Molecular docking and its application in search of antisickling agent from *Carica papaya*

Dibya Ranjan Das¹, Dhanesh Kumar¹,2*, Pravind Kumar³, Bisnu Prasad Dash¹

¹Department of Bioscience & Biotechnology, Fakir Mohan University, VyasaVihar, Nuapadhi, Balasore 756020, India  
²Department of Plant Science, School of Life Sciences, University of Hyderabad, Hyderabad 500046, India  
³Department of Biotechnology & Bioinformatics, School of Life Sciences, University of Hyderabad, Hyderabad 500046, India

**ABSTRACT**

Understanding the protein–ligand interaction is a fundamental step for drug discovery in numerous pharmaceutical enterprises. The integration of computational and experimental process can reduce the time and cost for the advancement of novel medications. Molecular docking is one of the modern drug designing strategies, which explore the competence of a ligand by computing the minimum binding energy. Docking is utilized in virtual screening of enormous databases of compounds for hit identification and assessing the impact of chemical modifications during lead optimization. Sickle cell disease is a serious issue that affects people worldwide. Leaf extracts and ripe fruits of *Carica papaya* have been identified for curing sickle cell disease. Molecular docking approach using ArgusLab 4.0.1 was used to study the interaction between 24 different phytocompounds of *C. papaya* and sickle cell protein [2 Deoxyhemoglobin S (2HBS)] to identify the best phytocompounds for curing sickle cell disease. Three phytocompounds, Xanthoangelol D, N-[(4R)-4-(3-fluorophenyl)-6-oxo-4,5-dihydro-H-pyrimidin-2-yl]-3-methoxybenzamide, and Carpaine showed the highest inhibitory activity against the 2HBS protein which may become potent anti-protein drugs for the treatment of sickle cell disease with the support of further studies.

1. **INTRODUCTION**

The structures of protein complexes are basics for understanding the sub-atomic systems of protein–protein interaction, analyzing genetic disorders, and controlling binding affinity. The difficulties with respect to time and cost for X ray-crystallography and nuclear magnetic resonance methodologies create a gap between the experimentally determined complex structures and sequences accessible for protein buildings [1]. The inadequacy of complex structures can be mitigated by molecular docking, which gives a quick and proficient alternative in the field of bioinformatics.

Molecular Docking is a tool to predict the energetically favored conversion of one molecule to the next molecule when bound to form a relatively stable complex with the overall least energy. Docking expects to accomplish an optimized conformation and relative direction between protein–ligand. It helps us to see how small ligands bind to different proteins, for example, transport protein, signal receptors, catalysts, viral proteins, and so forth so as to suppress or actuate their function. For molecular docking, it is basic to recognize the ligands that can bind protein at an explicit dynamic site.

Docking is utilized in in-silico screening of huge databases of compounds for hit ID and assessing the impacts of chemical alterations during lead optimization [2,3]. Hit identification is the significant objective of high throughput screening (HTS) that intends to recognize dynamic compounds (hits) by screening enormous quantities of various chemical compounds against chosen targets [4]. For most of the disease-related proteins, experimental structures are not accessible thus; ligand docking can be performed on hypothetical models [5]. The inaccuracies in a homology modeling can be overcome by utilizing computational docking techniques which decide the gross basic highlights of a complex, but find it exceedingly difficult to effectively anticipate high-goals structures of such protein–protein complex, indicating
the need to form new docking calculations which include the essential degrees of freedom to make up for the errors.

Thus, the present study intends to portray the leading edge of protein–ligand docking and its application looking for appropriate antisickling agents from Carica papaya.

2. DOCKING PROCEDURE

Whole docking procedure can be described in following simple steps:

2.1. Search for Protein and Ligand Structure

The search for convenient target protein and ligands is essential for performing docking. So, one has to search whether the target protein is deposited in the Protein Data Bank (PDB) database (http://pdb.org) or Swiss UniProt knowledge base (http://expasy.org/sprot). If the target protein is not present in the database but similar sequences are there, then homology modeling can be done by using the Swiss model repository (http://swiss model.expasy.org/repository/), modeller tools—I-TASSER, etc. Next, the ligand can be found from the PubChem database (http://pubchem.org) or Zinc (http://blaster.docking.org/zinc).

2.2. Search for Protein–Ligand Binding Site

The search for potential protein–ligand binding site can be achieved by using two approaches, i.e., (a) first, identify the protein–ligand binding site and then dock the ligand (b) dock ligand directly onto the complete receptor structure; blind docking [6].

