Iraqi Journal of Biotechnology, May 2019, Vol. 18, No. 2, 40-48

# Clone, purify and 1D Nuclear Magnetic Resonance Spectroscopy of the BRCT Domain of *E. coli* DNA Ligase LigA

**Arqam Alomari**[1] , **Mark Allen**[2]

[1]Basic Sciences / Agriculture and Forestry College / University of Mosul / Iraq.
[2]Institution of Biological and Biomedical Sciences / School of Biological Science / University of Portsmouth / UK.

**Received:** February 27, 2019 / **Accepted:** April 7, 2019

**Abstract:** DNA ligases are essential enzymes in all domains of life. The role of these enzymes are to bind DNA ends, nicked DNA and join broken nucleic acid strands. The tertiary structure of *E. coli* DNA ligase LigA has four main domains: nucleotidyltransferase, oligomer-binding, Helix-hairpin-Helix and BRCT domain. One of the main objectives of my previous study was to explore the potential DNA ligase LigA as possible antibiotic targets by using a molecular docking programme called Molecular Operating Environment (MOE) (*in silico*) just in the first three domains of the tertiary structure of NAD+-dependent DNA ligase of *E. coli* LigA protein. Fortunately, it was found that four compounds out of the eight (5-Azacytidine, Geneticin, Chlorhexidine and Imidazolidinyl Urea) did inhibit the activity of *E. coli* LigA protein *in silico*, *in vitro* and then *in vivo* experiments after purify the native LigA protein. Importantly, the tertiary structure of this small BRCT domain has not been solves before. It does not appear in the crystal structure of 2OWO (PDB) that solved by Nandakumar *et al*., 2007. In this paper, the project was carried out to clone, express and purify the 88 amino acid of the forth domain (BRCT) and doing initial NMR experiment (1D NMR spectrum) to check the folding of this domain. It was found and determined that the BRCT domain of *E. coli* DNA ligase LigA is folded accurately, which is increased and supported the possibility as antibiotic target in all domains of *E. coli* LigA protein in the future.

**Keyword: DNA ligases, LigA of *E. coli*, BRCT domain.**

**Corresponding author:** (Email: arqam.alomari@gmail.com).

## Introduction:

Ligases are essential enzymes that repair breaks in DNA that occur as a result of damage and DNA replication or genetic recombination (1,2). Based on the competent cofactor, DNA ligases are classified as ATP or NAD+ dependent (3). They are essential tools for catalysing the formation of a phosphodiester bonds between a 5'-phosphorylated DNA termini and 3'-OH of DNA (4). **(**Figure 1) shows the crystal structure of $NAD^+$-dependent DNA LigA that has four main domains: the first one is a nucleotidyltransferase (NTase domain) (blue, 70-316 amino acid). The NTase domain has further domain called the Ia sub-domain (light blue, 1-69 amino acids). The second domain called an oligomer-binding (OB domain) (yellow, 317-404 amino acids). The OB domain also has further domain called a tetracystein zinc (Zn) (purple, 405-432 amino acids).

