Med. Chem. Res.17, 189-198.(2008)

NAD$^+$ DEPENDENT DNA LIGASE (RV3014C) FROM *M. TUBERCULOSIS*: STRATEGIES FOR INHIBITOR DESIGN

Divya Dube[1], Vandna Kukshal[1], Sandeep Kumar Srivastava[1], Rama Pati Tripathi[2] & Ravishankar Ramachandran[1]*

[1]Molecular & Structural Biology Division, [2]Medicinal & Process Chemistry Division,
Central Drug Research Institute, P.O. Box 173, Lucknow – 226001, India
Email: ravi_anitha@yahoo.com

Abstract: NAD$^+$ -dependent DNA ligases (LigA) are essential enzymes found only in bacteria and some virus species. This makes them attractive drug targets. Based on the crystal structure of the NAD$^+$ binding domain of the *M. tuberculosis* enzyme (*Mtu*LigA) and virtual screening we have earlier identified several novel classes of inhibitors for this enzyme. These inhibitors bind to the adenylation domain and compete with the co-factor NAD$^+$. Recently we identified that the BRCT-domain is essential for the enzyme activity of *Mtu*LigA. We used virtual screening to identify compounds from the CAP database that should potentially bind to the BRCT domain. These will now be evaluated as inhibitors of the enzyme with a novel mechanism of action. Challenges faced in designing specific and potent inhibitors of the enzyme which can distinguish between the human ATP-dependent ligase and *Mtu*LigA are additionally discussed in this report. Proposed strategies for the design of potent inhibitors with desired properties are also outlined.

Introduction

Tuberculosis is one of the worst disease scourges afflicting mankind. *Mycobacterium tuberculosis*, the aetiological agent of the disease, kills more than two million people every year (Duncan 1998). Moreover, multi drug resistant (MDR) varieties have been identified. These factors necessitate the identification of novel therapies based on different mechanisms of action (Duncan 1998). DNA ligases use either ATP or NAD$^+$ as co-factors to catalyze the joining of breaks in double stranded DNA (Lehman 1974; Engler & Richardson 1982). NAD$^+$ -dependent ligases are conserved enzymes in bacteria and have recently drawn attention as novel drug targets (Srivastava *et al.,* 2005a; 2005b). No existing drug is known to target these enzymes and conceptually therefore designed inhibitors should also be useful against drug resistant strains of the pathogen. We had recently solved the crystal structure of the adenylation domain of the enzyme (Srivastava *et al.,* 2005a) (**Fig. 1**) and have used it in virtual

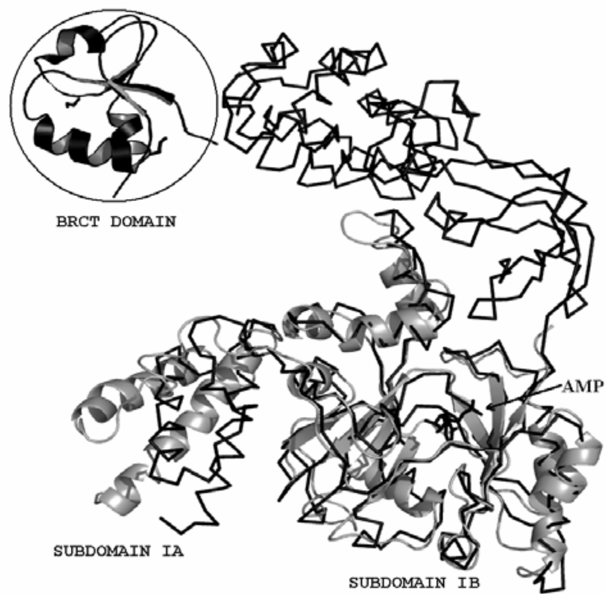


Figure 1. Crystal structure of the adenylation domain *of Mtu*LigA (cartoon representation) superposed onto crystal structure of full length *Tfi*LigA (ribbon representation). Modeled BRCT domain of *Mtu*LigA used in the reported virtual screening experiments is also depicted separately.