2.3. Development of Docking Program

Molecular docking has turned into an essential part in many drug discovery programs, particularly for virtual screening of phytochemicals as potential drug molecules. The first docking software was developed in the mid-1980s by Irwin Kuntz, University of California and attempts are still in the process to improve the docking calculations. Ongoing advancements in docking tools predict the capacity of an enzyme by distinguishing its natural substrates [7]. Table 1 gives the characteristic features of some important docking programs.

The successful predictions of protein complexes were done by identifying the protein of interest belonging to a specific super family so that the search for potential substrates and types of reactions are restricted to a specific area. However, various methodologies were utilized to precisely rank the docked molecules by using different programs such as:

- **DOCK 3.5.x**—This program used the idea that enzymes catalyze reactions by restricting the transition state superior to the substrate, docking transition state-like molecules ought to give a superior signal than docking substrates by using the knowledge of amidohydrolase super family performed hydrolysis reactions. The protein was kept rigid [8].

- **Glide**—The program distinguishes the enzymes having a place with a specific subgroup of enolase super family permitted tapering the arrangement of potential substrates, and precision in the positioning was improved by refining and rescoring the docked complex with an increasingly confounded material science-based scoring capacity and permitting receptor side chains to move [9].

2.4. Mechanism of Docking

2.4.1. Search algorithms

The principle target of the search algorithm is to locate every single imaginable direction and conformation of the protein combined with the ligand. It is classified into two types-

- **Systematic methods or Direct methods**—There are three subtypes of systematic methods as follows:
  - **Conformational search**—Here, the structural parameter of ligand, i.e., torsional (dihedral) translational and rotational degrees of freedom are altered gradually [10].
  - **Fragmentation**—Multiple fragments are either dock and tried to interface them with bonds, or anchor fragments is

| Program | Current version | Needs ligand set up | Needs protein set up | Input format | Method for conformational search of the ligand | Scoring function | Can deal with receptor flexibility | Allows to dock several ligands with a single set up | Free for academic research |
|---------|-----------------|---------------------|---------------------|--------------|---------------------------------------------|-----------------|-------------------------------|---------------------------------|--------------------------|
| AutoDock | V4.2.6          | Yes                 | Yes                 | pdbqt, pdbq mol2, pdb | Stochastic (GA), Stochastic (MC) | AMBER derived (FF) | Yes                           | No                                   | Yes                      |
| GOLD™   | V5.2 Suite      | Yes                 | Yes                 | sd (lig), mol (lig) mol2 (lig/prot) PDB (lig/prot) | Stochastic (GA) | GOLD score (FF) Chem score (EM) User defined score | Upto 10 user defined residues | Yes                      | No                       |
| eHITS   | V6.2            | No                  | No                  | mol sd/sdf, pdb mol2, tma tmb | Systematic (F) | eHITS score hybrid KB-EM user trained score | No                                   | Yes                      | Yes                      |
| MVD™    | V5.5            | No                  | No                  | pdb, mol2 Mol, sd/sdf Mol, omdml | Stochastic (TS), Stochastic (GA) | Mol dock score (EM) Mol dock score grid (EM) | Yes (only with the mol dock score (grid) | Yes                      | No                       |
| Argus lab | V 4.0.1         | Yes                 | Yes                 | pdb, mol, xml, agl, mol2, log | Stochastic (GA) | A score (FF) | Yes                                   | No                      | Yes                      |
| Glide™  | V4.5            | Yes                 | Yes                 | sd, pdb Maestro (mae) mol2 | Systematic pose generation with stochastic optimization | Sp 4.5 Glide score (EM) Xp 4.5 glide score (EM) | Yes through prime™ | Yes                      | No                       |

Abbreviations: EM = empirical, F= fragmentation method, FF = force field, GA = genetic algorithm, KB = knowledge-based, MC = Monte Carlo, TS: tabu search.
2.4.2. Scoring functions

Scoring functions helps in the evaluation of which ligand configuration and rotation is most favorable with respect to the receptor (protein) and by using virtual screening ranked them (ligand) according to their binding affinity [12]. The scoring function is categorized into four major groups:

- **Force field-based**—It measures the binding affinity by adding the contribution of bond-like angle bonding, and torsional deviation and non-bonded interaction like van der Waal forces, hydrogen bonding, or Coulomb electrostatics in an ace capacity [13]. Tools—GoldScore, AutoDock, DOCK, etc.

- **Empirical function**—It depends on multiple linear relapse investigation of a preparation set of complex structures by utilizing protein–ligand complex with known binding affinities, containing functional groups and sort of interaction. For example, N-O hydrogen bond, O-O hydrogen bond, salt scaffold, aromatic ring stacking, etc. [14]. Tools—LUDI score, ChemScore, AutoDock scoring, etc.

