ω	168.7	−179.2	10.5
227	φ	**−109.8**	**−71.7**	**38.1**
227	ψ	−64.0	−50.8	13.2
229	φ	**−95.7**	**−65.7**	**30.0**
229	ψ	**10.6**	**−41.1**	**51.7**
230	φ	**−127.0**	**−75.1**	**51.9**
230	ψ	**−66.1**	**−44.3**	**21.8**
232	φ	**−71.5**	**−98.9**	**27.4**
232	ψ	−35.4	−50.1	14.7
392	ψ	147.6	136.5	11.1
395	ψ	−48.7	−37.7	11.0
396	ψ	−32.2	−43.1	10.8

substrate-binding and cofactor-binding domains that functions as a hinge in the conformational transition. A comparison of the main-chain torsion angles of the open and closed subunits identifies the Asn222–Leu233 and His392–Val396 segments as the hinge regions, where the Ramachandran torsion angles differ by 13 and 7° on average, with maximum changes of 52 and 11°, respectively (Table 5). This indicates that the first hinge region is far more flexible and plays a key role in the conformational transition.

The fact that the Ado molecules were modelled with an occupancy of 0.7 implies that 30% of the $A/B/C$ subunits in the crystal lattice are apparently not occupied by the Ado ligand. Nevertheless, these subunits seem to be uniformly in the closed conformation, as there was no indication whatsoever in the electron density of dual conformation or disorder. Also, a C^α comparison with LlSAHase in a 1:1 complex with Ado (see §3.10) shows that BeSAHase subunits $A/B/C$ adopt a fully closed conformation. This would suggest that while binding of Ado (or a similar ligand) induces the closed conformation by necessity, this conformational state may also be assumed without a ligand (Zheng *et al.*, 2015), for instance upon stabilization by crystal-packing interactions.

3.5. The molecular gate to the active site

Reddy *et al.* (2008) showed that in the SAHase from *M. tuberculosis* His363 acts as a 'molecular gate' which opens and closes an access channel leading to the active site upon domain movement by rotating its side chain by ~180° around the backbone. In the BeSAHase structure, movement of the His342 side chain (corresponding to His363 in *M. tuberculosis* SAHase) is coupled with a rotation of the Phe343 side chain, and these residues together form a two-part access-channel gate (Fig. 3) which is closed upon adenosine binding (subunits A, B and C) and open in subunit D (Fig. 2). Strictly speaking, the access channel of BeSAHase is shut in the subunits in the closed conformation, because its closure in protomers $A/B/C$ seems to be insensitive to whether the Ado ligand is physically present (70%) or not (30%). In the recently reported structures of the plant SAHase from *L. luteus* (LlSA-Hase; Brzezinski *et al.*, 2012) the gate formed by the side chains of His350 (corresponding to His342 in BeSAHase) and Phe351 (corresponding to Phe343) was closed or open depending on the ligand molecule bound in the

Figure 3
BeSAHase subunits in the closed (left) and open (right) conformation. The domains are coloured as follows: blue, substrate-binding domain; red, cofactor-binding domain; yellow, C-terminal oligomerization domain. The side chains of His342-Phe343, forming the molecular gate at the passage to the active site, are highlighted in the enlargements.

active site, but in a way that is inconsistent with the situation in BeSAHase. With adenosine or 3′-deoxyadenosine (cordycepin) bound, the side chains of the molecular gate were swung open, while access to the active site was blocked by these side chains only in the complex with adenine. It is interesting to note that the opening of the access gate with the Ado molecule still in the active site (as in LlSAHase) is accompanied by disruption of one of the hydrogen bonds (from the gate His residue to O5′) that anchor the product molecule in the active site. In general, only plasmodial and plant SAHases have the molecular gate open in Ado-complexed forms, while in all eukaryotic and bacterial enzymes the access channel is closed upon Ado binding. The only exception is SAHase from *M. tuberculosis* in complex with SAH (Reddy *et al.*, 2008; PDB entry 3dhy), in which the side chain of His363 is shifted to a different position to avoid a steric clash with the Hcy moiety. The picture emerging from these structures suggests that upon substrate (SAH) binding the enzyme assumes the closed conformation but the access channel cannot be fully closed because of collision with the Hcy fragment. As the reaction progresses, the Hcy product dissociates from the active site in concert with full closure of the gate. This is the conformational state of the Ado-bound subunits of the BeSAHase structure. In the next phase, the gate opens (as in LlSAHase) to allow release of the Ado product, which is accompanied by the transition of the enzyme to the final open conformation, as seen in the 'empty' subunit *D* of the BeSAHase structure.

3.6. Quaternary interactions between the subunits

Since the BeSAHase tetramer should be considered as a dimer of dimers, the majority of the intramolecular interactions occur between the intimate dimer-forming subunits *AB* (and also *CD*). Almost all of the interactions between these subunits are formed by residues from the cofactor-binding (Tyr234–Gly391) and C-terminal (Leu427–Tyr473) domains. The exceptions are the hydrophobic interactions involving His203 from the substrate-binding domain. In the *AB* dimer there are eight such contacts out of the 271 hydrophobic interactions (six involving His203 from subunit *A* and two His203 from subunit *B*). In the *CD* dimer, five of the 259 hydrophobic interactions are formed exclusively by His203 from subunit *C*. The borderline asymmetry within the dimers is explained by the (generally slight) conformational differences between the individual subunits and the hydrophobic interaction cutoff distance set at 3.9 Å in the *PDBsum* server (Laskowski, 2009).

In the *AB* dimer there are 26 hydrogen bonds between the protomers. Among them there are four that are not present in
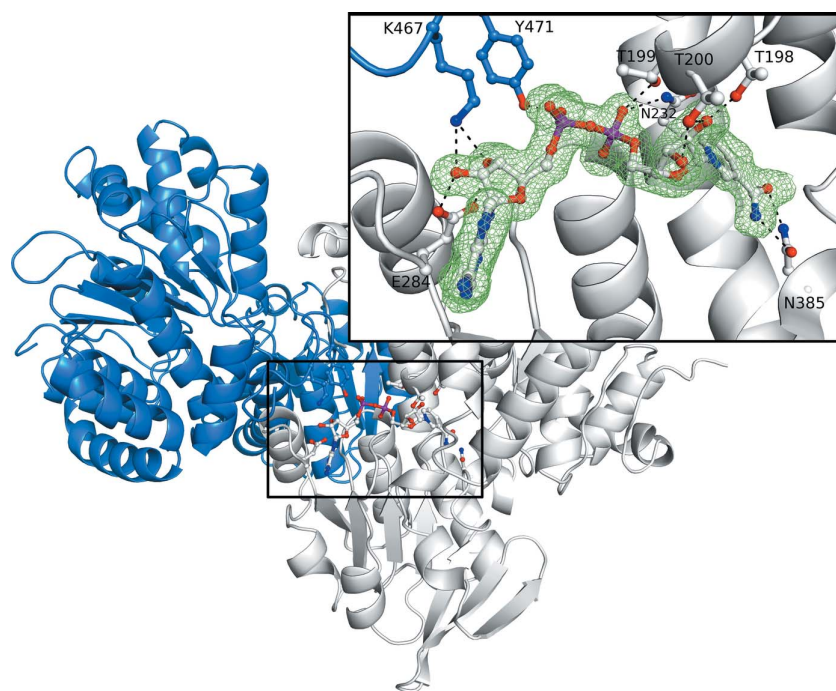

Figure 4
The intimate dimer of BeSAHase subunits *A* (grey) and *B* (blue) formed with the participation of the NAD$^+$ cofactor, shown in an $F_o - F_c$ OMIT electron-density map (calculated without the contribution of the NAD$^+$ atoms to F_c) contoured at the 3σ level (corresponding to 0.17 e Å^{-3}). For clarity, only selected hydrogen bonds (dashed lines) involving protein side chains are shown. All contacts are listed in Table 7.

the *CD* dimer: Glu296*A* O$^{\varepsilon2}$···O$^\eta$ Tyr234*B*, Arg451*A* N$^{\eta1}$···O Asp321*B*, Arg472*A* O···N$^\zeta$ Lys229*B* and Tyr473*A* O···N$^\zeta$ Lys229*B*. Also, two interactions in the *A*···*B* direction are unique and are not present in the opposite *B*···*A* direction: Arg452*A* N$^{\eta1}$···O Asp321*B* and Arg472*A* O···N$^\zeta$ Lys229*B*. In the *CD* dimer there are 24 hydrogen bonds between the subunits, among which the Lys320*C* N$^\zeta$···O$^{\delta2}$ Asp453*D* and Arg472*C* N$^{\eta2}$···O Asp223*D* interactions are unique to this dimer. The Lys229*C* N$^\zeta$···O Tyr473*D*, Tyr234*C* O$^\eta$···O$^{\varepsilon2}$ Glu296*D*, Lys320*C* N$^\zeta$···O$^{\delta2}$ Asp453*D* and Arg472*C* N$^{\eta2}$···O Asp223*D* interactions are present only in the *C*···*D* direction. As seen from this hydrogen-bond inventory, the conformational state of the subunits does not significantly affect the number of interactions between them in the intimate dimers and the presence of Ado molecules in the active site is not required for dimer formation.

Among the numerous hydrogen bonds and hydrophobic contacts, the most crucial for dimer formation seem to be the interactions formed by the small dimerization domain at the C-terminus. Two residues from this domain, Lys467 and Tyr471, are involved in cofactor binding in the complementary subunit (Fig. 4). Lys467 forms two hydrogen bonds using its N$^\zeta$ atom to O2B and O3B of the NAD$^+$ molecule. Tyr471 O$^\eta$ forms a hydrogen bond to the O1A atom of the cofactor (for NAD$^+$ atom numbering, see Fig. 5). This way, the complete cofactor-binding site is created with the participation of both protomers forming the tight dimer. Ault-Riché *et al.* (1994) showed that mutation of Lys426 in human SAHase (corresponding to Lys467 in BeSAHase) to alanine or glutamate

Table 6
Interaction interfaces between the subunits forming the BeSAHase tetramer, calculated with *PDBsum* (Laskowski, 2009).

N_r, number of interface residues; N_h, number of hydrogen bonds; N_{nbc}, number of nonbonded contacts.

Protein chains	N_r	N_h	N_{nbc}	Interface area (Å^2)
A:B	51:52	26	271	2975:2960
C:D	49:49	24	259	2832:2818
A:C	31:31	26	187	1685:1693
B:D	28:28	26	173	1593:1608
A:D	7:7	0	23	370:367
B:C	8:8	0	20	386:386

resulted in inactive, monomeric protein without the cofactor.

The numbers of interactions between subunits that do not form tight dimers (*e.g. A–C, A–D etc.*) are smaller than in the cases described above. Contacts between subunits that are juxtaposed in Fig. 1 (*A–C* and *B–D*) are formed by residues from all domains. Nevertheless, there are 26 hydrogen bonds in each pair involving amino-acid residues from the substrate-binding and cofactor-binding domains. However, the number of hydrophobic contacts is much smaller than in the tight dimers, with 187 hydrophobic interactions in the *A–C* pair and 173 in the *B–D* pair, with the involvement of only one residue from the C-terminal domain: Ile443 from subunit *A* or *B*.

The remaining pairs within the homotetramer, *i.e. A–D* and *B–C*, form no hydrogen bonds. There are only 23 hydrophobic

Figure 5
The molecule of NADH (top) from the PDB structure 1het (Meijers *et al.*, 2001) and NAD$^+$ (bottom) from the present BeSAHase structure (subunit *A*). The difference in the puckering of the nicotinamide ring indicates the oxidation state of each molecule, shown in $2F_o - F_c$ electron density contoured at 1σ, which corresponds to 0.72 e Å^{-3} for NADH and 0.48 e Å^{-3} for NAD$^+$. The chemical formulae illustrate the principal electronic forms of the molecules and the numbering of some key atoms of NAD$^+$.

Table 7
Hydrogen-bond interactions between the cofactor NAD$^+$ molecule and protein atoms.

Because of the similarity of all four cofactor-binding sites, only values for NAD$^+$ bound in subunit *A* are shown. The last three rows show interactions with amino-acid residues from the complementary subunit (*B*) of the tight dimer. The NAD$^+$ atom-labelling scheme is introduced in Fig. 5.

Residue No. and atom	NAD$^+$ atom	Distance (Å)
Thr198 O$^{\gamma1}$	O2D	2.76
Thr199 O$^{\gamma1}$	O1N	2.73
Thr200 N	O2D	3.21
Thr200 O$^{\gamma1}$	O2D	3.34
Thr200 O$^{\gamma1}$	O3D	2.67
Asn232 N$^{\delta2}$	O1N	3.32
Val265 N	O2N	2.87
Glu284 O$^{\varepsilon1}$	O3B	2.78
Glu284 O$^{\varepsilon2}$	O2B	2.66
Ile340 O	N7N	2.81
His342 N	O3D	3.09
Asn385 O$^{\delta1}$	N7N	3.10
Asn385 N$^{\delta2}$	O7N	2.92
Lys467(*B*) N$^{\xi}$	O2B	3.10
Lys467(*B*) N$^{\xi}$	O3B	2.84
Tyr471(*B*) O$^{\eta}$	O1A	2.54

interactions in the former case and 20 in the latter. These contacts involve residues from the substrate-binding domains only. Detailed information about the protein–protein interactions within the homotetramer is summarized in Table 6.

3.7. Crystal packing and interactions between the tetramers

BeSAHase crystallized in space group $P2_12_12$ with four protein molecules forming the complete 222 pseudosymmetric tetramer in the asymmetric unit (Fig. 1). Although this space group contains point symmetry compatible with the molecular symmetry of the enzyme, it is not utilized by the tetramer in the asymmetric unit. The pseudo-twofold molecular axis that relates subunits *A* to *D* (and *B* to *C*) is nearly parallel (3.3° inclination) to the [100] direction. The remaining molecular axes are not aligned with crystallographic directions.

