A combined computational and experimental investigation of the [2Fe-2S] cluster in biotin synthase

Fuchs, Michael G. G.; Meyer, Franc; Ryde, Ulf

Published in:
Journal of Biological Inorganic Chemistry

DOI:
10.1007/s00775-009-0585-6

2010

Document Version:
Peer reviewed version (aka post-print)

A Combined Computational and Experimental Investigation of the [2Fe–2S] Cluster in Biotin Synthase

Michael G. G. Fuchs, Franc Meyer, Ulf Ryde

Michael G. G. Fuchs, Franc Meyer

Georg-August-Universität Göttingen, Institut für Anorganische Chemie, Tammannstrasse 4, D-37077 Göttingen, Germany, e-mail: franc.meyer@chemie.uni-goettingen.de

Ulf Ryde (✉)

Department of Theoretical Chemistry, Lund University, Chemical Centre, P. O. Box 124, SE-221 00 Lund, Sweden, e-mail: Ulf.Ryde@teokem.lu.se

2017-03-19

Abstract

Biotin synthase (BioB) was the first example of what is now regarded as a distinctive enzyme class within the radical SAM (*S*-adenosylmethionine) superfamily which use Fe/S clusters as the sulphur source in radical sulphur insertion reactions. The crystal structure showed that this enzyme contains a [2Fe–2S] cluster with a highly unusual arginine ligand, besides three normal cysteine ligands. However, the crystal structure is at such a low resolution that neither the exact coordination mode nor the role of this exceptional ligand has been elucidated yet, although it has been shown that it is not essential for enzyme activity. We have used quantum refinement of the crystal structure and combined quantum mechanical and molecular mechanical (QM/MM) calculations to explore possible coordination modes and their influences on cluster properties. The investigations show that the protonation state of the arginine ligand has little influence on cluster geometry so that even a positively charged guanidinium moiety would be in close proximity to the Fe atom. Nevertheless, the crystallised enzyme most probably contains a deprotonated (neutral) arginine coordinating via the NH group. Furthermore, the Fe···Fe distance seems to be independent of the coordination mode and is in perfect agreement with other structurally characterised [2Fe–2S] clusters. The exceptionally large Fe···Fe distance found in the crystal structure could not be reproduced.

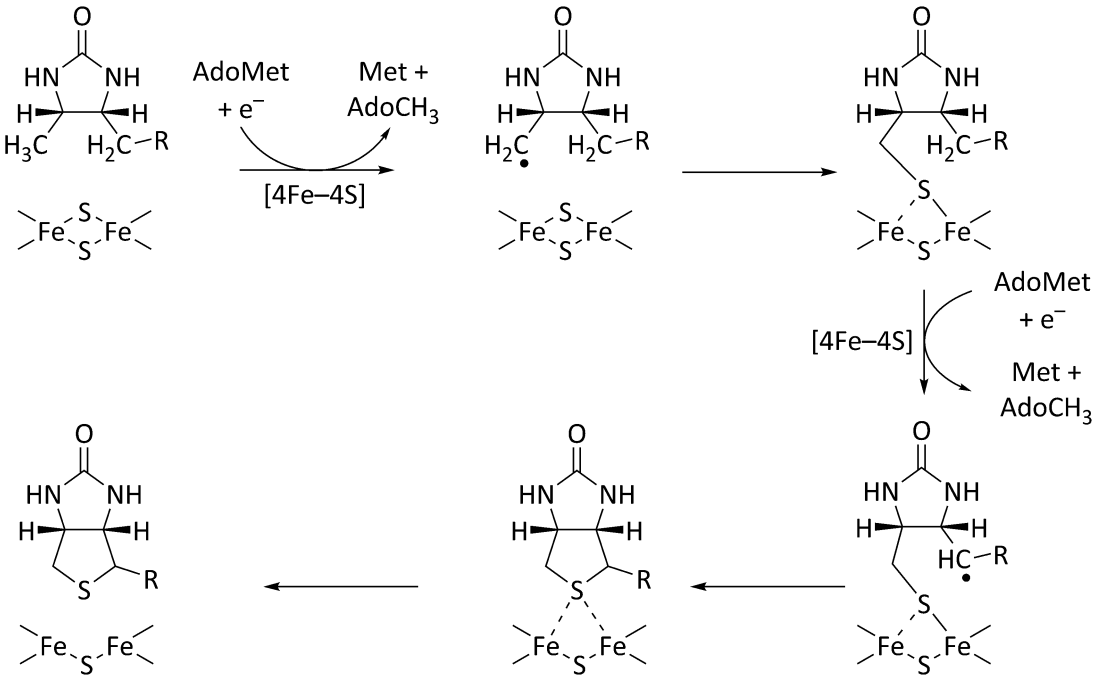
Keywords

Biotin Synthase, Fe/S Cluster, Radical SAM Enzyme, QM/MM, Quantum Refinement

Introduction

Biological Fe/S clusters are versatile cofactors in enzymes, well-known mainly for their ability to act as electron-transfer sites. In the past few years, an increasing number of other fascinating functions of Fe/S clusters have been discovered [1]. Among those, their use as the source for S atoms in biological radical reactions has been striking with regards to the complex mechanisms of Fe/S cluster assembly [2]. As the Fe/S cluster is destroyed during this reaction, these enzymes are regarded as suicide enzymes and their respective clusters as substrates rather than cofactors [3]. To date, Fe/S clusters have been identified as the source of S atoms in reactions catalysed by four different enzymes belonging to a distinctive class [4] within the radical SAM (*S*-adenosylmethionine) superfamily [5]: biotin synthase (BioB) [6], lipoyl synthase (LipA) [7], a tRNA-methylthiotransferase (MiaB) [8], and a ribosomal methylthiotransferase (RimO) [9].

While the latter three enzymes contain [4Fe–4S] clusters as the assumed sulphur source, biotin synthase (BioB) contains a [2Fe–2S] cluster which has been shown to be destroyed during catalytic turnover [10]. Of the four enzymes, BioB is the most extensively investigated and also the only one whose crystal structure has been solved [11]. A mechanism for its reaction was first published in 2001 [10] and has been closely investigated since (Scheme 1) [12].



Scheme 1. Proposed biotin synthase mechanism (AdoMet: *S*-adenosylmethionine) [10].

The proposed mechanism is supported by the close proximity of the [2Fe–2S] cluster to the dethiobiotin molecule found in the crystal structure. The closest bridging sulphide of the cluster is situated 4.6 Å away from C9 of dethiobiotin, one of the two C atoms to which sulphur is attached in the course of the reaction (Figure 1).