- **Knowledge-based**—It gives separation ward pair potential to particles, functional groups by statistical analyzing of a set of complex structures [15]. Tools—PMF, DrugScore, etc.

- **Consensus**—It basically consolidates the scores or rankings got from various scoring capacities in different ways.

2.5. Software’s for Docking

There are various softwares used for docking and some of the important ones are mentioned as follows:

2.5.1. Dock (http://dock.compbio.ucsf.edu/)

The Dock is a structure-based design developed in the 1980s by Irwin Kuntz and co-worker, University of California. Firstly Dock version 1 (Dock 1) calculated the quality of ligand and receptor complex on steric overlap [16]. Later on, Dock 2 provides more sophisticated control on the sampling algorithm, which provides better timing and accuracy [17]. Dock 3 includes an extremely thorough and careful parameter for an input of ligand and receptors. The first virtual screening of enzyme thymidylate synthase was performed by using Dock 3 [18]. Dock 4 was developed with an improved graph matching algorithm for ligand orientation [19]. Dock 6 (version 6.0–6.8) is being used for performing a docking with improved sampling, scoring, optimization, and numerous bug fixes [20]. Latest version Dock 6.9 provides new ligand searching method DOCK_DN, i.e., De Novo design using program based assembly [21].

2.5.2. Autodock (http://autodock.scripps.edu)

AutoDock is a broadly utilized non-business docking program. It was the first docking programming that ties ligand with full compliance adaptability. The Autodock was created by Prof. Arthur J. Olson at the Department of molecular biology of The Scripps Research Institute in La Jolla (CA, USA). The software comprises two consecutively connected programs, i.e., AutoDock and AutoGrid [22].

AutoGrid is utilized to compute the non-covalent vitality of interaction between the firm piece of the receptor and a probe atom that is available in the ligands that will be docked for the receptor. The principle job of AutoDock is to control the docking procedure of the chosen ligand through the cross section volume.

There are several versions of AutoDock like AutoDockFR that re-enacts fractional receptor adaptability by permitting countless unequivocally indicated receptor side-chains to investigate their conformational space while scanning for vivaciously positive restricting postures for a given ligand. Researcher made higher docking progress rates by utilizing receptor adaptability in the binding site of receptor adaptations that are tentatively decided without the ligand present, for example, Apo conformations [23].

AutoDock Vina can achieve two orders of magnitude speedup in comparison to AutoDock 4. With the help of AutoDock Vina, accuracy level increases twice as compared to previously published software and naturally ascertains the framework maps and bunches the outcomes in a manner straightforward to the client [24].

2.5.3. Argus lab 4.0.1 (http://www.arguslab.com)

Argus lab is a molecular modeling software developed by Mark Thomson, Department of Energy at Pacific Northwest National Laboratory, USA. Argus lab is based on algorithms for modeling solvent effects by combining quantum mechanics with classical mechanics. This program can be used for drug designing, generating graphics, and molecular modeling.
2.5.4. Genetic optimization for ligand docking (GOLD™) (http://www.ccdc.cam.ac.uk/products/life_sciences/gold)

GOLD’s principle qualities are that: (i) spine and side chain adaptability can be incorporated into the computations, (ii) the program utilizes client characterized scoring capacities and can adjust likewise, (iii) the vitality capacities depend on both conformational and non-reinforced contact data. (iv) an assortment of imperative alternatives can be characterized for the docking, (v) crystallographic water particles in the ligand binding site can be well-out during the docking, (vi) it can handles metal atoms consequently on the off chance that they are set up effectively in the protein information record, and (vii) virtual screening high throughput screening results can be broke down and post-prepared effectively with the sidekick programs SILVER™ or GoldMine™.

Latest version of GOLD Suite 5.2 is a combination of three components, i.e., Gold 5.2 to carry out protein–ligand docking, Hermes 1.6 for comprehensive protein visualisation, and Gold Mine 1.5 for analysis of docking grades.

2.5.5. MVD (Molecule & allegro virtual docker)

MVD™ (MolDock) is developed by Molegro Aps [25]. MVD’s special qualities are: (i) it automatically allocates charges, bond orders, hybridization, and add hydrogen to the given structures, (ii) it naturally foresees potential binding destinations in the receptor, (iii) it manages receptor side chain adaptability by considering incited fit interaction, (iv) it docks in predetermined vitality lattices (which accelerates the figuring’s), (v) it manages client characterized imperatives during docking, (vi) it can profit by the utilization of layouts (for example, pharmacophores) during docking, and (vii) it can convey the computations on various PCs.

