Journal of Molecular Graphics and Modelling 54 (2014) 123–130

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

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

Homology modeling, molecular dynamics, and site-directed mutagenesis study of AlkB human homolog 1 (ALKBH1)

Pavel Silvestrov[a], Tina A. Müller[b], Kristen N. Clark[b], Robert P. Hausinger[b], G. Andrés Cisneros[a,*]

[a] Department of Chemistry, Wayne State University, Detroit, MI 48202, United States
[b] Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, MI 48824, United States

ARTICLE INFO

Article history:
Accepted 21 October 2014
Available online 27 October 2014

ABSTRACT

The ability to repair DNA is important for the conservation of genetic information of living organisms. Cells have a number of ways to restore damaged DNA, such as direct DNA repair, base excision repair, and nucleotide excision repair. One of the proteins that can perform direct repair of DNA bases is *Escherichia coli* AlkB. In humans, there are 9 identified AlkB homologs, including AlkB homolog 1 (ALKBH1). Many of these proteins catalyze the direct oxidative dealkylation of DNA and RNA bases and, as such, have an important role in repairing DNA from damage induced by alkylating agents. In addition to the dealkylase activity, ALKBH1 can also function as an apyrimidinic/apurinic lyase and was proposed to have a distinct lyase active site. To our knowledge, no crystal structure or complete homology model of ALKBH1 protein is available. In this study, we have used homology modeling to predict the structure of ALKBH1 based on AlkB and Duffy-binding-like domain crystal structures as templates. Molecular dynamics simulations were subsequently performed on the predicted structure of ALKBH1. The positions of two disulfide bonds or a zinc-finger motif and a disulfide bond were predicted and the importance of these features was tested by mutagenesis. Possible locations for the lyase active site are proposed based on the analysis of our predicted structures and previous experimental results.

1. Introduction

Living cells depend on the conservation of genetic information stored in DNA. However, DNA is subject to various chemical reactions that can modify its structure and result in loss of its function. To ensure that genetic information is preserved and correctly propagated, cells have various pathways to protect the genetic code. For this purpose, cells can directly repair modified DNA bases, respond by regulating DNA transcription, stop their growth cycle, or undergo apoptosis [1].

Alkylation of DNA bases is an example of modifications that can result in DNA damage and mutations. There are various endogenous and exogenous factors that can induce alkylation [2–4]. Eukaryotic cells can repair alkylation damage by a number of ways: direct de-alkylation of DNA bases, base excision repair, and nucleotide excision repair [5]. Direct dealkylation can be performed by such *Escherichia coli* proteins as Ada and AlkB [6–8].

AlkB repairs DNA bases by catalyzing an oxidative dealkylation reaction. In humans, there are 9 identified AlkB homologs: ALKBH1 through ALKBH8 (alternatively named ABH1-ABH8) and FTO. The crystal structures of AlkB, ALKBH2, and ALKBH3 revealed an interesting mechanism: these enzymes flip out a nucleotide into the active site, which contains the Fe(II) cation coordinated by two histidines, an aspartate, and α-ketoglutarate [9–11]. It has been shown that enzymes in this family vary in the base eversion mechanism [12]. The reaction also requires molecular oxygen. In the course of the reaction the damaged DNA base is restored, α-ketoglutarate is transformed into succinate plus carbon dioxide, and formaldehyde is released [9]. AlkB and its homologs share a characteristic double-stranded beta-helix or "jelly-roll" fold [9–11]. The residues in the active site of known structures of AlkB and its homologs are strictly conserved [10,11,13]. Analysis of a crystal structure of AlkB suggested that the molecular oxygen necessary for the reaction diffuses to the active site of the enzyme through a putative oxygen tunnel [13]. QM/MM calculations and energy decomposition analysis performed for AlkB revealed a total of 9 residues around its active site that are important for the catalytic activity of the enzyme [14]. AlkB is a dynamic protein and when it has no bound cofactors it is mobile and can have different conformations [15].

* Corresponding author. Tel.: +1 313 577 2571; fax: +1 313 577 8022.
E-mail address: andres@chem.wayne.edu (G.A. Cisneros).

http://dx.doi.org/10.1016/j.jmgm.2014.10.013
1093-3263/

AlkB and its homologs differ in their affinity to DNA and RNA substrates and act on different nucleotides [16]. AlkB can repair 1-methyladenine and 3-methylcytosine and can act on both single- and double-stranded DNA [7,17]. ALKBH3 was shown to have a preference for single-stranded DNA, and ALKBH2 for double-stranded DNA [18,19]. AlkB and ALKBH3 can also repair 1-methyladenine and 3-methylcytosine in RNA [20]. AlkB, ALKBH2, and ALKBH3 repair 3-methylthymine and 1-ethyladenine in DNA [21–23]. 3-Ethylcytidine in DNA and 1-methylguanine in RNA also can be substrates for AlkB [22,24]. ALKBH1 catalyzes demethylation of single-stranded DNA and RNA and repairs 3-methylcytosine, but it does not act on 1-methyladenine [25].

In addition to having the DNA repair function, AlkB and its homologs could be involved in the regulation of DNA methylation and DNA expression [26]. Advances in understanding the cellular functions of dealkylases could result in the development of new medications. Alkylating agents have been used as chemotherapeutic drugs to treat cancer. Interfering with the ability of cancer cells to remove alkyl groups from DNA bases could increase the potency of the drugs used in chemotherapy [27,28].

ALKBH1 has unique structural and functional characteristics. Phylogenetic analysis revealed that evolutionarily ALKBH1 is closer to AlkB than to other AlkB homologs [29]. However, ALKBH1 is almost two times larger than AlkB. In addition to the dealkylase function, ALKBH1 also has an apyrimidinic/apurinic (AP) lyase activity, and the lyase active site has been proposed to be distinct from the dealkylase active site [30–32]. Upon performing the lyase catalysis, ALKBH1 covalently binds to the 5′ DNA product [31]. In addition, it has been proposed that ALKBH1 is implicated in the regulation of methylation patterns of histone H2A and is necessary for cell differentiation and neural development [33]. According to experimental studies by Müller et al., ALKBH1 is not involved in the base excision repair pathway [34].

