Journal of Computational Intelligence in Bioinformatics
ISSN 0973-385X Volume 5, Number 2 (2012) pp. 95-103

http://www.ripublication.com/jcib.htm

Docking Studies for Screening Anticancer Compounds of Allamanda cathartica

Reena B. and Dr. Krishnamoorthy M.

Department of Post Graduate Studies and Research in Biosciences, Mangalore University, Mangalagangothiri, Karnataka 574 199, India
E-mail: reenasatheeskumar@gmail.com

Abstract

Cancer is a disease in which abnormal cells divide without control. Cancer harms the body when damaged cells divide uncontrollably to form lumps or tumors. Cancerous cells start an infinite division, due to the mutation that takes place in the cell cycle controlling elements like cyclin dependent kinases (CDK). These cyclin dependent kinases are group of protein involved in regulating the cell cycle mechanism. Identification of novel compounds to inhibit the activity of CDK can be a beneficial approach in the treatment to combat cancer. For the present study the compounds present in the leaves of the plant Allamanda cathartica have been subjected for the docking analysis with the protein CDK1. The compounds chosen for the study includes Allamandin, β-amyrin, Plumericin, isoplumericin, β sitosterol and ursolic acid. The 3D structures of these compounds were subjected for the docking studies against cyclin dependent kinase protein (CDK1) of Saccharomyces cerevisiae. The 3D structuresof the compounds were obtained from the Pubchem database. The 3D structure of CDK1 was modeled using swisspdb viewer. The docking studies were performed using Arguslab.

Keywords: Allamanda cathartica, Anticancer compounds, Molecular modeling, docking.

Introduction

Cancer

Cancer is the uncontrolled growth of abnormal cells in the body [1, 2]. All cancers occur due to abnormalities in DNA sequence. Throughout life, the genome within

cells of the human body is exposed to mutagens and suffers mistakes in replication. These mutations influence DNA sequence in each cell to alter from its normal type to cancer type. These somatic mutations alters the function of a critical gene, providing growth advantage to the cell in which it has occurred, resulting in the emergence of an expanded clone derived from this cell. Acquisition of additional mutations and consequent waves of clonal expansion result in the evolution of the mutinous cells that invade surrounding tissues and metastasise [3]. One in three people in the world develop cancer and one in five die of the disease. Cancer is therefore the commonest genetic disease. The identification of genes that are mutated has been a central aim of cancer research since the advent of recombinant DNA technology [3].

Cancer Biology and its related studies have now leaded the way to identify a solution to take control of the mutated genes. On the other side the symptoms produced by the disease has devastated the future of human life [3]. The Cancer Genome Project has identified human gene sequence variants and mutations that play critical role in the development of human cancers [3].

The evolution in the field of Cheminformatics has brought in various approaches to identify the therapeutic properties of chemical compounds stored in the chemical databases. For the present study chemical compounds from the leaves of the plant Allamanda cathartica stored in the pubchem database were used.

Cyclin-dependent kinases that control the cell cycle

The cyclin-dependent kinases (Cdks) are a family of serine/threonine protein kinases whose members are small proteins (~34–40 kDa) composed of little more than the catalytic core shared by all protein kinases. Most of the known cyclin-CDK complexes regulate the progression through the cell cycle. Animal cells contain at least nine CDKs, four of which, Cdk1, 2, 3, and 4, are directly involved in cell cycle regulation [4, 5, 6]. In mammalian cells, different Cdks are active and required at different phases of the cell cycle. The expression of the Cdk subunit is generally constant throughout the cell cycle and these proteins have specific activity when its cyclin partner is expressed. The role of the Cdks is to control cell cycle progression through phosphorylation of proteins that function at specific cell cycle stages. Cell cycle defects are often mediated by alterations in cyclin-dependent kinase (CDK) activity. The advancement in the field of sequencing has revealed the sequences of these cell cycle control elements. The treatment strategy lies in blocking the activity of CDK proteins which over regulate the cell cycle mechanism.