2.6. Applications of Molecular Docking

- The investigation of the movement space is finished by an inspecting calculation and the steadiness of confirmation of the complex is assessed by utilizing a scoring or vitality work that measures the binding affinity of the complex.
- Drug configuration dependent on the peptides or peptidomimetics is quickly picking up footing in the pharmaceutical business. These compounds are getting to be prominent as a result of their low lethality and high explicitness. Enthusiasm for these compounds has likewise expanded with the advancement of refined assembling methods. The quantity of peptides approved by the United States Food and Drug Administration is expanding at a yearly pace of 8% and it is anticipated that the market for peptide-based medications will be tremendous.
- In expansion to pharmacodynamics information, e.g., strength, proclivity, viability, selectivity, pharmacokinetic properties, ADMET: absorption, distribution, metabolism, excretion, and toxicity have likewise been concentrated through the use of these techniques.
- Covalent medications have exhibited to be lucky options in a few remedial zones, for example, malignant growth, diabetes, and contagious, cardiovascular, gastrointestinal and neurologic ailments. Recent studies suggest that around 33% of enzyme modulators present in the market are covalent inhibitors. Covalent ligands irreversibly inactivate their target due to which recuperation of the biological functions requires the re-synthesis of the target molecule [27].

3. MOLECULAR DOCKING AS A TOOL FOR IDENTIFYING POTENTIAL DRUGS TO PREVENT SICKLE CELL ANEMIA

Sickle cell anemia is an autosomal recessive hereditary disease caused due to alteration in hemoglobin gene due to which the red blood cells (RBCs) become sickle/bow-shaped rather than round-shaped (Fig. 1) [28]. Sickle cell disease is known to affect the population living mostly in Africa, India, the Mediterranean, and Middle Eastern [29,30]. Blood of sickle cell patients contains a surprisingly enormous number of juvenile cells and many long, lean, bow-shaped erythrocytes that look like the sharp edge of a sickle. At the point when hemoglobin from sickle cells (Hbs) is deoxygenated, it brings about polymerization and misshapes the RBCs into a sickle shape [31].

Figure 1: Difference between normal and sickle red blood cells.
Nucleotide sequencing of β-globin mRNA from the sickle β-globin gene uncovered that the typical codon GAG at position β6 has been replaced by GUG [32]. This change decides the addition of valine at this position rather than glutamic acid as in the case of normal HbA [33]. The sickness and characteristic appear in individuals of African lineage, Mediterranean nations, India, and the Middle-East however once in a while in individuals of European and white lineage also. People suffer from many health problems such as anaemia, bacterial infection, brain damage, thrombosis, liver damage, strokes, etc. [34].

3.1. Antisickling Property in Plants

Various plant species show antioxidant treatment vital for the counteractive action of thick cell arrangement and free radicals preventing oxidative cell damage leading to prolongation of red blood cell life [35]. Table 2 enlists some of the important plants used in cure and prevention of sickle cell disease.

3.2 Carica papaya Linn.

Carica papaya is a long herbaceous plant, with plentiful smooth latex coming to 16–20 ft. in stature, the stem up to 30 cm thick, evergreen, erect, a perennial plant in habitat (Fig. 2). Carica papaya Linn. belongs to family Caricaceae which is known for its medicinal property like antimicrobial, antibacterial, antiulcer, anticancer, antioxidant, and antifungal. Papaya is being used since the past time for treating jaundice and sickle cell disease by the tribes of Nigeria. Many scientists reported that ripe fruit and leaf of C. papaya are used for treating sickle cell in humans [42–44].

3.3. Sickle Cell Target Protein (2HBS Protein)

A computational approach for drug discovery is done by the protein–ligand docking, in which an active site of a protein is taken as the target and inhibited by a ligand that may be a synthetic or a natural compound as a drug. There are several well-known target sites for drug-resistant treatment/control. Here, the target protein is employed for the evaluation and effectiveness of plant chemical compound obtained from C. papaya.

4. METHOD FOR DOCKING OF SICKLE CELL PROTEIN WITH AN ANTISICKLING AGENT

The discussed antisickling compounds and target protein were taken for docking study; these were downloaded from PubChem and PDB. Downloaded 3D protein structure and compound's relevant information are listed in Table 3. Furthermore, conversion of the (.sdf) file to the (.pdb) file of phyto-compounds was performed in online SMILE translator and the removal of the

Table 2: Phytocompound used in managing sickle cell disease and its modes of action [36].