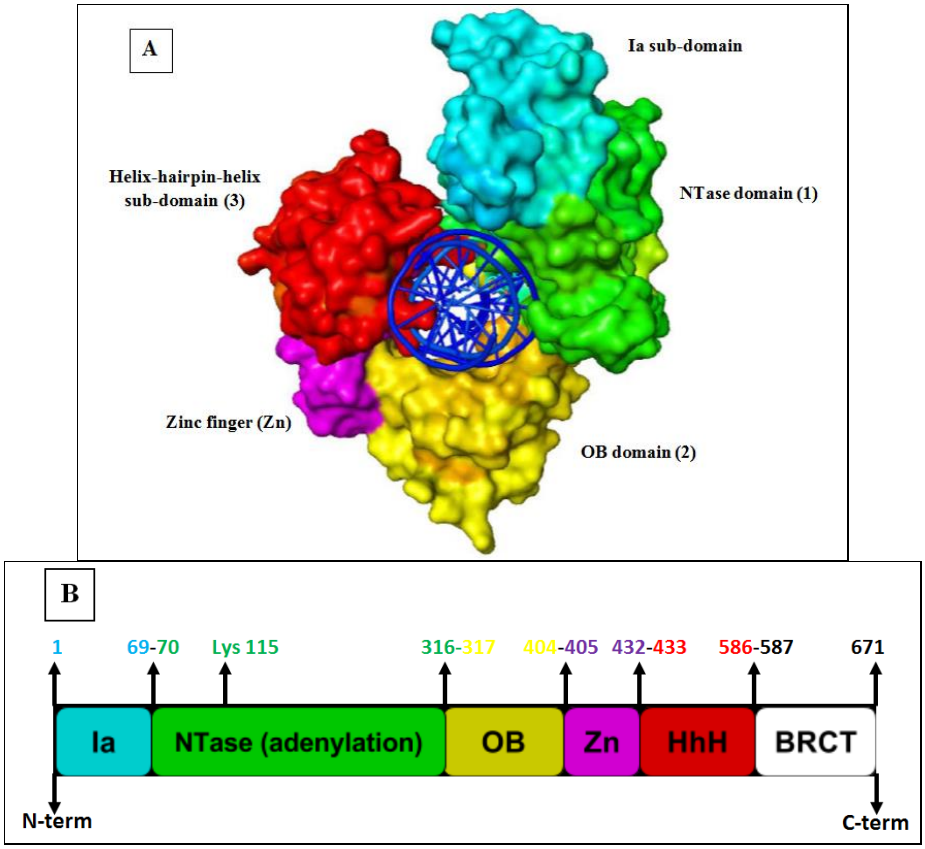


**Figure (1): The Crystal Structure of *E. coli* DNA LigaseA (PDB 1X9N- 2OWO). Panel A)** shows the four domains of the crystal structure of *E coli* DNA LigaseA with double strand of DNA in the centre by Pymol programme. The fourth domains are coloured as follow: The first domain I has two subdomains: Ia and NTase, which coloured blue and green colour, respectively. Second Domain OB (oligo-binding) with yellow colour. Domain 3 made up from two subdomains (Zn and HhH), which coloured with purple and red colour, respectively. Because of the high flexibility and small size of the BRCT domain of *E. coli* LigA, it becomes not shown in this structure when solved by X-ray. This Figure created manually in PyMOL using PDB code 2OWO. **Panel B)** shows the Domain layout of LigA. N-terminal residue (1) located at the Ia domain, the C-terminal residue is located in BRCT domain (671 residue) (18).

A third domain called Helix-hairpin-Helix (HhH domain) (red, 433-586 amino acid). The fourth domain is BRCT domain (587-671 amino acids) present on LigA (but not shown in the electron structure in Figure 1 above) (5,6). Ten individual residues that identified as critical for *E.coli* LigA function were presented in NTase domain: Lys115 (the site of covalent AMP attachment), Asp117, Gly118, Glu173, Arg200, Arg208, Arg277, Asp285, Lys290, and Lys314(7). NTase (adenylation) domain has a lysine residue that located in 115 of chain position and is so important for cofactor binding with $NAD^+$ and reacts with $NAD^+$ in eubacteria and then adenylated with an AMP group to form a covalent ligase-(lysyl-N)-AMP intermediate(8). Later, this AMP group is transferred to the DNA nick and eventually release NMN (9,10). The OB fold (second domain, 317-404 amino acids) is connected to the adenylation domain in the tertiary structure of LigA protein. This domain plays a crucial role for touching the un-nicked (template) strand of DNA and binding over the nick to help to form LigA. AppDNA

(LigA+ AMP+ phosphate on the DNA). The Ia domain is a subdomain of the first Domain. It is unique to $NAD^+$-dependent ligases is required for the reaction of LigA with $NAD^+$ to form the ligase-AMP intermediate(11). This domain is consider as essential for the reaction with the nicked DNA, and deleting of the Ia domain completely abolishes the ligation function for the protein(12,13). Alanine substations were tested within Ia domain of *E.coli* DNA ligase LigA and have five amino acid residues that are implicated specifically in adenylation and AMP transfer: $Y^2$, $H^{23}$, $D^{32}$, $Y^{35}$ and $D^{36}$, which reduce the activity LigA enzyme more than 80% (14). The third domain (HhH motif) called helix-hairpin-helix is extremely important and is to be wanted for the reaction and for non-specific DNA-binding (15). The scanning mutational analysis of the helix-hairpin-helix motif by using an alanine revealed that there are four glycine ($Gly^{455}$, $Gly^{489}$, $Gly^{521}$, and $Gly^{553}$), which are necessary constituents of the HhH domain and bind to the phosphate backbone across the minor groove at the outer margins of the LigA-DNA interface (16). The subdomain (Zn motif) belongs to the second and third domain has four cysteine residues and plays a crucial role in biding the OB and helix-hairpin-helix domains and for recognising the nick site of the DNA interaction (17,18). The forth domain in *E.coli* DNA ligase LigA is called BRCT domain and is found in all $NAD^+$ and some ATP-ligases. Importantly, the tertiary structure of this small BRCT domain has not been solves before. It does not appear in the crystal structure of 2OWO that solved by Nandakumar with his group in 2007 (18). It has a hydrophobic nature to the fold and it is about 80-100 amino acids in length, which is required for effective nick sealing *in vitro* (13).

