International Journal of Research Studies in Biosciences (IJRSB)
Volume 2, Issue 11, December 2014, PP 75-84
ISSN 2349-0357 (Print) & ISSN 2349-0365 (Online)
www.arcjournals.org

In *Silico* Drug Search for Better Treatment for Cancer: L-Asparaginase

Mundaganur.Y. D.

Department of Zoology
Miraj Mahavidyalaya, Miraj

Mundaganur .D.S

Department of Zoology
Willindon College, Sangli

Ashokan Kannarath

Department of Zoology P.V.P. College, K.Mahankal,
Sangli, Maharashtra, India
akvsangli@gmail.com

Abstract: *L-Asparaginase is required by normal cells for the utilization of the amino acid arginine. When the normal cell turns to cancerous, it cannot synthesis arginine, but it can extract the amino acid from external supply in the plasma and tissue. In such instance if we supply the L-asparaginase as drug it will destroy the supplied arginine and the cancer cell will die due to starvation. The source of L-asparaginase is different. It may be fungi, bacteria, plants and animals. Here we made an attempt to study the possible better source of L-asparaginase for the treatment of cancer by docking method. The docking study reveals that L-asparaginase is better as drug from 12 different sources we have studied. Erwinia carotovora's L-asparaginase shows -62. 83 energy this is least from all the selected 12 sources of organisms. This followed by Campylobacter jejuni an all other sources shown more or less same energy level higher that these two sources. The other parameters studied like Lipinski data and other drug parameters show that asparaginase from Erwinia carotovora is better than other sources in the sense of drug candidates. We also studied physico-chemical parameters, trans-membrane sequence prediction, secondary structure prediction of the sequence extracted from the source, 3D model preparation and its validation. The conclusion drawn from the study should be validated by further study and validation in animal models and human.*

Keywords: *L-asparaginase, Protein sequence, Docking, Lipinski, 3D model.*

1. INTRODUCTION

L-Asparaginase is required by normal cells for the utilization of the amino acid arginine. When the normal cell turns to cancerous, it cannot synthesis arginine, but it can extract the amino acid from external supply in the plasma and tissue. In such instance if we supply the L-asparaginase as drug it will destroy the supplied arginine and the cancer cell will die due to starvation. L-asparaginase catalyzes the hydrolysis of L-asparagine into L-aspartate and ammonia (Fig1).

The toxicity of L-asparaginase to lymphocytes has been reported (Schrek, *et al* 1967; Vita and Wataru, 2006). It is observed that guinea pig serum has anti tumor property to some tumor (Kidd, 1953). It was subsequently substantiated by many workers in the field of cancer research (Broom 1953; De Lowery, 1966; Hill, 1967).Target is the molecular or cellular structure naturally existing and involved in the pathology to which the drug is to be act. Today 435 human genome products have been identified as drug target approved by FDA (Rask-Andersen et al, 2011). The drug likeness can be assessed by ligand efficiency (LE) (Hopkins et al, 2004) and lipophilic efficiency (LiPE) (Ryckmans et al, 2009; Leeson and Springthorpe, 2007). Traditionally many drug that synthesized by one organism (Microbes) and affect on other organism (Human) is discovered (Roger, et al, 2006). Even today the intrusion of combinatorial chemistry did not form a hurdle for the natural material as a starting material for the drug development (Feher and Schmidt, 2003; Newman and Cragg, 2007). The target molecule, the ligand used here is Aspragine.L-

Asparaginase obtained from various sources like bacteria, fungi, plants and animals. L-asparaginase as drug for cancer treatment is usually obtained from fungal source like ***Sclerocium rolfsii*** Sacc. Here an attempt is made to find out other source which may be better than the present one.