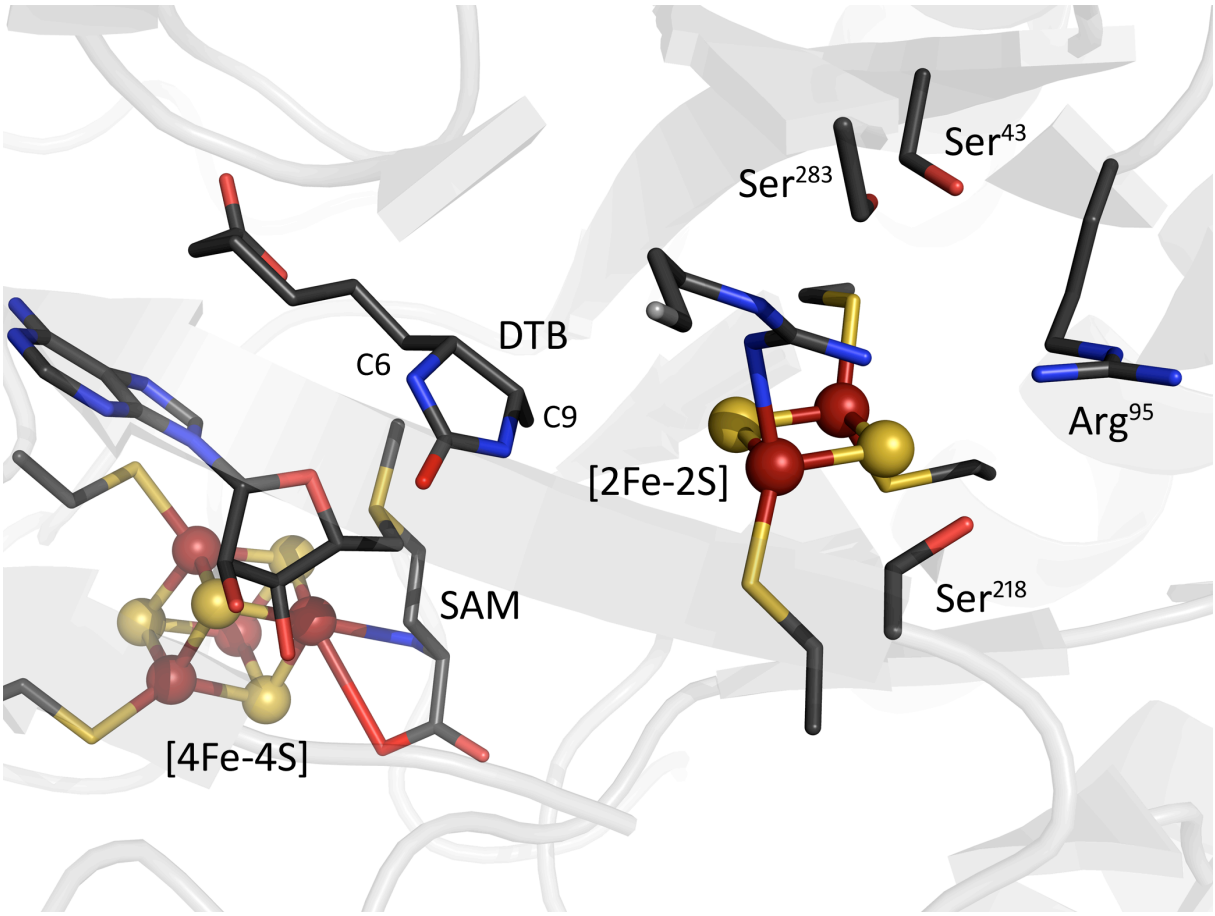


Figure 1. Detail of biotin synthase crystal structure (SAM: *S*-adenosylmethionine, DTB: dethiobiotin).

Surprisingly, the [2Fe–2S] cluster carries one strictly conserved arginine ligand in addition to three cysteine ligands typical for Fe/S clusters. As arginine is highly exceptional as a ligand coordinating a metal ion in biological systems [13], its possible importance has been much discussed since publication of the crystal structure. Mutation experiments showed that this arginine ligand is not essential for the catalytic reaction. It has been proposed that it may play electronic, mechanistic or structural roles, possibly related to its bidentate nature or its positive charge in the protonated state [14].

A basic question to be answered is the charge of the arginine guanidine group in the active enzyme. Arginine is usually protonated at physiological pH (the pK_a in water solution is

approximately 12 [15]), thereby bearing a positive charge. However, a positively charged guanidine group could not be regarded as a true ligand from a coordination chemist's point of view – for that, it should rather be deprotonated and neutral in BioB. This uncertainty is reflected by suggestions of both NH [13,16] and NH_2 [17] coordination in the literature.

In addition to the protonation issue, the role of the second, non-coordinating NH/NH_2 group is to be ascertained. Secondary bonding interactions [18] are conceivable, as well as an involvement in hydrogen bonds with the protein backbone or with the cluster S atoms. Furthermore, the crystal structure shows an Fe···Fe distance of 3.3 Å that is significantly longer than what is found in any other known [2Fe–2S] clusters (approximately 2.7 Å) [19].

As these issues cannot be solved with the published structural data alone, owing to the low resolution of the crystal structure (3.4 Å), and the importance of the unusual arginine ligand remains elusive, theoretical methods seem to be a promising strategy to answer the above-mentioned questions. This work focuses on the structural properties of the highly unusual [2Fe–2S] cluster of BioB in order to improve our understanding of its importance for the enzyme mechanism.

Materials and Methods

QM/MM calculations

The QM/MM calculations were performed with the ComQum software [20,21] utilising Turbomole 5.9 [22] for the quantum mechanical (QM) calculations and Amber 9 [23] for the molecular mechanical (MM) calculations. The QM calculations were performed using the BP86 functional [24,25] and the def2-SV(P) basis sets [26], which have given reasonable results for [2Fe–2S] clusters in previous calculations [18,27]. For the MM calculations, we used the Amber-99 force field [28,29]. For the [4Fe–4S] cluster and the dethiobiotin and *S*-

adenosylmethionine ligands, we used force-field parameters previously determined in our group [30-32].

In the QM/MM approach, the protein and solvent are split into three subsystems: The QM region (system 1) contains the most interesting atoms and is relaxed by QM methods. System 2 consists of the residues closest to the QM system and is optimised by MM. The remaining part of the protein and the surrounding solvent molecules (system 3) are kept fixed at the crystallographic coordinates. In the QM calculations, system 1 is represented by a wavefunction, whereas all the other atoms are represented by an array of partial point charges, one for each atom, taken from MM libraries. Thereby, the polarisation of the quantum chemical system by the surroundings is included in a self-consistent manner. When there is a bond between systems 1 and 2 (a junction), the quantum region is truncated by hydrogen atoms, the positions of which are linearly related to the corresponding carbon atoms in the full system (the hydrogen link-atom approach) [20]. In order to eliminate the non-physical effect of placing point charges on atoms in the MM region bound to junction atoms (i.e. the closest neighbours of QM system), those charges are zeroed, and the resulting residual charges are smoothly distributed [20].