screening to identify novel inhibitory classes of compounds (Srivastava *et al.,* 2005a; 2005b & 2007). In this report we summarize the ongoing effort in identification of new inhibitors of the enzyme; the challenges faced and proposed strategies to improve both the potency and specificity of a designed inhibitor for *Mtu*LigA. Recently, using a specific LigA deficient *E. coli* strain we demonstrated (Srivastava *et al.,* 2007) that the BRCT domain of the enzyme is essential for bacterial viability. We undertook virtual screening using the modeled domain from *Mtu*LigA and have identified compounds with the potential to bind to this domain. These molecules will be evaluated in an ongoing program.

Results and Discussion

As part of a long range program to develop anti-TB therapies based on *Mtu*LigA inhibition, we have in the first instance searched for diverse compound families which inhibit *Mtu*LigA with several fold specificity compared to ATP-dependent ligases including for the human DNA ligase I (**Table 1**).

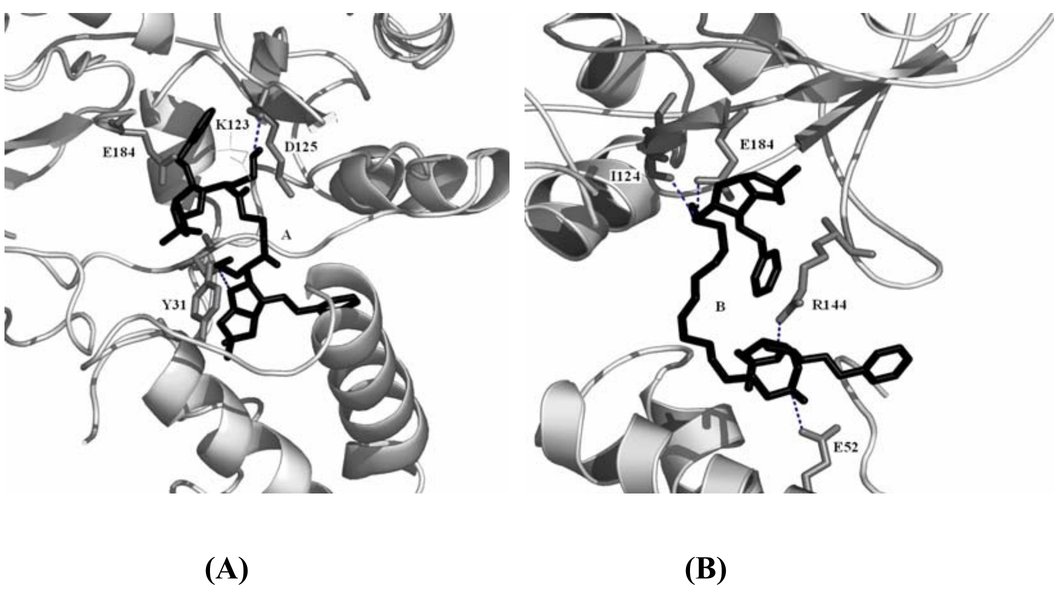
Table 1. MtuLigA inhibitors with corresponding IC_{50} values in µM. Representative structures from each compound class are depicted.

S.No	Class	Structure	IC_{50} (µM)
1.	Arylamino Compounds		46.0 ± 2.5
2.	Pyrido-chromanone	R_1=H,F R_2=H,F R_3=H,NH_2 R_4= -CN, -$CONH_2$	0.04-0.1
3.	Glycosyl ureides		9.65 ± 0.5 4.0 ± 0.3
4.	Glycosylamines		46.2 ± 1
5.	Tetracyclic indole derivatives		13.5 ± 0.6

The compound classes numbered 3-5 in the table were identified by our group. These compounds possess IC_{50} values in the low µM range. Bacterial growth inhibition studies using specific LigA deficient strains suggest that their observed antibacterial activity is most likely due to inhibition of the LigA in the bacteria (Srivastava *et al.,* 2005a; 2005b).