In the budding yeast Saccharomyces cerevisiae, the cell-cycle events are controlled by a single essential Cdk called Cdk1 [4, 5, 6]. Cdk function has been remarkably well conserved during evolution. It is possible, for example, for yeast cells to proliferate normally when their gene for Cdk1 is replaced with the human one. This evidence clearly illustrates that Cdk function in the cell-cycle control system, has remained fundamentally unchanged over hundreds of millions of years of eukaryotic evolution [4, 5, 6].

Allamanda cathartica

Allamanda, also known as angel's trumpet, golden trumpet, yellow bell, is an evergreen, vine-like woody shrub [7]. It may reach a free-standing height of 2 m and an extension of 5 m or more. The species also climbs a few meters into the crowns of tall brush and low trees. Older plants often have multiple stems from the root crown and long stems with relatively few branches. Bark of lower stems is brown and furrowed. Twigs are green or yellow green. Stems and twigs exude a milky sap when cut. The leathery, yellow-green to dark green leaves grow in whorls of three or four, or are sometimes opposite. Leaves are 6 to 16 cm long, obovate to oblong-lanceolate, pointed at both ends and have entire margins and short petioles. Inflorescences are few-flowered, axillary cymes that grow near the ends of branches. The bright yellow flowers are 5 to 7.5 cm across. Flowers of cultivated varieties are often larger and may be colored white, cream, pink, or orange. Capsules, which rarely occur in cultivated varieties, are subglobose, 4 to 6 cm in diameter, and densely prickled. They contain many tan, flattened, winged seeds. There are 2n = 18 chromosomes [7].

Medicinal uses

The decoction of leaves is taken as a cathartica and for biliousness. A decoction of leaves and flowers is taken as a cure for anuria. A decoction of leaves and flowers together with monkey-scala-simiae is used as a treatment for impotency [8]. The leaves contain compounds like iso-plumericin, plumericin, plumieride in addition to beta amyrin, beta-sitosterol and ursolic acid [8].

Methodology

Homology Modeling of CDK1_yeast protein

Three-dimensional (3D) protein structures provide valuable insights into the molecular basis of protein function, allowing an effective design of experiments, such as site-directed mutagenesis, studies of disease-related mutations or the structure based design of specific inhibitors [9]. Proteins in the same families frequently have noticeable similarities and thus share three-dimensional architecture, which allows a structural description of all proteins in a family even when only the structure of a single member is known [10]. This concept forms the base for homology modeling of proteins whose structures have not been determined. In this present study the CDK1_yeast protein of Saccharomyces cerevisiea also does not have the 3D structure which is of prime importance to carry out the docking study. Thus the CDK1_yeast protein was subjected for homology modeling to determine its 3D structure. The CDK1 protein sequence was obtained from the genbank database. The sequence selected for the current analysis was CDK1_YEAST Cyclin-dependent kinase 1 Saccharomyces cerevisiae (strain ATCC 204508 / S288c) made of 298 aminoacids. This protein sequence belongs to the organism Saccharomyces cerevisiae S288c. The CDK1_yeast protein of saccharomyces cerevisiae is very essential for the start and completion of the cell cycle mechanism. The protein sequence was retrieved in the fasta format and its 3D structure was determined using Swisspdb.