A complete structure of ALKBH1 could provide insights into the mechanism and functions of the enzyme. To our knowledge, there is currently no complete folded model of ALKBH1. Vidyarthy et al. used a homology model approach to predict the structure of a region of ALKBH1 that has high sequence identity with AlkB [35]. Because ALKBH1 is substantially larger than AlkB, almost half of its sequence is not covered by the AlkB template. In this study we searched for proteins with known structures that could be used as additional templates along with AlkB. The complete tertiary structure of ALKBH1 was then predicted using the Rosetta suite of programs and molecular dynamics simulations were performed with the AMBER software. Specific residues were predicted to participate in disulfide or zinc-finger domains, and site-directed mutagenesis was carried out to assess the importance of these features. In addition, the structure was analyzed and a search for a possible lyase active site was performed.

2. Methods

2.1. Template search

With 389 residues, the ALKBH1 protein is too large for *ab initio* structure prediction, and thus a homology modeling approach was applied to predict its structure. AlkB and its homologs with known structures can serve as templates for the structure prediction of ALKBH1. Phylogenetic analysis revealed a close evolutionary relationship between AlkB and ALKBH1 [29]. Based on these findings, AlkB (PDB ID: 3I3Q) was selected as a template. A search for additional templates was performed since ALKBH1 is larger than AlkB and has sections in its sequence that are not covered by the AlkB template. A BLAST search for sequences similar to ALKBH1 did not provide peptides with known structures that could be used as templates. A further search was performed in MODBASE, a database of annotated protein structures [36]. The search resulted in 11 structures with various degrees of sequence identity to ALKBH1. These structures were used as leads for further sequence alignment analysis. In particular, the structure of a Duffy-binding-like domain (DBL) from *Plasmodium knowlesi* (PDB ID: 2C6J) was found to have 29% sequence identity with ALKBH1. Sequence alignments were performed with the T-Coffee server [37]. As shown in Fig. 1, DBL aligns to areas in the ALKBH1 sequence that are different from those that align to AlkB. There are also regions in ALKBH1 that align to both the DBL and the AlkB templates. DBL and AlkB together cover 88% of the sequence of ALKBH1. These two proteins were selected to serve as templates for the ALKBH1 structure prediction.

2.2. Homology modeling

The comparative modeling application of the Rosseta 3.4 suite of programs was used to predict the structure of ALKBH1 [38]. This program was chosen as it has shown successful results in protein structure prediction [39]. Based on the model of ALKBH1 developed by Vidyarthy et al. [35] and our sequence alignment analysis, the sequence of ALKBH1 was divided into 5 consecutive sections. The 1st (residues 1–74), 3rd (residues 187–293) and 5th (residues 369–389) sections align with the AlkB sequence, while the 2nd (residues 75–186) and 4th (residues 294–368) sections align with the sequence of DBL. Two structures were modeled: one based on the AlkB template (AlkB part) and the other based on the DBL template (DBL part).

To run Rosetta, the DBL and AlkB sections of the ALKBH1 sequence were aligned to the sequences of the respective templates (See supplementary Fig. S1). Three residues were introduced into the sequence of the AlkB part of ALKBH1 to compensate for the excluded DBL sections: threonine and asparagine in place of the 2nd sections and alanine in place of the 4th section. These residues were selected to match those in the aligned regions of the AlkB template. Also, there was a region in the ALKBH1 sequence that aligned to both the AlkB and DBL templates and thus was included in both modeled structural units. Introduced residues and overlapping regions were subsequently removed from the modeled structures with preference given to the AlkB model as it had overall higher alignment scores.

Ten thousand structures were generated for each of the two parts. The generated models were grouped in clusters according to structure deviations and ranked by energy score. For each of the two parts of ALKBH1, models from the largest cluster with the lowest energy were selected to ensure that the predicted structures are more energetically favorable and are consistent with the entropic effect of protein folding. For the AlkB part, there was only one cluster, indicating high target template sequence identity. For the DBL part, the largest cluster contained 78% of the generated structures and included the lowest energy structure. The selected structures of the DBL and AlkB parts of ALKBH1 were spliced together and oriented to form a tertiary ALKBH1 structure by forming 4 connections in Chimera (Fig. 2) [40].

2.3. Energy minimization and molecular dynamics in AMBER

The loops that link the two modeled parts of ALKBH1 were minimized in vacuum. 1000 cycles of the minimization were run with the *sander* program in AMBER12 [41] with the first 700 cycles using the steepest decent method. The *ibelly* option was set to allow only atoms in the linking loops to move. Subsequently, gas minimization for the whole structure was run for 1000 cycles. The resulting structure was solvated using TIP3P water [42], and neutralized by adding 28,644 water molecules and one sodium ion with the *tleap* tool