LigA-inhibitor interactions

No co-crystal structure with any of the compounds is known so far. However detailed kinetic studies coupled with mutations strongly suggest that all compound classes, except for aryl amino compounds, are competitive inhibitors and apparently bind to the NAD^+ site. Aryl amino compounds however are suggested to have a different mode of interaction with the enzyme. They also exhibit some unwanted DNA intercalation (Ciarrocchi et al., 1999). The highest affinity among the compound classes is exhibited by pyridochromanones (Srivastava *et al.,* 2007). The other classes of compounds exhibit IC_{50} values in the low micro molar range. Detailed molecular docking studies with Glycosyl uriedes and amines (Srivastava *et al.,* 2005a; 2005b). (**Fig. 2**) demonstrated that mimicking the interactions of NAD^+ with the enzyme improves the specificity/distinguishing ability of the compounds for LigA. These studies also suggested that the inhibitors should interact with conserved residues in the binding site.



(A) **(B)**

Fig. 2 Glycosyl amines: compound A can distinguish between *MtuLigA* and ATP-dependent ligase whereas compound B whose modeled interactions extend beyond the NAD^+ binding site cannot. Some interacting residues are indicated for clarity. Please see reference (Srivastava *et al.,* 2005) for detailed information.

BRCT domain and inhibitor development

LigA is a highly modular enzyme with the BRCT domain occurring at the C-terminus of the enzyme (**Fig. 1**). More work is necessary to actually elucidate the role of this domain in detail. Its deletion resulted in 3-fold reduction in activity of *E. coli* LigA (Wilkinson *et al.*, 2005) whereas the *T. filiformis* LigA retains activity even after its deletion (Jeon *et al.,* 2004). In contrast, no activity, both *in vitro* and *in vivo*, is observed in the case of the *M. tuberculosis* enzyme. Even in the enzyme from viral sources, the same situation exists, this domain is absent in the active enzymes from *A. Moorei* (Sriskanda *et al,* 2001) and *M. Sanguinipes* (Lu *et al.,* 2004), while the LigA from mimivirus possesses a BRCT domain, essential for its activity (Benarroch & Shuman 2006).

In the quest for identifying better LigA inhibitors, there are two issues at stake *viz.* the identified inhibitors should in the first instance be able to distinguish between NAD^+ and ATP -dependent ligases. In the second instance, the inhibitor should possess high affinity (specificity) for a particular pathogen/bacterial species in contrast to being a general anti-bacterial compound. The present crop of inhibitors mostly block the binding of the co-factor and thereby the enzyme activity. Residues in its binding site are well conserved (Doherty & Suh 2000). In fact, 5 out of 6 conserved sequence motifs form part of the site (**Fig. 3**).

```
1B04    EIKIDGLA -44- LEARGEAF -45- DLFVYGLADAEA -51- DGIVI -20- AIAYKFPAEEV
1TAE    EIKIDGLA -44- VEVRGECY -46- TFLYTV-ADFGP -52- DGIVI -20- AIAYKFPPEEA
1V9P    EHKVDGLS -45- LEVRGEVY -47- TFYALGLGLEES -52- DGVVV -20- ALAYKFPAEEK
1ZAU    EIKIDGVA -51- LEVRGEVF -46- MICHGL-GHVEG -51- DGVVV -20- AIAYKYPPEEA
1X9N    EYKYDGQR -42- FILDTEAV -32- CLYAFDLIYLNG -52- EGLMV -17- WLKLKKDYLDG
          ↑I             II              III             IV            V
```

Figure 3. Conserved Sequence motifs in DNA ligases. The adenylation domain contains five of six conserved sequence motifs in NAD^+ ligases. The alignment includes LigA encoded by *M. tuberculosis* (Pdb code: 1ZAU), *T. filiformis* (1V9P), *B. stearothermophilus* (1B04), *E. faecalis* (1TAE) and human DNA ligase I (1X9N). The numbers of amino acid residues separating the motifs are indicated. The active site lysine is indicated by an arrow.