## Material and Methods:

### BRCT Gene from DH5alpha Cells by PCR:

The complete coding sequence of the open reading frame of BRCT domain of *E.coli* Ligase A was obtained from the NCBI database. Forward and reverse PCR primer pairs (~40 bases) were designed to amplify the BRCT domain of *E.coli* Ligase A genes directly. Standard PCR reactions of 35 cycles of 98°C denaturing stage for 30 seconds, a 55°C annealing stage for 30 seconds, and a 72°C elongating step for 1 minutes were carried out using 1.0 µL Deep Vent polymerase (Thermofisher scientific). DH5α *E. coli* cells were used as the template DNA. The PCR reaction products generated were analysed by 2% agarose gel. The PCR product generated were purified using Qiagen PCR purification kit to obtain pure BRCT domain.

Forward oligo was: GGGGGATCCatcgttatcaacgcggaagagattgacagcc

Reverse oligo was: GGGGAATTCagctacccagcaaacgcagcatttccgcttgc

atcgttatcaacgcggaagagattgacagcc: Forward primer of BRCT domain.

agctacccagcaaacgcagcatttccgcttg: Revers primer of BRCT domain.

GGATCC: BamHI restriction enzyme site.

GAATTC: EcoRI restriction enzyme site.

GGG: Three nucleotides to hold the restriction enzymes in the both primer.

**DNA and Plasmid Restriction Digests:**

The pure BRCT domain DNA that produced by the PCR was digested with BamH1 and Eco1 (NEB) at 37°C for 1 hour. The sample of pure BRCT domain (about 1 µg) was digested with 2 µl of 20 units for each enzyme above using 1× restriction digest buffer (NEB) for each enzyme as well and adding free water until reach to the appropriate volume (50 µl) and then subsequently run on a 2% agarose gel. The expression vector plasmid DNA (pHLTV), which is supplied by my colleague Dr Mark/Portsmouth University/UK, was similarly digested and run on a 2% agarose gel as well as above.

**Gel Extraction and Ligation:**

The BamHI and EcoRI cut BRCT gene and the cut vector plasmid DNA (pHLTV) that has six Histidine were gel extracted from the agarose gel using a Zymoclean gel extraction kit (Zymo research). The cut BamHI and EcoRI BRCT gene was ligated with BamHI-EcoRI cut pHLTV. A ratio of 4:1 (insert: vector) was used (5 µL total), with 15 µL of Rapid Ligation Kit (Caltech). The pHLTV-BRCT gene ligation mixtures were incubated on ice with competent *E.coli* DH5α cells (NEB) or 30 minutes, heat-shocked at 42°C for 30 seconds to facilitate plasmid uptake by the cells, and incubated at 37°C for one hour to recover. They were then plated onto agar plates containing ampicillin (100 µg/mL) and incubated overnight at 37°C until colonies grew. (Figure 2) shows the brief summary about the pHLTV plasmid that used to ligate the BRCT gene with all the details.