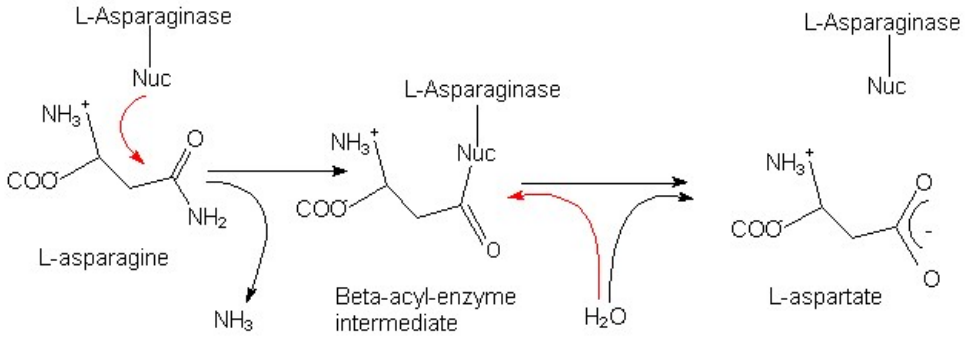
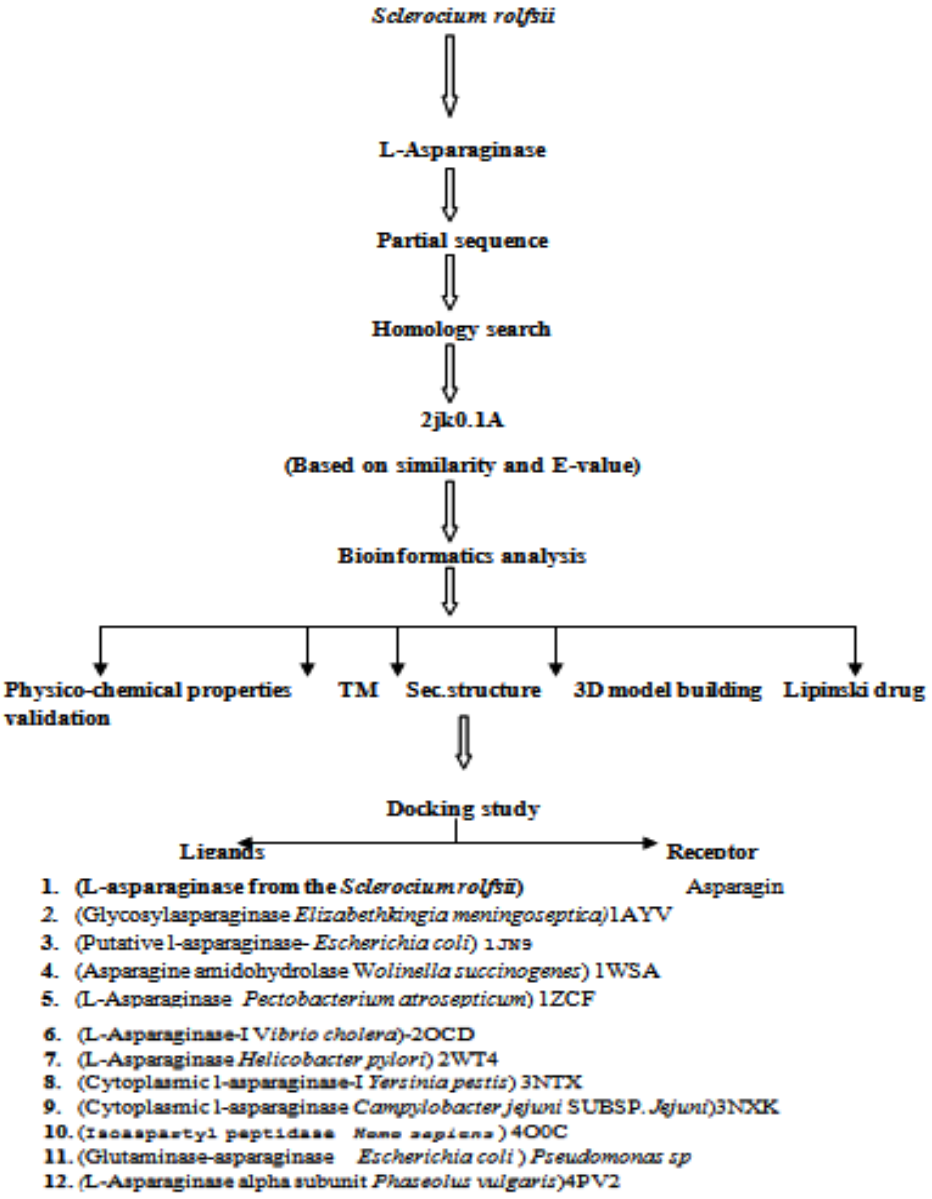


Fig1: Schematic illustration of the reaction mechanism of L-asparaginases. The proposed covalent intermediate is formed through nucleophilic attack by the enzyme.