Given the conserved nature of the co-factor binding site, expectedly most of the inhibitors exhibit some degree of general anti-bacterial activity too. Better inhibitor development is focused on improving specificity of the compounds for *Mtu*LigA. Two such approaches are discussed below; one approach deals with utilizing active site water

while the other approach involves the development of inhibitors, which can bind to other regions of the molecule, in this case the BRCT domain, and block subsequent catalytic steps.

Strategies in inhibitor development

Water clusters in the active site of an enzyme offer a lot of promise in this context. It is well recognized that inhibitors designed to mimic the interactions of displaced active site water can have improved affinities of up to 20 times (Ravishankar *et al.*, 1997; 1999). A superposition of the available LigA structures have led to identification of conserved water clusters in the NAD$^+$ site (**Fig. 4a**).

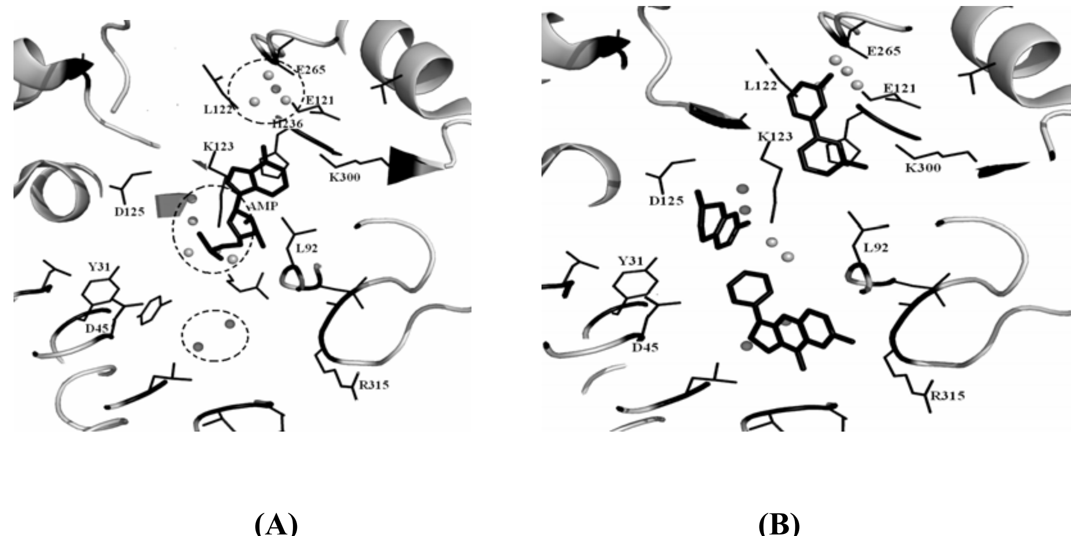


Figure 4. (A) Conserved water clusters in the active site of LigA. AMP is shown as sticks and the water clusters which are networked amongst them are indicated with dashed circles. Some interacting residues are shown for clarity. (B) Some fragments identified by virtual screening from a commercial library which are predicted to displace water clusters. These fragments will now be 'stitched' together and evaluated.

Virtual screening using a commercial small compound library, CAP database, (M/s Accelrys) identified several fragments predicted to displace and mimic the interactions of active site water. These will now be synthesized and tested.