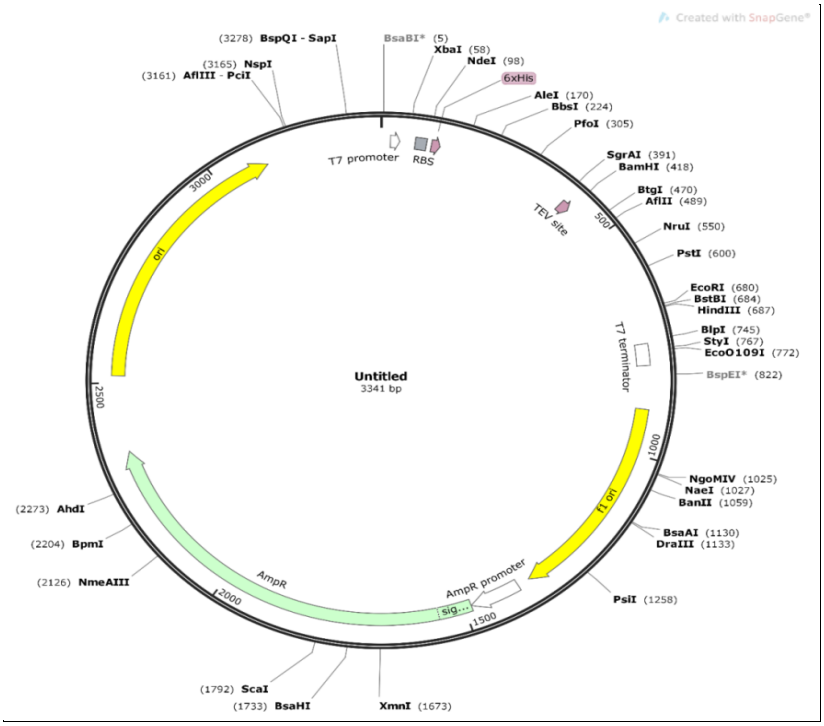


**Figure (2): The Brief Summary of pHLTV Plasmid. This figure is showing all the details of the circle plasmid of pHLTV that used to clone and express the BRCT domain in *E.coli* LigA, which has six Histidine amino acids (6-CAC).**

## Results:

### Colony PCR for BRCT Domain:

Six colonies obtained after ligation were amplified in 3 mL of 2YT media (16 g/L Tryptone, 10 g/L Yeast extracted and 5 g/L NaCl) supplemented with 100µg/ml ampicillin. The DNA from each was purified using a QIAGEN miniprep kit. To test for the presence of the BRCT DNA in the plasmid, a PCR was performed with the original oligos by using 1µL of the plasmid DNA obtained in place of the original DH5α cells. The PCR products obtained were analysed by 2.0% agarose gel to ascertain the presence of the correct insert. (Figure 3) is showing agarose gel (2%) of colony-PCR reaction, the second, third and fourth lanes show a PCR product with a size of ~270 bp, showing the BRCT domain of *E. coli* LigA gene was presented.

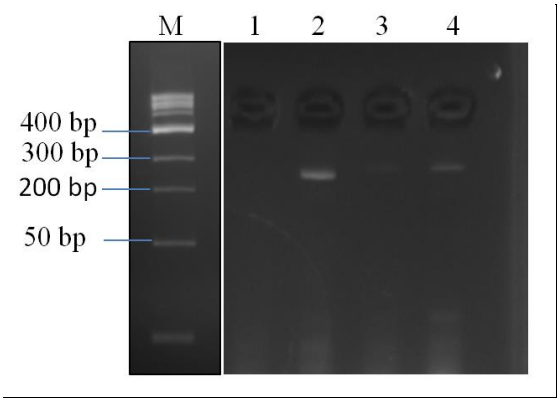


**Figure (3): Agarose Gel of Colony PCR for Native BRCT Domain of *E. coli* LigA. Lane M is DNA marker (NEB), showing sizes of DNA between 25 bp to 1000 bp. Lane 2-4 Showing agarose gel (2%) of colony-PCR reaction with a size of ~270 bp, showing the native BRCT domain of *E. coli* LigA gene.**

### Gene Sequencing:

A sample of each plasmid containing a BRCT insert was sent for DNA sequencing (Beckman-Coulter/ Genewiz). The sequencing primers was the T7 terminator primer used by Beckman-Coultier. Sequencing data were opened in SnapGene software and the DNA sequence checked manually. (Figure 4) shows an example of the successful DNA sequencing of the BRCT gene of *E.coli* DNA LigA.

### Plasmid Transformation:

Recombinant plasmids ~20-100 ng of pHLTV-BRCT gene was transferred into tubes containing 50 µl of competent cell C41 (DE3) (lucigen). The mixtures were incubated on ice for half an hour. They were transferred to a heat block at 42°C for exactly half a minute, and then left them on ice for 2 minutes. 2YT outgrowth medium (1 mL) was added to each tube of the mixture and they incubated at 37°C for 1 hour. ~50 µl of growth mixture was then spread onto a 2YT agar plate containing 50 µg/ml ampicillin and grown overnight at 37°C (agar plates not showing).

