An Elsevier Indexed Journal ISSN-2230-7346

Journal of Global Trends in Pharmaceutical Sciences

HOMOLOGY MODELLING AND DOCKING STUDIES OF BIOACTIVE COMPOUNDS FROM MEDICINAL PLANTS

Udaya Kumar Nagapatla[1*], P. Jaya Chandra Reddy[2], V. Raghu Moorthy[3]

[1]Research Scholar, Department of Pharmacy, Jawaharlal Nehru Technological University, Kakinada, Kakinada, A.P, India.

[2]Krishna Teja Pharmacy College, Chadalawada Nagar, Renigunta Road, Tirupati, A.P, India.

[3]Govt Polytechnic, Obulapuram, Kadapa, Andhra Pradesh, India.

*Corresponding author E-mail:-Uday307@gmail.com

ARTICLE INFO

Key Words

Nigella hispanica, Docking, antidiabetic, antibacterial,anticanceractivity and BCL

ABSTRACT

Nigella hispanica leaves oil contains an abundance of conjugated linoleic (18:2) acid, thymoquinone, nigellone (dithymoquinone), melanthin, nigilline, damascenine, and tannins. In this study we identified the compounds in the leaves of Nigella hispanica and their anti oxident activity through insilico studies. Homology modeling and flexible docking of Bcl2L10 has been studied in*silico* approach. Blast result was found to have similarity with Bcl2L10 of 48% identity with 2KUA. With the aid of Molecular dynamics and Molecular simulations studies it was identified that the generated structure was reliable. Active site of Bcl2L10 was identified by CASTP. Large potential drugs were designed for identifying molecules that can likely bind to protein target of interest. This structure was used to identify better inhibitor using docking studies. The drug derivatives were docked to the Bcl2L10 structure into the active site. The compounds were designed using chemsketch software and structure of the Cancer protein BCL-2 was collected from PDB databank. The active site was identified using Castp server. The results indicated that ASP-65, CYS-66, CYS-88 in BCL-2 were important determinant residues in binding as they have strong hydrogen bonding with compounds. These compunds were isolated using column chromatography based on the rate of separation. Then these extracts were used for antibacterial, antifungal, antioxidant, anti diabeticand anticancer activity

INTRODUCTION

Homology modeling, also known as comparative modeling refers to constructing an atomic-resolution model of the "*target*" protein from its amino acid sequence and an experimental three-dimensional structure of a related homologous protein (the "*template*"). Homology modeling relies on the identification of one or more known protein structures likely to resemble the structure of the query sequence, and on the production of an alignment that maps residues in the query sequence to residues in the template sequence. The sequence alignment and template structure are then used to produce a structural model of the target. Because protein structures are more conserved than DNA sequences, detectable levels of sequence similarity usually imply significant structural similarity. Suppose you want to know the 3D structure of a target protein that has not been solved empirically by X-ray crystallography or NMR. You have only the sequence. If an empirically determined 3D structure is available for a sufficiently similar protein (50% or better sequence identity would

be good), you can use software that arranges the backbone of your sequence identically to this template. This is called "homology modeling". It is, at best, moderately accurate for the positions of alpha carbons in the 3D structure, in regions where the sequence identity is high. It is inaccurate for the details of side chain positions, and for inserted loops with no matching sequence in the solved structure. The optimization procedure is a variable target function method that applies the conjugate gradients algorithm to positions of all non-hydrogen atoms. The query sequence fromHomo sapiens was submitted to domain fishing server Signal transducer and activator of transcription 4 prediction. The predicted domain was searched to find out the related protein structure to be used as a template by the BLAST (Basic Local Alignment Search Tool) program against PDB (Protein Databank). Sequence that showed maximum identity with high score and less e-value was aligned and Bile salt export pump was used as a reference structure to build a 3D model for. The sequence of Bile salt export pump was obtained from UNIPROT. The co-ordinates for the structurally conserved regions (SCRs) for Signal transducer and activator of transcription 4 were assigned from the template using multiple sequence alignment, based on the Needleman-Wunsch algorithm. The structure having the least modeller objective function, obtained from the modeller was improved by molecular dynamics and equilibration methods using NAMD 2.5 software using CHARMM27 force field for lipids and proteins along with the TIP3Pmodel for water. The energy of the structure was minimized with 1, 00, 00 steps. A cutoff of 12 Å (switching function starting at 10 Å)for van der Waals interactions was assumed. No periodic boundary conditions were included in this study. An integration time step of 2 fs was used, permitting a multipletime-stepping algorithm to be employed in which interactions involvingcovalent bonds were computed every time step, short-range non bondedinteractions were computed every two time steps and long-rangeelectrostatic forces were computed every four time steps. The pair list of the non bonded interaction was recalculated every ten time steps with a pair list distance of 13.5 Å. The short-range non bonded interactions were defined as van der Waals and electrostatics interactions between particles within 12 Å. A smoothing function was employed for the van der Waals interactions at a distance of 10 Å. CHARMM27 [force-field parameters were used in all simulations in this study. The equilibrated system was simulated for 1 ps with a 500 kcal/mol/Å2 restraint on the protein backbone under 1 atm constant pressure and 310 K constant temperature (NPT) and the Langevin damping coefficientwas set to 5 ps unless otherwise stated. Finally, the structure having the least energy with low RMSD (Root Mean Square Deviation) was used for further studies. In this step, the quality of the initial model was improved. The final structure obtained was analyzed by Ramachandran's map using PROCHECK (Programs to check the Stereo chemical Quality of Protein Structures) and environment profile using ERRAT graph (Structure Evaluation server). This model was used for the identification of active site and for docking of the substrate with the enzyme.