Template Identification

Homology modeling starts from selection of homologues with known structures from the PDB. If the query sequence has high sequence identity (>30%) to the structure, the homology detection is quite straightforward which is usually done by comparing the query sequence with all the sequences of the structures in the PDB [10]. The most popular software used for database similarity search is BLAST [11]. The program searches sequence databases for homologue sequences using the optimal local alignments to the query. For modeling the structure of the CDK1_yeast protein template identification was carried out using protein-protein BLAST (Blastp). The homologue search was done against the structure database PDB (protein Databank).The default parameters were set for the template search used by Blastp. The protein 1GZ8 (Human cyclin dependent Kinase 2 complexed with the inhibitor 2 amino-6-(3'methyl-2'oxo) Butoxypurine, isolated from Homosapiens and 3QHW_C (structure of cyclin A transition state mimic protein) isolated from Homosapiens were the two protein that showed similarity with the CDK1 of Saccharomyces cerevisiae. 1GZ8 showed a sequence identity of 62.1% when pair wise aligned with CDK1. On the other end 3QHW_C showed 62.0% of sequence identity with CDK1. Since 1GZ8 proteins had maximum sequence identity with the query protein, 1GZ8 was selected as the template for modeling the query. For modeling procedures the 3D structure of the template was downloaded and saved as 1Gz8.pdb

Modeling CDK1 using Swiss Model

The aminoacid sequence of CDK1_yeast was loaded in SwissModel. Swiss-model (http://swissmodel.expasy.org) is a server for automated comparative modeling of three dimensional (3D) protein structures. SWISS-MODEL is a fully automated protein structure homology-modeling server, accessible via the ExPASy web server, or from the program DeepView [9]. 3D structure of the template protein retrieved from the Protein Data Bank (PDB) was also loaded in the Swiss Model. Both the protein sequence and the template structure were superimposed together using the magic fit option of Swiss Model. The modeling procedure was started using the menu submit model request. A temporary project pdb file containing both the CDK1_yeast protein aminoacid sequence and structure of 1GZ8 is taken as the input by the swissmodel server. The server performs the modeling and the outputs are been send to the email id. The output contains the 3D structure of the CDK1_yeast protein, along with the structure verifications. The modeled structure of the CDK1 protein is saved as a separate file with the file extension as pdb.

Active compounds of Allamanda cathartica

The active compounds present in the leaves of Allamanda cathartica are β-amyrin, plumericin, Allamandin, iso-plumericin, ursolic acid. 3D structures of these compounds were retrieved from the Pubchem database. The structures were stored separately as pdb molecules.

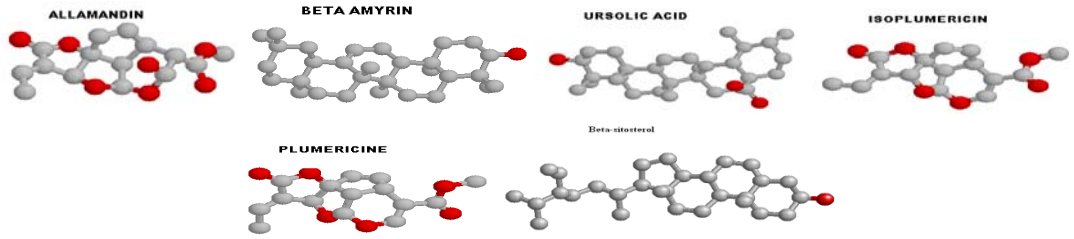

Figure 1: Structures of active compounds present in the leaves of Allamanda cathartica

Docking of active compounds with CDK1_yeast

Docking is a method which predicts the preferred orientation of one molecule to a second when bound to each other to form a stable complex. For the present study, the compounds of the plant leaf Allamanda cathartica was subjected for docking with the CDK 1 protein using ArgusLab. ArgusLab is a molecular modeling, graphics, and drug design program for Windows operating systems.

Binding site identification

Binding site identification is the very important aspect before carrying the docking analysis. The binding site is predicted based on identifying the pockets or the spaces that are capable to enable an external ligand to bind with the protein molecule. Binding site prediction was done using the online tool Q-Site finder. Many pocket sites in the given CDK1 molecule were identified. The binding sites selected for the present docking calculations were Val 14, Tyr 19, Val 22, Ala 38, Lys 40, Glu 58, Leu 62, Val 71, Phe 88, Glu 89, Phe 90, Leu 91, Asp 92, Leu 93, Asp 94, Arg 97, Asn 141, Leu 143, Gly 153, Asp 154, Phe 155. The binding site selected comprised the site volume of 343 Å3, where the total protein volume was 29022 Å3.