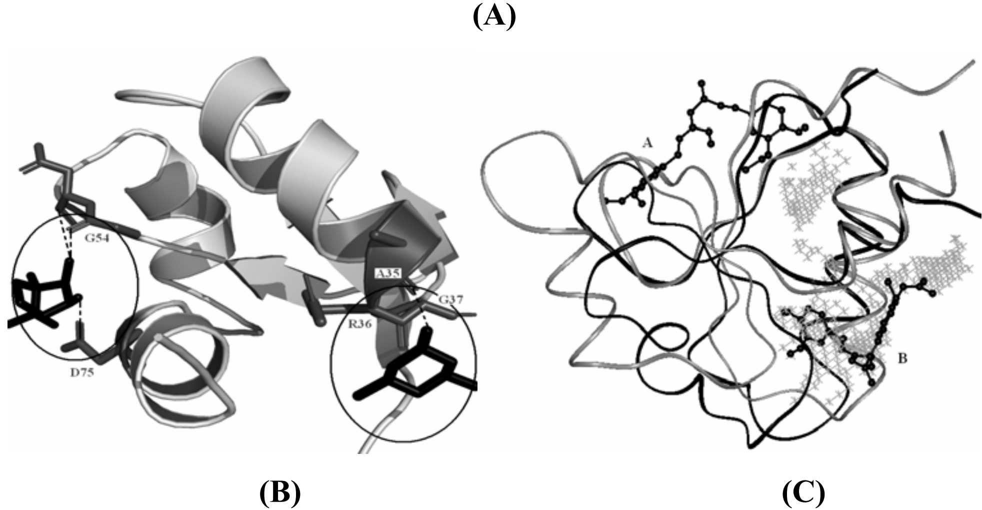
We had earlier identified that the BRCT domain of *Mtu*LigA is essential for activity (Srivastava *et al.,* 2007) in contrast to some other characterized LigA including those from *E. coli* and *T. filiformis*. Inhibitors designed to prevent BRCT domain

interactions with the adenylation site or DNA would conceptually be novel ones which block other steps of the enzyme's action. Previously, mutations in a homologous enzyme (Feng *et al*, 2004) had identified two glycine residues as being important for its activity. We accordingly modeled the *Mtu*LigA BRCT domain and used the structure in virtual screening experiments to identify potential binders. Compounds designed to bind to a pocket near the corresponding glycine residues in the model (**Fig. 5**) should be candidates for further testing by *in vitro* assays.

```
MtuBrct:  ERDESVPRTLAGLTIVVTGSLTGFSRDDAKEAIVARGGKAAGS
TthBrct:  ---------LKGLTFVITGELSR-PREEVKALLRRLGAKVTDS

MtuBrct:  VSKKTNYVVAGDSPGSKYDKAVELGVPILDEDGFRRLLADGPA
TthBrct:  VSRKTSYLVVGENPGSKLEKARALGVPTLTEEELYRLL-----
```

(A)



(B) **(C)**

Fig 5. (A) Sequence alignment of BRCT domains *of Mtu*LigA & *Thermus thermophilus* LigA (PDB code: 1L7B). Two conserved and essential glycine residues are circled and indicated.

(B) Cartoon representation of the modelled *Mtu*LigA BRCT domain. The essential glycine residues were used as the center-of-search in *in silico* screening calculations. Some potential inhibitors in both regions are also depicted.

(C) Potential binding sites on the modeled *Mtu*LigA BRCT domain (black ribbon) identified through 'Binding Site Analysis' (InsightII, *M/s Accelrys, San Diego, CA, USA*). The sites are indicated by grey crosses. The modeled BRCT domain of human pol λ (grey) is also shown to highlight possible differences in the structures of the BRCT domain. These differences are to be exploited in the design of species specific inhibitors. Docked monoacetylcurcumin is shown in ball-and-stick representation.

Indeed, such compounds could be identified from the CAP database and are being procured for testing. In another report (Takeuchi *et al.,* 2006) specific inhibitors of the human pol λ BRCT domain could be identified. Our modeling and docking simulations involving a curcumin derivative, monoacetylcurcumin (**Fig. 5c**) suggests that the same compound should bind to different regions of the BRCT domain in the respective proteins.