Docking Studies: The sequence of BCL-2 was obtained from UNIPROT. The structure was collected from Protein Dara Bank. After that the unnecessary chains and hetero atoms were removed using SPDBV software. Later hydrogens were added to the protein and used for active site identification. Finally, the structure having the least energy with low RMSD (Root Mean Square Deviation) was used for further studies. In this step, the quality of the initial model was improved. The final structure obtained was analyzed by Ramachandran's map using PROCHECK (Programs to check the Stereo chemical Quality of Protein Structures) and environment profile using ERRAT graph (Structure Evaluation server). This model was used for the identification of active site and for docking of the substrate with the enzyme.

Active site Identification

Active site of BCL-2 was identified using CASTp server. A new program, CASTp, for automatically locating and measuring protein pockets and cavities, is based on precise computational geometry methods, including alpha shape and discrete flow theory. CASTp identifies and measures pockets and pocket mouth openings, as well as cavities. The program specifies the atoms lining pockets, pocket openings, and buried cavities; the volume and area of pockets and

cavities; and the area and circumference of mouth openings.

Docking method: Docking was carried out using GOLD (Genetic Optimization of Ligand Docking) software which is based on genetic algorithm (GA). This method allows as partial flexibility of protein and full flexibility of ligand. The compounds are docked to the active site of the BCL-2. The interaction of these compounds with the active site residues are thoroughly studied using molecular mechanics calculations. The parameters used for GA were population size (100), selection pressure (1.1), number of operations (10,000), number of island (1) and niche size (2). Operator parameters for crossover, mutation and migration were set to 100, 100 and 10 respectively. Default cutoff values of 3.0 A° (dH-X) for hydrogen bonds and 6.0 A° for vanderwaals were employed. During docking, the default algorithm speed was selected and the ligand binding site in the BCL-2 was defined within a 10 A° radius with the centroid as CE atom of ALA220. The number of poses for each inhibitor was set 100, and early termination was allowed if the top three bound conformations of a ligand were within 1.5A° RMSD. After docking, the individual binding poses of each ligand were observed and their interactions with the protein were studied. The best and most energetically favorable conformation of each ligand was selected.

Gold Score fitness function:

Gold Score performs a force field based scoring function and is made up of four components: 1. Protein-ligand hydrogen bond energy (external H-bond); 2. Protein-ligand vander Waals energy (external vdw); 3.Ligand internal vander Waals energy (internal vdw); 4.Ligand intramolecular hydrogen bond energy (internal- H- bond). The external vdw score is multiplied by a factor of 1.375 when the total fitness score is computed. This is an empirical correction to encourage protein-ligand hydrophobic contact. The fitness function has been optimized for the prediction of ligand binding positions.

GoldScore = S (hb_ext) + S (vdw_ext) + S (hb_int) + S (vdw_int)

Where S (hb_ext) is the protein-ligand hydrogen bond score, S (vdw_ext) is the protein-ligand van der Waals score, S (hb_int) is the score from intramolecular hydrogen bond in the ligand and S (vdw_int) is the score from intramolecular strain in the ligand.

RESULTS AND DISCUSSION

Homology Modeling of Bcl2L10Protein

A high level of sequence identity should guarantee more accurate alignment between the target sequence and template structure. In the results of BLAST search against PDB, 2KUA (chain A) has a high level of sequence identity and the identity of the reference protein with the BCL2L10 are 48%. Structurally conserved regions (SCRs) for the model and the template were determined by superimposition of the two structures and multiple sequence alignment.