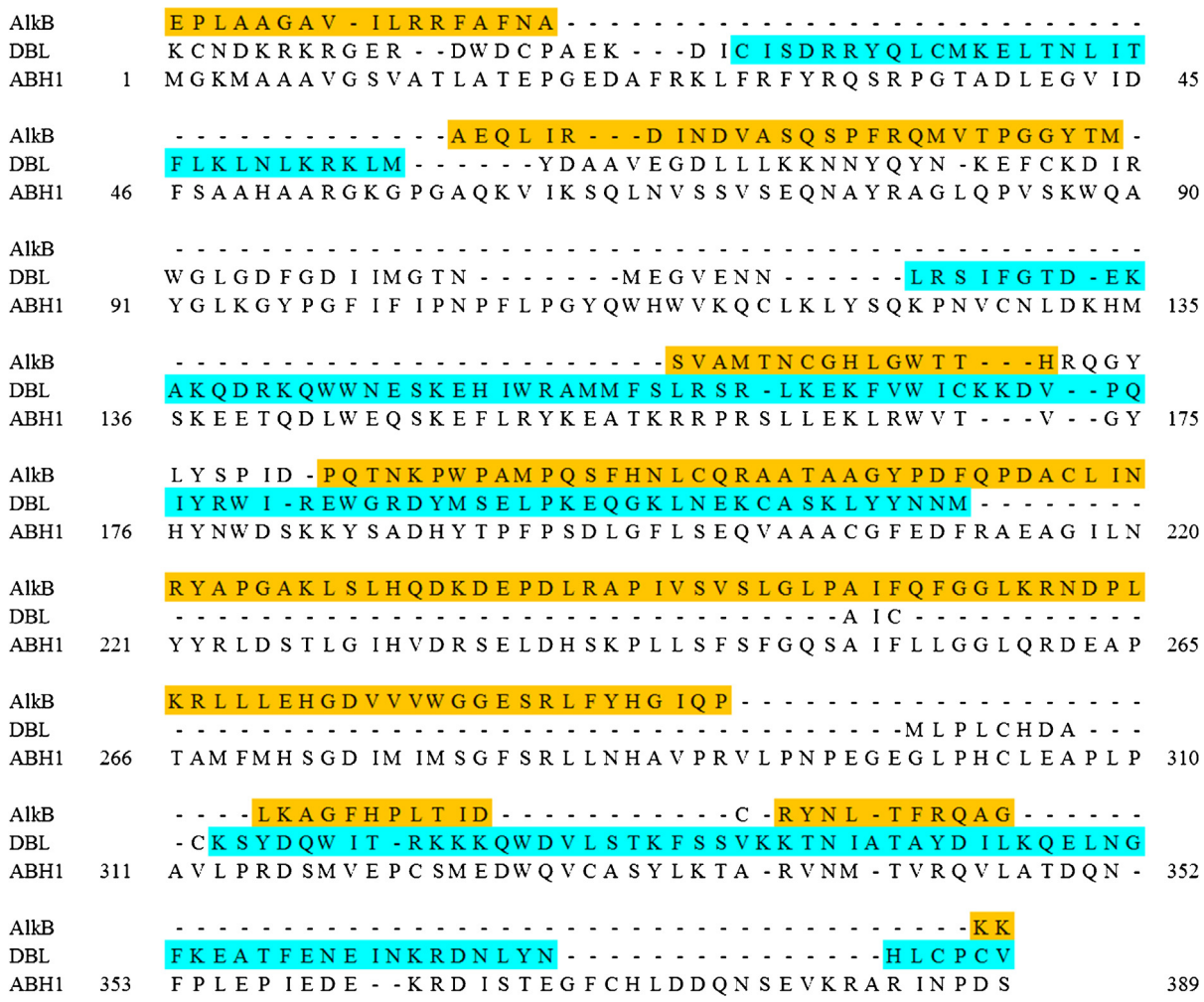

Fig. 1. Sequence alignments of AlkB and DBL with ALKBH1. Alignments with average and high scores as per T-coffee reports are color coded: orange for AlkB and cyan for DBL. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

resulting in a total of 92,039 atoms in the final system. Molecular dynamics (MD) simulation was first run with the ALKBH1 structure restrained under constant pressure conditions until the density of water reached 1 g cm^{-3}. Subsequently, the MD simulation was run under constant volume conditions with the restraints on ALKBH1 being gradually released. The unrestrained simulation was run for 100 ns using the ff99SB force field at 300 K temperature with the Langevin thermostat, a cut-off for non-bonded interactions of 8 Å and the sPME method for long-range electrostatics, and a time step of 1 fs. MD simulations were performed with the *pmemd* program in AMBER12 [41]. Two simulations were run: one with two pairs of cysteines specified as disulfide bridge forming residues, and another without such specification. Disulfide bonds were predicted based on our analysis of ALKBH1 structure and reports obtained from the DIpro server developed by Cheng et al. [43].

2.4. Site-directed mutagenesis

Mutants were created according to the QuickChange site directed mutagenesis protocol (Qiagen) using the respective plasmids and appropriate primers (Supplementary table S1). Plasmids were isolated with the Qiagen miniprep kit according to the manufacturer's instructions and the introduction of the mutations was confirmed by sequencing of the DNA.

2.5. Cell growth and protein purification

Plasmids were transformed into *E. coli* CodonPlus RIPL cells and the wild-type and variant proteins were expressed and purified as previously described [30]. Fractions were examined by sodium dodecyl sulfate polyacrylamide gel electrophoresis for their protein content and purity. The desired samples were pooled and subjected to buffer exchange into 20 mM Tris, pH 8, containing 1 mM EDTA (TE) by using a disposable G25 column (GE Healthcare) according to the manufacturer's instructions.

2.6. Apurinic/apyrimidinic (AP)-site lyase activity assays

AP lyase activity assays were essentially performed as described previously [31]. Briefly, single-stranded oligonucleotide 5′-AACTTCGTGCAGGCATGGTAG(dU)TTGTCTACT-3 (10 μM) was incubated with uracil DNA glycosylase (UDG) in UDG buffer for 30 min at 37 °C to create the AP site. For double-stranded DNA substrates, the complementary oligonucleotide 5′-AGTAGACAAG(dU)GACCATGCCTGCACGAAGTT-3′ was annealed prior to the UDG treatment by placing the sample in a boiling water bath and cooling it to room temperature.

Standard activity assays were performed with 5 μM protein and 1 μM DNA in TE buffer. The reactions were incubated for 1 h at 37 °C. When appropriate, methoxyamine was then added to the sample

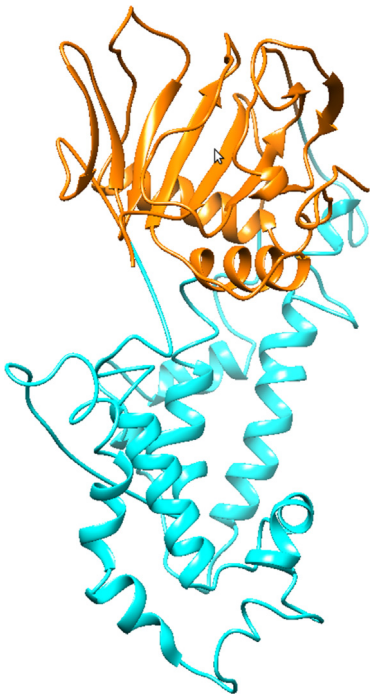

Fig. 2. Reconstituted ALKBH1 structure. The part colored in orange corresponds to the AlkB template, in cyan—to the DBL template. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

to a final concentration of 10 mM to reduce unspecific cleavage of the AP site and the samples were subjected to proteinase K treatment (2 U/sample) for 30 min at 65 °C. DNA products were analyzed by native or denaturing PAGE and the oligonucleotides were visualized either by staining with ethidium bromide or by using a 5′- or 3′-FAM labeled substrate and detecting the fluorescence using a phosphorimager.

3. Results and discussion

3.1. Overall predicted structure

The obtained structures of each of the two parts of ALKBH1 were analyzed in regard to differences from and similarities to their respective templates. Superposition of the AlkB template (PDB ID: 3I3Q) with the predicted structure of the respective part of ALKBH1 revealed that the alkylase active site residues, two histidines and an aspartate (H231, D233, and H287), are structurally conserved (Fig. 3a). When the predicted AlkB part of ALKBH1 was superposed onto AlkB with cofactors (PDB ID: 3BI3), ALKBH1 residues Y222, L228, S245, F254, and N220 located in proximity to the active site were found to be also structurally conserved.