Conclusions

LigA are important enzymes with good potential for development as novel drug targets. The existing inhibitors, several of which were identified by our group, have shown that these are capable of distinguishing between the human and pathogen enzymes; this being an important step in specific inhibitor development. Given the conserved nature of the NAD^+ binding site amongst the enzyme from different bacterial species, these inhibitors also show general anti-bacterial activity. One approach to improve their specificity and affinity for an enzyme from a specific bacterial species is to utilize the spatial dispositions of active site water, where inhibitors designed to mimic the interactions of displaced water oxygen should be better than the first generation. We had also identified earlier that the BRCT domain of *Mtu*LigA is essential for enzyme activity and for bacterial viability. Compounds designed to bind to the domain to prevent the interactions of its key residues should have better specificity for a given pathogen because of the suggested finer variations in individual enzyme action. Towards this we have modeled the *Mtu*LigA BRCT domain and demonstrated that there are potential regions which can be exploited in novel inhibitor development. The modeling and docking results involving this domain also suggests structural differences which can also be exploited in rational inhibitor design.

Materials & Methods

Enzyme Assays:

All assays have been carried out using procedures previously reported by our group (Srivastava *et al.,* 2005a; 2005b; 2007).

Modeling of the BRCT domain:

Homology model for the BRCT domain was generated using MODELLER6v2 (Marti-Renom *et al.*, 2000) where *T. thermophillus* BRCT domain (PDB: 1L7B) was used as the template. The stereo-chemical quality of the model was verified using PROCHECK (Laskowski *et al.,* 1993) and WHAT IF (Vriend 1990). Prior to virtual screening experiments, all the hydrogens were added to the model. The model of human pol λ BRCT domain was also generated as reported by the group (Takeuchi *et al.,* 2006).

Databases used:

The commercially available Ludi/CAP database (M/s Accelrys Inc.) was used for all the virtual screening experiments.

Virtual screening protocol:

Ludi module interfaced with InsightII package (Accelrys Inc. 2000) was used to identify fragments and design ligands that would potentially overlap with the active site water. Ludi is based on the fragment approach whereby it saturates small fragments into the clefts of the target sites in such a way that hydrogen bonds can be formed with the enzyme and hydrophobic pockets are filled with hydrophobic groups. These fragments might be linked together by using Ludi in "Link mode".

Generating Novel Ligand de novo: The binding pockets within the modelled BRCT domain and adenylation domain were visually identified using InsightII (M/S Acclerys Inc.). The Ludi parameters were reviewed, modified and also the Ludi library was specified. The program was executed in the 'Targeted_Mode'. The 'Center of Search' parameter with a radius sphere of 5Å was specified based on two conserved glycine residues in the BRCT domain and also water clusters within the adenylation domain (Pdb: 1ZAU). The docked conformations of the identified 'Hits' were analyzed and Ludi scores tabulated within the modeling and simulation environment of InsightII. AUTODOCK v. 3.05 (Morris *et al.,* 1998) was also used in several instances for examining the poses of potential inhibitors with the enzyme.

Hardware used:

A computer cluster consisting of SGI ORIGIN350 servers and SGI OCTANES was used for the reported work.

Illustrations

Images were made using PyMol (http://www.pymol.org), Photoshop (Adobe Systems) and InsightII (M/s Accelrys Inc.).

Acknowledgements

The work was funded by *Council of Scientific and Industrial Research*, India grant CMM0017 and SMM0003. DD, VK & SKS acknowledge fellowships from CSIR. This is communication number 7135 from CDRI.

References

Benarroch D & Shuman S. (2006) Characterization of mimivirus NAD+-dependent DNA ligase. Virology 353: 135-43

Ciarrocchi G, MacPhee DG, Deady LW & Tilley L. (1999) Specific inhibition of the eubacterial DNA ligase by arylamino compounds. Antimicrob. Agents Chemother. 43: 2766-2772.

Doherty AJ & Suh SW. (2000) Structural and mechanistic conservation in DNA ligases. Nucleic Acids Res. 28: 4051–4058.

Duncan K. (1998) The impact of genomics on the search for novel tuberculosis drugs. Novartis Found. Symp. 217: 228-237.