The total energy is calculated as:

$$E_{\mathrm{QM/MM}} = E_{\mathrm{QM1+ptch}} - E_{\mathrm{MM1}} + E_{\mathrm{MM123}} \qquad (1)$$

where $E_{\mathrm{QM1+ptch}}$ is the QM energy of system 1 truncated by the hydrogen atoms and embedded in the set of point charges (but excluding the self-energy of the point charges). E_{MM1} is the MM energy of system 1, still truncated by hydrogen atoms, but without any electrostatic interactions. Finally, E_{MM123} is the classical energy of all atoms with normal atoms at the junctions and with the charges of the quantum system set to zero (to avoid double-counting of

the electrostatic interactions). By this approach, which is similar to the one used in the Oniom method [33], errors caused by the truncation of the quantum system should cancel.

The calculations were based on the crystal structure (PDB code 1R30) [11]. As the enzyme was crystallised as a homodimer with little difference in atom positions (less than 0.1 Å differences within the [2Fe–2S] cluster), only the A subunit was used for the investigations and only this subunit is discussed. Hydrogen atoms were added to the crystal structure and the protein was solvated in a sphere of water molecules with a radius of 36 Å using the Leap module in the Amber software suite. The protonation status of all residues was checked by the PROPKA program [34] and it was concluded that no residues have strongly perturbed pK_a values (thus, all arginine and lysine residues, except Arg^{260}, see below, were considered in their protonated state, whereas all aspartate and glutamate residues were considered in their deprotonated state). For the histidine residues, the protonation was decided from a detailed study of the solvent exposure and hydrogen-bond pattern. This procedure led to the following assignment: His^{34} and His^{107} were protonated on both nitrogen atoms, whereas His^{31} was protonated on $N^{\varepsilon 2}$ only and His^{152} was protonated on $N^{\delta 1}$ only. The cysteine residues coordinating the Fe/S clusters were assumed to be deprotonated. The [4Fe–4S] cluster, SAM and the dethiobiotin molecule found in the crystal structure were all included in the calculations. The total charge of the simulated system was −8 (neutral Arg) or −7 (protonated Arg). The positions of the added atoms were optimised by a 90 ps simulated-annealing molecular dynamics simulation, followed by 10 000 steps of conjugate gradient energy minimisation. All bond lengths involving hydrogen atoms were constrained by the SHAKE algorithm [35]. The water solvent was described explicitly using the TIP3P model [36]. The temperature was kept constant at 300 K using the Berendsen weak-coupling algorithm [37] with a time constant of 1 ps. The molecular dynamics time step was 2 fs. The non-bonded cut-off was 15 Å and the pair list was updated every 50 fs. In the QM/MM calculations, an

infinite cut-off was used instead.

The entire system was then divided into three subsystems: System 1 contained the [2Fe–2S] cluster and the relevant atoms of the four coordinating amino acids (Cys^{97}, Cys^{128}, Cys^{188} and Arg^{260}) and was treated with QM methods. The side chains were included as far as C^{β} for the Cys residues (replacing C^{α} by a H atom) and as far as C^{δ} for the Arg residue (replacing C^{γ} by a H atom). Thus, it consisted of $[(CH_3S)_3(CH_3NHCH(NH)NH_2)Fe_2S_2]^-$ for the calculations with neutral arginine or $[(CH_3S)_3(CH_3NHCH(NH_2)NH_2)Fe_2S_2]$ for the calculations with protonated arginine. System 2 included all residues with any atom within 6 Å of any atom in system 1 and was relaxed with MM methods. System 3 included the remaining protein atoms as well as the water molecules and was kept fixed at the crystallographic coordinates.

As both Fe atoms of the oxidized [2Fe–2S] cluster are in the Fe^{III} high-spin state ($S = 5/2$), two spin states are possible (the ferromagnetically, F, $S = 5$, or antiferromagnetically coupled states, AF, $S = 0$). The AF state always had a lower energy than the F state and it is also the one observed experimentally. Therefore, all presented results are AF energies. To ensure that the QM/MM energy differences are stable, the calculations were in general run forth and back between the relevant states until the energies were stable within 4 kJ/mol.

Similar calculations were also performed on one-electron reduced clusters, i.e., clusters containing one Fe^{II} and one Fe^{III} ion (net charge of QM system –1 or –2, depending on the protonation of the Arg model), on two-electron reduced clusters (net charge –2 or –3) and on clusters with one of the bridging S atoms removed (the one closest to dethiobiotin; net charge 0 or –1, so this is equivalent to remove a S^{2-} ion and reduce both iron ions to Fe^{II}), in all cases in the AF ($S = 1/2$ or $S = 0$) state.

For convenience, all discussed geometry optimisations were started from an initial optimisation with structure **2** (see below). In order to verify that this is acceptable, we tested

to what extent the optimised geometry depends on the starting geometry. In addition, the influence of the spin and oxidation states of the [2Fe–2S] cluster on structural properties was investigated. These explorative calculations were performed in vacuum (i.e., system 1 only), starting from the crystal geometry. Geometry optimisations for structure **2** in different oxidation states (Fe^{II}/Fe^{II}, Fe^{II}/Fe^{III} or Fe^{III}/Fe^{III}) and spin states (AF or F) showed that the final geometries of the intact clusters, especially the Fe···Fe distances and the orientation of the arginine residue, do not depend on starting geometries. The oxidised AF and F states were also tested for the other protonation states (structures **1**, **3** and **4**) with similar results. In all calculations, the AF state was energetically favoured (by 18 to 118 kJ/mol). Short Fe···Fe distances were found in all cases, although they were slightly longer for the F states (AF: 2.57–2.65 Å, F: 2.42–2.89 Å); no additional electronic states with larger Fe···Fe distances were detected. In order to verify this observation, the Fe···Fe distance was fixed to values between 2.5 and 3.5 Å (structure **2**, oxidised, AF state) and the rest of the geometry was optimised. Only one energy minimum was found at approximately 2.6 Å (61 kJ/mol more stable than the distance in the crystal structure) and no evidence for a second minimum close to the crystal structure distance was found.

Similar explorative calculations were performed with the hybrid B3LYP functional [38,39] (to examine the effect of another functional with exact exchange), giving similar results ($E_{AF} - E_F$ = −20 to −55 kJ/mol, d(Fe···Fe) = 2.55–2.72 (AF) and 2.70–2.96 Å (F) and an energy minimum at 2.8 Å, 36 kJ/mol lower than the crystal structure).