| Herbs | Phytocompound and its modes of action |
|-------|--------------------------------------|
| Fagara zanthoxyloides (root) | Three isomeric divanilloylquinic acids (burkinabin B, A, and C) were recognized with possible antisickling properties. Some workers have reported the antisickling properties of vanillic acid, coumarins, paraflurobenzoic acid, and parahydroxybenzoic. |
| C. papaya (unripe fruit and leaf) | Researchers have discovered that by fermenting unripe fruit of C. papaya at 2.5 mg/ml of water shows antisickling impacts of 87% inhibitory and 74% inversion action. They also found that methanol concentrate had 64% inhibitory and 55% inversion action while the chloroform concentrate was dormant. Phenylalanine, tyrosine, and glycine were believed to be capable of showing antisickling properties [37]. |
| Hymenocardia acida (leaf) | Scientists have related the anti-sickling action of anthocyanins present in H. acida [38]. |
| Cajanus cajan (seed) | Phenylalanine is an active ingredient of C. cajan seed which is used in the preparation of Ciklavit—an antisickling phytomedicine, developed in Nigeria [39]. |
| Khaya senegalensis (stem bark/leaf) | Some workers ascribed the antisickling impacts of limonoids found in K. Senegalensis [40]. |
| Niprisan: | The drug Niprisan was developed in Nigeria. Its ingredients Sorghum bicolor and Pterocarpus osun are rich source of red/orange flavonoids and possibly act as hematronics. It had been assumed that active component of Niprisan such as Pterocarpus guineense and clove helps in moderate, conceal, or lessen the recurrence of sickle cell anemia [41]. |

Figure 2: Carica papaya plant.
Table 3: Name of the phytochemicals with their 3D structure and docking scores against sickle cell protein (2HBS).

| Sl. no. | Compound name | Ligand structure | Docking score (kcal/Mol) & Information |
|---------|---------------|------------------|----------------------------------------|
| 1       | Glycine       | ![Glycine](image) | -5.58397 kcal/mol                      |
|         |               |                  | Compound ID: 750                        |
|         |               |                  | Molecular Weight: 75.067 g/Mol          |
|         |               |                  | Molecular Formula: C₂H₅NO₂                |
|         |               |                  | H-Bond Donor: 2                          |
|         |               |                  | H-Bond Acceptor: 3                       |
| 2       | Tyrosine      | ![Tyrosine](image) | -6.99467 kcal/mol                       |
|         |               |                  | Compound ID: 6057                        |
|         |               |                  | Molecular Weight: 181.191 g/mol          |
|         |               |                  | Molecular Formula: C₉H₁₁NO₃              |
|         |               |                  | H-Bond Donor: 3                          |
|         |               |                  | H-Bond Acceptor: 3                       |
| 3       | Tryptophan    | ![Tryptophan](image) | -8.06312 kcal/mol                       |
|         |               |                  | Compound ID: 6305                        |
|         |               |                  | Molecular Weight: 204.229 g/mol          |
|         |               |                  | Molecular Formula: C₁₁H₁₂N₂O₂             |
|         |               |                  | H-Bond Donor: 3                          |
|         |               |                  | H-Bond Acceptor: 4                       |
| 4       | Phenylalanine | ![Phenylalanine](image) | -8.2888 kcal/mol                       |
|         |               |                  | Compound ID: 6140                        |
|         |               |                  | Molecular Weight: 165.192 g/mol          |
|         |               |                  | Molecular Formula: C₉H₁₁NO₂              |
|         |               |                  | H-Bond Donor: 2                          |
|         |               |                  | H-Bond Acceptor: 3                       |
| 5       | Carpane       | ![Carpane](image) | -9.23601 kcal/mol                       |
|         |               |                  | Compound ID: 442630                      |
|         |               |                  | Molecular Weight: 478.718 g/Mol          |
|         |               |                  | Molecular Formula: C₂₈H₅₀N₂O₄            |
|         |               |                  | H-Bond Donor: 2                          |
|         |               |                  | H-Bond Acceptor: 6                       |
| 6       | Anthraquinone | ![Anthraquinone](image) | -9.07431 kcal/mol                       |
|         |               |                  | Compound ID: 6780                        |
|         |               |                  | Molecular Weight: 208.216 g/mol          |
|         |               |                  | Molecular Formula: C₁₄H₈O₂                |
|         |               |                  | H-Bond Donor: 0                          |
|         |               |                  | H-Bond Acceptor: 2                       |
| 7       | Formic Acid-2 | ![Formic Acid-2](image) | -5.70657 kcal/mol                       |
|         | MethylHex3yl  |                  | Compound ID: 54515185                   |
|         |               |                  | Molecular Weight: 144.214 g/mol          |
|         |               |                  | Molecular Formula: C₁₄H₂₉O₃              |
|         |               |                  | H-Bond Donor: 0                          |
|         |               |                  | H-Bond Acceptor: 2                       |
| 8       | N-Methyl      | ![N-Methyl](image) | -6.43191 kcal/mol                       |
|         | Asparticacid  |                  | Compound ID: 852                        |
|         |               |                  | Molecular Weight: 147.13 g/mol           |
|         |               |                  | Molecular Formula: C₇H₁₄NO₃              |
|         |               |                  | H-Bond Donor: 3                          |
|         |               |                  | H-Bond Acceptor: 5                       |