Residues that were not conserved were also analyzed. In the place of AlkB residues T208, N206, and R204, the predicted ALKBH1 structure had R382, V380, and S378, respectively. The arginine residue in the ALKBH1 structure in the place of T208 of AlkB overlaps with the α-ketoglutarate cofactor of AlkB when the two structures are superposed (Fig. 3b). A mirror difference in the location of AlkB residue R204, with the ALKBH1 structure having a serine there, does bring a guanidine group close to the active site in AlkB, but from a different side. Such structural variation could result in a different orientation of cofactors in AlkB and ALKBH1 proteins and could modulate substrate selectivity.

Superposition of the DBL template (PDB ID: 2C6J) with the respective predicted part of ALKBH1 and analysis of the residues that were shown to be necessary for recognition of cell surface receptors by DBL as described by Singh et al. [44], revealed that only one residue was conserved: K75 of DBL superposed with K87 of ALKBH1.

3.2. Analysis of cysteine residues

The ALKBH1 sequence has 7 cysteines, which could have structural and functional roles. Using DIpro, a disulfide bond predictor server [43], three possible disulfide bonds were identified from the sequence of ALKBH1: C118-C129, C330-C371, and C304-C322, listed in the order of decreasing probability. Analysis of the predicted structures of ALKBH1 revealed that the disulfide bond C118-C129 is possible without significant modifications of the tertiary structure of ALKBH1. However, C330-C371 and C304-C322 disulfide bridges would require substantial changes to the structure and thus are less feasible. The DBL protein, which was used as a template to predict the structure of a respective part of ALKBH1, has a total of 12 cysteines that form 6 disulfide bonds. Superposition of the DBL template and the DBL part of ALKBH1 revealed that the ALKBH1 residues C304 and C371 correspond to two disulfide bridge-forming cysteines in the DBL structure. Formation of a disulfide bridge between these residues requires a small change to the predicted tertiary structure of ALKBH1. Thus, we propose

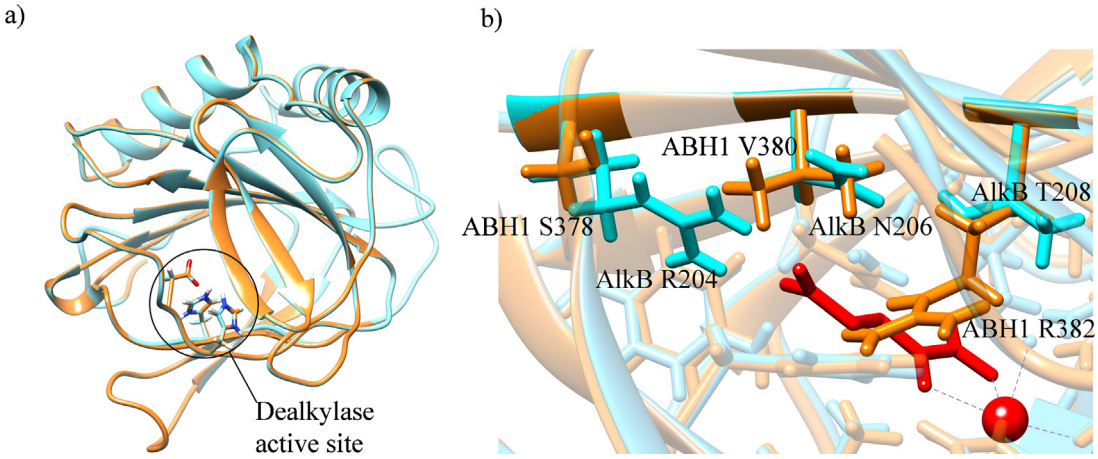

Fig. 3. Superposed structures of ALKBH1 (orange) and the AlkB template (cyan). (a) Dealkylase active site residues. (b) Non-conserved residues in proximity to the active site. Metal ion and α-ketoglutarate of the AlkB structure are shown in red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

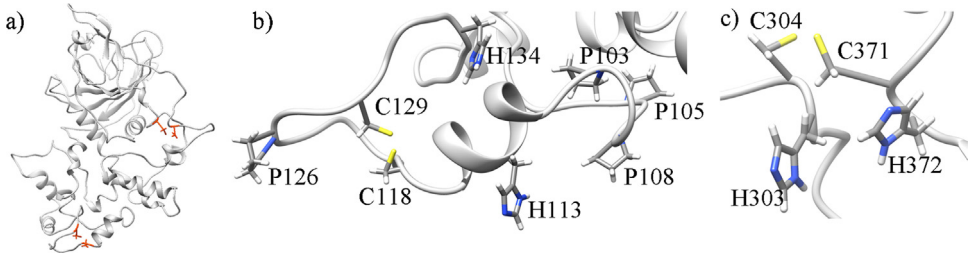

Fig. 4. (a) Location of the two proposed disulfide bridges in the structure of ALKBH1; cysteines are shown in orange. (b) C118-C129 disulfide bridge with nearby histidines and prolines (c) C304-C371 disulfide bridge with contiguous histidines. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

that ALKBH1 has two disulfide bridges: C118–C129 and C304–C371 (Fig. 4).

3.3. MD simulations

Two MD simulations were performed: with and without the two proposed disulfide bridges, each for 100 ns. With the disulfide bonds specified, the RMSD of the ALKBH1 structure stabilizes after 60 ns and overall is lower than the RMSD of the simulation without the disulfide bonds (Fig. 5a). All RMSD plots were calculated for the backbone atoms with the first structures after gas minimization as reference points. RMSD analysis of each of the two parts of ALKBH1 showed that the AlkB part has a stable RMSD of around 5 Å for both systems. It can be seen in Fig. 5b that the presence of the disulfide bridges has a strong stabilizing effect on the DBL part. Higher stability of the structure with the disulfide bonds is the result of lower degrees of freedom in comparison to the structure that did not have such constraint and could indicate that the ALKBH1 protein requires the formation of disulfide bonds in order to fold correctly and maintain its structural stability and function.