### Large Scale Induction of BRCT:

A large-scale protein expression of C41 (DE3) cells containing pHLTV-BRCT construct was conducted. A 10 mL aliquot of the overnight bacterial

culture of recombinant BRCT domain was added to 500 mL LB broth containing ampicillin 50 ug/ml and 4g glucose per litre has carbon 13 isotope was added to the media (only 1D NMR spectrum scanning), and the all the media put in an orbital shaker at 37 °C until an optical density (OD600) of 0.6.

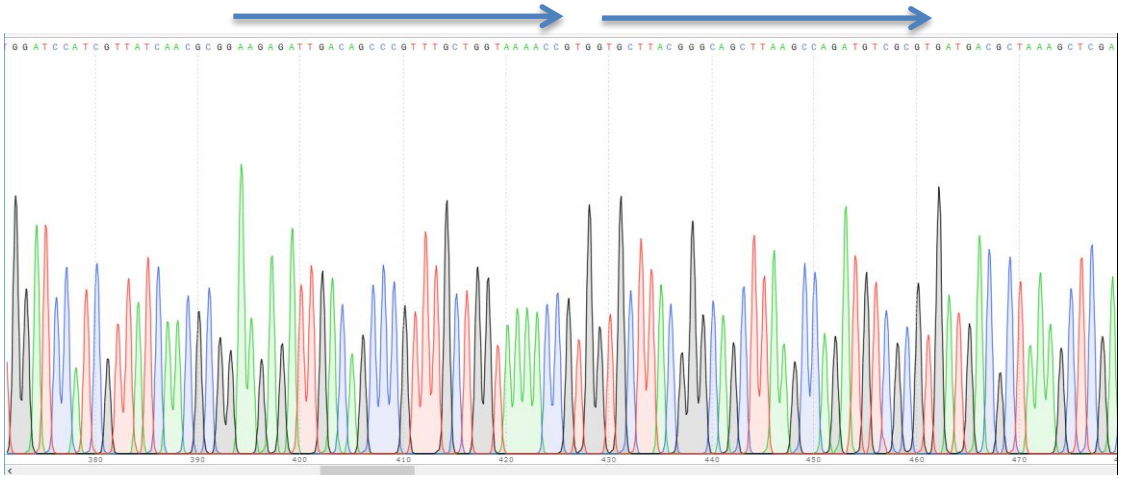

**Figure (4): Showing an Example of the Top Sequence of the Insert Fragment Including the Cloning Sites of the BRCT Domain of *E.coli* LigA as Follow:**
GGATCCATCGTTATCAACGCGGAAGAGATTGACAGCCCGTTTGCTGGTAAAACCGTGGTGCT
TACGGGCAGCTTAAGCCAGATGTCGCGTGATGACGCTAAAGCTCGACTGGTCGAACTGGGC
GCGACAGTCGCGAGCAGCGTGTCGAAGAAAACCGATCTGGTGATAGCGGGTGAAGCTGCAG
GATCTAAACTGGCGAAGGCGCAGGAACTGGGCATTGAAGTCATCGACGAAGCGGAAATGCT
GCGTTTGCTGGGTAGCTGAATTC. Blue arrows on the top of the figures explain the sequencing reading way.

Was reached (about 2 hours). Protein expression was induced with 1 mM isopropyl *B*-D-thiogalactoside (IPTG), and the culture incubated at 20ºC overnight until an $OD_{600}$ ~ 1.50-2.00 was reached. Bacterial cultures were harvested by centrifugation at 6000 rpm for 20 minutes at 4°C and the media discarded.

## Purification of Native BRCT Domain:

The cell pellet was re-suspended in 40 ml of Ni-NTA Buffer A (10 mM Tris, pH 8.0, 40mM NaCl, 10mM Imidazole), and disrupted by sonication for 3 minutes whilst cooled on ice, and the cell debris removed by centrifugation (20,000 rpm for 40 min). The supernatant was applied to a 50ml Ni-NTA column (Ni-Sepharose. Fast-flow – GE) and washed with 120 ml of Ni-NTA buffer to remove non-specifically bound proteins. The BRCT domain was eluted using Ni-NTA buffer B (10mM Tris, pH 8.0, 400mM NaCl, 250mM Imidazole), and the factions containing the BRCT domain (as judged by SDS-PAGE) were pooled. Lane 1 in (Figure 5) shows the expression of the BRCT domain of *E. coli* DNA ligase LigA that purified using 50ml Ni-NTA column (Ni-Sepharose Fast-flow–GE). TEV protease (supplied by Dr. Mark Allen (Portsmouth University)) was added in a ratio of 50:1 BRCT domain: TEV protease to cut and obtain the BRCT domain without 6-His (CAC), and the sample dialysed overnight using a 3.5kDa cut off membrane into Ni-NTA buffer A to remove the imidazole.