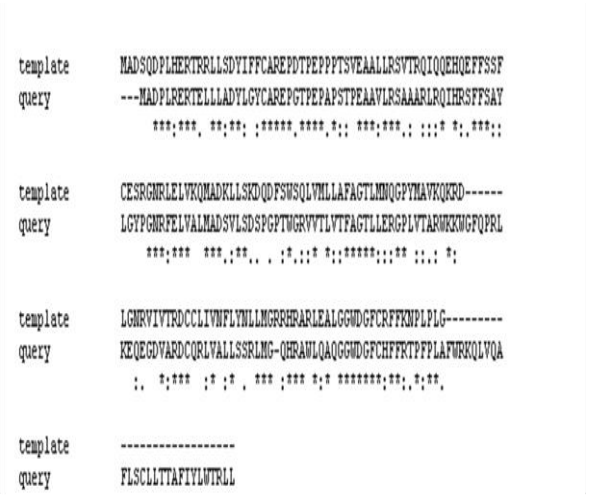


Figure 12: Clustal multiple sequence alignment

In the following study, we have chosen 2KUA as a reference structure for modeling domain. Coordinates from the reference protein 2KUA to the SCRs, structurally variable regions (SVRs), N-termini and C-termini were assigned to the target sequence based on the satisfaction of spatial restraints. In the modeller we will get a 20 PDB out of which we select a least energy. The energy unit will be in kilo joule. All side chains of the model protein were set by rotamers. The final stable structure of the BCL2L10 protein obtained is shown in Figure 2. By the help of SPDBV it is evident that Bile salt export pump domain has 8 helices. The final structure was further checked by verify3D graph and the results have been shown in Figure 12: The overall scores indicates acceptable protein environment

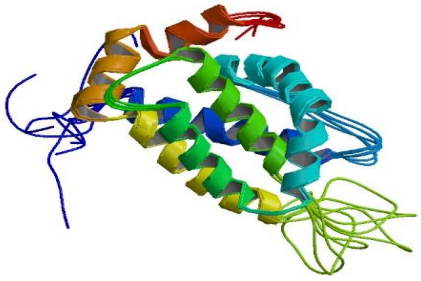


Figure 11: Structure of the BCL2L10 protein

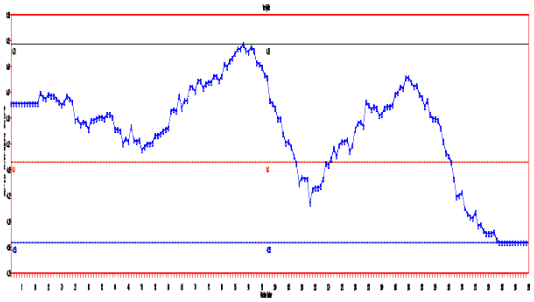


Figure 12: Graphical representation for verify 3D

Validation of BCL2L10

After the refinement process, validation of the model was carried outusing Ramachandran plot calculations computed with the PROCHECKprogram. The distributions of the Ramachandranplots of non-glycine, non-proline residues are summarized inTable 1. The RMSD (Root Mean Square deviation) deviation for covalent bonds and covalent angles relative tothe standard dictionary of BCL2L10 was -5.27 and -0.55 Å. Altogether 99.4 % of the residues of BCL2L10 was infavored and allowed regions. The overall PROCHECK G-factor of BCL2L10 was – 2.32 and verify3D environment profile was good.

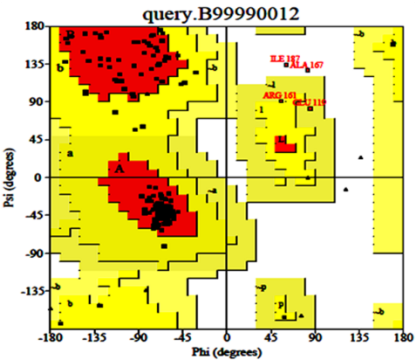


Figure 14: Ramachandran Plot using RAMPAGE server

Table 1: % of residue falling in the core region of the Ramachandran's plot

% of residue in most favored regions	87.3
% of residue in the additionally allowed zones	10.3
% of residue in the generously regions	1.2
% of residue in disallowed regions	1.2
% of non-glycine and non-proline residues	100.0

Superimposition of 2KUA with BCL2L10 domain: The structural superimposition of 2KUA template and BCL2L10 is shown in Figure 14. The weighted root mean square deviation of trace between the template and final refined models 0.82A^{o}. This final refined model was used for the identification of active site and for docking of the substrate with the domain BCL2L10.