Quadrupole splittings were calculated according to

$$\Delta E_Q = 1/2eQV_{zz} \cdot (1 + \eta^2/3)^{1/2} \quad (2)$$

where Q = 0.16 barn ($1.6 \cdot 10^{-29}$ m^2) for ^{57}Fe, $\eta = (V_{xx} - V_{yy}) / V_{zz}$ with $|V_{xx}| < |V_{yy}| < |V_{zz}|$, and 1 mm/s = $4.8075 \cdot 10^{-18}$ eV.

Quantum-refinement calculations

We also performed a set of quantum refinement calculations, using the software COMQUM-X [40]. It can be seen as a QM/MM calculation, in which the structures are restrained towards crystallographic raw data. In COMQUM-X, the MM program is replaced by the crystallographic refinement program CNS (Crystallography & NMR system) [41]. In crystallographic refinement, the coordinates, *B* factors, occupancies, etc. are improved by optimising the fit of the observed and calculated structure-factor amplitudes, typically estimated by the residual disagreement, the *R* factor. Because of the limited resolution normally obtained with X-ray diffraction of biomolecules, a MM force field is used to supplement the data for the whole protein [42]. This force field ensures that the bond lengths and angles make chemical sense. In COMQUM-X, this force field is replaced by more accurate QM calculations for a small, but interesting, part of the protein (system 1), in a manner completely analogous to the use of QM in QM/MM calculations. The junctions are handled in the same way as in COMQUM.

Thus, the COMQUM-X refinement takes the form of a minimisation using an energy function of the form

$$E_{\text{ComQum-X}} = E_{\text{QM1}} - E_{\text{MM1}} + E_{\text{MM123}} + w_{\text{A}}\, E_{\text{Xray}} \qquad (3)$$

Here, E_{MM1} and E_{MM123} have the same meaning as in Eqn. 1, whereas E_{QM1} is the energy of the QM system, without any point-charge model of the surroundings. E_{Xray} is a penalty function, describing how well the model agrees with the experimental X-ray data. We have used the default maximum likelihood refinement target using amplitudes (MLF) in CNS [43]. w_{A} is a weight factor, which is necessary because E_{Xray} is in arbitrary units, whereas the other

terms are in energy units. It should be emphasized that the w_A factor is nothing special for quantum refinement. On the contrary, it also has to be set in standard crystallographic refinement (which is obtained from Eqn. 3 with $E_{QM1} = E_{MM1} = 0$), although it is rarely discussed. The default behaviour of CNS is to determine w_A so that the E_{Xray} and E_{MM123} forces have the same magnitude during a short molecular dynamics simulation [44], i.e. that the crystallographic raw data and the MM force field has a similar influence on the structure. We tested nine different values for the w_A factor between 0 and 30. Unfortunately, we encountered convergence problems if we used the default value of w_A (4.87) for some of the structures (because the crystallographically preferred structure of the [2Fe–2S] cluster is so poor at this low resolution that it becomes incompatible with the QM calculations). Therefore, we present results only for the largest value of the w_A factor that gave converged structures for all models, viz. $w_A = 1$. The results are qualitatively the same if other values are used, regarding the preferred model and coordination mode of the arginine ligand.

Following crystallographic custom, no hydrogen atoms were included in the MM region of the COMQUM-X calculations, because hydrogen atoms are not discernible in the crystal structure. Therefore, polarisation of the quantum system by the surrounding protein is not included in COMQUM-X.

Finally, it should be noted that the MM force field used in CNS (protein_rep.param, dna-rna_rep.param, water.param, and ion.param) is based on a statistical survey of crystal structures [45], rather than the energy-based force field in Amber and in the QM calculations. Therefore, the CNS energy has to be weighted by a factor of 1/3 to be comparable with the QM and Amber MM energies [40].

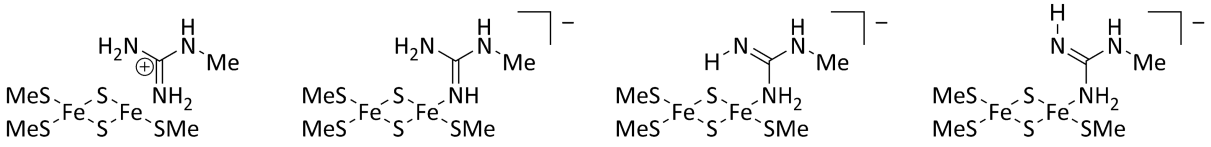
The quantum-refinement calculations were based on the same crystal structure as the QM/MM calculations (but both subunits were considered) [11] and the corresponding structure factors were downloaded from the PDB. Calculations were performed with the same

QM system as with QM/MM $[(CH_3S)_3(CH_3NHCH(NH_{1-2})NH_2)Fe_2S_2]$, as well as a QM system enlarged with a CH_3OH model of Ser[43] and a $CH_3NHCH(NH_2)_2$ model of Arg[95] (for both the intact oxidised cluster, as well as the one-electron reduced cluster without one of the bridging sulphide ions). The QM method and basis sets were the same as in the QM/MM calculations.

Results and Discussion

QM/MM calculations

Because of the importance of interactions of the [2Fe–2S] cluster and its four ligands with the surrounding protein and solvent, a QM/MM approach including the protein environment is the theoretical method of choice.



1 **2** **3** **4**

Structure 1. Conceivable coordination modes for arginine in the protonated (**1**) or neutral state (**2**–**4**) as represented in the QM system.

Four different structures were studied depending on the protonation state of Arg^{260}, as is illustrated in Structure 1. In the first (**1**), Arg^{260} is protonated, and therefore positively charged. In the other three, one of the terminal NH_2 groups of Arg^{260} is deprotonated. The three structures differ in whether the deprotonated NH group (**2**) or the protonated NH_2 group coordinates to Fe. In the latter case, the non-coordinating NH group can either have the H atom pointing towards the cluster (**3**) or away from the cluster (**4**).

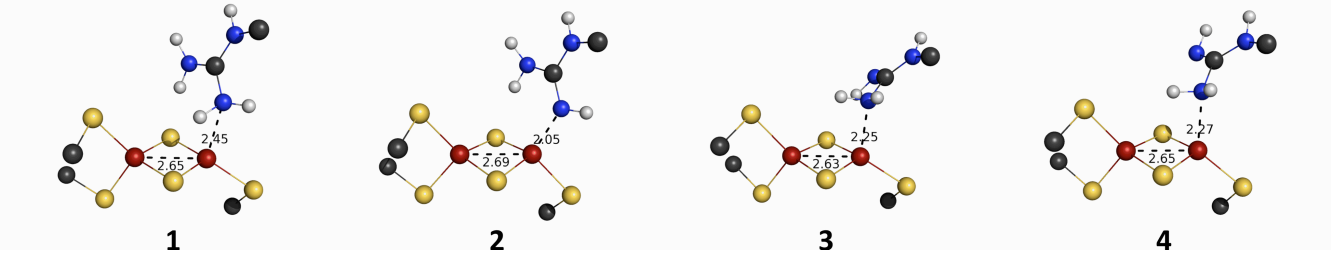


Figure 2. Optimised structures from the QM/MM calculations.