PISA (Krissinel & Henrick, 2007) analysis revealed that the largest interaction area between symmetrically equivalent protein molecules is less than 500 Å^2. This value is comparable with the interfaces between subunits with a diagonal disposition across the tetramer (*A–D* and *B–C*; Table 6).

3.8. The cofactor: mode of binding and oxidation state

The cofactor-binding site is highly conserved in all SAHases. Most of the interactions are provided by residues from the cofactor-binding domain. However, as described above, the C-terminal domain of the complementary subunit also participates in NAD$^+$ binding, at the same time providing the most important cohesive force for the intimate dimer. All cofactor interactions are listed in Table 7.

To support the conclusion about the oxidation state of the cofactor, we analyzed the puckering of the nicotinamide moiety. It has been shown by Meijers *et al.* (2001) that addition of hydride anion to the NAD$^+$ nicotinamide ring causes a loss of its aromatic character and its deformation into a boat conformation. To obtain an unbiased view, we refined the
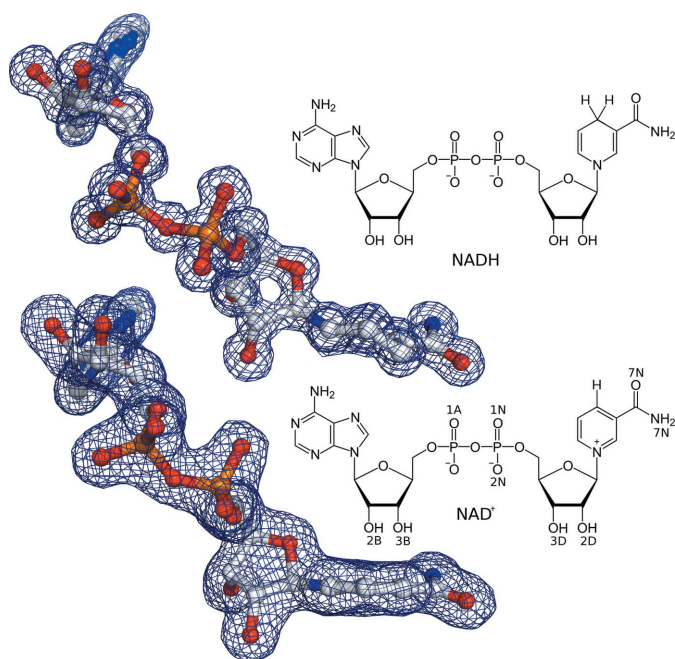
Table 8
Hydrogen-bond distances around the ammonium ions in subunits $A/B/C$.

Ligand atom	Bond length (Å)
Met390 O	3.00/2.98/2.94
His392 O	3.16/3.09/3.08
Water(A)	2.59/2.37/2.53
Water(B)	2.78/2.89/2.92
Water(C)	2.98/3.09/2.94

Table 9
R.m.s. deviations (Å) for C^{α} atoms of superposed subunits of BeSAHase and RnSAHase in closed and open conformations calculated with *ALIGN* (Cohen, 1997).

	BeSAHase, closed	BeSAHase, open	RnSAHase, closed
BeSAHase, open	2.32		
RnSAHase, closed	0.75	2.05	
RnSAHase, open	2.14	0.97	2.12

NAD^+ molecules without planarity restraints and found that the nicotinamide ring of all of these moieties in the BeSAHase crystal structure is flat (Fig. 5). Thus, it can be concluded that the cofactor is in its oxidized NAD^+ state. The question about the cofactor oxidation state has a biochemical and an electrochemical aspect. We tried to address the former aspect by incubating the protein with a cofactor of uniform oxidation state (NAD^+; see §2.1). The latter aspect is connected to the belief that the X-ray beam (through generation of electrons) provides a strongly reducing milieu. However, in the crystal of BeSAHase this evidently did not lead to a noticeable effect.

3.9. Ammonium/metal cations and their role in structure stabilization

In the Ado-bound subunits A, B and C, an ammonium cation was found in the loop Ala389–Pro393 (Fig. 6) that links the substrate-binding and cofactor-binding domains and is located near the active site of the enzyme. The last two resi-
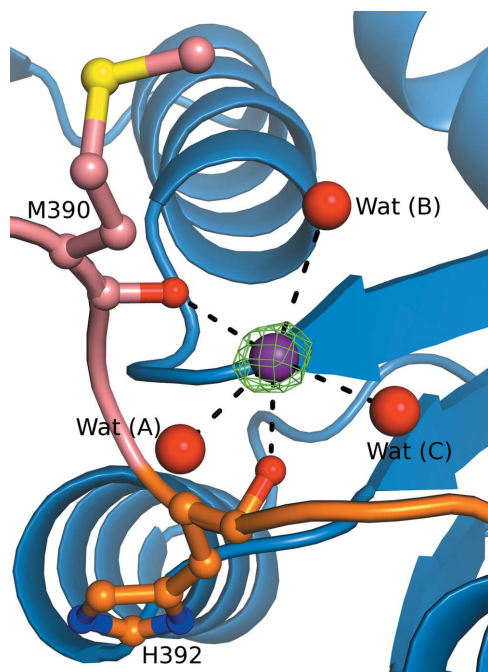

Figure 6
The ammonium ion (violet sphere) bound in subunit A, shown in $F_o - F_c$ OMIT electron density (calculated without the contribution of the NH_4^+ atoms to F_c) contoured at 5σ (corresponding to $0.27\ e\ \text{Å}^{-3}$). Hydrogen-bond (dashed lines) details are listed in Table 8. Water molecules are represented as red spheres. The domains are colour-coded as follows: blue, substrate-binding domain; salmon, cofactor-binding domain. Orange colour represents the second hinge region between these domains.

dues of this loop, His392 and Pro393, are also part of the second hinge region. This loop corresponds to the Ala401–Pro405 loop of LlSAHase from *L. luteus*, which was also shown to bind a sodium cation (Brzezinski *et al.*, 2012) in the closed form of the subunits. The ammonium ion was present at a relatively high concentration (0.2 *M*) in the crystallization buffer. Its identification is based on the hydrogen-bond distances (Table 8) and was additionally validated by the calcium bond-valence sum method (Müller *et al.*, 2003). In the pattern of interactions, the NH_4^+ ion forms hydrogen bonds to five O acceptors (including three water molecules), meaning that one of the NH donors forms a bifurcated hydrogen bond. No cationic species were found in the Ala389–Pro393 loop of the open subunit D, but a poorly bound water molecule is situated near the carbonyl group of Met390.

The presence of the ammonium ion only in those subunits that have Ado molecules bound in the active site and are in the closed conformation may suggest that it is important for stabilizing this conformation. It was shown by Yin *et al.* (2000) that the cofactor-binding and substrate-binding domains are mobile until ligand binding. Wang *et al.* (2006) in turn showed that the mutation of His353 in human SAHase (corresponding to His392 in BeSAHase that takes part in NH_4^+ binding) to alanine slowed down the domain mobility. The structure of BeSAHase sheds new light on this aspect, showing that the closed-conformation ligand-bound state is additionally stabilized by a cation bound in a loop linking the cofactor-binding and substrate-binding domains. Apparently, the ion bound in this loop can be either an alkali-metal cation, as in LlSAHase (Na^+) and in SAHase from *B. melitensis* (K^+), or NH_4^+, as in BeSAHase.

3.10. BeSAHase *versus* other SAHases: conformational analysis

Among the SAHase models deposited in the PDB, BeSA-Hase is the first structure in which two conformational states of the enzyme are present in one crystal. SAHase from *R. norvegicus* (RnSAHase) is the only protein for which structures in the open (Hu *et al.*, 1999; PDB entry 1b3r) and closed (Ado complex; Komoto *et al.*, 2000; PDB entry 1d4f) conformations have been observed, but in distinct crystal structures. Superpositions of the C^{α} traces (Table 9) show that the differences between the open and closed conformations are the same in both proteins, with corresponding r.m.s.d. values of 2.32 and 2.12 Å for BeSAHase and RnSAHase, respectively. Comparison of the closed and open subunits of the two enzymes shows that protomers in the same confor-

mational state are quite similar, with r.m.s.d. values of 0.75 and 0.97 Å, respectively. This shows that the open-closed conformational transition is structurally conserved among SAHases from different organisms. Also, when the closed subunits of BeSAHase are compared with the closed subunits of SAHases from other organisms, *L. luteus*, *B. melitensis* (both in complex with Ado) and from *H. sapiens* in complex with neplanocin, a high degree of structural conservation is revealed, with C$^\alpha$ r.m.s.d. values of 0.51, 0.33 and 0.64 Å, respectively.

4. Conclusions

The crystal structure of SAHase from *B. elkanii* was solved at a resolution of 1.74 Å, revealing a homotetramer in the asymmetric unit of space group $P2_12_12$. Although only NAD$^+$ molecules were deliberately added during protein preparation and crystallization, molecules of Ado, which is one of the products of the reaction catalyzed by SAHase, were found in the electron-density maps in three of the four subunits of the tetrameric enzyme. The *A*, *B* and *C* subunits, in which the Ado product is bound, have a closed conformation, in contrast to the 'empty' subunit *D*, which is open. In the closed conformation, the substrate-binding and cofactor-binding domains are brought together to allow the catalytic reaction to proceed. The structure clearly shows that transition between the open and closed states requires a significant conformational change of a hinge segment of five residues and is coupled to a conformational switch of two gating residues (His342 and Phe343), which control access to the active site. In addition, the closed conformation is stabilized by an NH$_4^+$ cation bound in a loop connecting the substrate-binding and cofactor-binding domains. The identity of the bound cation, consistent with the high concentration of NH$_4^+$ in the crystallization buffer, was confirmed by the pattern of hydrogen-bond distances and the calcium bond-valence sum method (Müller *et al.*, 2003). In other SAHases studied in the closed conformation, the cation-binding loop coordinates an alkali-metal ion. In each subunit, there is clear electron density indicating the presence of the NAD$^+$ cofactor, the oxidation state of which is confirmed by the flat (no puckering) nicotinamide ring. The cofactor-binding site of subunit *A* is complemented in a mutual fashion by residues from the C-terminal oligomerization domain of subunit *B*, and an analogous relation exists between subunits *C* and *D*. Owing to these interactions, the tetrameric enzyme should be considered as a dimer of two tight dimers (*AB* and *CD*).

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Ault-Riché, D. B., Yuan, C.-S. & Borchardt, R. T. (1994). *J. Biol. Chem.* **269**, 31472–31478.
Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
Brzezinski, K., Bujacz, G. & Jaskolski, M. (2008). *Acta Cryst.* F**64**, 671–673.
Brzezinski, K., Dauter, Z. & Jaskolski, M. (2012). *Acta Cryst.* D**68**, 218–231.
Cantoni, G. L. (1975). *Annu. Rev. Biochem.* **44**, 435–451.
Cantoni, G. L. & Chiang, P. K. (1980). *Natural Sulfur Compounds*, pp. 67–80. New York: Plenum.
Chiang, P. K. & Cantoni, G. L. (1979). *Biochem. Pharmacol.* **28**, 1897–1902.
Cohen, G. H. (1997). *J. Appl. Cryst.* **30**, 1160–1161.
De La Haba, G. & Cantoni, G. L. (1959). *J. Biol. Chem.* **234**, 603–608.
DeLano, W. L. (2002). *PyMOL*. http://www.pymol.org.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Fujioka, M. & Takata, Y. (1981). *J. Biol. Chem.* **256**, 1631–1635.
Guranowski, A. & Pawełkiewicz, J. (1977). *Eur. J. Biochem.* **80**, 517–523.
Hershfield, M. S. (1979). *J. Biol. Chem.* **254**, 22–25.
Hershfield, M. S., Kredich, N. M., Ownby, D. R., Ownby, H. & Buckley, R. (1979). *J. Clin. Invest.* **63**, 807–811.
Hershfield, M. S. & Kredich, N. M. (1978). *Science*, **202**, 757–760.
Hu, Y., Komoto, J., Huang, Y., Gomi, T., Ogawa, H., Takata, Y., Fujioka, M. & Takusagawa, F. (1999). *Biochemistry*, **38**, 8323–8333.
Huang, Y., Komoto, J., Takata, Y., Powell, D. R., Gomi, T., Ogawa, H., Fujioka, M. & Takusagawa, F. (2002). *J. Biol. Chem.* **277**, 7477–7482.
Jaskólski, M. (1984). *Acta Cryst.* A**40**, 364–366.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Keller, B. T. & Borchardt, R. T. (1988). *Antivirial Drug Development: A Multidisciplinary Approach*, edited by E. De Clerq & R. T. Walker, pp. 123–138. New York: Plenum.
Komoto, J., Huang, Y., Gomi, T., Ogawa, H., Takata, Y., Fujioka, M. & Takusagawa, F. (2000). *J. Biol. Chem.* **275**, 32147–32156.
Kredich, N. M. & Martin, D. W. Jr (1977). *Cell*, **12**, 931–938.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. (2008). *Nature Protoc.* **3**, 1171–1179.
Laskowski, R. A. (2009). *Nucleic Acids Res.* **37**, D355–D359.
Liu, S., Wolfe, M. S. & Borchardt, R. T. (1992). *Antiviral Res.* **19**, 247–265.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
McKeever, M. P., Weir, D. G., Molloy, A. & Scott, J. M. (1991). *Clin. Sci.* **81**, 551–556.
Meijers, R., Morris, R. J., Adolph, H. W., Merli, A., Lamzin, V. S. & Cedergren-Zeppezauer, E. S. (2001). *J. Biol. Chem.* **276**, 9316–9321.
Mitsui, S., Wakasugi, T. & Sugiura, M. (1993). *Plant Cell Physiol.* **34**, 1089–1096.
Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. (2009). *Acta Cryst.* D**65**, 1074–1080.
Müller, P., Köpke, S. & Sheldrick, G. M. (2003). *Acta Cryst.* D**59**, 32–37.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Nygård, O., Nordrehaug, J. E., Refsum, H., Ueland, P. M., Farstad, M. & Vollset, S. E. (1997). *N. Engl. J. Med.* **337**, 230–237.
Painter, J. & Merritt, E. A. (2006). *Acta Cryst.* D**62**, 439–450.
Palmer, J. L. & Abeles, R. H. (1976). *J. Biol. Chem.* **251**, 5817–5819.
Palmer, J. L. & Abeles, R. H. (1979). *J. Biol. Chem.* **254**, 1217–1226.
Reddy, M. C., Kuppan, G., Shetty, N. D., Owen, J. L., Ioerger, T. R. & Sacchettini, J. C. (2008). *Protein Sci.* **17**, 2134–2144.
Richards, H. H., Chiang, P. K. & Cantoni, G. L. (1978). *J. Biol. Chem.* **253**, 4476–4480.
Robinson, K., Mayer, E. & Jacobsen, D. W. (1994). *Cleve. Clin. J. Med.* **61**, 438–450.

research papers

Stępkowski, T., Brzeziński, K., Legocki, A. B., Jaskólski, M. & Béna, G. (2005). *Mol. Phylogenet. Evol.* **34**, 15–28.