The high RMSD observed can be partly explained by a hinge movement of the DBL part around the AlkB part that can be seen in the superposition of the starting structure onto snapshots throughout the simulation (Fig. 6). Loop reorganization and partial

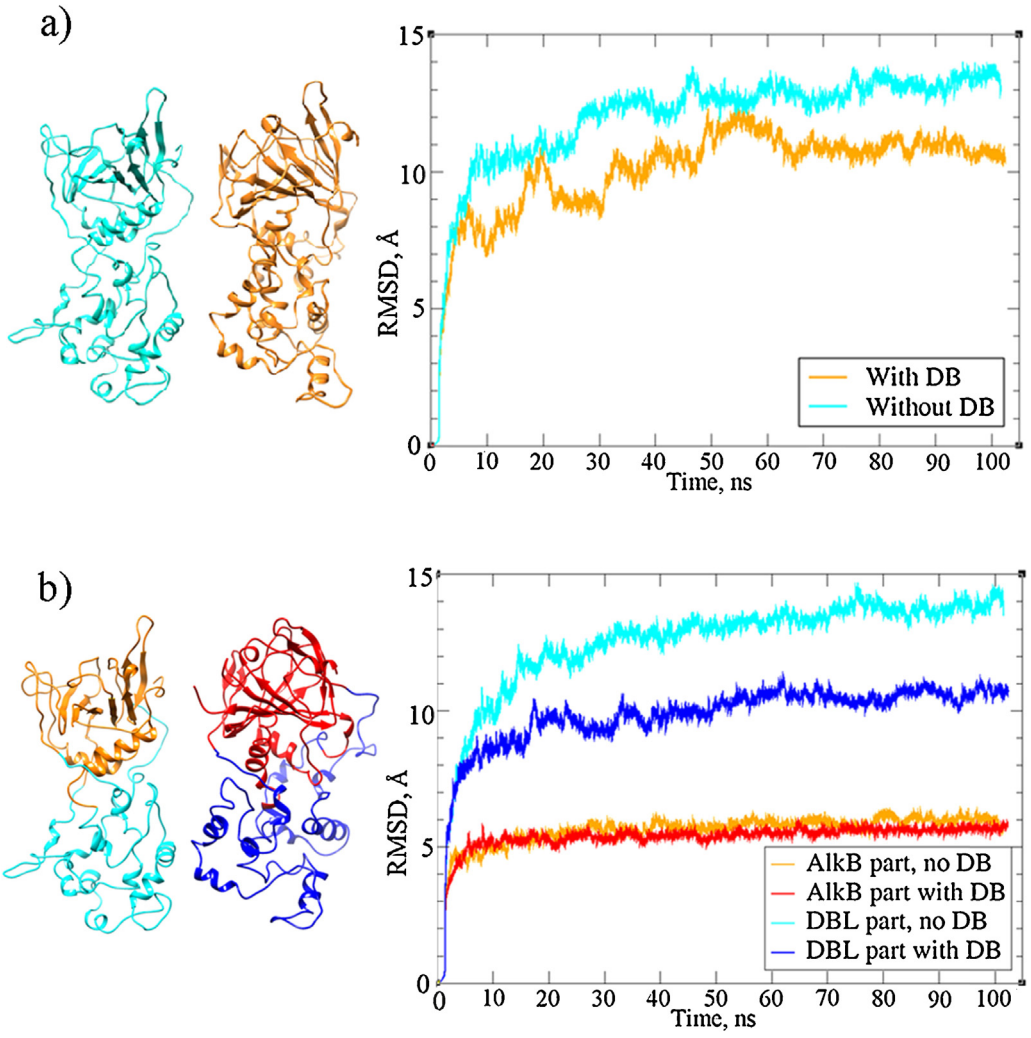

Fig. 5. RMSD plots for (a) ALKBH1 with and without disulfide bridges (DB); (b) AlkB and DBL parts of ALKBH1 with and without DB.

Fig. 6. Superposition of ALKBH1 at the beginning (cyan) of the unrestrained simulation and the final structure (orange) after 100 ns of MD simulation.

unfolding close to the flexible loops, particularly in the DBL part, also contributes to the high RMSD level. Here, we point out that the predicted AlkB and DBL parts to form the full ALKBH1 homology model were obtained separately and were spliced together without any structural information (as explained in Section 2.2). Thus, the relaxation of the overall structure observed as a rigid body rotation of the DBL part with respect to the AlkB part during the MD simulation is not unexpected. Indeed, it is observed that both systems (with and without disulfide bonds) show a rapid decrease in potential energy in the first 5 ns (see Fig. S2). This shows that the initial spliced structure is in a high energy conformation and the hinge motion results in a more stable structure.

3.4. Site-directed mutagenesis studies

Residues in proximity to the proposed disulfide bonds were further analyzed. For the C118-C129 bond, it was noticed that there is a histidine located 4 residues away from each of the two cysteines (H113 and H134, respectively) (Fig. 4b). It was also found that there are four prolines (P126, P103, P105, P108), three of which are between these two cysteines. The disulfide bond forms a closed loop containing residues 118 through 129. Given the location of C118, C129, H113, and H134, as well as the flexibility of adjacent regions, these four residues could alternatively form a 2Cys2His zinc finger domain that could potentially play a role in DNA sequence recognition. Interestingly, cysteines that may form the second proposed disulfide bridge, C371 and C304, have contiguous histidine residues, H372 and H303 (Fig. 4c).

Site-directed mutagenesis studies were carried out to test the importance of the two predicted disulfide bonds and the proposed zinc-finger motif. Thus, eight variant proteins (H113A, C118A, C129A, H134A, H303A, C304A, C371A, and H372A forms of ALKBH1) were produced in recombinant *E. coli* cells and purified. The AP lyase activity of these proteins was indistinguishable from the wild-type protein, thus ruling out the need for an intact disulfide or zinc finger for DNA binding or cleavage at the AP site. These studies do not, however, rule out the existence of the predicted disulfides or the presence of a zinc-finger domain in the native protein. In addition to the single-site variants, seven double-site substitutions (H113A/C118A, H113A/C129A, C118A/C129A, C129A/H134A, H303A/C304A, C304A/C371A and C371A/H372A) were examined. Most of these variants were unaffected in their AP lyase activity with the exception of the H113A/C118A and C129A/H134A proteins. The former showed 83 ± 12.6%, the latter 65 ± 9.5% of wild-type ALKBH1 activity using our standard assay conditions. A quadruple variant (H113A/C118A/C129A/H134A) was also created and analyzed, but showed no further reduction in its DNA cleaving activity than the double variant C129A/H134A (69 ± 8% of wild-type ALKBH1). Taken together, these data suggest that the residues in this region of the protein are important for AP lyase activity, which is in agreement with the previous findings that K133 is the most critical active site lysine (see below).