Optimised structures obtained from these calculations are shown in Figure 2. At a first glance, the geometry is quite similar in all cases. The Fe–N distances (Table 1) are in a range that agrees with the crystal structure (2.4 Å) considering its low resolution (3.4 Å) in all three structures with neutral arginine (**2**: 2.05 Å, **3**: 2.25 Å, **4**: 2.27 Å), as well as in the protonated structure (**1**: 2.45 Å). Judging from these distances, the Fe–N bond is strongest in **2**, which could be expected as the NH group has an sp^2-like, nucleophilic lone pair whereas the *p*-like lone pair of the NH_2 group is involved in π interactions within the guanidine group. The Fe···N distance in **1** is too long to assign a Fe–N bond in the protonated case. Nevertheless, the guanidinium group still is in close proximity to the [2Fe–2S] cluster.

Recently, Christianson et al. [13] performed a survey of metal–guanidine interactions in the Cambridge structural database. They found 150 such interactions in 45 different structures, but all except four of these involved a diguanidine moiety chelating a single metal, which is quite different from an arginine–metal coordination. They obtained metal–N distances of 1.84–2.08 Å (average 1.91±0.06 Å), but none of the complexes involved Fe. However, it is clear that only structure **2** gives a Fe–N distance that is similar to what is found in small inorganic metal–guanidine complexes. The survey of Christianson et al. also involved the only three protein crystal structures with metal–guanidine coordination, viz. biotin synthase, a H67R carbonic anhydrase I mutant [46], and an arginase L-arginine complex [47]. The latter

structure shows a Mn–N distance of 2.5 Å, whereas the Zn–N distance in the carbonic anhydrase mutant is 2.1 Å, i.e. the only protein structure that has metal–guanidine bond length similar to the small inorganic complexes. Christianson et al. assumed that the biotin synthase structure involved a deprotonated (neutral) arginine and did not consider any other possibility.

The Fe···Fe distance is quite independent of the coordination as well as the protonation state of the arginine group (2.63–2.69 Å) and is in perfect agreement with those of other [2Fe–2S] clusters that have been structurally characterised. None of the calculated structures reproduce the long Fe···Fe distance found in the BioB crystal structure (3.24–3.28 Å). As metal sites are often reduced during X-ray experiments, one- or two-electron reduced clusters containing Fe^{II} ions were also optimised, but they do not show any increased Fe···Fe distance. The main geometric change upon reduction is the dissociation of the NH_2 group (**1**, **3**, **4**). However, when one of the bridging S atoms was removed from the cluster (and the two Fe atoms thereby were reduced), the Fe···Fe distance increased in the structures with neutral arginine (**2**: 3.08 Å, **3**: 2.79 Å, **4**: 2.74 Å). In this case, only the protonated arginine (**1**) dissociates from the Fe atom. It is possible that one of the S atoms has been removed from the cluster (and inserted into dethiobiotin) in the crystal since the enzyme reaction is started by reduction of the [4Fe–4S] cluster, which could happen due to radiation damage during the measurement. It is not possible to decide whether or not there is a sulphur atom in the (dethio)biotin molecule in the crystal by examination of the electron density map due to the low resolution.

The non-coordinating NH/NH_2 group acts as a hydrogen-bond donor in the structures with a proton pointing towards the bridging sulphide (**1**, **2**, **3**). Additional hydrogen bonds can be found between Arg^{260} and four surrounding amino acids (Ser^{43}, Ser^{218}, Ser^{283} and Arg^{95}). Except for Ser^{283}, which accepts a hydrogen bond from the non-terminal NH^{ε} group in all cases, these residues seem to be quite flexible. Thus the positively charged Arg^{95} can act as

hydrogen bond donor towards the bridging sulphide of the [2Fe–2S] cluster (**1**, **2**) or the NH group of Arg^{260} (**3**, **4**). Ser^{218} acts as acceptor towards the NH_2 group of Arg^{95} in all structures and can in addition donate a hydrogen bond towards the bridging sulphide (**2**, **3**), whereas Ser^{43} can accept hydrogen bonds from the non-coordinating NH_2 group of Arg^{260} (**1**, **2**). In all cases, the arginine acts as a monodentate ligand; no evidence for secondary bonding interactions were found.

Table 1. Structural parameters from the crystal structure and from the QM/MM calculations.

	Fe–N /Å	Fe···Fe /Å
Crystal structure	2.40, 2.35	3.28, 3.24
1	2.45	2.65
2	2.05	2.69
3	2.25	2.63
4	2.27	2.65
1 + e^-	3.26	2.59
2 + e^-	2.11	2.66
3 + e^-	3.26	2.57
4 + e^-	3.42	2.58
1 + 2 e^-	3.19	2.62
2 + 2 e^-	2.18	2.68
3 + 2 e^-	3.24	2.62
4 + 2 e^-	3.23	2.60
1 − S	4.15	2.54
2 − S	2.01	3.08
3 − S	2.14	2.79

4 − S	2.15	2.74

Of the three structures with a neutral arginine, **2** has the lowest energy, 80 and 88 kJ/mol lower than that of structures **3** and **4**, respectively. This shows that coordination by the more nucleophilic NH group is preferred before NH_2 coordination. The energy of **1** cannot be compared directly because of the additional proton in the system. Comparing the protonated structure **1** with the best neutral structure **2**, only minor differences can be found, besides the Fe–N distance. The other relevant distances are similar, as are the hydrogen bonds close to the [2Fe–2S] cluster.

Mössbauer parameters of the [2Fe–2S] cluster in biotin synthase have been measured [48-50], showing a single quadrupole doublet with a quadrupole splitting of ΔE_Q = 0.51–0.53 mm/s. This is quite unexpected, because for a cluster containing two iron atoms with different coordination environment, two doublets would be expected. Quadrupole splittings (ΔE_Q) were calculated from the electric field gradients at the position of the iron atoms for the optimised geometries of structures **1**–**4** and are presented in Table 2.

Table 2. Experimental [48-50] and calculated quadrupole splittings (mm/s).