(Continued)
### Table 3: (Continued)

| Sl. no. | Compound name | Ligand structure | Docking score (kcal/Mol) & Information |
|---------|---------------|------------------|---------------------------------------|
|         | Cyclohexyl    | ![Cyclohexyl Ligand](image1) | −6.46558 kcal/mol                     |
|         | Isothiocyanate| ![Isothiocyanate Ligand](image2) | −8.36549 kcal/mol                     |
|         | Phenylalaninamide | ![Phenylalaninamide Ligand](image3) |                |
|         | 2-Methoxy-4-Vinylphenol | ![2-Methoxy-4-Vinylphenol Ligand](image4) | −6.41212 kcal/mol                     |
|         | 2-[(E)-4-(4-Chlorophenyl)-2-Butenyl] malonic acid diethyl ester | ![2-[(E)-4-(4-Chlorophenyl)-2-Butenyl] malonic acid diethyl ester Ligand](image5) | −8.50388 kcal/mol                     |
|         | (9R,10R)-1,5-dichloro-9,10-diphenylanthracene-9,10-diol | ![2-[(E)-4-(4-Chlorophenyl)-2-Butenyl] malonic acid diethyl ester Ligand](image6) | −7.4063 kcal/mol                     |
|         | 9b-(3-methylphenyl)-2,3-dihydro-[1,3]thiazolo[2,3-a]isoindol-5-one | ![9b-(3-methylphenyl)-2,3-dihydro-[1,3]thiazolo[2,3-a]isoindol-5-one Ligand](image7) | −9.02044 kcal/mol                     |
|         | Heneicosane   | ![Heneicosane Ligand](image8) | −7.24859 kcal/mol                     |
|         | N-[(4R)-4-(3-fluorophenyl)-6-oxo-4,5-dihydro-H-pyrimidin-2-yl]-3-methoxybenzamide | ![N-[(4R)-4-(3-fluorophenyl)-6-oxo-4,5-dihydro-H-pyrimidin-2-yl]-3-methoxybenzamide Ligand](image9) | −9.33596 kcal/mol                     |

(Continued)
Table 3: (Continued)

| Sl. no. | Compound name                                      | Docking score (kcal/Mol) & Information                                                                 |
|--------|----------------------------------------------------|---------------------------------------------------------------------------------------------------------|
|        | Papaverinol                                        | −7.427 kcal/mol  
Compound ID: 275192  
Molecular Weight: 355.39 g/mol  
Molecular Formula: $C_{20}H_{21}NO_5$  
H-Bond Donor: 1  
H-Bond Acceptor: 6 |
|        | Hexasiloxane, Dodecamethyl                         | −6.0461 kcal/mol  
Compound ID: 71338303  
Molecular Weight: 430.941 g/mol  
Molecular Formula: $C_{12}H_{38}O_5S_6$  
H-Bond Donor: 0  
H-Bond Acceptor: 5 |
|        | (E)-2-[[4-(2-imidazol-1-ylethoxy)phenyl]methyl]-3-propan-2-ylbut-2-enedioic acid | −6.223 kcal/mol  
Compound ID: 90184972  
Molecular Weight: 358.394 g/mol  
Molecular Formula: $C_{19}H_{22}N_2O_5$  
H-Bond Donor: 2  
H-Bond Acceptor: 6 |
|        | 2-[(4R)-6-fluoro-2,2,4-trimethyl-3,4-dihydroquinolin-1yl]acetohydrazide | −7.672 kcal/mol  
Compound ID: 100207507  
Molecular Weight: 265.332 g/mol  
Molecular Formula: $C_{14}H_{20}FN_3O$  
H-Bond Donor: 2  
H-Bond Acceptor: 4 |
|        | Etretinate                                          | −8.772 kcal/mol  
Compound ID: 5282375  
Molecular Weight: 354.49 g/mol  
Molecular Formula: $C_{23}H_{30}O_3$  
H-Bond Donor: 0  
H-Bond Acceptor: 3 |
|        | Xanthoangelol D                                    | −10.5994 kcal/mol  
Compound ID: 11302670  
Molecular Weight: 354.402 g/mol  
Molecular Formula: $C_{21}H_{22}O_5$  
H-Bond Donor: 3  
H-Bond Acceptor: 5 |
|        | Oleic Acid                                          | −8.38791 kcal/mol  
Compound ID: 445639  
Molecular Weight: 282.468 g/mol  
Molecular Formula: $C_{18}H_{34}O_2$  
H-Bond Donor: 1  
H-Bond Acceptor: 2 |
|        | D-Glucitol, Hexaacetate                            | −7.8862 kcal/mol  
Compound ID: 23613  
Molecular Weight: 434.394 g/mol  
Molecular Formula: $C_{18}H_{22}O_{12}$  
H-Bond Donor: 0  
H-Bond Acceptor: 12 |
Das, et al.: Molecular docking and its applications 2020;8(01):105-116