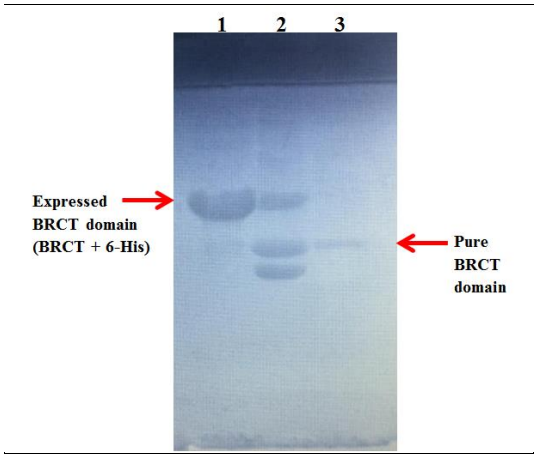


**Figure (5): Purification of Native BRCT Domain by Using Ni-NTA column (Ni-Sepharose Fast-flow–GE) as follow:**

- **Lane 1** shows the corresponding samples of the expression of BRCT domain (shown by blue arrows) were run on an 10% (w/v) SDS-PAGE gel that pooled by Ni-NTA column (Ni-Sepharose Fast-flow–GE). Lane 2 shows three samples: uncleaved BRCT domain (BRCT + 6-His), cleavage of BRCT domain and 6-His CAC. Lane 3 shows only pure and native BRCT domain (without 6-His CAC) of *E. coli* LigA that purified using Ni-NTA column, which used in NMR experiment.
- **Lane 2** in **Figure 5** shows some sample of uncleaved BRCT domain (BRCT + 6-His), cleavage of BRCT domain (pure BRCT, without 6-His) and 6-His CAC. The sample was reapplied to the Ni-NTA column and the material eluted through collected and fractions containing BRCT domain were collected (as judged by SDS-PAGE). Lane 3 in **Figure 5** shows just the pure BRCT domain of LigA.

The collected fractions of pure BRCT domain were concentrated and dialysed into the a buffer suitable for NMR scanning (20mM potassium phosphate, pH 6.5, 100 mM NaCl). (Figure 6) shows 1 D NMR spectrum of BRCT domain of *E. coli* DNA ligase LigA. The domain result in this figure appears that the structure of BRCT domain is folded depending on the location of the amides and methyl groups in the peak below. As we can see that the amides group are located between 10-11 ppm and the methyl groups are located between -1 to 1 ppm in the figure, which are identical to the standard form that confirm that the structure of the BRCT domain is folded correctly.

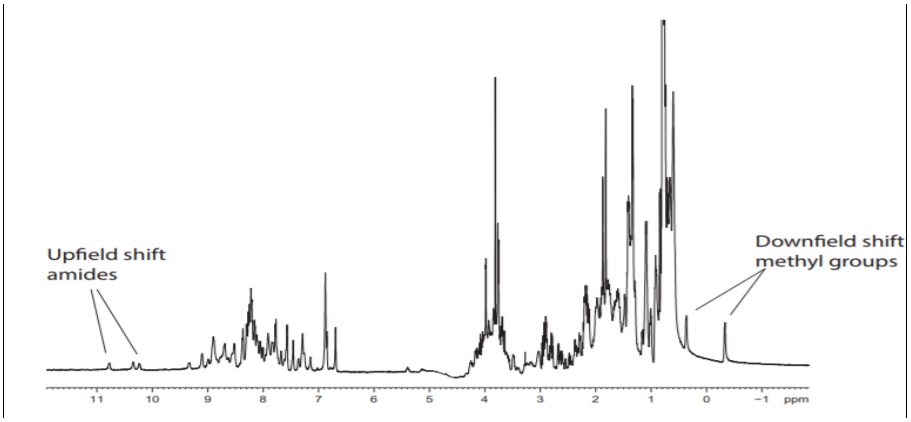


**Figure (6): Is showing 1D-1H NMR spectrum of the a 9-kDa BRCT domain of *E. coli* LigA ; The folded structure of the domain results in an expanded NMR spectrum with a number upfield shift amides and downfield shift methyl groups.**