2. METHODS AND MATERIALS



2.1. The Drug

L-Asparaginase

Organism used– ***Sclerocium rolfsii*** Sacc. (Hedge, 2010) is a polyphagous soil born fungal pathogen infecting over 500 plant species worldwide causing huge losses. The partial sequence of the L-asparaginase was extracted from NCBI data bank. The 3D models of the drug were prepared by using SWISS MODEL platform. This drug was considered as receptor. The receptor was docked with different ligands as shown in the study plan.

2.2. Study Plan

2.2.1. Study of the physico-chemical parameters of the protein purified

The physic-chemical parameters were studied by using Protparam tools from Expasy suit.

2.2.2. Extinction-coefficients

The extinction coefficient indicates how much light a protein absorbs at a certain wavelength.

E (Prot) = Numb (Tyr)*Ext (Tyr) + Numb (Trp)*Ext (Trp) + Numb (Cystine)*Ext (Cystine)

The absorbance (optical density) can be calculated using the following formula:

Absorb (Prot) = E (Prot) / Molecular weight

2.3. Instability-Index-(II)

The instability index provides an estimate of the stability of your protein in a test tube (Guruprasad,, et al, 1990, Ashokan and Pillai). Using the weight values of the proteins it is possible to compute an instability index (II) which is defined as:

$$II = \frac{10}{L} * \sum_{i=1}^{i=L-1} DIWV(x[i] \times [i+1])$$

Where: L is the length of sequence

DIWV(x[i] x [i+1]) is the instability weight value for the dipeptide starting in position

A protein whose instability index is smaller than 40 is predicted as stable, a value above 40 predicts that the protein may be unstable.

2.4. Aliphatic Index

The aliphatic index of a protein is defined as the relative volume occupied by aliphatic side chains (alanine, valine, isoleucine, and leucine).

Aliphatic index = X (Ala) + a * X (Val) + b * (X (Ile) + X(Leu))

Where X (Ala), X (Val), X (Ile), and X (Leu) are mole percent (100 X mole fraction) of alanine, valine, isoleucine, and leucine.

The coefficients a and b are the relative volume of valine side chain (a = 2.9) and of Leu/Ile side chains (b = 3.9) to the side chain of alanine.

2.5. GRAVY (Grand Average of Hydropathy)

The GRAVY value for a peptide or protein is calculated as the sum of hydropathy values of all the amino acids, divided by the number of residues in the sequence.

2.6. Trans-Membrane Sequence

Trans-membrane sequences were predicted by using TMHMM bioinformatics tool. A TMHMM prediction service is available at http://www.cbs.dtu.dk/services/TMHMM (Krogh eta l, 2001; Sonnhammer et al 1998)

2.7. Secondary Structure Prediction

Secondary structure of the protein sequence was predicted by using SOPMA. This method is based on a pattern recognition Support Vector Machine with a dedicated kernel and multiple alignments.

More than 200 proteins that probably have coiled-coil domains were identified in GenBank, including alpha- and beta-tubulins, flagellins, G protein beta subunits, some bacterial transfer RNA synthetases, and members of the heat shock protein (Hsp70) family (Combet et l, 2000).

2.8. Homology search BLASTp

Basic local alignment **search tool** (**BLAST**), directly approximates alignments that optimize a measure of local similarity, the maximal segment pair (MSP) **score**. MSP scores allow an analysis of the performance of this method as well as the statistical significance of alignments it generates. In addition to its flexibility and tractability to mathematical analysis, BLAST is an order of magnitude faster than existing sequence comparison tools of comparable sensitivity (Altschul, et al, 1990; Camacho et al, 2008).

2.9. Building 3D Model of the Protein

SWISS model was used to predict the 3D structure of the sequence. The SWISS-MODEL template library provides annotation of quaternary structure and essential ligands and co-factors to allow for building of complete structural models, including their oligomeric structure. The accuracy of the models generated by SWISS-MODEL is continuously evaluated by the CAMEO system. SWISS-MODEL is available at http://swissmodel.expasy.org/ (Biasini et al, 2014; Arnold et al, 2006; Kiefe et al, 2009; Guex et al, 2009**.**

2.10. Docking study

Docking study was performed by using HEX. *Hex* is an interactive molecular graphics program for calculating and displaying feasible docking modes of pairs of protein and DNA molecules (Ritchie et al, 2000). The various parameters used for the docking study are the following