	ΔE_Q (S_4)	ΔE_Q (S_3N)
Experimental	0.53	0.53
1	0.33	1.27
2	0.33	0.59
3	0.45	1.14
4	0.43	1.23

While the quadrupole splittings calculated for the S coordinated iron atom are roughly the same in all four cases (0.33–0.45 mm/s), the other iron atom exhibits very different values

depending on the exact coordination mode. While ΔE_Q is quite large with the NH_2 coordinating (**1**: 1.27, **3**: 1.14, **4**: 1.23 mm/s) it is relatively small in case of NH coordination (**2**: 0.59 mm/s). Taking into account that calculated quadrupole splittings are usually too low in similar cases [18] and that the accuracy of calculated ΔE_Q (i.e., the amount by which they are lower than experimental values) seems to dependend on the coordination [18,27], the neutral state with the NH group coordinating (**2**) fits the experimental data best. As biological samples usually exhibit weak Mössbauer signals due to their low iron content, especially when another Fe/S cluster is present, it seems reasonable that the experimentally found doublet is the sum of two doublets with similar quadrupole splittings.

Thus, we can conclude that QM/MM calculations predict a neutral arginine with NH-coordination (structure **2**). Nevertheless, the reason for the experimentally found Fe···Fe distance as well as the significance of the unusual arginine ligand remains elusive.

Quantum-refinement calculations

Therefore, we studied the enzyme also by quantum refinement, which is standard crystallographic refinement, using the original experimental structure factors, but replacing the MM force field (which is used to supplement the crystallographic raw data and give accurate bond lengths and angles) for the active site by more accurate QM calculations. This will allow us to study what realistic structures of the [2Fe–2S] site actually fit into the electron density. In particular, we will be able to test what protonation state (structures **1**–**4**) fits the crystallographic raw data best. Two sizes of the QM system were tested (with or without models of Ser^{43} and Arg^{95}), as well as models of both the oxidised state with an intact

cluster and the two-electron reduced state with either an intact cluster or with one of the bridging S atoms removed.

The results are summarised in Table 3. It can be seen that all re-refined structures of the intact cluster give Fe···Fe distances (2.58–2.77 Å) that are appreciably shorter than in the crystal structure and therefore similar to those obtained in the QM/MM calculations. In the structures without one of the S atoms, the Fe···Fe distance is longer (2.85–2.99 Å), but not as long as in the crystal structure. However, it should be noted that both the R_{free} and residue (real-space) R factors are slightly lower for the original crystal structure than for any of the re-refined structures. This indicates a misfit between the crystal structure and the tested QM systems, which may indicate that we still have not yet tested the correct QM system or that the crystal structure is a mixture of several different structures, which is expected if the metal site is reduced during data collection.

The structure of the cluster also depends on details on the refinement protocol. Unfortunately, the original publication [11] does not provide such details and we have not been able to obtain them from the authors. Therefore, we tested to re-refine the structure of the [2Fe–2S] cluster with standard crystallography (i.e. with $E_{QM1} = E_{MM1} = 0$ in Eqn. 3) and with different treatments of the Fe–S interactions in the MM force field (i.e. in the E_{MM123} term). As can be seen in Table 3, the results are insensitive to whether Fe–S bonds are included with zeroed force constants (the preferred method to allow the site to be determined entirely by the experimental data; protocol i) in Table 3) or if no Fe–S bonds are defined (so that there is van der Waals interactions between all Fe and S ions; protocol ii) in Table 3). This indicates that the default w_A factor is so large that the MM force field has only minor influence on the structure of the [2Fe–2S] site. However, it can also be seen that the structure of the re-refined [2Fe–2S] site is quite different from the original crystal structure, showing that details in the refinement still differ. In particular, the Fe···Fe distance in our re-refined structure (2.97–

2.99 Å) is appreciably shorter than in the original crystal structure (3.24 Å). The re-refined results are similar to the quantum-refined results with a cluster without one S atom.

Table 3. Results of the quantum refinement calculations with $w_A = 0.1$. Distances between the Fe ions and the ligands (Å) are given, as well as the R_{free}, residue R factors, ΔE_{QM1} and Δr_{QM}, which are the differences in the energy and Fe–ligand distances of the QM system optimised in the crystal and in vacuum. Four sets of calculations are presented: with the small QM system or with the QM system enlarged by Ser^{43} and Arg^{95}, for the oxidised state (Fe^{III}_2) or for the two-electron reduced state (Fe^{II}_2), and for the intact [2Fe–2S] cluster or for the cluster with one sulphide ion removed. In addition, the data from the crystal structure (both subunits) are presented, as well as a standard crystallographic re-refinement of the [2Fe–2S] cluster with two different treatments of the Fe–S interactions (see the text).

	Fe···Fe	Distance to Fe1 (Å)				Distance to Fe2 (Å)				R_{free}	Residue	ΔE_{QM1}	Δr_1
	Å	N	S1	S2	S3	S2	S3	S4	S5		*R*	kJ/mol	Å
Small QM system, oxidised state, intact cluster													
1	2.58	3.29	2.24	2.21	2.19	2.25	2.22	2.25	2.25	0.3042	0.236	60.8	0.27
2	2.69	2.09	2.32	2.23	2.21	2.26	2.22	2.33	2.31	0.3028	0.197	41.6	0.13
3	2.69	2.28	2.30	2.20	2.21	2.24	2.23	2.33	2.30	0.3032	0.212	95.3	0.29
4	2.69	2.37	2.31	2.18	2.21	2.22	2.24	2.33	2.30	0.3035	0.213	127.1	0.36
Large QM system, oxidised state, intact cluster													
1	2.60	3.34	2.22	2.22	2.18	2.29	2.21	2.23	2.23	0.3044	0.243	178.5	0.91
2	2.74	2.10	2.29	2.25	2.20	2.31	2.20	2.31	2.29	0.3028	0.199	89.8	0.20
3	2.72	2.33	2.26	2.20	2.21	2.27	2.23	2.31	2.28	0.3031	0.228	130.9	0.31
4	2.67	3.18	2.24	2.19	2.17	2.26	2.23	2.28	2.29	0.3048	0.285	133.0	0.87
Large QM system, reduced state, intact cluster													
1	2.77	3.55	2.27	2.36	2.17	2.49	2.31	2.28	2.35	0.3042	0.250	125.0	0.65
2	2.67	2.09	2.31	2.27	2.24	2.34	2.28	2.38	2.34	0.3029	0.206	75.2	0.11
3	2.77	3.08	2.29	2.23	2.21	2.32	2.30	2.33	2.36	0.3043	0.265	204.9	0.96
4	2.66	3.25	2.25	2.19	2.33	2.36	2.33	2.33	2.40	0.3038	0.276	191.1	1.03
Large QM system, without one S atom													
1	2.88	2.41	2.23	2.22		2.27		2.27	2.24	0.3034	0.212	192.4	0.26
2	2.85	2.06	2.29	2.22		2.24		2.30	2.27	0.3040	0.196	161.6	0.20
3	2.99	2.24	2.24	2.20		2.25		2.30	2.26	0.3030	0.193	195.3	0.33
4	2.98	2.20	2.25	2.18		2.25		2.32	2.27	0.3030	0.196	156.4	0.14
Crystal structure re-refined without QM													
i	2.99	2.33	2.37	2.19	2.10	2.11	2.16	2.19	2.25	0.3004	0.138		1.18
ii	2.97	2.33	2.38	2.19	2.13	2.12	2.17	2.19	2.27	0.3004	0.140		1.11
Crystal structure													
A	3.24	2.35	2.32	2.23	2.22	2.22	2.23	2.28	2.30	0.3003	0.140	244.4	1.00