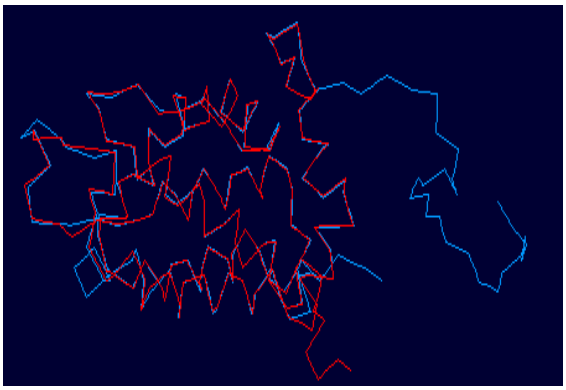


Figure 18: Compare the structure of BCL2L10 (represented in Red color) with 2KUA (represented with Blue color)

Active site Identification of BCL2L10:

After the final model was built, the possible binding sites of BCL2L10 was searched based on the structural comparison of template and the model build and also with CASTP server and was shown in Figure 19. Since, BCL2L10 from Humanand the 2KUA are well conserved in both sequence and structure; their biological function should be identical. Infact from the structure-structure comparison of template, final refined model of BCL2L10 domain using SPDBV program. It was found that secondary structures are highly conserved and the residues, PHE5, LYS7, CYS9,

TYR10GLY11,TYR12, CYS13, ILE14, PHE28, LEU30, TYR31, ALA61, VAL62, THR63, PHE82, VAL230, LEU281.

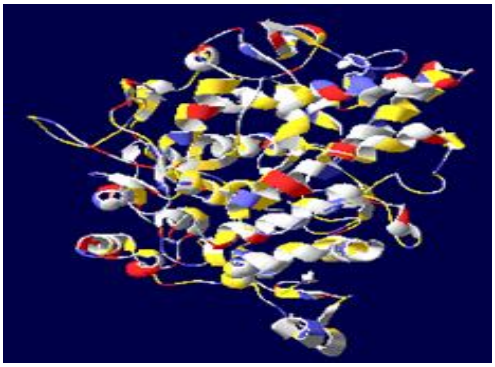
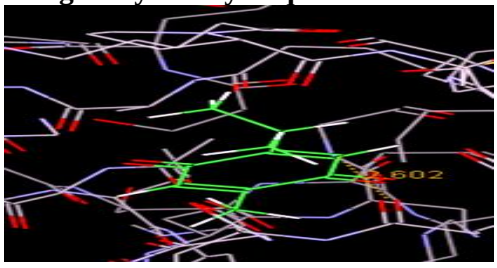


Fig 1: structure of BCL-2

Docking of inhibitors with the active site of BCL2:

Docking of the inhibitors with BCL-2 was performed using GOLD 3.0.1, which is based on genetic algorithimn. This program generates an ensemble of different rigid body orientations (poses) for each compound conformer within the binding pocket and then passes each molecule against a negative image of the binding site. Poses clashing with this 'bump map' are eliminated. Poses surviving the bump test are then scored and ranked with a Gaussian shape function. We defined the binding pocket using the ligand-free protein structure and a box enclosing the binding site. This box was defined by extending the size of a cocrystalized ligand by 4A. This dimension was considered here appropriate to allow, for instance, compounds larger than the cocrystallized ones to fit into the binding site. One unique pose for each of the best-scored compounds was saved for the subsequent steps. The compounds used for docking was converted in 3D with SILVER. To this set, the substrates corresponding to the modelled protein were added. Docking of best inhibitor with the active site of protein showed the activity of the molecule on protein function (Fig 4 and table 2).

Docking study of Thymoquinone



Fitness	S(hb_ext)	S(vdw_ext)	S(hb_int)	S(int)	Ligand name
27.77	0.00	24.24	0.00	-5.56	thymoquinone

CONCLUSION

BCL-2 is one of the regulator protein , which specifically binds with the cell division proteins. In this work, we have collected 3D model of BCL-2 domain, from human using the SPDBV software and obtained a refined model after energy minimization. The final refined model was further assessed by ERRAT & PROCHECK program, and the results show that this model is reliable. The stable structure is further used for docking of substrate with the compounds of Nigella hispanica. Docking results indicate that conserved amino-acid residues in BCL-2 main play an important role in maintaining a functional conformation and are directly involved in donor substrate binding. The interaction between the domain and the inhibitors proposed in this study are useful for understanding the potential mechanism of domain and the inhibitor binding. As is well known, hydrogen bonds play important role for the structure and function of biological molecules. In this study it was found that ASP-61, CYS-66, CYS-88, TYR-112, LEU-113, LEU-115, ASP-117, ILE-118, GLN-136 of BRCA1 are important for strong hydrogen bonding interaction with the inhibitors. To the best of our knowledge ASP-65, CYS-66, CYS-88, TYR-112, LEU-113 are conserved in this domain and may be important for structural integrity or maintaining the hydrophobicity of the inhibitor-binding pocket. Compounds from Nigella hispanica leaves were extracted through column chromatography. The extracts of Nigella hispanica leavesproved excellent antimicrobial anti diabetic and anti cancer activity.