Many workers have also reported beta carotene, Vit. C, gamma terpinene, citric acid, methionine, lycopene, tartaric acid, alanine, and sucrose have antisickling properties in inhibited by using nutritional antioxidants in combination with Vit. C and E [47]. Many workers have also reported beta carotene, Vit. C, gamma terpinene, citric acid, methionine, lycopene, tartaric acid, alanine, and sucrose have antisickling properties [48].

4.1. Target Protein Selection for Docking Attempts
The sickle cell protein, 2 Deoxyhemoglobin S (2HBS) (Fig. 3) was downloaded from Protein Data Bank and phytochemicals were taken from the PubChem database. PyMOL and Discovery studio software is used for analysis and interaction study.

4.2. Phytocompound Selection for Docking Attempts
Carica species contains various secondary metabolites including cyanogenic glycosides, amines and alkaloids, gums, fatty acids, terpenes (counting basic oils, phytosterol, triterpene genins, diterpenes, and saponins), nonprotein amino acids, hydrolyzable tannins, consolidated tannins, and flavonoids. The plant is more extravagant in glycine, carpine, anthraquinone, phenylalanine, tryptophan, and tyrosine. The plant chemical compounds like N-methyl aspartic acid, oleic acid, Hexasiloxane dodecamethyl-, 2-Methoxy-4-vinylphenol, and Cyclohexyl isothiocyanate, D-Glucitol, Hexaacetate, Phenylalanin amide, and Papaverinol, etc., are found in leaf and unripe fruits. Papaya latex contains four cysteine compounds like papain, chymopapain, glycyl endopeptidase, and caracain. Reports are there on the impact of amino acids on gelatine kinetics and solubility of sickle cells [45]. Aromatic amino acids like tyrosine, tryptophan, and phenylalanine were significantly more energetic as anti-sickling [46]. The dense cell formation due to polymerization of the sickle cell can be inhibited by using nutritional antioxidants in combination with Vit. C and E [47]. Many workers have also reported beta carotene, Vit. C, gamma terpinene, citric acid, methionine, lycopene, tartaric acid, alanine, and sucrose have antisickling properties [48].

4.3. Docking with 2HBS
For this study, Argus lab 4.0 was used to dock ligand (phytochemicals from C. papaya) into the active site of 2HBS protein. Argus lab has been reported as an effective software having the ability to predict bound configuration and binding energies of the ligand with macromolecular targets rapidly and precisely. Polar hydrogen atoms were added to the ligands and its non-polar hydrogen particles were combined. All the adaptable bonds were set to be adaptable. Protein–ligand docking was finished utilizing the GA strategy. The grid box with a dimension of $20 \times 20 \times 25$ and $0.400 \text{ Å}$ frameworks separating was to cover the protein binding site and suit ligands to move unreservedly. Subsequent to docking searches were finished, the best configuration was looked for the most populated bunch with the minimum binding vitality. The connection of docked protein–ligand complex configuration, including hydrogen bond and different interactions were broke down utilizing Discovery Studio Visualizer 16.1.0.

4.4 Steps Involved in Docking of 2HBS

4.4.1. Downloading of protein structures and ligands
- The sickle cell protein (2HBS) was downloaded from protein data bank using protein id and opened in the Argus Lab program. The molecule Tree view tool of 2HBS (positioned on the left part of the monitor) was expanded, plus the residue/misc folder was opened up. The selection of various ligands was taken from previously reviewed articles and then downloaded from Pubchem sites in .sdf format. Ligands were converted to .PDB format using an online smile translator.