Correlation type -------Shape only

FFT mode----------3D

Sampling method-------Range angles

Grid dimension----------0.6

Receptor range---------180

Ligand range----------180

Twist range----------360

Distance range---------40

Score threshold--------0.0

Steric scan----------18

2.11. Lipinski Validation of the Drug

We have tested the drug status of the L-asparaginase by using Lipinski Rule of Five**.**

Lipinski-Rule-of-Five

Lipinski rule of 5 helps in distinguishing between drug like and non drug like molecules. It includes the the following rules

- Molecular mass less than 500 Dalton
- High lipophilicity (expressed as LogP less than 5)
- Less than 5 hydrogen bond donors
- Less than 10 hydrogen bond acceptors
- Molar refractivity should be between 40-130

3. RESULT AND DISCUSSION

The partial sequence was subjected for homology search. Homology search by BLAST-p from NCBI is given 22 most probable homologues. Out of these one that is most suitable in the sense of E value and similarity percentage was selected as 2jk0.1A. This is the –Asparaginase from *Erwinia carotovora.* The FADTA sequence from this organism was used for the further study. The physic-chemical parameters studied include molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity (GRAVY).The amino acid composition (Table 2) shows highest percentage serine (Ser-12%) followed by alanine (Al-11.5%) and leucine (Le-11.5%) then threonine (Thr-9.5%) and least by pyrolysine (0%) and Selenocysteine(0%) . The pyrolysine and selenocystein are proteinogenic amino acids that are added during post translational biosynthetic mechanism. Post translational modification is absent in prokaryotic hence these are absent in our sequence. Serine is important in metabolism in that it participates in the biosynthesis of purines and pyrimidines. It is the precursor to several amino acids including glycine and cysteine, and tryptophan in bacteria. It is also the precursor to numerous other metabolites, including sphingolipids and folate, which is the principal donor of one-carbon fragments in biosynthesis. Serine plays an important role in the catalytic function of many enzymes. It has been shown to occur in the active sites of chymotrypsin, trypsin, and many other enzymes and L-asparaginase. The L-isomer of alanine is one of the 20 amino acids encoded by the genetic code. It is classified as a non-polar amino acid. L-Alanine is second only to leucine in rate of occurrence, accounting for 7.8% of the primary structure in a sample of 1,150 proteins. D-Alanine occurs in bacterial cell walls and in some peptide antibiotics.

The various physical parameters studied for the sequence (Table3) shows that the enzyme is very stable as the instability index (II) is 24 and half life is greater than 20 (>20). The enzyme can be purified at pH greater than or equal to 5.3 as the sequence shows iso-electric point (pI) 5.3. A statistical analysis shows that the aliphatic index (AI), which is defined as the relative volume of a protein occupied by aliphatic side chains (alanine, valine, isoleucine, and leucine), of proteins of thermophilic bacteria is significantly higher than that of ordinary proteins. The index may be regarded as a positive factor for the increase of thermo stability of globular proteins. The high aliphatic index shows that the protein-enzyme is more thermo stable. The Grand average of hydropathicity (GRAVY) of the linear polypeptide sequence is calculated as the sum of hydropathy values of all amino acids, divided by the number of residues in the sequence. Increasing positive score indicates greater hydrophobicity. The gravy in the case L-asparaginase is below .5 (0.138). Hence the enzyme is more soluble in water. The physical parameter of the enzyme reveals that the enzyme is more stable and thermo liable. Purification of L-asparaginase need pH about 5.3 and is less hydrophobic in nature. The molecular weight (20915) indicates it is smaller in nature; it may be due to the fact that the sequence we obtained is partial.

Table 2. *Amino acid composition*

Amini acid	%	Amini acid	%
Al	11.5	Leu	11.5
Ar	2.5	lys	4
Asn	6	Met	3.5
Asp	4.5	Phe	2.5
Cys	0.5	Pro	3.5
Gln	0.5	ser	12
Glu	4	Thr	9.5
Gly	8	Try	0.5
His	1.5	tyr	3
Ile	6	Pyl	0
Val	5	sec	0

Table 3. *Physical parameters of the sequence*

MolWt	pI	Total (-)vely charged residues	Total (+)vely charged residues	Total No.Atoms	H.life	II	AI	GRAVY
		Asp+ Glu	Arg+Lys		30hrs			
20915	5.3	17	13	2942	>20hr Y	24	94.25	0.138

The trans-membrane sequence analysis shows that the proteins have very little transmembarne sequence (Fig 2).