Takata, Y., Yamada, T., Huang, Y., Komoto, J., Gomi, T., Ogawa, H., Fujioka, M. & Takusagawa, F. (2002). *J. Biol. Chem.* **277**, 22670–22676.

Tanaka, N., Nakanishi, M., Kusakabe, Y., Shiraiwa, K., Yabe, S., Ito, Y., Kitade, Y. & Nakamura, K. T. (2004). *J. Mol. Biol.* **343**, 1007–1017.

Turner, M. A., Yuan, C.-S., Borchardt, R. T., Hershfield, M. S., Smith, G. D. & Howell, P. L. (1998). *Nature Struct. Biol.* **5**, 369–376.

Wang, M., Unruh, J. R., Johnson, C. K., Kuczera, K., Schowen, R. L. & Borchardt, R. T. (2006). *Biochemistry*, **45**, 7778–7786.

Yang, X., Hu, Y., Yin, D. H., Turner, M. A., Wang, M., Borchardt, R. T., Howell, P. L., Kuczera, K. & Schowen, R. L. (2003). *Biochemistry*, **42**, 1900–1909.

Yin, D., Yang, X., Hu, Y., Kuczera, K., Schowen, R. L., Borchardt, R. T. & Squier, T. C. (2000). *Biochemistry*, **39**, 9811–9818.

Zheng, Y., Chen, C.-C., Ko, T.-P., Xiao, X., Yang, Y., Huang, C.-H., Qian, G., Shao, W. & Guo, R.-T. (2015). *J. Struct. Biol.* **190**, 135–142.

3

Manszewski T, Szpotkowski K, Jaskólski M. (2017)

Crystallographic and SAXS studies on S-adenosyl-L-homocysteine hydrolase from

Bradyrhizobium elkanii. *IUCrJ*, in press.

Crystallographic and SAXS studies of S-adenosyl-L-homocysteine hydrolase from Bradyrhizobium elkanii

Tomasz Manszewski,[a] Kamil Szpotkowski[a] and Mariusz Jaskolski[a,b]*

[a]Center for Biocrystallographic Research, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan, Poland, and [b]Department of Crystallography, Faculty of Chemistry, A. Mickiewicz University, Poznan, Poland. *Correspondence e-mail: mariuszj@amu.edu.pl

S-Adenosyl-L-homocysteine hydrolase (SAHase) from the symbiotic bacterium *Bradyrhizobium elkanii* (BeSAHase) was crystallized in four ligand complexes with (i) mixed adenosine (Ado) and cordycepin (Cord; 3'-deoxyadenosine), (ii) adenine (Ade), (iii) Ado and (iv) mixed 2'-deoxyadenosine (2'-dAdo) and Ade. The crystal structures were solved at resolutions of 1.84, 1.95, 1.95 and 1.54 Å, respectively. Only the Ade complex crystallized with a dimer in the asymmetric unit, while all of the other complexes formed a crystallographically independent tetrameric assembly. In the Ado/Cord complex, adenosine is found in three subunits while the fourth subunit has cordycepin bound in the active site. In the Ade and Ado complexes only these ligand molecules are present in the active sites. The 2'-dAdo/Ade complex has Ade bound in two subunits and 2'-dAdo bound in the other two subunits. The BeSAHase fold adopted a closed conformation in the complexes with Ado, Ade and 2'-dAdo, and a semi-open conformation when cordycepin occupied the active site. An SAHase-specific molecular gate, consisting of residues His342 and Phe343, behaves differently in the different complexes, but there is no simple correlation with the ligand type. Additional small-angle X-ray scattering (SAXS) experiments confirm the tetrameric state of the protein in solution. The main conclusions from this work are (i) that the SAHase subunit does not simply oscillate between two discrete conformational open/closed states in correlation with the absence/presence of a ligand in the active site, but can also assume an intermediate form for some ligands; (ii) that the shut/open state of the molecular gate in the access channel to the active site is not correlated in a simple way with the open/closed subunit conformation or empty/occupied status of the active site, but that a variety of states are possible even for the same ligand; (iii) that a cation (typically sodium) coordinated in an intersubunit loop rigidifies a molecular hinge and thus stabilizes the closed conformation; (iv) that BeSAHase in solution is a tetramer, consistent with the model derived from crystallography.

1. Introduction

S-Adenosyl-L-methionine (SAM)-dependent methyltransferases are inhibited by S-adenosyl-L-homocysteine (SAH), a byproduct of the methylation reaction (Chiang & Cantoni, 1979; Cantoni & Chiang, 1980; Liu *et al.*, 1992). SAM is the most commonly used substrate for cellular methylation processes of a variety of compounds (Cantoni, 1975; Richards *et al.*, 1978). SAH, which is formed upon methyl-group transfer from SAM, must be removed from the reaction equilibrium (which is shifted far towards SAH formation) by its hydrolase (SAHase; De La Haba & Cantoni, 1959), which splits it into Ado and homocysteine (Hcy). In this way, SAHase is a key regulator of SAM-dependent methylation reactions and of Hcy levels, which, when elevated, are a risk factor for neurodegenerative diseases, neural tube defects and cardiovascular diseases (Jakubowski, 2006).

Files: m/lz5013/lz5013.3d m/lz5013/lz5013.sgml **LZ5013** FA IU-1715/11(22)2 1715/11(22)2 ()   LZ5013 PROOFS M:FA:2017:4:3:0:0-0
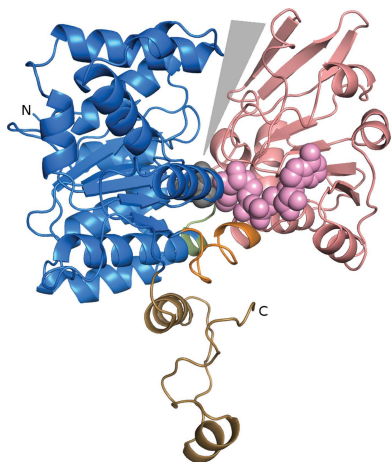
The mechanism of SAH hydrolysis catalyzed by SAHase was first elucidated by Palmer & Abeles (1976, 1979). It assumes the formation of a 3'-keto derivative *via* oxidation of SAH by NAD$^+$, followed by H3' proton abstraction from the ribose moiety by the N$^\zeta$ atom of a catalytic lysine residue and by Hcy elimination. After a Michael-type addition of a water molecule, the final product is reduced by NADH, which is regenerated to NAD$^+$.

SAHases are present in all kingdoms of life, although their sequences show notable differences, usually in the form of kingdom-specific insertions or deletions (Stępkowski *et al.*, 2005). Most SAHases function as homotetramers, although the enzymes from plants have been shown to be active as dimers (Guranowski & Pawełkiewicz, 1977; Brzezinski *et al.*, 2008). For enzymatic activity, the protein requires one NAD$^+$ cofactor molecule bound in each subunit near the active site (Palmer & Abeles, 1979; Fujioka & Takata, 1981).

A typical SAHase protomer is usually divided into three closely connected domains: an N-terminal substrate-binding domain, a cofactor-binding domain and a small C-terminal dimerization domain (Brzezinski *et al.*, 2012). The substrate-binding and cofactor-binding domains are separated by a bipartite hinge (Manszewski *et al.*, 2015), allowing them to oscillate between two conformational states, closed and open, during the catalytic cycle. The former state is assumed when the domains are close together and is thought to be stabilized by a ligand molecule (substrate, product or inhibitor) bound in the active site (Hu *et al.*, 1999; Yin *et al.*, 2000) and an alkali-metal or ammonium cation coordinated in a metal-binding loop near the active site (Brzezinski *et al.*, 2012; Manszewski *et al.*, 2015). However, the closed conformation has also been observed for a ligand-free protein (Zheng *et al.*, 2015). In the open conformation, which is considered to be the characteristic state of the ligand-free enzyme, the principal domains are spatially separated, forming a distinct substrate-access channel leading to the active site. Access to the active site is then regulated by two side chains (His and Phe) forming a molecular gate (Reddy *et al.*, 2008; Manszewski *et al.*, 2015).

Several crystal structures of SAHases from different organisms are available in the Protein Data Bank (PDB; Berman *et al.*, 2000); for example, mammalian [*Homo sapiens*, PDB entries 1a7a (Turner *et al.*, 1998) and 1li4 (Yang *et al.*, 2003); *Rattus norvegicus*, PDB entries 1b3r (Hu *et al.*, 1999), 1k0u (Huang *et al.*, 2002), 1ky5 (Takata *et al.*, 2002) and 1d4f (Komoto *et al.*, 2000)], plant (*Lupinus luteus*, PDB entries 3ond, 3one and 3onf; Brzezinski *et al.*, 2012), bacterial (*Mycobacterium tuberculosis*, PDB entry 3ce6; Reddy *et al.*, 2008) and protozoan (*Plasomodium falciparum*, PDB entry 1v8b; Tanaka *et al.*, 2004).

In our previous studies of BeSAHase from *Bradyrhiziobium elkanii*, which is the nitrogen-fixing bacterial symbiont of soybean (PDB entry 4lvc; Manszewski *et al.*, 2015), we described the enzyme in complex with Ado molecules that were bound in three of the four subunits of the tetramer. In this complex, the Ado-bound subunits were found in the closed conformation stabilized by an ammonium cation bound in the metal-binding loop, while the ligand-free subunit was in an open conformation and without an ammonium cation. Moreover, the molecular gate (MG) leading to the active site, formed by residues His342 and Phe343, was found to be shut when Ado was bound in the active site and open in the ligand-free subunit. As a continuation of these studies, here we present four crystal structures of BeSAHase in complex with different ligand molecules, namely as a mixed 3:1 complex with adenosine (Ado) and cordycepin (Cord; 3'-deoxy-adenosine), in complexes with adenine (Ade) and with adenosine (Ado), and as a mixed 1:1 complex with 2'-deoxy-adenosine (2'-dAdo) and Ade. The set of ligands used in our study is the same as in the study of LlSAHase from *L. luteus* (Brzezinski *et al.*, 2012) to facilitate comparison of the active-site architecture and protein conformation. The presented crystal structures shed new light on the question of subunit conformation in relation to the ligand type and on the details of the MG mechanism, showing that, contrary to the tentative view held to date, no simple correlation exists. In addition, small-angle X-ray scattering (SAXS) measurements conclusively demonstrate that the crystallographic model of the BeSAHase tetramer is consistent with the protein structure in solution.

2. Materials and methods

2.1. Purification

The purification of BeSAHase was carried out as described previously (Manszewski *et al.*, 2015) only for the mixed Ado/Cord complex (procedure P1). When it became apparent that the above ligand-exchange procedure *via* incubation does not necessarily lead to a complex with the desired ligand (see §3.1.1), a modification of the procedure of Yuan *et al.* (1993) was implemented as a step before size-exclusion chromatography to exchange the cofactor and prepare ligand-free enzyme for the subsequent incubation reaction. In this modified approach (procedure P2), 5 ml BeSAHase in buffer *A* (50 m*M* Tris pH 8.0, 500 m*M* NaCl, 20 m*M* imidazole) at a concentration of 12 mg ml^{-1} (estimated according to Bradford, 1976) was mixed with 10 ml saturated ammonium sulfate solution pH 3.3 and stored on ice for 10 min. The mixture was centrifuged, and the precipitate was dissolved in 5 ml buffer *A* and mixed again with ammonium sulfate solution as above. The pellet was again dissolved in 5 ml buffer *A* and mixed with 10 ml saturated ammonium sulfate solution pH 7.0. Finally, the pellet was dissolved in 5 ml buffer *A* and the protein concentration was adjusted to 8 mg ml^{-1}. Ultimately, NAD$^+$ was added in a 12-fold molar excess and the mixture was stored at 4°C for 12 h.

2.2. Crystallization

2.2.1. BeSAHase complex with adenosine/cordycepin. BeSAHase solution obtained without the cofactor-exchange step (procedure P1) at a concentration of 12 mg ml^{-1} was incubated with an eightfold molar excess of cordycepin for 24 h. The mixture was submitted for crystallization in the High-Throughput Crystallization Facility at EMBL Hamburg

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the last resolution shell.