## Discussion:

In 1967, Lehman with his group and other laboratories were first discovered *E.coli* DNA ligase LigA (19,20). As we know that *E.coli* DNA ligases are very necessary and important enzymes inside the cell into two pathways: replication and repair. BRCT domain is not present in all ligases, it is not looked at to be an essential compound to ligase activity (13). However, scanning mutational analysis confirmed that the deletion of BRCT domain of *E.coli* LigA revealed that is required in its totality for effective nick sealing *in vitro* (16). In 2005, Wilkinson and his group compared the ligation activity (*in vitro*) of the full-length protein of *E.coli* DNA ligase LigA, and they discovered that the deletion of BRCT domain from LigA reduced by 3-fold ligation activity of the protein and also reduced DNA binding by gel-shift experiment (21). In *E. coli* DNA ligase there is another protein for ligation called LigB protein, and this native protein does not has the BRCT domain in its structure. Sriskanda & Shuman 2001, stated that LigB is less active than LigA on their nicked DNA substrate, and they established that the slow activity of LigB is too weak to seal the Okazaki fragments during the replication (22). Alomari 2017, suggested that the lack of BRCT domain of the LigB protein that are present in LigA lead to back-up or support and help LigA protein to do its work or does not play a major role in the cell (23). Sriskanda *et al*., in 1999, found that $NAD^{+}$-dependent *E.coli* ligase LigA can support the growth and functionally substitute for the essential DNA ligase (ATP) in yeast *Saccharomyces cerevisiae called* Cdc9 ligase, Whereas LigB cannot do that in Cdc9 (24).

Furthermore, Feng with his group in 2004, they stated that the mutation of *Thermu*s species AK16D of LigA in BRCT domain (G671I mutant) showed a low ligation activity and lack of accumulation of the AMP-DNA intermediate (24). Therefore, the importance of the forth domain of *E. coli* ligase LigA (BRCT domain) led to clone, express and purify this domain and doing initial structural experiment as 1 D nuclear magnetic resonance spectroscopy in this paper, which is open the door in the future to solve whole structure of *E. coli* DNA ligase LigA by 2D or 3D NMR analysis and will be increased the possibility for LigA protein as new drug target. During the test of NMR by 1 D spectra in the result of this paper we can confirm and state that the forth domain is folded correctly and accurately.

## References:


1. Lehnman, I.R. (1974). DNA ligase: structure, mechanism, and function. *Science*, 186(4166): 790-797.
2. Mills, S.D.; Eakin, A.E.; Buurman, E.T.; Newman, J.V.; Gao, N.; Huynh, H., *et al*. (2011). Novel bacterial NAD+-dependent DNA ligase inhibitors with broad-spectrum activity and antibacterial efficacy *in vivo*. *Antimicrobial agents and chemotherapy*, 55(3): 1088-1096.
3. Srivastava, S.K.; Dube, D.; Tewari, N.; Dwivedi, N.; Tripathi, R.P. and Ramachandran, R. (2005). Mycobacterium tuberculosis NAD+-dependent DNA ligase is selectively inhibited by glycosylamines compared with human DNA ligase I. *Nucleic acids research*, 33(22): 7090-7101.
4. Lohman, G.J.; Chen, L. and Evans, T.C. (2011). Kinetic characterization of single strand break ligation in duplex DNA by T4 DNA ligase. *J. Biol. Chem*., 286: 44187-44196.