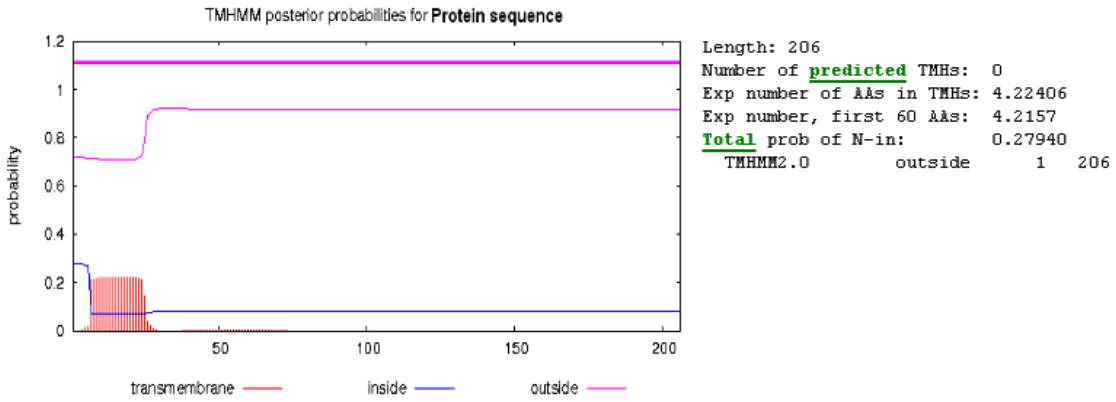


Fig 2. *Trans-membrane position of amino acid sequences predicted*

Trans-membrane proteins are polytopic proteins that aggregate and precipitate in water. They require detergents or nonpolar solvents for extraction; although some of them (beta-barrels) can be also extracted using denaturing agents. All transmembrane proteins are integral membrane proteins (IMPs), but not all IMPs are trans-membrane proteins. Thus the enzyme can be extracted with ease as it contain less amount of trans-membrane protein

The secondary structure prediction revelas (Table 4) that alpha helix is predominant than other helix, turns and coils. Hence the protein is more stable in nature.

Table 4. *Secondary structure prediction of L-Asparaginase*

Alpha helix	(Hh):	54 is 26.87%
3_{10} helix	(Gg):	0 is 0.00%
Pi helix	(Ii):	0 is 0.00%
Beta bridge	(Bb):	0 is 0.00%
Extended strand	(Ee):	59 is 29.35%
Beta turn	(Tt):	21 is 10.45%
Bend region	(Ss):	0 is 0.00%
Random coil	(Cc):	67 is 33.33%
Ambigous states	(?) :	0 is 0.00%
Other states	:	0 is 0.00%

The docking study to find out better drug candidate was performed by using L-asparaginase from different source with different nature of the asparaginase. The different source for the enzyme selected was*;*

1. Elizabethkingia meningoseptica 1AVY

Elizabethkingia meningoseptica is a Gram negative rod shaped bacteria widely distributed in nature (e.g. fresh water, salt water, or soil). It may be normally present in fish and frogs but are not normally present in the human microflora.

2. Escherichia coli ybik 1JN9

Escherichia coli *commonly abbreviated* ***E. coli)*** is a Gram-negative, facultatively anaerobic, rod-shaped bacterium of the genus Escherichia that is commonly found in the lower intestine of warm-blooded organisms (endotherms)

3. Wolinella succinogenes 1wsa

Wolinella succinogenes belongs to the epsilon subclass of Proteobacteria along with its close relatives Helicobacter pylori, Helicobacter hepaticus, and Campylobacter jejuni. H.pyori and C. jejuni are of the groups Helicobacteraceae and Campylobacteraceae respectively which are harmful pathogens in humans and animals.

4. Pectobacterium carotovorum 1ZCF

It (syn Erwinia carotovora) is a rod shaped bacterium – originally isolated from carrot- is a plant pathogen and opportunistic human pathogen, causative agent of soft rot and blackleg potato (P.atrosepticum) diseases

5. Vibrio cholera *2OCD*

It is a Gram-negative, comma-shaped bacterium. Some strains of V. cholerae cause the disease cholera. *V. cholerae* is a facultative anaerobic organism

6. Helicobacter pylori *2wt4*

Helicobacter pylori previously named **Campylobacter pylori**, is a Gram-negative, microaerophilic bacterium found in the stomach.