Docking analysis

CDK1 was subjected to docking analysis with the active compounds using the tool ArgusLab. CDK1 structure was loaded to the ArgusLab tool. The binding sites identified were selected and labeled as the pocket sites. The individual chemical structure was also loaded to the tool and its atoms were selected as the ligand molecule. By using the menu option set up a Dock Calculations the docking process was initiated. The parameters were set as docking engine ArgusDock, calculation type as Dock and the ligand as Flexible type.

CDK1 was docked with Allamandin with the parameters set up leading to flexible docking. Different poses were examined by ArgusLab to fit the ligand allamandin to the CDK1 molecule. Among the different poses the best pose was fixed to the CDK1 molecule. Allamandin was docked with the CDK1 protein with the energy -821186 kcal/mol.

CDK1 was docked with Isoplumericin with the parameters set up leading to flexible docking. Different poses were examined by ArgusLab to fit the ligand isoplumericin to the CDK1 molecule. Among the different poses the best pose was

fixed to the CDK1 molecule. Isoplumericine was docked with the CDK1 protein with the energy -5.20993kcal/mol.

CDK1 was docked with β-sitosterol with the parameters set up leading to flexible docking. Different poses were examined by ArgusLab to fit the ligand β-sitosterol to the CDK1 molecule. Among the different poses the best pose was fixed to the CDK1 molecule. β-sitosterol was docked with the CDK1 protein with the energy -13.9086kcal/mol.

Results and Discussion

Three Dimensional structure of CDK1_yeast protein molecule

CDK1_yeast protein structure was modeled by using the Swisspdb viewer software. CDK1_yeast protein and the template protein structure showed a sequence identity of 62.1%. Thus the sequence identity that existed between CDK1_yeast and 1GZ8 was enough to model the target protein structure. The given alignment shows the aminoacid equivalent between the CDK1_yeast and 1GZ8. The sequence shows the arrangement of the alpha helices and beta sheet structures in the protein. The aminoacid sequence of CDK1_yeast protein retrieved from genbank database in the fasta format is given below in the figure

>sp|P00546|CDK1_YEAST Cyclin-dependent kinase 1 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=CDC28 PE=1 SV=1

MSGELANYKRLEKVGEGTYGVVYKALDLRPGQGQRVVALKKIRLESEDEGVPSTAIREIS
LLKELKDDNIVRLYDIVHSDAHKLYLVFEFLDLDLKRYMEGIPKDQPLGADIVKKFMMQL
CKGIAYCHSHRILHRDLKPQNLLINKDGNLKLGDFGLARAFGVPLRAYTHEIVTLWYRAP
EVLLGGKQYSTGVDTWSIGCIFAEMCNRKPIFSGDSEIDQIFKIFRVLGTPNEAIWPDIV
YLPDFKPSFPQWRRKDLSQVVPSLDPRGIDLLDKLLAYDPINRISARRAAIHPYFQES

Figure 2: Aminoacid sequence of CDK1_yeast protein.