Engler MJ & Richardson CC. (1982) The Enzymes (Boyer, P. D, ed) 15: 3–29, Academic Press, New York

Feng H, Parker JM, Lu J & Cao W. (2004) Effects of deletion and site-directed mutations on ligation steps of NAD^+-dependent DNA ligase: a biochemical analysis of BRCA1 C-terminal domain. Biochemistry 43: 12648- 12659.

Jeon HJ, Shin HJ, Choi JJ, Hoe HS, Kim HK, Suh SW & Kwon ST. (2004) Mutational analyses of the thermostable NAD+-dependent DNA ligase from Thermus filiformis. FEMS Microbiol. Lett. 237: 111-118.

Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R &Thornton JM.(1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J. Biomol. NMR. 8: 477-86.

Lehman IR. (1974) DNA ligase: structure, mechanism, and function. Science 186: 790-797.

Lu J, Tong J, Feng H, Huang J, Afonso CL, Rock DL, Barany F & Cao W. (2004) Unique ligation properties of eukaryotic NAD+-dependent DNA ligase from Melanoplus sanguinipes entomopoxvirus. Biochim. Biophys. Acta. 1701: 37-48

Marti-Renom MA, Stuart A, Fiser A, Sánchez R, Melo F & Sali A. (2000) Comparative protein structure modeling of genes and genomes. Anu. Rev. Biophys. Biomol. Struct. 29: 291-325

Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK & Olson AJ. (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J. Comput. Chem. 9: 1639–1662

Ravishankar R, Ravindran M, Suguna K, Surolia A & Vijayan M. (1997) Crystal structure of the peanut lectin-T-antigen complex. Carbohydrate specificity generated by water bridges. Current Science 72: 855-861. (also: Current Science, (1999) 76: 1393.

Ravishankar R, Suguna K, Surolia A & Vijayan M. (1999) The crystal structures of the complexes of peanut lectin with Methyl-β-galactose and N-acetyllactosamine and a comparative study of carbohydrate binding in Gal/GalNAc specific legume lectins. Acta Crystallogr. D 55: 1375-1382.

Sriskanda V, Moyer RW & Shuman S. (2001) NAD+-dependent DNA ligase encoded by a eukaryotic virus. J. Biol. Chem. 276: 36100-36109.

Srivastava SK, Tripathi RP & Ravishankar R. (2005) NAD^+ -dependent ligase (Rv3014c) from M. tuberculosis H37Rv: Crystal structure of the adenylation domain and identification of novel inhibitors. J. Biol. Chem. 280, 30273-30281.

Srivastava SK, Dube D, Tewari N, Dwivedi N, Tripathi RP. & Ravishankar R. (2005) Mycobacterium tuberculosis NAD^+ -dependent DNA ligase is selectively inhibited by glycosylamines compared with human DNA ligase I. Nucleic Acids Res. 33: 7090-7101.

Srivastava SK, Dube D, Vandna K, Jha AK, Hajela K & Ravishankar R. (2007) NAD+ -dependent DNA ligase (Rv3014c) from Mycobacterium tuberculosis: Novel structure-function relationship and identification of novel inhibitors. Proteins 69: 97-111.

Takeuchi T, Ishidoh T, Iijima H, Kuriyama I, Shimazaki N, Koiwai O, Kuramochi K, Kobayashi S, Sugawara F, Sakaguchi K, Yoshida H, Mizushina Y. (2006) Structural relationship of curcumin derivatives binding to the BRCT domain of human DNA polymerase lambda. *Genes Cells*. 11: 223-35.

Vriend G. (1990) WHAT IF: a molecular modeling and drug design program. J. Mol. Graph 8: 52-56.

Wilkinson A, Smith A, Bullard D, Lavesa-Curto M, Sayer H, Bonner A, Hemmings A & Bowater R. (2005) Analysis of ligation and DNA binding by Escherichia coli DNA ligase (LigA). Biochim. Biophys. Acta 1749: 113-122