Complex	Ado/Cord	Ade	Ado	Ade/2′-dAdo
Data collection				
Beamline	P14, EMBL/DESY	14.2, BESSY	14.2, BESSY	P13, EMBL/DESY
Wavelength (Å)	1.223	0.918	0.918	0.969
Temperature (K)	100	100	100	100
Space group	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$
Unit-cell parameters				
a (Å)	108.2	100.7	106.4	107.5
b (Å)	176.3	103.1	173.6	174.5
c (Å)	102.3	90.7	97.4	96.6
Resolution (Å)	47.82–1.84	45.87–1.95	46.90–1.95	43.63–1.54
	(1.95–1.84)	(2.06–1.95)	(2.07–1.95)	(1.63–1.54)
Total reflections	1276255	283603	481479	2486875
Unique reflections	168719	69284	130113	268793
Multiplicity	7.6 (5.1)	4.1 (4.1)	3.7 (3.7)	9.3 (9.1)
Completeness (%)	98.6 (91.7)	99.1 (97.3)	98.9 (96.0)	99.6 (98.3)
$\langle I/\sigma(I)\rangle$	16.62 (2.01)	10.06 (2.17)	13.01 (1.93)	15.41 (2.19)
R_{merge}†	0.098 (1.005)	0.131 (0.739)	0.087 (0.740)	0.096 (1.021)
Refinement				
Working/test reflections	167476/1243	68276/1008	129104/1009	267784/1009
R/R_{free}†	0.148/0.192	0.171/0.209	0.170/0.209	0.145/0.176
No. of atoms				
Protein	14702	7291	14527	14680
Ligand	75	20	76	56
NAD⁺	176	88	176	176
Water	1490	570	896	2156
Na⁺ ions	3	2	4	5
Acetate ions	1	—	—	2
PEG molecules	3	—	5	3
Ethylene glycol	—	5	—	—
Br⁻ ions	—	2	—	—
$\langle B\rangle$ factors (Å²)				
Protein	27.4	30.1	35.1	21.6
Ligand	21.4	31.8	27.4	16.5
NAD⁺	19.3	19.6	28.0	15.5
Na⁺ ions	25.0	36.9	27.1	16.4
Water	35.9	33.2	38.0	33.8
R.m.s.d. from ideality for bonds (Å)	0.012	0.017	0.016	0.015
Ramachandran statistics (%)				
Favoured	97.3	96.7	95.9	97.5
Outliers	2.7	3.3	4.1	2.5
PDB code	5m5k	5m65	5m66	5m67

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the average intensity of reflection hkl. † $R = \sum_{hkl}||F_{obs}| - |F_{calc}||/\sum_{hkl}|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively. R_{free} is calculated analogously for the test reflections, which were randomly selected and excluded from the refinement.

(Mueller-Dieckmann, 2006) using vapour diffusion in sitting drops at 292 K. Single crystals suitable for X-ray diffraction experiments were obtained in 0.2 M sodium acetate, 16% PEG 4000, 0.1 M Tris pH 8.5 in two weeks.

2.2.2. BeSAHase complexes with adenine, adenosine and adenine/2′-deoxyadenosine. BeSAHase solution obtained with the additional purification step (procedure P2) described above was concentrated to 12 mg ml⁻¹ and divided into three samples. Each sample was incubated for 24 h with an eightfold molar excess of Ade, Ado or 2′-dAdo, respectively, and crystallized by vapour diffusion in sitting drops at 292 K. Crystals of the Ade complex were obtained in condition No. 14 of the Morpheus crystal screen (Molecular Dimensions) consisting of 90 mM Halogens Mix (sodium fluoride, chloride and bromide), 40% ethylene glycol, 20% PEG 8000, 0.1 M imidazole, 0.1 M MES pH 6.5. Crystals of the Ado complex

were obtained in 0.3 M sodium acetate, 16% PEG 4000, 0.1 M Tris pH 9.0. The third sample crystallized as an Ade/2′-dAdo complex using 0.3 M sodium acetate, 14% PEG 4000, 0.1 M Tris pH 8.0.

2.3. Data collection, structure solution and refinement

Low-temperature X-ray diffraction data were collected on BESSY beamline 14.2 for the complexes with Ade and Ado, both to a resolution of 1.95 Å. Low-temperature X-ray diffraction data for the Ado/Cord and Ade/2′-dAdo complexes were collected on beamlines P14 and P13 at PETRA III, EMBL/DESY, Hamburg to resolutions of 1.84 and 1.54 Å, respectively. All crystals belonged to the orthorhombic system, space group $P2_12_12$. The diffraction data were processed and scaled with the *XDS* system (Kabsch, 2010). Data-collection statistics are presented in Table 1.

All structures were solved by molecular replacement with *Phaser* (McCoy *et al.*, 2007) using chain A of PDB entry 4lvc for BeSAHase (Manszewski *et al.*, 2015) as a model. Three of the four structures are roughly isomorphous and contain the complete BeSAHase tetramer in the asymmetric unit. The asymmetric unit of the Ade complex consists of a dimeric assembly and the complete tetramer is created by the crystallographic twofold axis along z. The *PHENIX* package (Adams *et al.*, 2010) was used for refinement of all of the structural models with maximum-likelihood targets and with TLS parameters assigned to rigid-body segments predicted by the *TLSMD* server (Painter & Merritt, 2006). *Coot* (Emsley & Cowtan, 2004) was used for manual model rebuilding between rounds of automatic refinement, as well as for building the ligand molecules and for validation of the solvent structure. The final structure-refinement statistics are shown in Table 1.

The Engh & Huber (2001) parameters were taken as stereochemical restraints for the protein chains and *elBOW* (Moriarty *et al.*, 2009), as implemented in *PHENIX* (Adams *et al.*, 2010), was used to create restraint libraries for all of the active-site ligands and cofactor molecules. All structures were standardized with *ACHESYM* (Kowiel *et al.*, 2014).

Atomic coordinates and structure factors have been deposited in the Protein Data Bank (PDB) with accession codes 5m5k (Ado/Cord complex), 5m65 (Ade), 5m66 (Ado)

and 5m67 (2′-dAdo/Ade). Raw diffraction images were deposited in the RepOD repository (ICM, University of Warsaw) with the following DOIs: https://doi.org/10.18150/repod.1236363 (Ado/Cord), https://doi.org/10.18150/repod.7716153 (Ade), https://doi.org/10.18150/repod.8491539 (Ado) and https://doi.org/10.18150/repod.3824734 (2′-dAdo/Ade).

2.4. SAXS studies

Small-angle X-ray scattering (SAXS) patterns were collected on beamline P12 at the PETRA III storage ring at EMBL/DESY, Hamburg. 20 µl of the protein sample at a concentration of 1.0, 2.0 and 4.0 mg ml^{-1} and 20 µl of the corresponding matching buffer were loaded into a 96-well plate. Automated loading of the SAXS samples into the cuvette was achieved using a Hamilton syringe robot (Hura *et al.*, 2009). SAXS data were collected over an *s* range of 0.0088–5.0 nm^{-1} and overlays of the merged data sets were used to detect concentration-dependent scattering in the lowest *s* region. The *ATSAS* package (Konarev *et al.*, 2006) was used to process all SAXS data. Integration, scaling and buffer subtraction were accomplished using *PRIMUS* (Konarev *et al.*, 2003). The resulting scattering curves were used for all calculations and reconstructions. The structural parameters of the protein were calculated using the *PRIMUS* package. The radius of gyration (R_{g}) was computed using the Guinier approximation and the distance distribution function $p(r)$ was calculated using the indirect Fourier transformation method as implemented in *GNOM* (Svergun, 1992). *CRYSOL* (Svergun *et al.*, 1995) was applied for evaluation of the solution scattering patterns using the crystallographic model of BeSAHase in complex with Ado. *Ab initio* modelling with fourfold symmetry was performed with *DAMMIN* (Svergun, 1999). The molecular weight of the analyzed protein was calculated from the extrapolated $I(0)$ values in comparison to a standard protein (bovine serum albumin) sample using the equation

$$\mathrm{MM_{p}} = \frac{I_{0p}}{c_{p}} \frac{\mathrm{MM_{st}}}{(I_{0st}/c_{st})}, \tag{1}$$

where $\mathrm{MM_{p}}$ and $\mathrm{MM_{st}}$ are the molecular weights of the studied and standard proteins, I_{0p} and I_{0st} are the scattering intensities at zero angle of the studied and standard protein solutions, respectively, and c_{p} and c_{st} are the concentrations of studied and standard protein, respectively. The SAXS data-collection statistics are presented in Table 2.

3. Results and discussion

3.1. Overall structure of BeSAHase

The overall structure of BeSAHase has been described in detail in our earlier paper (Manszewski *et al.*, 2015). Briefly,

Table 2
SAXS data-collection and scattering-derived parameters.

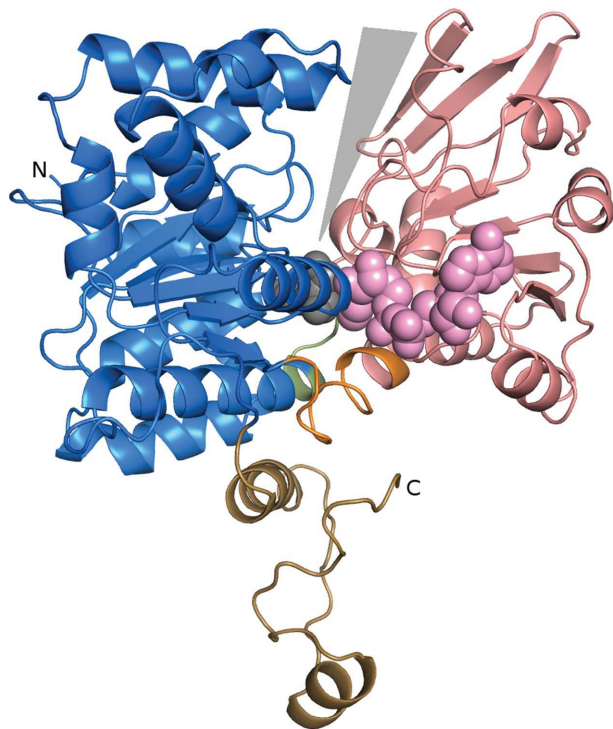| Data collection			
Instrument	P12, PETRA III	P12, PETRA III	P12, PETRA III
Wavelength (Å)	1.24	1.24	1.24
q range (nm^{-1})	0.0088–5	0.0088–5	0.0088–5
Exposure time (s)	1	1	1
Concentration (mg ml^{-1})	1	2	4
Temperature (K)	293	293	293
Structural parameters			
$I(0)$ (arbitrary units) [from $P(r)$]	54.82 ± 1	62.37 ± 1	68.92 ± 1
R_{g} (Å) [from $P(r)$]	39.7	39.2	38.2
$I(0)$ (arbitrary units) (from Guinier)	54.95 ± 1	62.45 ± 1	68.97 ± 1
R_{g} (Å) (from Guinier)	39.8	39.3	38.3
D_{max} (Å)	117	116	102.5
Porod volume estimate (Å^{3})	258549	258931	259165
Dry volume calculated from sequence (Å^{3})	251864	251864	251864
Molecular-mass determination			
Contrast (Δρ × 10^{10} cm^{-2})	3.047	3.047	3.047
Molecular mass M_{r} [from $I(0)$] (kDa)	220	220	220
Calculated monomeric M_{r} from sequence (kDa)	52.05	52.05	52.05
Software used			
Primary data reduction	*PRIMUS*		
Data processing	*PRIMUS*		
Ab initio analysis	*DAMMIN*		
Validation and averaging	*DAMAVER*		
Computation of model intesities	*CRYSOL*		
Three-dimensional graphics representation	*PyMOL*		

Figure 1
Subunit of BeSAHase in the closed conformation complexed with 2′-deoxyadenosine. The domains are colour-coded as follows: blue, substrate-binding domain; salmon, cofactor-binding domain; sand, C-terminal dimerization domain. The interdomain hinge regions are highlighted in orange (Asn222–Leu233) and green (His392–Val396). The grey triangle highlights the active-site access channel.

each subunit contains the native sequence of 473 amino-acid residues. The recombinant protein has six extra artifactual (GIDPFT–) residues at the N-terminus introduced by the cloning vector, but they are not visible in the electron-density maps of any of the models. BeSAHase protomers are divided into three domains, as illustrated in Fig. 1: a substrate-binding domain composed of residues Met1–Val221 and Met397–Val426, a cofactor-binding domain composed of residues Tyr234–Gly391, and a small C-terminal dimerization domain composed of residues Leu427–Tyr473. The first two domains are joined by two hinge regions comprised of residues Asn222–Leu233 and His392–Val396. It is notable that both hinge elements (demarcated according to main-chain torsions upon open–closed transition) contain α-helical turns of both conjoined subunits.

As described in §2.2, two types of protein preparations were used for the cocrystallization experiments. In one case (P2) the protein was precipitated with ammonium sulfate to remove any bound ligands prior to size-exclusion chromatography, and was then incubated with the desired cofactor (NAD$^+$) and target ligands (adenosine or adenine or 2′-deoxyadenosine). In the other scenario (P1) no ligand removal (ammonium

sulfate treatment) was used and the protein for the crystallization experiment was directly incubated with cordycepin.

All of the ligand and cofactor molecules in the presented complexes were modelled at full occupancy.

3.1.1. Adenosine/cordycepin complex. The structure of BeSAHase in the crystal obtained by cocrystallization with cordycepin of a protein sample that had not been treated with ammonium sulfate (see §2.2.1) is a tetramer with three subunits (A, B and C) complexed with Ado molecules and one subunit (D) with Cord bound in the active site. The first five residues (MNAKP) of the genuine BeSAHase sequence of chains A, B anad D and two residues of chain C could not be modelled in electron density because of disorder. One NAD$^+$ molecule is clearly defined by electron density in each subunit. There are three sodium cations near the active site of subunits A, B and C. 1490 water molecules plus three PEG molecules and one acetate ion from the crystallization buffer were located in the solvent region.