REFFERENCES

1. Zohary, Daniel; Hopf, Maria (2000). Domestication of plants in the Old World (3 ed.). Oxford University Press. p. 206. ISBN 0-19-850356-3.
2. Hammond, Edward (2012). "Food giant Nestlé claims to have invented stomach soothing use of habbat al-barakah (Nigella sativa)". Briefing

Paper. Third World Network. Retrieved 23 April 2013.

3. Mohammad HosseinBoskabady, BatoolShirmohammadi (2002). "Effect of Nigella Sativa on Isolated Guinea Pig Trachea". Arch Iranian Med 5 (2): 103–107.
4. Banni S (June 2002). "Conjugated linoleic acid metabolism". Current Opinion in Lipidology 13 (3): 261–6. doi:10.1097/00041433-200206000-00005. PMID 1204539.
5. Talbott SM, Hughes K (2007). "Conjugated linoleic acid". The Health Professional's Guide to Dietary Supplements (Lippincott Williams & Wilkins). pp. 14–. ISBN 978-0-7817-4672-4.
6. Rainer L, Heiss CJ (June 2004). "Conjugated linoleic acid: health implications and effects on body composition". J Am Diet Assoc (Review) 104 (6): 963–. doi:10.1016/j.jada.2004.03.016. PMID 15175596.
7. Onakpoya IJ, Posadzki PP, Watson LK, Davies LA, Ernst E (March 2012). "The efficacy of long-term conjugated linoleic acid (CLA) supplementation on body composition in overweight and obese individuals: a systematic review and meta-analysis of randomized clinical trials". Eur J Nutr (Systematic review) 51 (2): 127–34. doi:10.1007/s00394-011-0253-9. PMID 21990002.
8. Ha YL, Grimm NK, Pariza MW (1987). "Anticarcinogens from fried ground beef: heat-altered derivatives of linoleic acid". Carcinogenesis 8 (12): 1881–7. doi:10.1093/carcin/8.12.1881. PMID 3119246.
9. Pariza MW (June 2004). "Perspective on the safety and effectiveness of conjugated linoleic acid". Am. J. Clin. Nutr. 79 (6 Suppl): 1132S–1136S. PMID 15159246.
10. "Fatty Acid Profiles of Liver, Adipose Tissue, Speen, and Heart of Mice Fed Diets Containing T10, C-12-, and C9, T11-Conjugated Linoleic Adic".
11. Tricon S, Burdge GC, Kew S et al. (September 2004). "Opposing effects of cis-9,trans-11 and trans-10,cis-12 conjugated linoleic acid on blood lipids in most healthy humans". Am. J. Clin. Nutr. 80 (3): 614–20. PMID 15321800.
12. Ulf Risérus, MMed; Samar Basu, PhD; Stefan Jovinge, MD, PhD; GunillaNordinFredrikson, PhD; Johan Ärnlöv, MD; BengtVessby, MD, PhD (September 2002). "Supplementation With Conjugated Linoleic Acid Causes Isomer-Dependent Oxidative Stress and Elevated C-Reactive Protein". American Heart Association Journals 106 (15): 1925–9. doi:10.1161/01.CIR.0000033589.15413.48. PMID 12370214. 01.CIR.0000033589.15413.48v1. Retrieved 2007-02-19
13. II International Congress on CLA from Experimental Models to Human Application
14. Kuhnt K, Kraft J, Moeckel P, Jahreis G (April 2006). "Trans-11-18 : 1 is effectively Delta9-desaturated compared with trans-12-18 : 1 in humans". Br J Nutr. 95 (4): 752–761. doi:10.1079/BJN20051680. PMID 16571155.
15. Banni S, Angioni E, Murru E, Carta G, Melis M, Bauman D, Dong Y, Ip C (2001). "Vaccenic acid feeding increases tissue levels of conjugated linoleic acid and suppresses development of premalignant lesions in rat mammary gland". Nutr Cancer 41 (1–2): 91–7.