7. Yersinia pestis *3NTX*

It (formerly Pasteurella pestis) is a Gram-negative, rod-shaped coccobacillus, a facultative anaerobic bacterium that can infect humans and animals.

8. Campylobacter jejuni *3NXK*

It is a species of bacteria commonly found in animal feces. It is curved, helical-shaped, non-spore forming, Gram-negative, and microaerophilic.

9. Human *4O0C*

10. Plants **4PV2**

11. Erwinia carotovora

It is a rod shaped bacterium that was named after the crop of carrots from which it was first isolated. The bacterium infects a variety of vegetables and plants including carrots, potatoes, cucumbers, onions, tomatoes, lettuce and ornamental plants like iris

12. Yersinia pestis (formerly Pasteurella pestis)

It is a Gram-negative, rod-shaped coccobacillus, a facultative anaerobic bacterium that can infect humans and animals.

Table 5. *Energy level for docking Receptor (Aspargain) and ligands (L-Asparaginase from different sources*

Sl.No	Source	PDB Id	Energy
1	*Escherichia coli*	1JN9	-114.57
2	*Elizabethkingia meningoseptica*	1AVY	-147.60
3	*Erwinia carotovora*	1ZCF	-62. 83
4	*Helicobacter pylori*	2WT4	-145.03
5	Campylobacter jejuni	3NXK	-91.90
6	Pseudomonas	4PGA	-134.85
7	*Wolinella succinogenes*	1WSA	-131.05
8	*Vibrio cholera*	2OCD	-123. 09
9	*Plants*	*4PV2*	*-135.98*
10	*Human*	*4O0C*	*-134.90*
11	*Yersinia pestis*	*3NTX*	*-146.01*

The free energy level of docking (Table5) shows that *Erwinia carotovora* has performed docking with least free energy (-62.83). That is followed by *Campylobacter jejuni* with -91.90 and all other selected organisms shows more or less same energy level >100. Thus as a drug L-Asparaginase from *Erwinia carotovora* may be ideal for the treatment of cancer and other purpose. *Erwinia carotovora* is a rod shaped bacterium that was named after the crop of carrots from which it was first isolated. The bacterium infects a variety of vegetables and plants including carrots, potatoes, cucumbers, onions, tomatoes, lettuce and ornamental plants like iris. Only the genome of *Erwinia carotovora subsp. atroseptica strain SCRI1043* (Eca 1043) has been sequenced. The sequencing shows that the genome of Erwinia carotovora subsp. atroseptica is a single 5,064,019 base pair circular chromosome with 50.97 percent of guanine and cytosine, approximately 4,491 coding sequences (CDSs), seven rRNA operons, 76 tRNAs, and around 25 stable noncoding RNAs. It causes cell death through plant cell wall destruction by creating an osmotically fragile cell.