3.1.2. Adenine complex. The incubation of ligand-free BeSAHase (§2.2.2) with Ade resulted in crystals with a dimer in the asymmetric unit, from which the complete tetramer is created by crystallographic symmetry. The first four residues
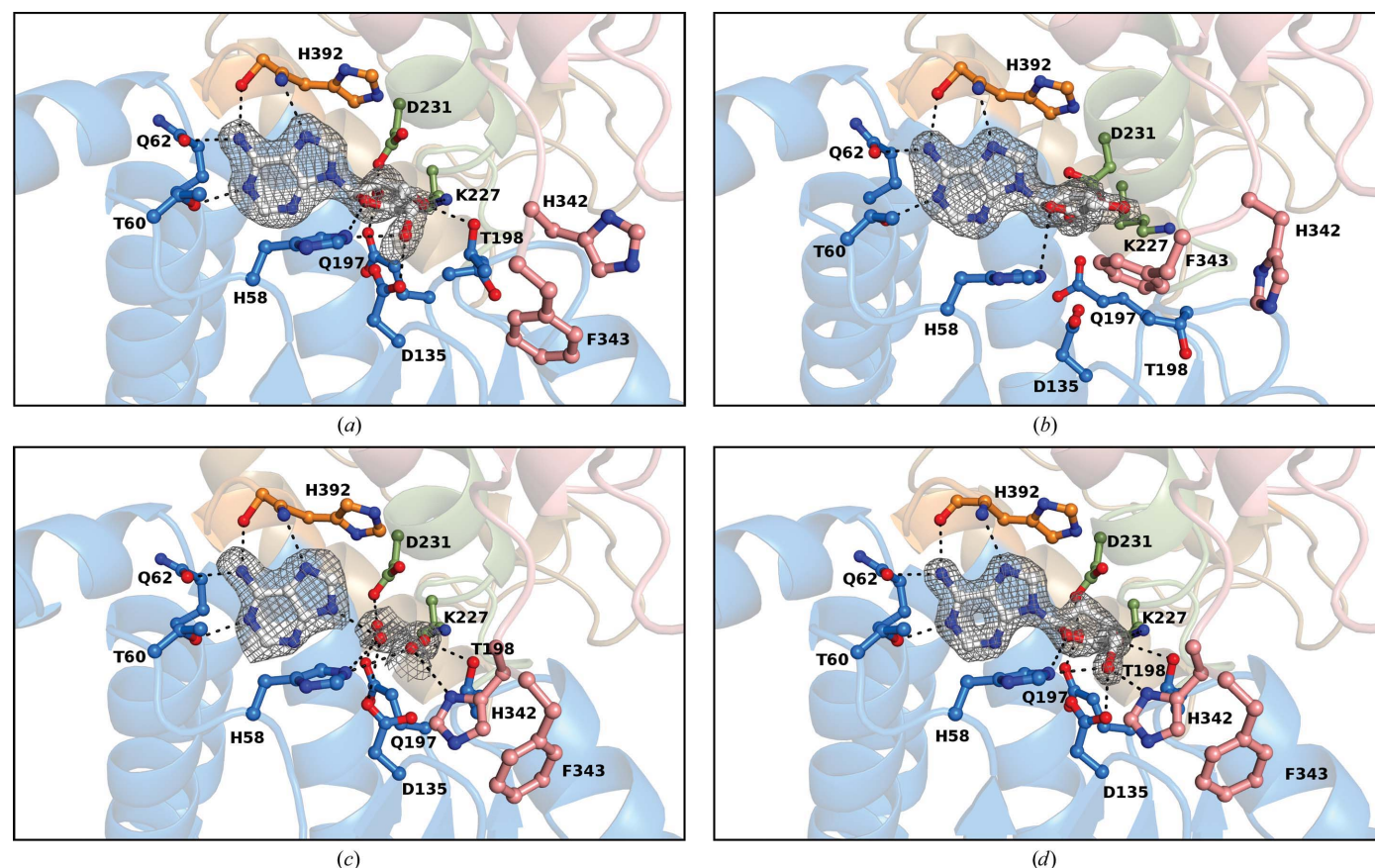
Figure 2
Comparison of BeSAHase active sites: (a) the Ado-complexed subunit A from the Ado/Cord complex, (b) the Cord-complexed subunit D from the Ado/Cord complex, (c) the Ade-complexed subunit A from the Ade complex and (d) the 2′-dAdo-complexed subunit C from the Ade/2′-dAdo complex. Since the mode of binding of all Ado molecules in all respective complexes is identical, only the ligand from the Ado/Cord structure is shown. Similarly, as the mode of binding of all Ade ligands is identical, only Ade from the Ade complex is shown. The ligands and water molecules are shown in $F_o - F_c$ OMIT electron-density (calculated without the contribution of the nucleosides, Ade and water atoms to F_c). The maps are contoured at 4σ, 3σ, 3σ and 4σ, respectively.

Table 3
Polar interactions between BeSAHase atoms and molecules in the active site.

Since the interactions of the same ligand molecules bound in the same complex are very similar, only values for the ligands bound in subunit *A* (Ado and Ade) or subunit *C* (2′-dAdo) are listed. Distances are in Å.

	Atom in											
	Ado/Cord complex				Ade complex		Ado complex		Ade/2′-dAdo complex			
	Ado		Cord						Ade		2′-dAdo	
Residue and atom	Atom	Distance	Atom	Distance	Atom	Distance	Atom	Distance	Atom	Distance	Atom	Distance
His58 $N^{\varepsilon 2}$	O4′	3.23	O4′	3.44	Wat3	3.46	O4′	2.95	Wat3	3.38	O4′	3.07
His58 $N^{\varepsilon 2}$	O5′	2.81	—	—	Wat4	3.38	O5′	2.66	Wat4	2.90	O5′	2.79
Thr60 $O^{\gamma 1}$	N1	2.67	N1	2.76	N1	2.77	N1	2.65	N1	2.72	N1	2.72
Gln62 $O^{\varepsilon 1}$	N6	2.91	N6	2.92	N6	2.87	N6	2.84	N6	2.86	N6	2.86
Asp135 $O^{\delta 1}$	O5′	2.93	—	—	Wat4	3.21	O5′	2.66	Wat4	2.77	O5′	2.73
Glu197 $O^{\varepsilon 2}$	O2′	2.51	—	—	Wat1	2.77	O2′	3.12	Wat1	2.37	Wat1	2.33
Thr198 $O^{\gamma 1}$	O3′	3.14	—	—	Wat2	2.90	O3′	2.91	Wat2	3.05	O3′	3.00
Lys227 N^{ζ}	O3′	2.97	—	—	Wat2	2.72	O3′	3.03	Wat2	2.96	O3′	2.94
Asp231 $O^{\delta 2}$	O2′	2.55	O2′	2.47	Wat1	2.62	O2′	2.52	Wat1	2.61	Wat1	2.48
His342 $N^{\delta 1}$	—	—	—	—	Wat4	3.39	—	—	Wat4	2.65	O5′	2.57
His392 N	N7	3.03	N7	2.89	N7	2.77	N7	2.86	N7	2.85	N7	2.88
His392 O	N6	3.01	N6	3.06	N6	2.93	N6	3.03	N6	3.03	N6	3.04

Table 4
Conformation (°) of the nucleoside molecules.

The torsion angles defining the nucloside conformation are as follows: χ, C4—N9—C1′—O4′; γ, O5′—C5′—C4′—C3′; ν_0, C4′—O4′—C1′—C2′; ν_1, O4′—C1′—C2′—C3′ *etc.* Since the conformation of the same ligand molecules bound in the same complex is very similar, only values for the ligands bound in subunit *A* (Ado and Ade) or subunit *C* (2′-dAdo) are listed. The amplitude (τ_m) and the phase angle (*P*) of pseudorotation (with estimated standard deviations in parentheses, given in units of the last significant digit) were calculated by the method of Jaskólski (1984).

	Ado/Cord complex			
	Ado	Cord	Ado	2′-dAdo
Glycosidic bond χ	−106.95 (*anti*)	−112.78 (*anti*)	−103.30 (*anti*)	−98.71 (*anti*)
γ	−163.2	−58.1	−161.6	−165.3
ν_0	−34.4	−25.2	−41.3	−40.6
ν_1	13.8	5.0	25.5	35.9
ν_2	10.1	15.9	−1.3	−18.3
ν_3	−30.3	−32.3	−22.9	−5.2
ν_4	41.0	35.9	41.0	30.4
P	75.4 (8)	62.7 (5)	163.9 (8)	116.8 (10)
τ_m	40.4 (6)	35.0 (3)	42.7 (6)	41.1 (7)
Sugar pucker	O4′-*endo*-C4′-*exo* ($^{O}_{4}T$)	C4′-*exo* ($_{4}T^{O}$)	C2′-*endo* (^{2}E)	C1′-*exo* ($_{1}T^{O}$)

of chain *A* and the first five residues of chain *B* are not visible in electron density. In each subunit, one Ade molecule is found in the active site, accompanied by a sodium cation. Also, one NAD$^+$ cofactor molecule is found in each protomer. Moreover, 570 water molecules, plus two bromide anions and five ethylene glycol molecules from the crystallization buffer, were modelled in the electron density of the solvent region.

3.1.3. Adenosine complex. Ligand-free BeSAHase incubated with Ado yielded crystals with a tetrameric assembly in the asymmetric unit. In each chain, the first five residues are not present in electron density. Moreover, residues Asn415–Ser418 of chain *D* and Asn415–Lys419 of chain *C* could not be modelled because of poor electron density. All four subunits have an Ado molecule bound in the active site. Also, four sodium cations are found in the metal-binding loop nearby.

High-quality electron density defines one NAD$^+$ molecule in each subunit. In addition, 896 water molecules and five PEG molecules from the crystallization buffer were found in the structure.

3.1.4. Adenine/2′-deoxyadeosine complex. Ligand-free BeSAHase incubated with 2′-deoxyadenosine gave the best diffracting crystals (1.54 Å resolution) and the structure revealed that the tetramer is composed of two subunits (*C* and *D*) complexed with the nucleoside used for incubation (2′-dAdo) and of two subunits (*A* and *B*) with adenine (Ade), which is a product of 2′-dAdo hydrolysis (Abeles *et al.*, 1980, 1982), in the active site. The first five genuine residues (MNAKP) of chains *A*, *B* and *C* and the first three residues of chain *D* could not be modelled because of disorder. Additionally, residues Asn415–Lys419 in chain *B* are not visible in electron density. As in the above structures, the NAD$^+$ cofactor molecule was found in each protomer. 2156 water molecules, five sodium cations, three PEG molecules and two acetate ions were modelled in the structure.

3.2. The ligand nucleosides and adenine molecules

BeSAHase was incubated and crystallized with four different ligands: cordycepin (3′-deoxyadenosine), adenine, adenosine and 2′-deoxyadenosine. The latter three incubations were preceded by precipitation with ammonium sulfate to remove all ligands bound in the overexpression step. In each structure, the ligand molecules used for incubation were unambiguously identified (at least in some subunits) in difference electron density (Fig. 2). However, in two cases the desired ligands were found only in some subunits, whereas the remaining subunits were occupied either by a competing ligand from the expression system (Ado) or by a product of ligand hydrolysis (Ade). The polar interactions between the protein and the ligand molecules are listed in Table 3. In all structures the adenine moiety is bound in the same mode. All N atoms except N3 form hydrogen bonds with both main-chain and side-chain atoms of BeSAHase. In the case of the

sugar moieties, however, there are differences in the hydrogen-bond networks resulting from the different ribose O-atom patterns as well as from the protein conformation. The conformational details of all of the nucleoside ligands, including the pseudorotation parameters of the sugar rings, are listed in Table 4.

3.2.1. Adenosine. Ado molecules were found in two of the four structures presented here. In the Ado/Cord complex, which was obtained using protein untreated with ammonium sulfate and incubated with cordycepin, it is serendipitously found (Fig. 2a) in three of the four subunits. This indicates that the adenosine molecules were sequestered by the protein in the overexpression step and retained during purification. Since the protein:cordycepin molar ratio was 1:8 and the original Ado ligand was apparently replaced in only 25% of the cases, this may suggest that the affinity of BeSAHase for cordycepin is over an order of magnitude lower than that for adenosine, leading to the conclusion that cordycepin would not be a good candidate as an inhibitor of SAHase. However, inferences about complex-formation equilibria in solution from crystallization effects should be regarded with caution.

The intended Ado complex was obtained from ligand-free protein incubated with an eightfold molar excess of the nucleoside. In this case, all four subunits have the active site occupied by Ado.

The interactions between the Ado molecules and the protein in these two complexes are the same and are similar to those described previously (Manszewski *et al.*, 2015). There is, however, one important difference: in the present complexes the side chain of His342, which is an element of the molecular gate (see §3.3.1), is shifted away from the active site and the hydrogen bond between the His342 $N^{\delta 1}$ atom and the O5′ atom of Ado is not formed. The sugar rings have an O4′-*endo*-C4′-*exo* pucker in Ado from the Ado/Cord complex and a C2′-*endo* pucker in Ado molecules from the Ado complex.

3.2.2. Cordycepin. A cordycepin molecule was found in subunit *D* of the Ado/Cord complex (Fig. 2b). Owing to the conformational changes of this subunit when compared with subunits *A*, *B* and *C*, and to the absence of the ribose O3′ atom, the hydrogen-bond network around the cordycepin ligand is significantly different. Specifically, the hydrogen bonds between O3′ and the Thr198 $O^{\gamma 1}$ and Lys227 N^{ζ} atoms that are observed in the Ado-occupied subunits are absent. The backbone conformation of subunit *D* is sufficiently different (see §3.3) as to place the side chain of Glu197 in a different direction, so that the hydrogen bond between the O2′ atom of cordycepin and Glu197 $O^{\varepsilon 2}$ cannot be formed (the distance between these atoms is 4.1 Å). Also, the hydrogen bonds created with the participation of the O5′ atom are not observed in chain *D* of the Ado/Cord complex , because the γ torsion angle (defined by the O5′—C5′—C4′—C3′ atoms of the ribose moiety; Table 4) of the cordycepin ligand is twisted by ∼100° with respect to that of the Ado molecules. The Cord sugar ring puckering is C4′-*exo*, which is in good agreement with the puckering of the ribose moiety of cordycepin bound to plant SAHase (LlSAHase from *L. luteus*; Brzezinski *et al.*, 2012; PDB entry 3onf).