Alignment:

```
CDK1 sac  5          LANY KRLEKVGEGT YGVVYKALDL RPGQGQRVVA LKKIRLESED
1GZ8      1          menf qkvekigegt ygvvykarnk ltge---vva lkkir-----

CDK1 sac               ss sssssssss  ssssssssss         ss sssss
1GZ8                   ss sssssssss  ssssssssss    s   sss sssss


CDK1 sac  49    EGVPSTAIRE ISLLKELKDD NIVRLYDIVH SDAHKLYLVF EFLDLDLKRY
1GZ8      44    --vpstaire isllkelnhp nivklldvih ts-nklylvf eflhqdlkkf

CDK1 sac           hhhhhh hhhhh         sssssss s   ssssss s   ssshhh
1GZ8               hhhhhh hhhhh         sssssss s   ssssss s   ssshhh


CDK1 sac  99    MEGIPKDQPL GADIVKKFMM QLCKGIAYCH SHRILHRDLK PQNLLINKDG
1GZ8      91    mdasaltg-i plpliksylf qllqglafch shrvlhrdlk pqnllinteg

CDK1 sac        hhh         hhhhhhhhh hhhhhhhhhh hh           sssss
1GZ8            hhh         hhhhhhhhh hhhhhhhhhh hh           sssss
```

```
CDK1_sac  149   NLKLGDFGLA RAFGVPLRAY THEIVTLWYR APEVLLG -- ----------
1GZ8      140   aikladfgla rafgvpvrty thevvtlwyr apeillg-yy stavdiwslg

CDK1_sac        sssss    h hh    sss    sss        hhhh
1GZ8            sssss    h hh    sss    sss       hhhhh       hhhhhhhhh


CDK1_sac        ---------- ---------- ---------- ---------- ----------
1GZ8      191   cifacmvtrr alfpgdscid qlfrifrtlg tpdevvwpgv tempdykpsf

CDK1_sac
1GZ8            hhhhhhhhh          hhh hhhhhhhhh


CDK1_sac        ---------- ---------- ---------- ---------- ----------
1GZ8      241   pkwarqdfsk vvppldedgr sllsqmlhyd pnkrisakaa lahpffqdvt

CDK1_sac
1GZ8                  hh h     hhhh hhhhhh           hhhh h    hh


CDK1_sac        --------
1GZ8      291   kpvphlrl

CDK1_sac
1GZ8
```

Figure: 3: Alignment of CDK1_yeast protein sequence with 1GZ8 protein.

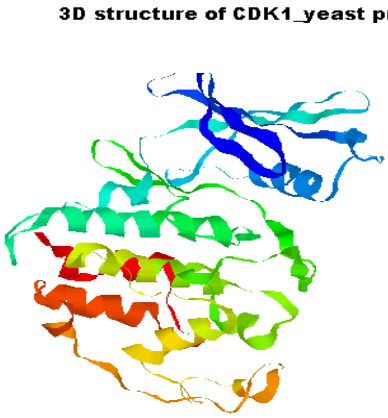

Figure 4: Ribbon structure illustration of CDK1_yeast protein structure.

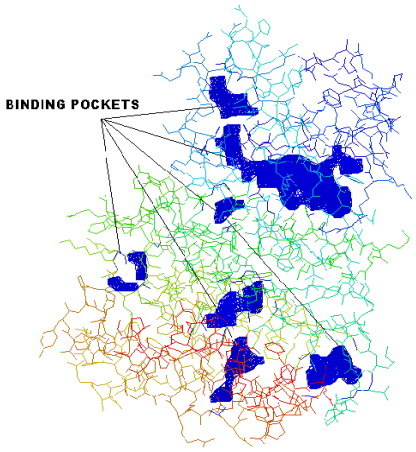

Figure 5: Binding pockets of CDK1_yeast protein

Docking studies

The docking studies show the binding of the active compounds Allamandin and Isoplumericin β-sitosterol with the protein CDK1_yeast. The other active compounds belonging to the plant Allamanda cathartica did not show a binding pose with the target protein structure. From the docking studies we can conclude that Allamandin, Isoplumericin and β-sitosterol prove to possess the ability to bind with the CDK1_yeast. Thus these two compounds can be considered as the lead molecules to be optimized and used as antiproliferative agents by inhibiting the activity of CDK1-yeast protein.