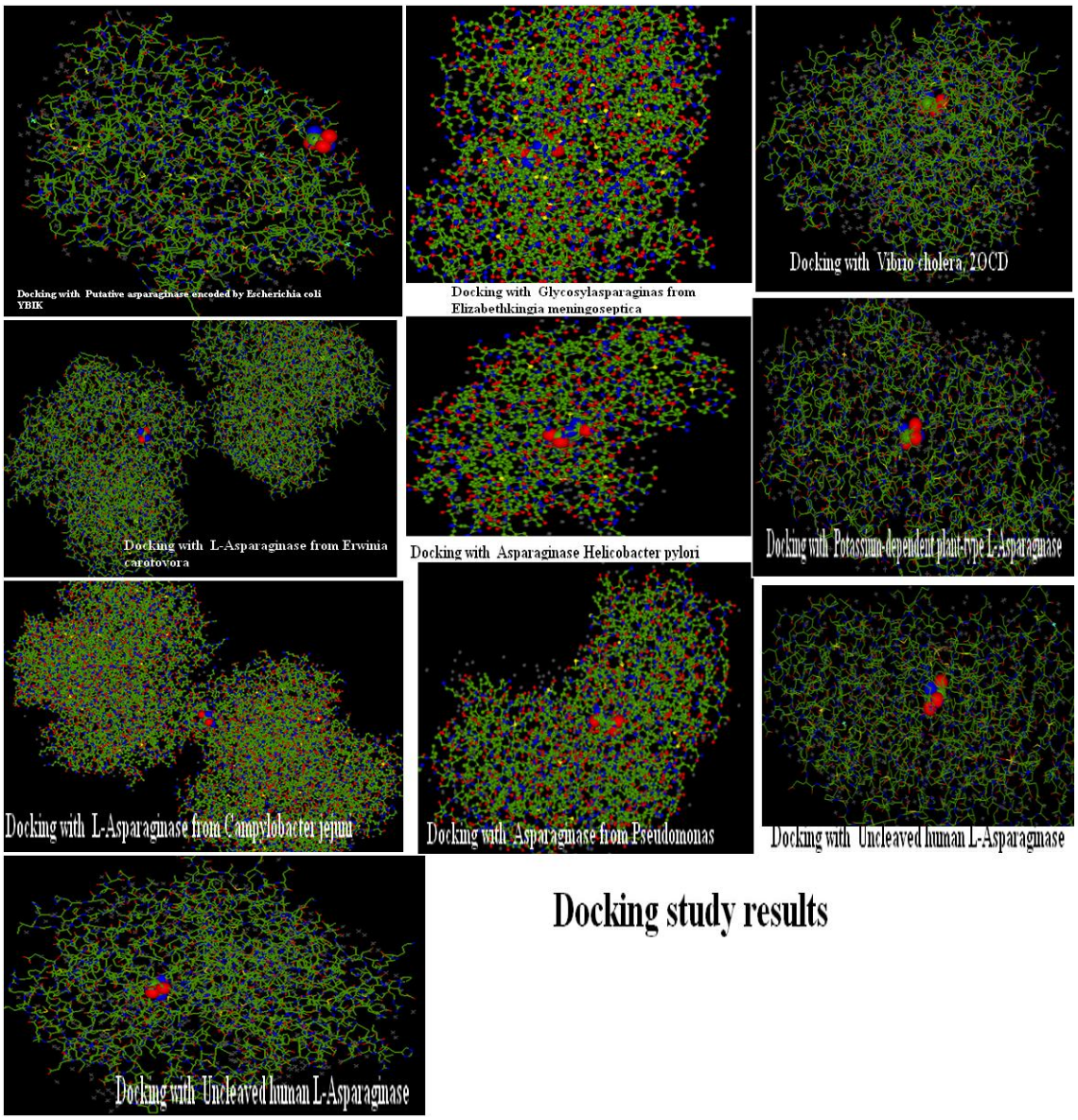


Fig 3. *Docking study results*

4. CONCLUSION

The docking study of asparaginase from different sources as ligands and asparagine as receptor reveals that asparaginase from *Erwinia carotovora* will be a most likely drug candidate in the sense of efficiency of drug- receptor interaction. The availability of the species is one more advantage in the commercial production of the enzyme. Antitumor properties of the asparaginase

are well established so far. By the analysis of the energy needed for the docking of the receptor and ligand shows that out of the many microbial, plant and animal source of the enzyme, asparaginase from *Erwinia carotovora* may be better in curing the various type of cancer. A further study concentrating its molecular interaction and metabolic pathway is needed to substantiate the range of success in the anti-cancer property of the enzyme.

REFERENCES

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) "Basic local alignment search tool." J. Mol. Biol. 215:403-410.

Arnold K., Bordoli L., Kopp J., and Schwede T. (2006). The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. Bioinformatics, 22,195-201.

Ashokan.K.V and Pillai.M.M. (2010). In Silico characterization of Cu, Zn superozide dismutase of Dinococus radiodurans, Bioinformatics :A journal of Bioinformatics and it s applications.5 (1&2): 58-69. (ISSN: 0973-6581).

Broome,J. D.: Evidence that the L-asparaginase of guinea pig serumis responsible for its antilym phoma effects.I. Properties of the L-asparaginase of guinea pig serumin relation to those of theantilym phoma substance.J. Exp. Med. 118:99-120. 1963.

Camacho C., Coulouris G., Avagyan V., Ma N., Papadopoulos J., Bealer K., & Madden T.L. (2008) "BLAST+: architecture and applications." BMC Bioinformatics 10:421.

Combet C., Blanchet C., Geourjon C. and Deléage G. (2000). Network Protein Sequence Analysis, TIBS 25(3):147-150.

De Lowery, W. C.: Henson.D.; Cornet,J.,and Sel lin. H.: Toxic andantineo plastic effects of L-aspara ginase. Study of mice with lymphoma and normal monkeys and report on a child with leukemia. Cancer. 19: 1813-1819. 1966.