3.5. Analysis of lysine residues

There are 21 lysines in the sequence of ALKBH1. Experimental studies have shown that ALKBH1 can function as an AP lyase, with lysine residues likely to act as nucleophiles in catalysis [31]. Also, it has been shown that residues in the N-terminus of ALKBH1, possibly lysines, can covalently bind to a cleaved DNA product [31]. Mutagenesis and activity assay studies have revealed that the lyase

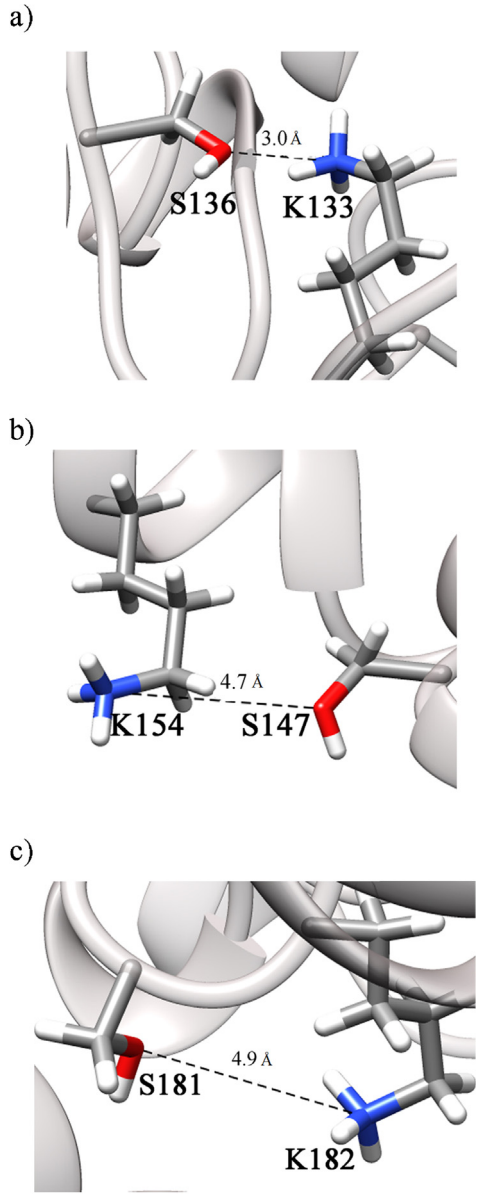

Fig. 7. Serines in proximity to active lysine residues. Snapshots at the 100th ns of the simulation are shown for the three pairs: (a) K133 and S136 (b) K154 and S147 (c) K182 and S181. See supplementary Fig. S3 for the dynamics of the distances throughout the simulation.

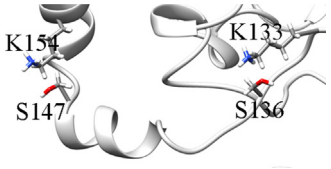

Fig. 8. Region between residues K154 and K133.

function of ALKBH1 is lowest in the K133A and K25A variants and is further decreased in a double mutant K133A/K154A [31]. It was proposed that K133 could function as a primary lyase nucleophile, with other lysines, such as K154, being able to substitute for it in the K133A variant [31]. The K182A variant also has reduced activity, though to a lesser extent than does K133A ALKBH1 [31].

Our analysis of the snapshot of the predicted structure of ALKBH1with disulfide bridges at the 100th ns of the MD simulation revealed that K133, K154, and K182 are located within 5 Å of serine residues S136, S147, and S181, respectively (see Fig. 7). The distances between the residues were identified as the distances between the oxygen atoms of the hydroxyl groups of serines and the nitrogen atoms of the amino groups of lysines as these atoms could have electrostatic interactions. K133 and K182 adopt conformations where they could be activated by proximate serines for a significant amount of time during the simulation. Conversely, K154 exhibits a relatively long distance to S147 for most of the simulation time, and only gets closer at the end of the simulation. Serine residues having higher pK_as could act as nucleophiles or increase the nucleophilicity of nearby lysines enabling them to perform the lyase catalytic function. In addition, it has been shown that serines can form a covalent bond with DNA cleavage products [45]. Serines could also facilitate cleavage by forming a hydrogen bond with the phosphate group of DNA, thus directing lysines to the right location for the nucleophilic attack [46].

Proposed compensation of the lyase activity in the K133A variant by K154 could be explained by the proximity of the two lysines to each other (see Fig. 8) [31]. Substituting K133 could result in a change of the orientation of K154 and S147, which could result in an increase in their catalytic ability. Alternatively, K133 and K154 could play a role in DNA recognition and non-covalent binding prior to the cleavage of the DNA molecule. When the positively charged lysines are mutated to alanines, the affinity of ALKBH1 to the negatively charged DNA backbone could be lowered as a result of decreased electrostatic interactions between the two molecules. If the C118, C129, H113, and H134 residues form a zinc finger domain that has a DNA binding function, other residues in this region of the ALKBH1 structure, such as K133 and K154, could also play a role in binding to DNA.

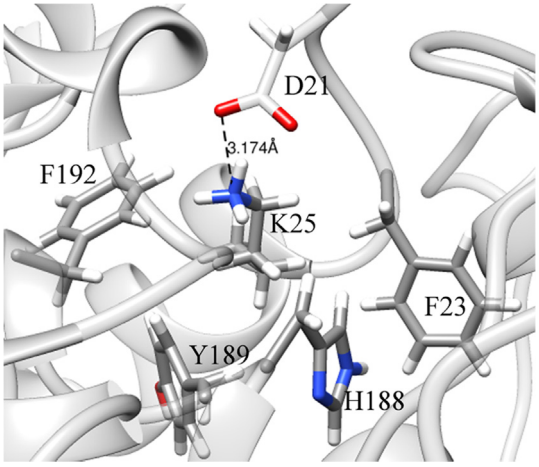

Fig. 9. Residues in proximity to K25. Shown is a snapshot at the 100th ns of the simulation with disulfide bridges. See supplementary Fig. S4 for the dynamics of distances.

The amino group of K25 is close to a carboxylate oxygen of D21 throughout the course of the simulation (Fig. 9) and the resulting structure revealed that K25 is located in a hydrophobic region formed by F23, Y189, F192, and H188. Such an environment could significantly increase the reactivity of K25 as a nucleophile, thus making K25 a possible candidate to form and maintain a covalent bond with DNA. The K25 and D21 pair resembles the active site of endonuclease III, which has active site lysine and aspartate residues located by a hydrophobic pocket [47]. Analysis of distances from the nitrogen of the amino group of K25 to the nearby residues revealed that the distance to F192 was on average the shortest (6.6 Å), and was more stable than distances to the other residues in the region (Fig. 9).