5. Wang, L.K.; Zhu, H. and Shuman, S. (2009). Structure-guided mutational analysis of the nucleotidyltransferase domain of Escherichia coli DNA ligase (LigA). *Journal of Biological Chemistry*, 284(13): 8486-8494.
6. Sriskanda, V. and Shuman, S. (2002a). Role of nucleotidyl transferase motif V in strand joining by chlorella virus DNA ligase. *J. Biol. Chem*., 277: 9661-9667.
7. Zhu, H. and Shuman, S. (2005). Structure-guided mutational analysis of the nucleotidyltransferase domain of Escherichia coli NAD+-dependent DNA ligase (LigA). *J. Biol. Chem*., 280: 12137-12144 175.
8. Singleton, M.R.; Hakansson, K.; Timson, D.J. and Wigley, D.B. (1999). Structure of the adenylation domain of an NAD+-dependent DNA ligase. *Structure*., 7: 35-42.
9. Zhao, A.; Gray, F.C. and MacNeill, S.A. (2006). ATP-and NAD+-dependent DNA ligases share an essential function in the halophilic archaeon Haloferax volcanii., *Molecular microbiology*, 59(3): 743-752.
10. Little, J.W.; Zimmerman, S.B.; Oshinsky, C.K. and Gellert, M. (1967). Enzymatic joining of DNA strands, II. An enzyme-adenylate intermediate in the dpn-dependent DNA ligase reaction. *Proceedings of the National Academy of Sciences*, 58(5): 2004-2011.
11. Benarroch, D. and Shuman, S. (2006). Characterization of mimivirus NAD+-dependent DNA ligase. *Virology*, 353(1): 133-143.
12. Jeon, H.J.; Shin, H.J.; Choi, J.J.; Hoe, H.S.; Kim, H.K.; Suh, S.W., *et al*. (2004). Mutational analyses of the thermostable NAD+-dependent DNA ligase from Thermus filiformis. *FEMS microbiology letters*, 237(1): 111-118.
13. Timson, D.J. and Wigley, D.B. (1999). Functional domains of an NAD+-dependent DNA ligase. *Journal of molecular biology*, 285(1): 73-83.
14. Sriskanda, V. and Shuman, S. (2002b). Conserved residues in domain Ia are required for the reaction of Escherichia coli DNA ligase with NAD+. *J. Biol. Chem*., 277: 9695-9700.
15. Doherty, A.J. and Suh, S.W. (2000). Structural and mechanistic conservation in DNA ligases. *Nucleic acids research*, 28(21): 4051-4058.
16. Lim, J.H.; Choi, J.; Kim, W.; Ahn, B.Y. and Han, Y.S. (2001). Mutational analyses of Aquifex pyrophilus DNA ligase define essential domains for self-adenylation and DNA binding activity. *Archives of biochemistry and biophysics*, 388(2): 253-260.
17. Lee, J.Y.; Chang, C.; Song, H.K.; Moon, J.; Yang, J.K.; Kim, H.K., *et al*. (2000). Crystal structure of NAD(+)-dependent DNA ligase: modular architecture and functional implications. *EMBO J*., 19: 1119-1129.
18. Nandakumar, J.; Nair, P.A. and Shuman, S. (2007). Last stop on the road to repair: structure of E. coli DNA ligase bound to nicked DNA-adenylate. *Molecular cell*, 26(2): 257-271.
19. Olivera, B.M.; Hall, Z.W.; Anraku, Y.; Chien, J.R. and Lehman, I. R. (1967). On the mechanism of the polynucleotide joining reaction. *Cold Spring Harb*. *Symp. Quant. Biol*., 33: 27-31.
20. Gellert, M. (1967). Formation of covalent circles of lamda DNA by E.coli extracts. *Proc. Nat. Acad. Sci*. 57, 148-155.
21. Wilkinson, A.; Smith, A.; Bullard, D.; Lavesa-Curto, M.; Sayer, H.; Bonner, A., *et al*. (2005). Analysis of ligation and DNA binding by Escherichia coli DNA ligase (LigA). *Biochim Biophys Acta*., 1749(1): 113-122.
22. Sriskanda, V. and Shuman, S. (2001). A second NAD+-dependent DNA ligase (LigB) in *Escherichia coli*. *Nucleic acids research*, 29(24): 4930-4934.
23. Alomari, A. (2017). Biophysical and Kinetic Analysis of Escherichia coli DNA Ligase Activity and Inhibition (Doctoral dissertation, University of Portsmouth/UK).
24. Sriskanda, V.; Schwer, B.; Ho, C.K. and Shuman, S. (1999). Mutational analysis of Escherichia coli DNA ligase identifies amino acids required for nick-ligation in vitro and for in vivo complementation of the growth of yeast cells deleted for CDC9 and LIG4. *Nucleic acids research*, 27(20): 3953-3963.
25. Feng, H.; Parker, J.M.; Lu, J. and 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(39): 12648-12659.