Feher M, Schmidt JM (2003). "Property distributions: differences between drugs, natural products, and molecules from combinatorial chemistry". J Chem Inf Comput Sci 43 (1): 218–27.

Guex, N., Peitsch, M.C., Schwede, T. (2009). Automated comparative protein structure modeling with SWISS-MODEL and Swiss-Pdb Viewer: A historical perspective. Electrophoresis, 30(S1), S162-S173.

Guruprasad, K., Reddy, B.V.B. and Pandit, M.W. (1990) Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. Protein Eng. 4,155-161

Hedge,Y.R.,Tippeshi,L.,Chavan,s.,Sreedevi,C.K.,Chav han,M..,Venugopal,S., Rao,L., Rasalkar, C.K .Evaluation of fungicides against Sclerotium rolfsiicausing wiltof Stevia rebaudiana.J.Plant Dis Si.5(2010):254-256.

Hill. J. M. et. al.: L-asparaginase therapy for leu kemia and other malignantneoplasms. Remissionin

Hopkins AL, Groom CR, Alex A (May 2004). "Ligand efficiency: a useful metric for lead selection". Drug Discov. Today 9 (10): 430–1.

humanleukemia.J.A.M.A. 202: 882-888, 1967.

Kidd, J. G.: Regression of transplanted lyni phomas induced in vivo by meansof normal guinea pig serum. Course of transplanted cancers of various kinds in mice and rats given guinea pig serum, horse serum, orrabbitserum.J.Exp.Med.98:565-582,1953.

Kiefer F, Arnold K, Künzli M, Bordoli L, Schwede T (2009). The SWISS-MODEL Repository and associated resources. Nucleic Acids Research. 37, D387-D392.

Krogh. A ,Larsson B, vonHeijne G, and. Sonnhammer E. L. L(2001). Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *Journal of MolecularBiology*, 305(3):567-580,January2001.

Leeson PD, Springthorpe B (November 2007). "The influence of drug-like concepts on decision-making in medicinal chemistry". Nat Rev Drug Discov 6 (11): 881–90.

Marco Biasini, Stefan Bienert, Andrew Waterhouse, Konstantin Arnold, Gabriel Studer, Tobias Schmidt, Florian Kiefer, Tiziano Gallo Cassarino, Martino Bertoni, Lorenza Bordoli, Torsten Schwede. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Research; doi: 10.1093/nar/gku340.

Newman DJ, Cragg GM (March 2007). "Natural products as sources of new drugs over the last 25 years". J. Nat. Prod. 70 (3): 461–77.

Rask-Andersen M, Almén MS, Schiöth HB (August 2011). "Trends in the exploitation of novel drug targets." Nature Reviews Drug Discovery 8 (10): 549–90

Ritchie, D.W. and Kemp, G.J.L.(2000). Protein Docking Using Spherical Polar Fourier Correlations, PROTEINS: Struct. Funct. Genet. 39: 178-194.

Roger, Manuel Joaquín Reigosa; Reigosa, Manuel J.; Pedrol, Nuria; González, Luís (2006), Allelopathy: a physiological process with ecological implications, Springer, p. 1, ISBN 1-4020-4279-5.

Ryckmans T, Edwards MP, Horne VA, Correia AM, Owen DR, Thompson LR, Tran I, Tutt MF, Young T (August 2009). "Rapid assessment of a novel series of selective CB(2) agonists using parallel synthesis protocols: A Lipophilic Efficiency (LipE) analysis". Bioorg. Med. Chem. Lett. 19 (15): 4406–9.

Schrek R.D, Dolowy W.C and Ammeraal R.N (1967). L-asparaginase: toxicity to normal and leukemic humanlymphocytes. Scienec , 155 (3760): 329-330.

Sonnhammer E. L.L., von Heijne G and Krogh, A. (1998). A hidden Markov model for predicting transmembrane helices in protein sequences. In J. Glasgow, T. Littlejohn, F. Major, R. Lathrop, D. Sankoff, and C. Sensen, editors, *Proceedings of the Sixth* International Conference *on Intelligent Systems for Molecular Biology*, pages 175-182.

Vita.J.L and Wataru W.S. (2006). Toxicity of L-asparaginase in children with advanced leukemia, Cancer 30(2): 339-347.