3.2.3. Adenine. Ade molecules were found in two structures. In the complex formed by the incubation of BeSAHase with Ade, all subunits have this ligand bound in the active site (Fig. 2c). In the Ade/2′-dAdo complex, two (*A* and *B*) of the four subunits are complexed with Ade, which is a product of 2′-deoxyadenosine hydrolysis (Abeles *et al.*, 1980, 1982). In both structures, despite the absence of the sugar moiety, the architecture of the active site is preserved, as observed previously by Brzezinski *et al.* (2012). In a pattern of molecular mimicry, four water molecules imitate the positions of the O2′, O3′, O4′ and O5′ atoms of an Ado molecule and create hydrogen bonds to the corresponding side-chain atoms of the protein. Since there are seven different Ado molecules among the structures reported here that can be used as a reference for the ribose-mimicking water molecules in the Ade complexes, the deviations in atom positions of the water molecules in each Ade complex have been averaged over the Ado templates. In each case the reported values are given for O2′, O3′, O4′ and O5′, respectively. For the water molecules in subunits *A* and *B* of the Ade complex these differences are 0.23 and 0.32, 0.26 and 0.20, 0.77 and 1.01, and 1.08 and 1.11 Å, respectively. The second model that contains Ade molecules is the complex obtained by cocrystallization with 2′-dAdo. The Ade molecules are bound in subunits *A* and *B* and the corresponding water molecules are superposable with Ado ribose O atoms within 0.65 and 0.69, 0.39 and 0.29, 0.45 and 0.34, and 0.21 and 0.39 Å, respectively. It is of note that one of these water molecules always replaces the O2′ atom, even though no such atom is present in the 2′-dAdo ligand of the Ade/2′-dAdo complex. The hydrogen bonds created by these water molecules are the same in each case, although their distances vary within ∼0.4–0.6 Å. This is because the positions of the corresponding water molecules differ, on average, by 0.55, 0.33, 0.59 and 0.87 Å for the molecules that mimic the positions of O2′, O3′, O4′ and O5′, respectively.

In contrast to the protein–nucleoside interactions of the abovementioned complexes, in the Ade complexes there is an additional hydrogen bond between His342 $N^{\delta 1}$ and the water molecule that mimics the O5′ position. A similar His342 $N^{\delta 1}\cdots$O5′ interaction was found in the Ado complex of BeSAHase reported previously (Manszewski *et al.*, 2015; PDB entry 4lvc). In the Ade/2′-dAdo complex there is also an additional water molecule in the Ade-complexed subunits that is superposable within 0.41 Å with the C1′ atom of the 2′-dAdo molecules and does not interact with any protein atoms.

3.2.4. 2′-Deoxyadeosine. In the Ade/2′-dAdo complex, subunits *C* and *D* have 2′-dAdo in the active site (Fig. 2d). A comparison of these subunits with those complexed with Ade (chains *A* and *B*) reveals that the geometry of the hydrogen-bond networks between the ligand and protein atoms is nearly the same. Superposition of a 2′-dAdo molecule on Ade shows that all of the O atoms of 2′-dAdo overlap with water molecules within 0.19 Å (O3′), 0.47 Å (O4′) and 0.26 Å (O5′). Moreover, another water molecule is present in the 2′-dAdo-complexed subunits, which mimics the position of the O2′ atom of an Ado molecule and creates two hydrogen bonds,

with Glu197 $O^{\varepsilon 2}$ and Asp231 $O^{\delta 2}$, that are characteristic of the Ado complexes.

3.3. Ligand type *versus* protein conformation

A simplified view holds that the conformation of SAHase subunits switches from open to closed upon ligand binding (Hu *et al.*, 1999), although it was shown by Zheng *et al.* (2015) that the closed conformation can also be maintained without ligand molecules in the active site. In our previous BeSAHase structure (Manszewski *et al.*, 2015), the protein was found in two different conformational states within one tetramer: the three Ado-complexed subunits were in the closed conformation, while one ligand-free subunit was open.

Analysis of subunit conformation in the present structures, carried out by means of r.m.s.d. calculations for superposed C^{α} atoms (Table 5), reveals that, except for the subunit complexed with cordycepin, all other BeSAHase subunits are in the closed conformation. The r.m.s.d. values between these subunits and a closed subunit from the earlier BeSAHase

Table 5
R.m.s. deviations (Å) for C^{α} atoms of superposed subunits of BeSAHase complexed with different ligands.

The closed and open subunits from the BeSAHase model with PDB code 4lvc (Manszewski *et al.*, 2015) are included for comparison. Since the values for the same ligand molecules bound in the same complex are very similar, only data for the ligands bound in subunit A (Ado and Ade) or subunit C (2′-dAdo) are listed. Values were calculated with *ALIGN* (Cohen, 1997).

Model†	Ado/Cord	Ado/**Cord**	Ade	**Ado**	Ade/**2′-dAdo**	**Ade**/2′-dAdo	4lvc closed
Ado/**Cord**	1.68						
Ade	0.30	1.54					
Ado	0.26	1.66	0.30				
Ade/**2′-dAdo**	0.20	1.69	0.32	0.22			
Ade/2′-dAdo	0.19	1.75	0.35	0.21	0.14		
4lvc closed	0.18	1.71	0.31	0.28	0.21	0.16	
4lvc open	2.32	1.29	2.17	2.27	2.33	2.38	2.32

model with PDB code 4lvc (chain A; Manszewski *et al.*, 2015) range from 0.16 Å for an Ade-complexed subunit of the Ade/2′-dAdo complex to 0.31 Å for the Ade complex. The most interesting situation is found in subunit D of the Ado/Cord
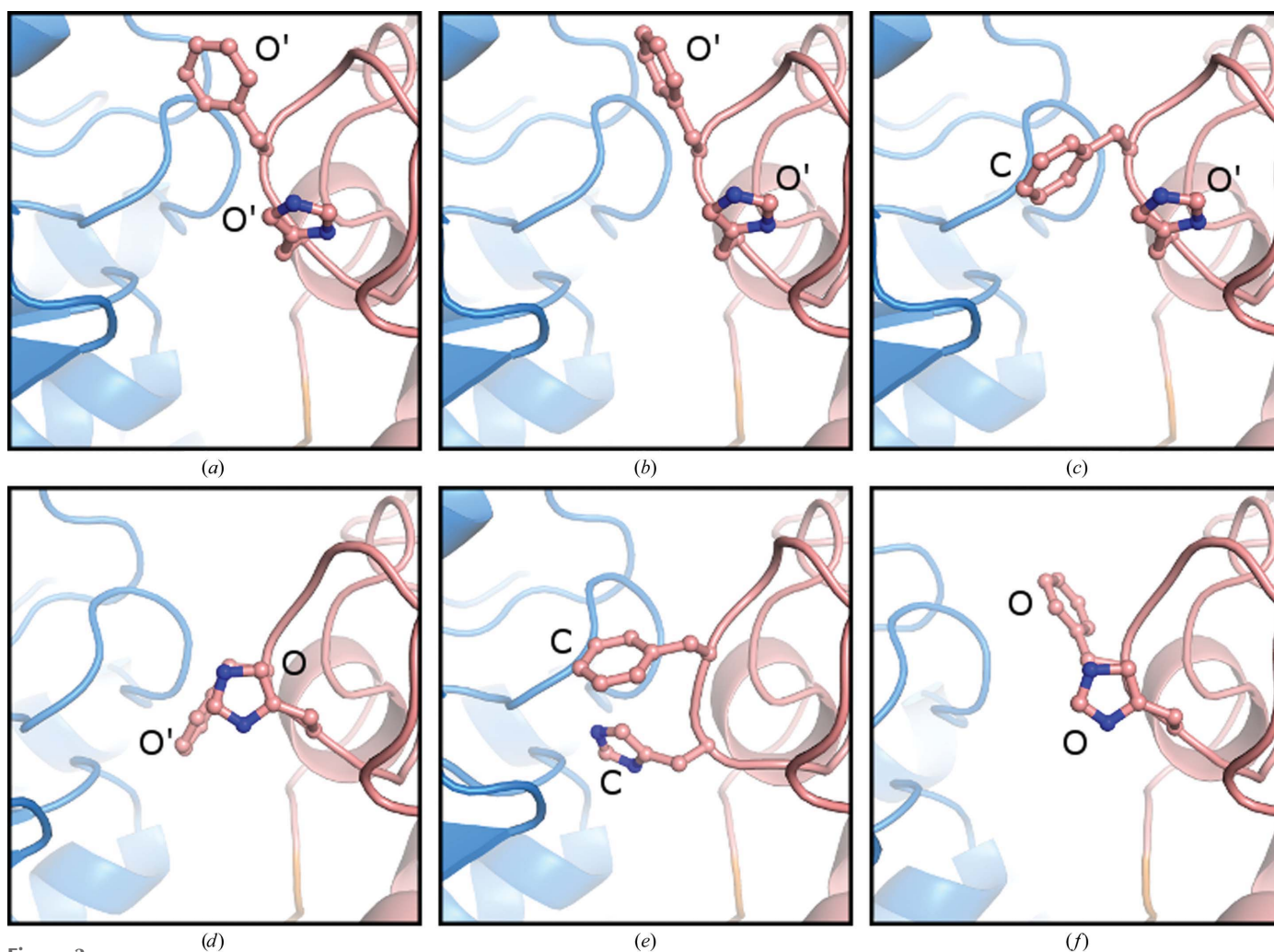

Figure 3
Comparison of the side-chain conformations of the MG residues in (*a*) the Ado-complexed subunit of the Ado/Cord complex, (*b*) subunit *A* of the Ado complex, (*c*) subunit *B* of the Ado complex, (*d*) the Cord-complexed subunit of the Ado/Cord complex, (*e*) subunit *A* of the Ade complex and (*f*) the open subunit (*D*) of PDB entry 4lvc. The side chains of the molecular-gate residues His342 and Phe343 are shown in ball-and-stick representation and annotated as C (closed state), O (open state) or O′ (variant open state; see text).

complex, where cordycepin is bound. The r.m.s.d. values between this subunit and the closed/open subunits of PDB entry 4lvc are 1.71 and 1.29 Å, respectively. This shows that BeSAHase in complex with cordycepin adopts an semi-open conformation that is halfway between the closed and open forms. This is in contrast to the cordycepin complex of LlSAHase (Brzezinski *et al.*, 2012), which is much closer to the
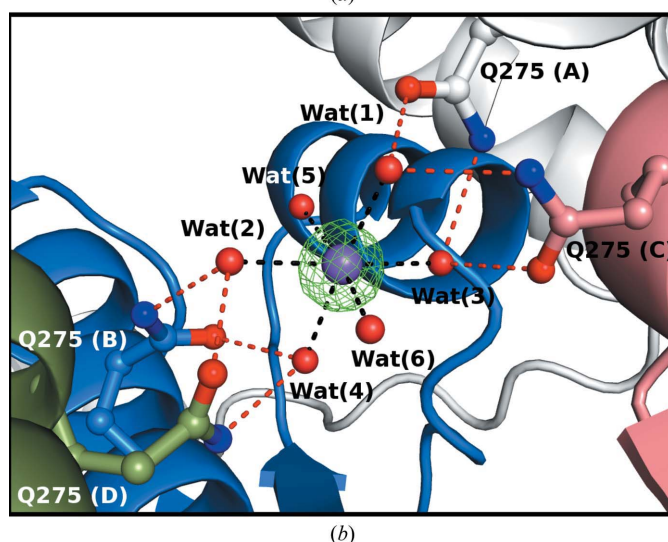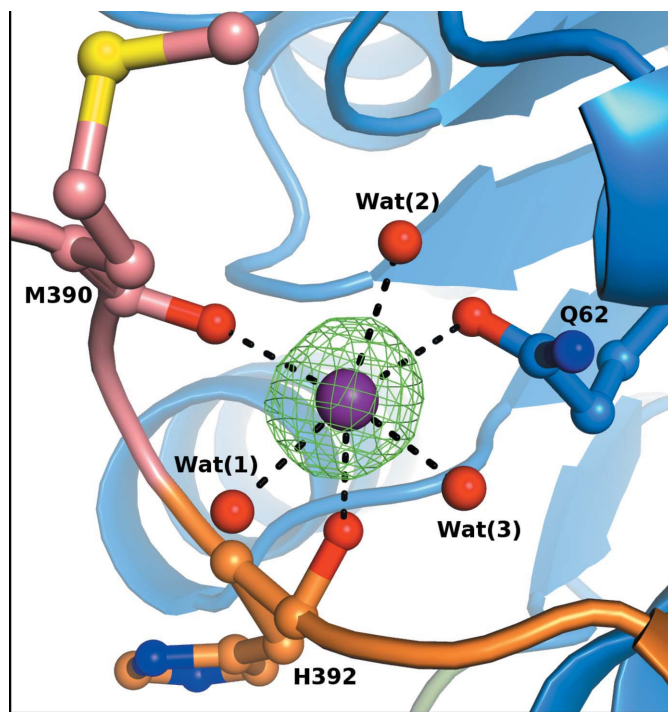
Table 6
Details of sodium-ion coordination (Å) near the active site.

In each structure, the interactions are nearly identical in each subunit; therefore, only values for the ions bound in subunit A are listed.