B	3.28	2.40	2.30	2.26	2.23	2.22	2.23	2.25	2.27	1.17

Among the four tested QM systems (models **1**–**4**), it is clear that the one with a deprotonated Arg^{260} and the NH group coordinated to Fe (model **2**) fits the crystallographic data best for all structures with an intact cluster: It has the lowest R_{free} and residue (real-space) R factors, and it also gives the lowest strain energy (ΔE_{QM1}, i.e. the energy difference of the QM system when optimised in the crystal or in vacuum) as well as the lowest difference in geometry when optimised in the crystal or in vacuum (Δr_1 in Table 3). Thus, all these four criteria point out the same structure as the best one, showing that the results are conclusive. In particular, it is clear that model **2** fits the crystallographic data appreciably better than the structure with a protonated Arg^{260} (model **1**). In fact, Arg^{260} dissociates from Fe in the latter model, giving Fe–N distances of 3.29–3.55 Å, when re-refined with $w_A < 1$. However, when $w_A = 1$ or higher, the Fe–N distance is shortened to 2.47–2.26 Å, showing that the crystal structure prefers shorter values. The two deprotonated models coordinating through the NH_2 group (models **3** and **4**) give intermediate fits to the crystal structure. Figure 3 shows that the quantum refined structure of model **2** fits the electron density equally well as the original crystal structure, and it also shows the rather poorly defined electron density at this low resolution.

For the structure without one of the S atoms, the results are somewhat different. Then, the various quality criteria give different results: Model **3** gives the lowest R_{free} and residue R factors and, whereas model **4** gives the smallest strain energy and difference in geometry. This indicates that this is not the correct model of the protein. This is also supported by the higher R_{free} factor (0.3030), compared to the best values obtained for the oxidised models (0.3028). However, the difference is not very large, indicating that the crystal structure might actually be a mixture of oxidised and reduced structures.

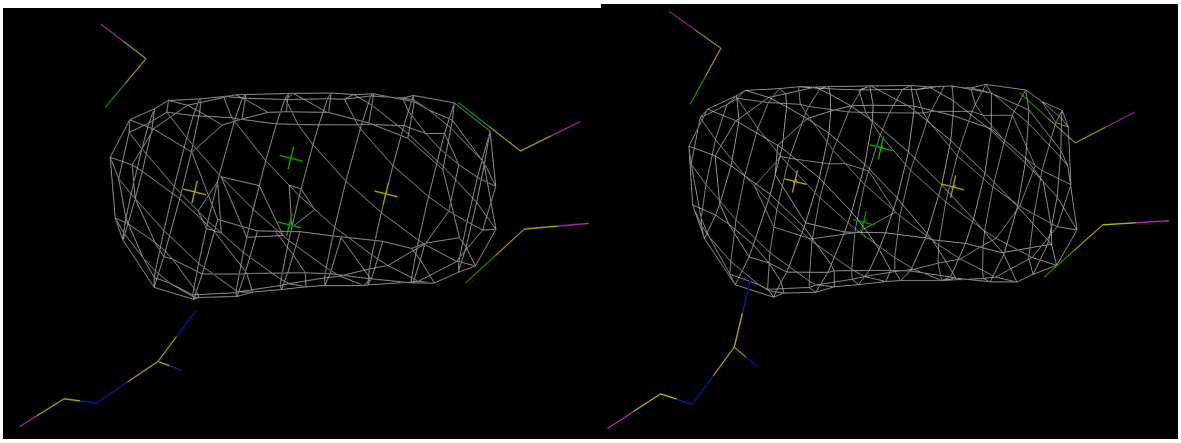


Figure 3. Electron-density omit $2f_o - f_c$ maps of the [2Fe–2S] cluster in the original crystal structure (left) and the quantum-refined model **2** (right) at the 2.0 σ level.

Conclusions

We have studied the structure of biotin synthase with both quantum refinement and QM/MM methods. This gives us the opportunity to interpret the crystal structure as much as its low resolution (3.4 Å) allows us. Several interesting results are obtained. First, it is quite clear that the Arg^{260} ligand is deprotonated in the crystal structure, because such structures fit the crystallographic raw-data best. Likewise, both the QM/MM energies and the quantum refinement strongly indicate that it is more favourable for the deprotonated Arg^{260} to coordinate to Fe via the deprotonated NH group, rather than by the NH_2 group, even if hydrogen bonds with the surrounding residues are considered. These conclusions are supported by calculated Mössbauer parameters which also fit the experimental data best for this coordination mode. Finally, it also seems clear that the Fe···Fe distance in the [2Fe–2S] cluster is not as long as the crystal structure indicated. Instead, it is most likely similar to what is found in all normal [2Fe–2S] clusters, i.e. approximately 2.7 Å. The reason for the long bond in the crystal structure may either be the low resolution or that the structure is a mixture of different states of the [2Fe–2S] cluster, e.g. caused by a successive reduction of the cluster

during data collection. Clearly, more accurate crystal structures are needed, as well as further theoretical investigations of the reaction intermediates, which should help to understand this fascinating enzyme.

Acknowledgements

Financial support by the Fonds der Chemischen Industrie (Kekulé fellowship for M.G.G.F.), the Deutsche Forschungsgemeinschaft and the Swedish research council (International Research Training Group GRK 1422 "Metal Sites in Biomolecules: Structures, Regulation and Mechanisms"; see www.biometals.eu) is gratefully acknowledged. The investigation has also been supported by computer resources of Lunarc at Lund University

References

1. Beinert, H, Holm, RH, Münck, E (1997) Science 277:653-659.

2. Johnson, DC, Dean, DR, Smith, AD, Johnson, MK (2005) Annu Rev Biochem 74:247-281.

3. Gibson, KJ, Pelletier, DA, Turner, IM, Sr. (1999) Biochem Biophys Res Commun 254:632-635.

4. Booker, SJ, Cicchillo, RM, Grove, TL (2007) Curr Opin Chem Biol 11:543-552.

5. Layer, G, Heinz, DW, Jahn, D, Schubert, W-D (2004) Curr Opin Chem Biol 8:468-476.

6. Tse Sum Bui, B, Florentin, D, Fournier, F, Ploux, O, Méjean, A, Marquet, A (1998) FEBS Letters 440:226-230.

7. Cicchillo, RM, Booker, SJ (2005) J Am Chem Soc 127:2860-2861.

8. Hernández, HL, Pierrel, F, Elleingand, E, García-Serres, R, Huynh, BH, Johnson, MK, Fontecave, M, Atta, M (2007) Biochemistry 46:5140-5147.