4. Conclusions

We have developed a homology model of ALKBH1 based on two templates with known structures: AlkB (PDB ID: 3I3Q) and DBL (PDB ID: 2C6J). Two structural units of ALKBH1, each corresponding to one of the templates, were calculated separately and subsequently combined to form the tertiary structure of ALKBH1. The Rosetta suite of programs was used to predict possible structures and to select the most feasible of them. Subsequently, MD simulations were run on the full structure using the AMBER software. Based on our final model, we propose two disulfide bridges exist in the structure of ALKBH1: C118-C129, and C304-C371. Alternatively, residues C118 and C129 together with H113 and H134 could form a zinc finger domain that could play a role in DNA recognition and binding. While site-directed mutagenesis studies could not distinguish between these interpretations, the importance of these residues was confirmed by the reduced AP lyase activity of the double and quadruple variants. The environments of lysine residues that were experimentally shown by Müller et al. [31] to be implicated in the lyase activity of ALKBH1 were analyzed. Residues K133, K154, K182 are located in proximity to serines S136, S147, S181, respectively. These serines could act as nucleophiles or enhance the nucleophilicity of the nearby lysines, they could form covalent or hydrogen bonds with DNA, they could play a role of directing lysines for a nucleophilic attack. K25 is located in proximity to D21 in a hydrophobic pocket formed by F192, F23, and Y189. We propose K25 as a possible nucleophile that forms a covalent bond with the DNA product.

Supporting information

Primers and plasmids used to create the ABH1 variants.

Sequence alignments of ALKBH1 parts to the respective templates.

Potential energy plots of the MD simulations.

Plots of the dynamics of the distances between selected serine and lysine residues.

Plots of the dynamics of the distances between K25 and proximal residues.

Acknowledgements

This work was supported in part by the National Institutes of Health (GM063584 to R.P.H. and GM108583-A1 to G.A.C.). Computing time from Wayne State's C&IT is gratefully acknowledged. P.S. thanks Wayne State University for a graduate research fellowship during this work.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jmgm.2014.10.013.

References

[1] A. Sancar, L.A. Lindsey-Boltz, K. Ünsal-Kaçmaz, S. Linn, Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints, Annu. Rev. Biochem. 73 (2004) 39–85.

[2] B. Rydberg, T. Lindahl, Nonenzymatic methylation of DNA by the intracellular methyl group donor *S*-adenosyl-L-methionine is a potentially mutagenic reaction, EMBO J. 1 (1982) 211–216.

[3] P. Taverna, B. Sedgwick, Generation of an endogenous DNA-methylating agent by nitrosation in *Escherichia coli*, J. Bacteriol. 178 (1996) 5105–5111.

[4] S.S. Hecht, DNA adduct formation from tobacco-specific *N*-nitrosamines, Mutat. Res. 424 (1999) 127–142.

[5] F. Drabløs, E. Feyzi, P.A. Aas, C.B. Vaagbø, B. Kavli, S. Bratlie, J. Pena-Diaz, M. Otterlei, G. Slupphaug, H.E. Krokan, Alkylation damage in DNA and RNA—repair mechanisms and medical significance, DNA Repair 3 (2004) 1389–1407.

[6] B. Sedgwick, Repairing DNA-methylation damage, Nat. Rev. Mol. Cell Biol. 5 (2004) 148–157.

[7] S.C. Trewick, T.F. Henshaw, R.P. Hausinger, T. Lindahl, B. Sedgwick, Oxidative demethylation by *Escherichia coli* AlkB directly reverts DNA base damage, Nature 419 (2002) 174–178.

[8] Y. Mishina, C. He, Oxidative dealkylation DNA repair mediated by the mononuclear non-heme iron AlkB proteins, J. Inorg. Biochem. 100 (2006) 670–678.

[9] C. Yi, G. Jia, G. Hou, Q. Dai, W. Zhang, G. Zheng, X. Jian, C.G. Yang, Q. Cui, C. He, Iron-catalysed oxidation intermediates captured in a DNA repair dioxygenase, Nature 468 (2010) 330–333.

[10] C.G. Yang, C. Yi, E.M. Duguid, C.T. Sullivan, X. Jian, P.A. Rice, C. He, Crystal structures of DNA/RNA repair enzymes AlkB and ABH2 bound to dsDNA, Nature 452 (2008) 961–965.

[11] O. Sundheim, C.B. Vagbo, M. Bjoras, M.M.L. Desousa, V. Talstad, P.A. Aas, F. Drablos, H.E. Krokan, J.A. Tainer, Slupphaug, Human ABH3 structure and key residues for oxidative demethylation to reverse DNA/RNA damage, EMBO J. 25 (2006) 3389–3397.

[12] O. Sundheim, V.A. Talstad, C.B. Vågbø, G. Slupphaug, H.E. Krokan, AlkB demethylases flip out in different ways, DNA Repair 7 (2008) 1916–1923.

[13] B. Yu, W.C. Edstrom, J. Benach, Y. Hamuro, P.C. Weber, B.R. Gibney, J.F. Hunt, Crystal structures of catalytic complexes of the oxidative DNA/RNA repair enzyme AlkB, Nature 439 (2006) 879–884.

[14] D. Fang, R.L. Lord, G.A. Cisneros, Ab initio QM/MM calculations show an intersystem crossing in the hydrogen abstraction step in dealkylation catalyzed by AlkB, J. Phys. Chem. B 117 (21) (2013) 6410–6420.

[15] B. Bleijlevens, T. Shivarattan, K.S. van den Boom, A. de Haan, G. van der Zwan, P.J. Simpson, S.J. Matthews, Changes in protein dynamics of the DNA repair dioxygenase AlkB upon binding of Fe^{2+} and 2-oxoglutarate, Biochemistry 51 (2012) 3334–3341.

[16] B. Sedgwick, P.A. Batesb, J. Paikc, S.C. Jacobsa, T. Lindahla, Repair of alkylated DNA: recent advances, DNA Repair 6 (2007) 429–442.

[17] P.O. Falnes, R.F. Johansen, E. Seeberg, AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli*, Nature 419 (2002) 178–182.

[18] P.O. Falnes, M. Bjoras, P.A. Aas, O. Sundheim, E. Seeberg, Substrate specificities of bacterial and human AlkB proteins, Nucleic Acids Res. 32 (2004) 3456–3461.

[19] Y. Mishina, C.J. Lee, C. He, Interaction of human and bacterial AlkB proteins with DNA as probed through chemical cross-linking studies, Nucleic Acids Res. 32 (2004) 1548–1554.