Ligand atom	Ado/Cord complex	Ade complex	Ado complex	Ade/2′-dAdo complex
Gln62 O$^{\varepsilon 1}$	3.16	3.31	2.92	2.83
Met390 O	2.71	2.63	2.46	2.48
His392 O	2.74	3.18	2.68	2.73
Wat1	2.46	2.38	2.30	2.43
Wat2	2.80	2.72	2.53	2.54
Wat3	2.60	2.81	2.69	2.53

standard closed form (0.44 Å) than to the Cord-complexed subunit of BeSAHase (1.66 Å).

3.3.1. The molecular gate. In the structure of *M. tuberculosis* SAHase (Reddy *et al.*, 2008), the side chain of His363 (which corresponds to His342 in BeSAHase) was recognized as a 'molecular gate' (MG) that opens or shuts an access channel leading to the active site. In the structure of LlSAHase (Brzezinski *et al.*, 2012) the MG was shown to actually consist of a tandem of His-Phe residues. Also, in the structure of BeSAHase with PDB code 4lvc (Manszewski *et al.*, 2015) the MG element was shown to consist of His342-Phe343 and appeared to be in the open/shut state in correlation with the open/closed conformation of the protein chain. This simple rule is not preserved in the present structures as the MG state varies even among subunits complexed with the same ligand (Fig. 3).

In the Ado-bound (closed) subunits *A*, *B* and *C* of the Ado/Cord complex (Fig. 3*a*), the side chains of the MG element are in an open conformation; that is, in a different state to in the Ado-bound closed subunits of the 4lvc structure (Manszewski *et al.*, 2015). The same situation is found in chains *A* and *D* of the Ado complex (Fig. 3*b*). In subunits *B* and *C* of the same complex (Fig. 3*c*), the side chain of His342 remains in the open state (and for this reason there is no hydrogen bond between the Ado O5′ atom and His342 N$^{\delta 1}$; see §3.2.1), but the side chain of Phe343 assumes a conformation that is compatible with the shut state of the MG. In the Cord-complexed subunit the side chain of His342 adopts exactly the same conformation as in the open subunit of BeSAHase 4lvc, but the side chain of Phe343, while remaining in the open conformation, is rotated by ∼180° around the C$^\alpha$—C$^\beta$ bond (Fig. 3*d*). The MG residues in all the Ade- and 2′-dAdo-complexed subunits (Fig. 3*e*) are in the shut conformation, which allows the formation of the Ado O5′···N$^{\delta 1}$ His342 hydrogen bond.

3.4. Sodium cations

In all structures presented here there are sodium cations identified in clear electron density (Fig. 4). The metal identification was confirmed by the *B* factors (similar to those of the coordinating ligand atoms) and by the coordination geometry (Tables 6 and 7), and was additionally validated using the *CheckMyMetal* (*CMM*) server (Zheng *et al.*, 2014). Except for the Cord-complexed subunit, which is metal-free, the Na$^+$ ions are present in the Ala389–Pro393 loop, which is located near

Figure 4
Coordination of sodium cations (purple spheres). (*a*) The Na$^+$ cation bound in subunit *A* of the Ade/2′-dAdo complex. The coordination bond (dashed lines) distances are listed in Table 6. (*b*) The Na$^+$ cation bound in the core of the Ade/2′-dAdo complex (the Na–Wat coordination-bond distances are listed in Table 7). Hydrogen bonds are represented by red dashed lines. In both panels, the $F_o − F_c$ OMIT electron density was calculated without the contribution of Na$^+$ to F_c and contoured at 5σ. The domains are colour-coded as in Fig. 1. Water molecules are represented as red spheres.

the active site. Additionally, a sodium ion was found in the core of the Ade/2′-dAdo complex near the interface of all four subunits.

3.4.1. Sodium cations near the active site. In the structure of BeSAHase with PDB code 4lvc (Manszewski *et al.*, 2015), an ammonium cation was modelled in the electron density in the loop Ala389–Pro393 that overlaps the His392–Val396 element of the interdomain hinge. The NH_4^+ ion was hydrogen-bonded by three water molecules and the carbonyl O atoms of Met390 and His392 (one of the NH donors was assumed to form a bifurcated hydrogen bond). The coordination sphere of the present sodium cations is comprised of the same five ligands as in the case of the NH_4^+ cation but is complemented into an octahedral pattern by the Gln62 $O^{\varepsilon 1}$ atom (Fig. 4*a*, Table 6). The absence of the sodium cation in the Cord-complexed subunit, which adopts an intermediate open/closed conformation, and the fact that BeSAHase in the open conformation also does not bind any ion (Manszewski *et al.*, 2015), suggest that cation binding in the Ala389–Pro393 loop rigidifies the interdomain hinge and stabilizes the protein in the closed conformation.

3.4.2. A sodium cation in the core of the adenine/2′-deoxyadeosine complex. Careful analysis of the electron-density maps for the highest resolution (1.54 Å) Ade/2′-dAdo complex structure revealed an additional sodium cation located right in the centre of the BeSAHase tetramer (Fig. 4*b*), where amino-acid residues from the cofactor-binding domains of all four subunits meet. The slightly distorted octahedral coordination sphere is created by six water molecules, four of which are hydrogen-bonded by the Gln275 $O^{\varepsilon 1}$ and $N^{\varepsilon 2}$ atoms. Since the interaction interfaces between the subunits of the BeSAHase tetramer, as calculated with *PDBsum* (Laskowski, 2009) and reported by Manszewski *et al.* (2015), are relatively large (∼2900 Å² between subunits forming a dimer and ∼1600 Å² between juxtaposed subunits), the binding of the core sodium cation appears to be artifactual as it does not seem to be important for tetramer stabilization.

3.5. Cofactor molecules

For its enzymatic activity, SAHases requires one molecule of nicotinamide adenine dinucleotide in its oxidized form (NAD^+) bound near the active site of each subunit. The mode of NAD^+ binding by BeSAHase has been described in detail in our previous work (Manszewski *et al.*, 2015). Briefly, residues of both the cofactor-binding and substrate-binding domains take part in the hydrogen-bond network formed between the protein and cofactor atoms. Moreover, the side chains of residues Lys467 and Tyr471 from the C-terminal domain from an adjacent subunit are involved in cofactor binding, high-lighting the role of the C-terminal domain in the dimerization process.

In the structures reported here, the cofactor was supplied to the protein using two different procedures (P1 and P2; see §2.1). Notwithstanding this difference, in each structure the cofactor was found in the oxidized (NAD^+) rather than the reduced (NADH) form. The main piece of evidence for the

Table 7
Details of sodium coordination in the core of the tetramer of the Ade/2′-dAdo complex.

Ligand	Bond length (Å)
Wat1	2.60
Wat2	2.74
Wat3	2.59
Wat4	2.60
Wat5	2.48
Wat6	2.44

cofactor oxidation state is the electron density of the nicotine moiety clearly indicating a planar ring. Meijers *et al.* (2001) showed that the addition of the hydride ion to the nicotine ring in the oxidized form leads to loss of aromatic character and geometrical deformation into a boat conformation, which are characteristic of the reduced (NADH) form. In the present study, the cofactor molecules were refined without planarity restraints to probe the geometry of the nicotine ring. Despite the less-than-atomic resolution of the diffraction data, the refinement in each case was stable and converged with a flat nicotine ring, as illustrated in Fig. 5 for the NAD^+ molecule from subunit *A* of the Ade/2′-dAdo complex.

3.6. SAXS studies

The global structure of BeSAHase in solution was reconstructed using small-angle X-ray scattering (SAXS) in order to confirm the oligomerization state in solution. SAXS data inform about the overall size and shape of macromolecules in solution and can be used to generate a global structural model at resolutions of about 12 Å or above.

The radius of gyration (R_g) of BeSAHase calculated from the Guinier approximation is largely independent of concentration: 38.3 Å at 4 mg ml^{-1} protein concentration, 39.3 Å at 2 mg ml^{-1} and 39.8 Å at 1 mg ml^{-1} (see Supplementary Fig. S1). The theoretical value of R_g, calculated using the coordi-
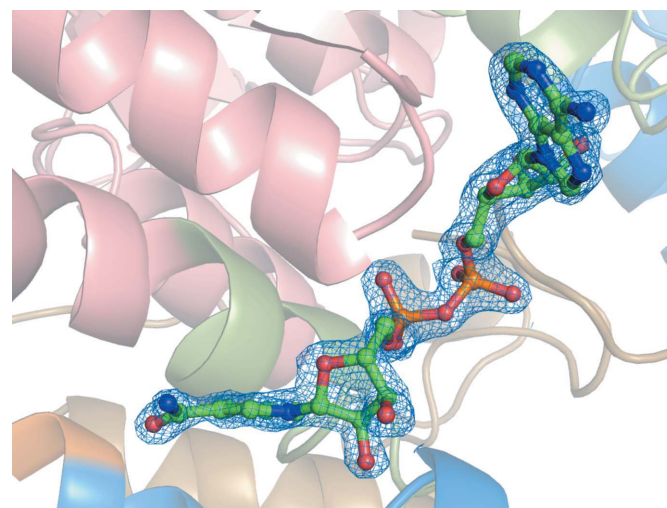

Figure 5
The NAD^+ cofactor molecule bound in subunit *A* of the Ade/2′-dAdo complex, shown in $2F_o - F_c$ electron density contoured at 1σ. The flat electron density of the nicotine ring confirms the oxidized state of the cofactor.

nates of the BeSAHase tetramer from the crystal structure of the Ado complex, is 36 Å, while the theoretical values calculated for the monomer and dimer are 22.5 and 30 Å, respectively, confirming that BeSAHase is tetrameric in solution. The presence of a tetramer in solution is confirmed by the χ^2 parameters calculated in *CRYSOL*, which are 753.2, 482.7 and 8.78 for the momoner, dimer and tetramer, respectively.

The experimental and theoretical SAXS curves are presented in Fig. 6(*a*). A comparison of the crystallographic structure with an *ab initio* model derived from the SAXS data, shown in Fig. 6(*b*), additionally corroborates the tetrameric quaternary state of the protein in solution.

4. Conclusions

Four crystal structures of *S*-adenosyl-L-homocysteine hydrolase from *B. elkanii* were solved in complex with different ligands, (i) mixed adenosine/cordycepin, (ii) adenine, (iii) adenosine and (iv) mixed adenine/2′-deoxyadenosine, to resolutions of 1.84, 1.95, 1.95 and 1.54 Å, respectively. The Ado molecules in structure (i) were sequestered by the protein during overexpression in *E. coli* cells. All other ligands [except for Ade, which in (iv) is a product of hydrolysis of 2′-dAdo by BeSAHase] were added to the protein solution after the purification step, which included a careful ligand-removal and cofactor-exchange procedure. The NAD$^+$ cofactor is found at full occupancy in the cofactor-binding domain of all of the subunits and also in case (i), where no cofactor exchange was applied. The various active-site ligands were all modelled at full occupancy in unambiguous electron-density maps and also in complexes (i) and (iv) where subsets of subunits of the tetrameric enzyme are charged with different cargo. The architecture of the active site is conserved, even when Ade, a ligand without a sugar ring moiety, is bound. This is because well defined water molecules mimic the pattern of the ribose O atoms in largely conserved hydrogen-bond interactions. The conformation of BeSAHase is affected by the ligand type in the active site and in all cases except for the cordycepin complex the subunit is closed. When cordycepin is bound in the active site, the protein molecule assumes a semi-open conformation that is intermediate between the canonical open and closed states. The side chains of His342 and Phe343 (both from the cofactor domain) form a molecular gate that controls passage through an active-site access channel formed between the substrate-binding and cofactor-binding domains. There is no apparent correlation between the ligand type and the open/shut state of the molecular gate, which can even vary with the same ligand molecule bound in the active site. There is also no unambiguous correlation between the open/closed conformation of the protein subunit and the open/shut state of the molecular gate. Moreover, the gate can also assume 'ambiguous' states, with one side chain in the open conformation and the other in the shut conformation. Certainly, further studies are needed to elucidate the actual mechanism of the molecular gate. A sodium cation is coordinated in a metal-binding loop near the active site of all subunits in the closed conformation, and this
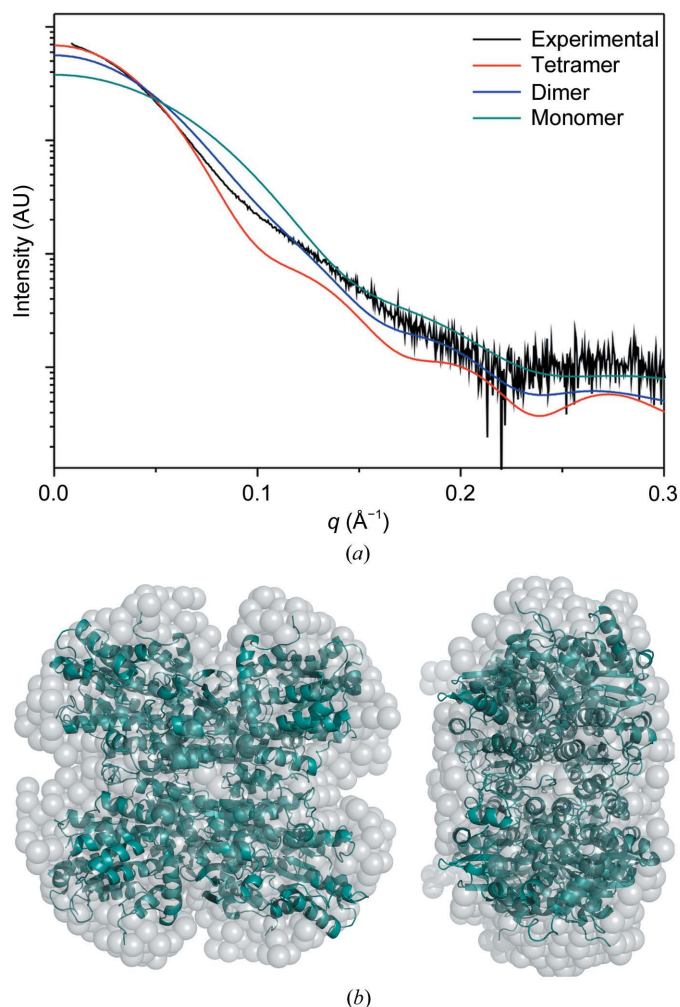

(*a*)

(*b*)

Figure 6
(*a*) Comparison of the experimental SAXS curve for BeSAHase (collected at a protein concentration of 4 mg ml^{-1}) with theoretical SAXS scattering curves calculated for the tetramer, dimer and monomer using the crystallographic coordinates from the present study. (*b*) Molecular envelope of BeSAHase obtained with *DAMMIN* (Svergun, 1999), with the crystallographic model of the Ado complex superposed. The two views are related by a 90° rotation around the vertical axis in the plane of the drawing.

cation binding seems to be the most constant structural feature that can be correlated with the conformational state of the SAHase subunit. The metal-binding loop overlaps a hinge region between the substrate and cofactor domains. Metal (or ammonium) coordination in this region helps to fix the molecular conformation in the closed state. The identity of the metal ion was confirmed by the pattern of Na–O distances (including *CMM* tests), by the electron density and by successful refinement. Small-angle X-ray scattering at three different protein concentrations confirmed the tetrameric state of BeSAHase in solution.