9. Anton, BP, Saleh, L, Benner, JS, Raleigh, EA, Kasif, S, Roberts, RJ (2008) Proc Natl Acad Sci U S A 105:1826-1831.

10. Ugulava, NB, Sacanell, CJ, Jarrett, JT (2001) Biochemistry 40:8352-8358.

11. Berkovitch, F, Nicolet, Y, Wan, JT, Jarrett, JT, Drennan, CL (2004) Science 303:76-79.

12. Farrar, CE, Jarrett, JT (2009) Biochemistry 48:2448-2458.

13. Di Costanzo, L, Flores Jr., LV, Christianson, DW (2006) Proteins 65:637-642.

14. Broach, RB, Jarrett, JT (2006) Biochemistry 45:14166-14174.

15. Ullmann, GM, Knapp, E-W (1999) Eur Biophys J 28:533-551.

16. Taylor, AM, Farrar, CE, Jarrett, JT (2008) Biochemistry 47:9309-9317.

17. Jarrett, JT (2005) Arch Biochem Biophys 433:312-321.

18. Ballmann, J, Dechert, S, Bill, E, Ryde, U, Meyer, F (2008) Inorg Chem 47:1586-1596.

19. Venkateswara Rao, P, Holm, RH (2004) Chem Rev 104:527-559.

20. Ryde, U (1996) J Comput-Aided Mol Des 10:153-164.

21. Ryde, U, Olsson, MHM (2001) Int J Quantum Chem 81:335-347.

22. Ahlrichs, R, Bär, M, Häser, M, Horn, H, Kölmel, C (1989) Chem Phys Lett 162:165-169.

23. Case, DA, Darden, TA, Cheatham III, TE, Simmerling, TL, Wang, J, Duke, RE, Luo, R, Merz, KM, Pearlman, DA, Crowley, M, Walker, RC, Zhang, W, Wang, B, Hayik, S, Roitberg, A, Seabra, G, Wong, KF, Paesani, F, Wu, X, Brozell, S, Tsui, V, Gohlke, H, Yang, L, Tan, C, Mongan, J, Hornak, V, Cui, G, Beroza, P, Mathews, DH, Schafmeister, C, Ross, WS, Kollman, PA. (2006) AMBER 9. University of California, San Francisco.

24. Becke, AD (1988) Phys Rev A 38:3098-3100.

25. Perdew, JP (1986) Phys Rev B 33:8822-8824.

26. Weigend, F, Ahlrichs, R (2005) Phys Chem Chem Phys 7:3297-3305.

27. Ballmann, J, Albers, A, Demeshko, S, Dechert, S, Bill, E, Bothe, E, Ryde, U, Meyer, F (2008) Angew Chem 120:9680-9684.

28. Cornell, WD, Cieplak, PI, Bayly, CI, Gould, IR, Merz, KM, Ferguson, DM, Spellmeyer, DC, Fox, T, Caldwell, JW, Kollman, PA (1995) J Am Chem Soc 117:5179-5197.

29. Wang, J, Cieplak, P, Kollman, PA (2000) J Comput Chem 21:1049-1074.

30. Rod, TH, Ryde, U (2005) J Chem Theory Comput 1:1240-1251.

31. Söderhjelm, P, Ryde, U (2006) J Mol Struct Theochem 770:199-219.

32. Weis, A, Katebzadeh, K, Söderhjelm, P, Nilsson, I, Ryde, U (2006) J Med Chem 49:6596-6606.

33. Svensson, M, Humbel, S, Froese, RDJ, Matsubara, T, Sieber, S, Morokuma, K (1996) J Phys Chem 100:19357-19363.

34. Li, H, Robertson, AD, Jensen, JH (2005) Proteins 61:704-721.

35. Ryckaert, JP, Ciccotti, G, Berendsen, HJC (1977) J Comput Phys 23:327-341.

36. Jorgensen, WL, Chandrasekhar, J, Madura, J, Klein, ML (1983) J Chem Phys 79:926-935.

37. Berendsen, HJC, Postma, JPM, van Gunsteren, WF, DiNola, A, Haak, JR (1984) J Chem Phys 81:3684-3690.

38. Becke, AD (1993) J Chem Phys 98:5648-5652.

39. Stephens, PJ, Devlin, FJ, Frisch, MJ, Chabalowski, CF (1994) J Phys Chem 98:11623-11627.

40. Ryde, U, Olsen, L, Nilsson, K (2002) J Comput Chem 23:1058-1070.

41. Brunger, AT, Adams, PD, Clore, GM, Delano, WL, Gros, P, Grosse-Kunstleve, RW, Jiang, J-S, Kuszewsk, JI, Nilges, M, Pannu, NS, Read, RJ, Rice, LM, Simonson, T, Warren, GL. (2000) Crystallography & NMR System CNS, Version 1.0. Yale University.

42. Kleywegt, GJ, Jones, TA (1997) Methods Enzymol 277:208-230.

43. Pannu, NS, Read, RJ (1996) Acta Cryst A 52:659-668.

44. Brünger, AT, Rice, LM (1997) Methods Enzymol 277:243-269.

45. Engh, RA, Huber, R (1991) Acta Cryst A 47:392-400.

46. Ferraroni, M, Tilli, S, Briganti, F, Chegwidden, WR, Supuran, CT, Wiebauer, KE, Tashian, RE, Scozzofava, A (2002) Biochemistry 41:6237-6244.

47. Bewley, MC, Jeffrey, PD, Patchett, ML, Kanyo, ZF, Baker, EN (1999) Structure 7:435-448.

48. Cosper, MM, Jameson, GNL, Hernández, HL, Krebs, C, Huynh, BH, Johnson, MK (2004) Biochemistry 43:2007-2021.

49. Tse Sum Bui, B, Benda, R, Schünemann, V, Florentin, D, Trautwein, AX, Marquet, A (2003) Biochemistry 42:8791-8798.

50. Tse Sum Bui, B, Marquet, A, Benda, R, Trautwein, AX (1999) FEBS Letters 459:411-414.