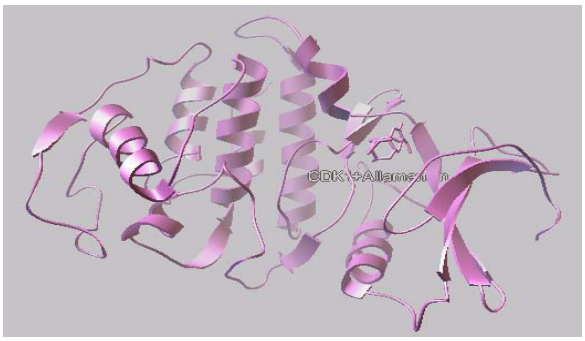

Figure 6: Allamandin docked with CDK1_yeast protein

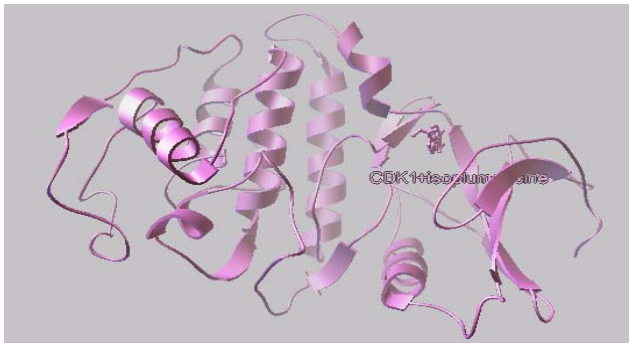

Figure 7: Isopumericine docked with CDK1_yeast protein.

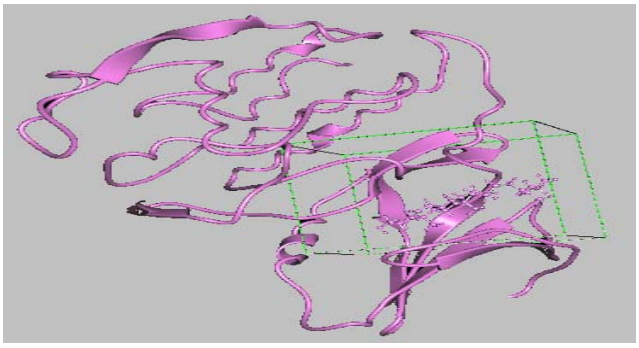

Figure 8: β-sitosterol docked with CDK1_yeast protein.

Conclusion

Allamandin, Isoplumericine and β-sitosterol are the compounds showing the ability to bind with the target protein CDK1_yeast protein. Thus these compounds can be further subjected for lead optimization and also for invitro studies for proving their drug action as an antiproliferative compounds.

References

[1] Moscow, J. A., Cowan, K. H., 2007, "Biology of Cancer", Cecil Medicine., chap 187.
[2] Thun, M.J., 2007, "Epidemiology of cancer", Cecil Medicine., 23rd ed, Elsevier, chap 185.
[3] The Cancer Genome Project.
[4] De Bondt, H.L., et al., 1993," Crystal structure of cyclindependent kinase 2", Nature, 363:595–602.
[5] Morgan, D.O., 1997, "Cyclin-dependent kinases: engines, clocks, and microprocessors", Cell Dev. Biol. 13:261–291.
[6] Ubersax, J.A., et a.,l 2003, "Targets of the cyclin-dependent kinase Cdk1", Nature, 425:859– 864.
[7] Howard, R.A., 1989, "Flora of the Lesser Antilles, Leeward and Windward Islands. Dicotyledoneae", Harvard University: vol 6, 658.
[8] LachmanWhite, D. A., Adams. C.A., Trotz. U.O.D., 1992," A guide to the medicinal plants of coastal Guyana".
[9] Torsten Schwede., Jürgen Kopp.,Nicolas Guex., and Manuel. C., Peitsch, 2003," SWISS-MODEL: an automated protein homology- Modeling server", Nucleic Acids Research. Vol. 31: 3381–3385.
[10] Zhexin Xiang., 2007, "Homology-Based Modeling of Protein Structure.Computaional methods for protein structure prediction and modeling", Vol.1: 319-350.
[11] Altschul, et al., 1997," Gapped BLAST and PSI-BLAST A new generation of protein database search programs", Nucl. Acids Res., 25, 3389-3402.