References

Abeles, R. H., Fish, S. & Lapinskas, B. (1982). *Biochemistry*, **21**, 5557–5562.

Abeles, R. H., Tashjian, A. H. & Fish, S. (1980). *Biochem. Biophys. Res. Commun.* **95**, 612–617.

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.

Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.

Brzezinski, K., Bujacz, G. & Jaskolski, M. (2008). *Acta Cryst.* F**64**, 671–673.

Brzezinski, K., Dauter, Z. & Jaskolski, M. (2012). *Acta Cryst.* D**68**, 218–231.

Cantoni, G. L. (1975). *Annu. Rev. Biochem.* **44**, 435–451.

Cantoni, G. L. & Chiang, P. K. (1980). *Natural Sulfur Compounds*, pp. 67–80. New York: Plenum.

Chiang, P. K. & Cantoni, G. L. (1979). *Biochem. Pharmacol.* **28**, 1897–1902.

Cohen, G. H. (1997). *J. Appl. Cryst.* **30**, 1160–1161.

De La Haba, G. & Cantoni, G. L. (1959). *J. Biol. Chem.* **234**, 603–608.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Engh, R. & Huber, R. (2001). *International Tables for Crystallography*, Vol. *F*, edited by M. G. Rossmann & E. Arnold, pp. 382–392. Dordrecht: Kluwer Academic Publishers.

Fujioka, M. & Takata, Y. (1981). *J. Biol. Chem.* **256**, 1631–1635.

Guranowski, A. & Pawełkiewicz, J. (1977). *Eur. J. Biochem.* **80**, 517–523.

Hu, Y., Komoto, J., Huang, Y., Gomi, T., Ogawa, H., Takata, Y., Fujioka, M. & Takusagawa, F. (1999). *Biochemistry*, **38**, 8323–8333.

Huang, Y., Komoto, J., Takata, Y., Powell, D. R., Gomi, T., Ogawa, H., Fujioka, M. & Takusagawa, F. (2002). *J. Biol. Chem.* **277**, 7477–7482.

Hura, G. L., Menon, A. L., Hammel, M., Rambo, R. P., Poole, F. L., Tsutakawa, S. E., Jenney, F. E. Jr, Classen, S., Frankel, K. A., Hopkins, R. C., Yang, S., Scott, J. W., Dillard, B. D., Adams, M. W. W. & Tainer, J. A. (2009). *Nat. Methods*, **6**, 606–612.

Jakubowski, H. (2006). *J. Nutr.* **136**, 1741S–1749S.

Jaskólski, M. (1984). *Acta Cryst.* A**40**, 364–366.

Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.

Komoto, J., Huang, Y., Gomi, T., Ogawa, H., Takata, Y., Fujioka, M. & Takusagawa, F. (2000). *J. Biol. Chem.* **275**, 32147–32156.

Konarev, P. V., Petoukhov, M. V., Volkov, V. V. & Svergun, D. I. (2006). *J. Appl. Cryst.* **39**, 277–286.

Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. (2003). *J. Appl. Cryst.* **36**, 1277–1282.

Kowiel, M., Jaskolski, M. & Dauter, Z. (2014). *Acta Cryst.* D**70**, 3290–3298.

Laskowski, R. A. (2009). *Nucleic Acids Res.* **37**, D355–D359.

Liu, S., Wolfe, M. S. & Borchardt, R. T. (1992). *Antiviral Res.* **19**, 247–265.

Manszewski, T., Singh, K., Imiolczyk, B. & Jaskolski, M. (2015). *Acta Cryst.* D**71**, 2422–2432.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Meijers, R., Morris, R. J., Adolph, H. W., Merli, A., Lamzin, V. S. & Cedergren-Zeppezauer, E. S. (2001). *J. Biol. Chem.* **276**, 9316–9321.

Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. (2009). *Acta Cryst.* D**65**, 1074–1080.

Mueller-Dieckmann, J. (2006). *Acta Cryst.* D**62**, 1446–1452.

Painter, J. & Merritt, E. A. (2006). *Acta Cryst.* D**62**, 439–450.

Palmer, J. L. & Abeles, R. H. (1976). *J. Biol. Chem.* **251**, 5817–5819.

Palmer, J. L. & Abeles, R. H. (1979). *J. Biol. Chem.* **254**, 1217–1226.

Reddy, M. C., Kuppan, G., Shetty, N. D., Owen, J. L., Ioerger, T. R. & Sacchettini, J. C. (2008). *Protein Sci.* **17**, 2134–2144.

Richards, H. H., Chiang, P. K. & Cantoni, G. L. (1978). *J. Biol. Chem.* **253**, 4476–4480.

Stępkowski, T., Brzeziński, K., Legocki, A. B., Jaskólski, M. & Béna, G. (2005). *Mol. Phylogenet. Evol.* **34**, 15–28.

Svergun, D. I. (1992). *J. Appl. Cryst.* **25**, 495–503.

Svergun, D. I. (1999). *Biophys. J.* **76**, 2879–2886.

Svergun, D., Barberato, C. & Koch, M. H. J. (1995). *J. Appl. Cryst.* **28**, 768–773.

Takata, Y., Yamada, T., Huang, Y., Komoto, J., Gomi, T., Ogawa, H., Fujioka, M. & Takusagawa, M. (2002). *J. Biol. Chem.* **277**, 22670–22676.

Tanaka, N., Nakanishi, M., Kusakabe, Y., Shiraiwa, K., Yabe, S., Ito, Y., Kitade, Y. & Nakamura, K. T. (2004). *J. Mol. Biol.* **343**, 1007–1017.

Turner, M. A., Yuan, C.-S., Borchardt, R. T., Hershfield, M. S., Smith, G. D. & Howell, P. L. (1998). *Nat. Struct. Biol.* **5**, 369–376.

Yang, X., Hu, Y., Yin, D. H., Turner, M. A., Wang, M., Borchardt, R. T., Howell, P. L., Kuczera, K. & Schowen, R. L. (2003). *Biochemistry*, **42**, 1900–1909.

Yin, D., Yang, X., Hu, Y., Kuczera, K., Schowen, R. L., Borchardt, R. T. & Squier, T. C. (2000). *Biochemistry*, **39**, 9811–9818.

Yuan, C.-S., Yeh, J., Liu, S. & Borchardt, R. T. (1993). *J. Biol. Chem.* **268**, 17030–17037.

Zheng, H., Chordia, M. D., Cooper, D. R., Chruszcz, M., Müller, P., Sheldrick, G. M. & Minor, W. (2014). *Nature Protoc.* **9**, 156–170.

Zheng, Y., Chen, C.-C., Ko, T.-P., Xiao, X., Yang, Y., Huang, C.-H., Qian, G., Shao, W. & Guo, R.-T. (2015). *J. Struct. Biol.* **190**, 135–142.
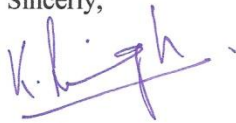
CO-AUTHORS'

STATEMENTS

Kriti Singh

University of British Columbia,

Vancouver, BC, Canada

sn.kriti@gmail.com

This letter is to certify my contribution to the following publication of which I was a co-author

Manszewski T, Singh K, Imiolczyk B, Jaskolski M. (2013) Structural enzymology at the legume-microbe interface: S-adenosyl-L-homocysteine hydrolase of rhizobia. *Biotechologia* **91**, 38-39.

I participated in writing the manuscript as well as in cloning, overexpression and purification of the S-adenosyl-L-homocysteine hydrolase protein. The optimization of crystallization conditions and crystallization of the purified protein was set up by me.

Sincerly,

Kriti Singh

Poznań, 20.02.2017

dr Barbara Imiołczyk

Zakład Krystalografii

Instytut Chemii Bioorganicznej PAN

ul. Z. Noskowskiego 12/14

61-704 Poznań

OŚWIADCZENIE

Oświadczam, że mój udział w publikacji

Manszewski T, Singh K, Imiolczyk B, Jaskolski M. (2013) Structural enzymology at the legume-microbe interface: S-adenosyl-L-homocysteine hydrolase of rhizobia. *Biotechologia* **91**, 38-39.

polegał na: klonowaniu, ekspresji, oczyszczaniu oraz krystalizacji SAHazy z Bradyrhizobium elkani w stopniu równym z Kriti Singh.

Barbara Imiołczyk

Barbara Imiołczyk

Poznań, 20.02.2017

prof. dr hab. Mariusz Jaskólski

Zakład Krystalografii

Instytut Chemii Bioorganicznej PAN

ul. Z. Noskowskiego 12/14

61-704 Poznań

OŚWIADCZENIE

Oświadczam, że mój udział w publikacji

Manszewski T, Singh K, Imiolczyk B, Jaskolski M. (2013) Structural enzymology at the legume-microbe interface: S-adenosyl-L-homocysteine hydrolase of rhizobia. *Biotechologia* **91**, 38-39.

polegał na koordynacji projektu oraz na współredagowaniu manuskryptu.
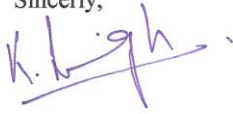
Prof. Mariusz Jaskólski

Kriti Singh
University of British Columbia,
Vancouver, BC, Canada
sn.kriti@gmail.com

This letter is to certify my contribution to the following publication of which I was a co-author

Manszewski T, Singh K, Imiolczyk B, Jaskolski M. (2015) An enzyme captured in two conformational states: Crystal structure of S-adenosyl-l-homocysteine hydrolase from Bradyrhizobium elkanii. *Acta Cryst.* **D71**, 2422-2432.

 I participated in cloning, overexpression and purification of the S-adenosyl-L-homocysteine hydrolase protein. I also set up the crystallization of the protein and optimized the conditions for crystallization. Moreover, I have contributed towards the writing of the manuscript.

Sincerly,

Kriti Singh

Poznań, 20.02.2017

dr Barbara Imiołczyk

Zakład Krystalografii

Instytut Chemii Bioorganicznej PAN

ul. Z. Noskowskiego 12/14

61-704 Poznań

OŚWIADCZENIE

Oświadczam, że mój udział w publikacji

Manszewski T, Singh K, Imiolczyk B, Jaskolski M. (2015) An enzyme captured in two conformational states: Crystal structure of S-adenosyl-l-homocysteine hydrolase from Bradyrhizobium elkanii. *Acta Cryst.* **D71**, 2422-2432.

polegał na: klonowaniu, ekspresji, oczyszczaniu oraz krystalizacji SAHazy z Bradyrhizobium elkani w stopniu równym z Kriti Singh.

Barbara Imiołczyk

Poznań, 20.02.2017

prof. dr hab. Mariusz Jaskólski

Zakład Krystalografii

Instytut Chemii Bioorganicznej PAN

ul. Z. Noskowskiego 12/14

61-704 Poznań

OŚWIADCZENIE

Oświadczam, że mój udział w publikacji

Manszewski T, Singh K, Imiolczyk B, Jaskolski M. (2015) An enzyme captured in two conformational states: Crystal structure of S-adenosyl-l-homocysteine hydrolase from Bradyrhizobium elkanii. *Acta Cryst.* **D71**, 2422-2432.

polegał na koordynacji projektu, analizie wyników i współredagowaniu manuskryptu.

Prof. Mariusz Jaskólski

Poznań, 20.02.2017

dr Kamil Szpotkowski

Zakład Krystalografii

Instytut Chemii Bioorganicznej PAN

ul. Z. Noskowskiego 12/14

61-704 Poznań

OŚWIADCZENIE

Oświadczam, że mój udział w publikacji

Manszewski T, Szpotkowski K, Jaskolski M. (2017) Crystallographic and SAXS studies on S-adenosyl-L-homocysteine hydrolase from Bradyrhizobium elkanii. *IUCrJ*, in press.

polegał na rejestracji danych SAXS oraz ich późniejszej analizie i opracowaniu wspólnie z doktorantem. Przygotowałem również fragment tekstu publikacji dotyczący badań SAXS.

Poznań, 20.02.2017

prof. dr hab. Mariusz Jaskólski

Zakład Krystalografii

Instytut Chemii Bioorganicznej PAN

ul. Z. Noskowskiego 12/14

61-704 Poznań

OŚWIADCZENIE

Oświadczam, że mój udział w publikacji

Manszewski T, Szpotkowski K, Jaskolski M. (2017) Crystallographic and SAXS studies on S-adenosyl-L-homocysteine hydrolase from Bradyrhizobium elkanii. *IUCrJ*, in press.

polegał na koordynacji projektu, analizie wyników i współredagowaniu manuskryptu.

Prof. Mariusz Jaskólski