[20] A. Aas, M. Otterlei, P.O. Falnes, C.B. Vågbø, F. Skorpen, M. Akebari, O. Sundheim, M. Bjørås, G. Slupphaug, E. Seeberg, H.E. Krokan, Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA, Nature 421 (2003) 859–863.

[21] P. Koivisto, P. Robins, T. Lindahl, B. Sedgwick, Demethylation of 3-methylthymine in DNA by bacterial and human DNA dioxygenases, J. Biol. Chem. 279 (2004) 40470–40474.

[22] P.Ø. Falnes, Repair of 3-methylthymine and 1-methylguanine lesions by bacterial and human AlkB proteins, Nucleic Acids Res. 32 (2004) 6260–6267.

[23] T. Duncan, S.C. Trewick, P. Koivisto, P.A. Bates, T. Lindahl, B. Sedgwick, Reversal of DNA alkylation damage by two human dioxygenases, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 16660–16665.

[24] J.C. Delaney, J.M. Essigmann, Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in AlkB *Escherichia coli*, Proc. Natl. Acad. Sci. U.S.A. 101 (2004) 14051–14056.

[25] M.P. Westbye, E. Feyzi, P.A. Aas, C.B. Vagbo, V.A. Talstad, B. Kavli, L. Hagen, O. Sundheim, M. Akbari, N.B. Liabakk, G. Slupphaug, M. Otterlei, H.E. Krokan, Human AlkB homolog 1 is a mitochondrial protein that demethylates 3-methylcytosine in DNA and RNA, J. Biol. Chem. 283 (2008) 25046–25056.

[26] A. Monfort, A. Wutz, Breathing-in epigenetic change with vitamin C, EMBO Rep. 14 (2013) 337–346.

[27] D. Fu, J.A. Calvo, L.D. Samson, Balancing repair and tolerance of DNA damage caused by alkylating agents, Nat. Rev. Cancer 12 (2012) 104–120.

[28] L.H. Hurley, DNA and its associated processes as targets for cancer therapy, Nat. Rev. Cancer 2 (2002) 188–200.

[29] M.A. Kurowski, A.S. Bhagwat, G. Papaj, J.M. Bujnicki, Phylogenomic identification of five new human homologs of the DNA repair enzyme AlkB, BMC Genomics 4 (2003) 1471–2164.

[30] T.A. Müller, K. Meek, R.P. Hausinger, Human AlkB homologue 1 (ABH1) exhibits DNA lyase activity at abasic sites, DNA Repair 9 (2010) 58–65.

[31] T.A. Müller, M.M. Andrzejak, R.P. Hausinger, A covalent protein–DNA 5′-product adduct is generated following AP lyase activity of human ALKBH1 (AlkB homologue 1), Biochem. J. 452 (2013) 509–518.

[32] H. Korvald, P.O. Falnes, J.K. Laerdahl, M. Bjoras, I. Alseth, The *Schizosaccharomyces pombe* AlkB homolog Abh1 exhibits AP lyase activity but no demethylase activity, DNA Repair 11 (2012) 453–462.

[33] R. Ougland, D. Lando, I. Jonson, J.A. Dahl, M.N. Moen, L.M. Nordstrand, R. Torbjorn, T. Rognes, J.T. Lee, A. Klungland, T. Kouzarides, E. Larsen, ALKBH1 is a histone H2A dioxygenase involved in neural differentiation, Stem Cells 30 (12) (2012) 2672–2682.

[34] T.A. Müller, K. Yu, R.P. Hausinger, K. Meek, ALKBH1 is dispensable for abasic site cleavage during base excision repair and class switch recombination, PLoS One 8 (6) (2013) e67403.

[35] A.S. Vidyarthy, S.S. Das, Homology modeling and function prediction of hABH1, involving in repair of alkylation damaged DNA, Interdiscip. Sci. Comput. Life Sci. 3 (2011) 175–181.

[36] U. Pieper, U. Webb, D.T. Barkan, D. Schneidman-Duhovny, A. Schlessinger, H. Braberg, Z. Yang, E.C. Meng, E.F. Pettersen, C.C. Huang, R.S. Datta, P. Sampathkumar, M.S. Madhusudhan, K. Sjolander, T.E. Ferrin, S.K. Burley, A. Sali, MODBASE, a database of annotated comparative protein structure models and associated resources, Nucleic Acids Res. 39 (2001) 465–474.

[37] C. Notredame, D.G. Higgins, J. Heringa, T-Coffee: a novel method for fast and accurate multiple sequence alignment, J. Mol. Biol. 302 (1) (2000) 205–217.

[38] D. Chivian, D. Baker, Homology modeling using parametric alignment ensemble generation with consensus and energy-based model selection, Nucleic Acids Res. 34 (17) (2006) e112.

[39] K.W. Kaufmann, G.H. Lemmon, S.L. DeLuca, J.H. Sheehan, J. Meiler, Practically useful: what the ROSETTA protein modeling suite can do for you, Biochemistry 49 (2010) 2987–2998.

[40] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera—a visualization system for exploratory research and analysis, J. Comput. Chem. 25 (13) (2004) 1605–1612.

[41] D.A. Case, T.A. Darden, T.E. Cheatham III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Goetz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, P.A. Kollman, AMBER 12, University of California, San Francisco, 2012.

[42] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926–935.

[43] J. Cheng, H. Saigo, P. Baldi, Large-scale prediction of disulphide bridges using kernel methods: two-dimensional recursive neural networks, and weighted graph matching, Proteins 62 (3) (2006) 617–629.

[44] S.K. Singh, R. Hora, H. Belrhali, C.E. Chitnis, A. Sharma, Structural basis for Duffy recognition by the malaria parasite Duffy-binding-like domain, Nature 439 (2006) 741–744.

[45] M. Katzman, J.P. Mack, A.M. Skalka, J. Lies, A covalent complex between retroviral integrase and nicked substrate DNA, Proc. Natl. Acad. Sci. U.S.A. 88 (11) (1991) 4695–4699.

[46] D. Strahs, C. Zhu, B. Cheng, J. Chen, Y. Tse-Dinh, Experimental and computational investigations of Ser10 and Lys13 in the binding and cleavage of DNA substrates by *Escherichia coli* DNA topoisomerase I, Nucleic Acids Res. 34 (6) (2006) 1785–1797.

[47] J.C. Fromme, G.L. Verdine, Structure of a trapped endonuclease III-DNA covalent intermediate, EMBO J. 22 (13) (2003) 3461–3471.