4.4.2. Creation of ligand and binding site groups
- Sickle cell protein inhibitor was selected from “1440 LK2” in the Tree View, which appeared yellow.
- The Edit/Hide Unselected menu choice was chosen to conceal all molecules that are not chosen. The only atoms appearing on the screen ought to be the sickle cell protein inhibitor.
- Center the sickle cell protein inhibitor in the window by choosing the View/Centre Molecule in the Window menu on the toolbar.
- While the inhibitor is chosen, hydrogen atoms were included by squeezing the H key on the toolbar.
- Right side click on the “1440 LK2” in the Tree View and the “Make a Ligand Group from this Residue” choice was chosen. Argus lab will build a cluster underneath the Groups file with a similar name “1 LK2” that is of sort Ligand.
- Left side clicks on either “1440 LK2” residue in the Residues/Misc organizer. This activity will again choose the atoms of the Ligand on the monitor.
- The selected residue was copied and pasted. In the Residues/ Misc organizer in the Tree View tool, a recently featured residue with a name “2480 LK2” will be observed.
- Ligand cluster of this new residue was made in a similar way as done beforehand. Right-tapped on “2480 LK2” and “Make a Ligand Group from this Residue” was selected.
- The binding site for the ligand cluster by right clicking on the ligand cluster in the Groups folder and select the “Make a Binding Site Group for this Group” menu choice. This will create a Binding Site that comprises all residues that have at least one atom within 3.5Å from any atom in the ligand cluster.
4.4.3. Docking the ligand into the definite binding site

- Dock Setting dialog box was brought up by choosing the Calculation/Dock a Ligand menu choice on the toolbar.
- The ligand to dock in the “Ligand” drop-box was chosen. Make a point to choose the cluster named “ligand” group.
- The “Calculate Size” key was clicked and a docking box custom made to the binding site was made which appeared on the monitor.
- Made sure that “GA Dock” is the chosen docking the Calculation type is “Dock” and the Ligand is Flexible.
- Docking calculation was started by clicking on “Start” button.
- Various docking scores were shown in the screen but the best score was taken into consideration.

5. DOCKING RESULT

The results were based on free energy binding and the lowest binding energy. The phytocompounds showing the most negative binding energies were considered to have the strongest binding affinity toward sickle cell protein (2HBS). Twenty-four phytocompounds from C. papaya (Table 3) are docked against sickle cell protein, i.e., 2HBS from which three compounds showed the best results. These compounds were Xanthoangelol D; N-[(4R)-4-(3-fluorophenyl)-6-oxo-4,5-dihydro-H-pyrimidin-2-yl]-3-methoxy benzamide, and Carpaine showing minimum energy values in protein–ligand interaction against sickle cell protein (2HBS). The docking of Xanthoangelol D formed a stable complex by the formation of hydrogen bond with ALA1126, GLU1004, ALA1125, PRO1122, PRO893, HIS856, LYS896, ALA217, HIS218, SER905, ASP220, LEU219, LEU904, LEU890, and PHE889 with a binding free energy of −10.5994 kcal/mol, N-[(4R)-4-(3-fluorophenyl)-6-oxo-4,5-dihydro-H-pyrimidin-2-yl]-3-methoxy benzamide formed hydrogen bond with GLY224, PHE226, THR228, THR225, PRO1122, ALA1007, VAL1003, GLU1004, THR1001, ASP214, LYS896, HIS856, LEU890, VAL1123, ALA1126, and HIS218 with an energy of −9.39596 kcal/mol, whereas that of Carpaine to 2HBS showed THR225, ASP214, ALA1126, THR1001, PRO1122, ASP220, ASN221, HIS906, LYS896, HIS856, DEF1152, and ALA217 establishing a hydrogen bond with Carpaine with a binding energy of −9.23601 kcal/mol (Fig. 4). Several workers have used docking methods in search of phytocompounds for possible treatment of diseases or degradation of environmental pollutants. Kurjogi et al. [49] used the docking method in search of possible natural drug for Staphylococcus aureus enterotoxins (SEs) like SEA and SEB and based on the minimum binding energy concluded that 28-Norolean-12-en-3-one, a compound present in flowering plants of Lardizabalaceae family acts as a good inhibitor for SEA, while Betulin, a natural triterpene isolated from bark of birch trees will act as good inhibitor for SEB-like enterotoxins. Satapute et al. [50] docked triazole fungicide propiconazole against the superoxide dismutase (SOD) and catalase enzyme of plasmid cured Pseudomonas aeruginosa and found that propiconazole binds more efficiently with catalase as compared to SOD and can degrade propiconazole.

6. CONCLUSIONS

Docking studies were carried out in order to find the inhibitory activity of the compounds found in leaf extract and ripe fruit of C. papaya. Among the 24, three compounds showed the best docking result based on the binding energy. Docking studies and binding free energy calculations of these 24 compounds revealed that Xanthoangelol D has minimum interaction energy (−10.5994 kcal/mol), followed by N-[(4R)-4-(3-fluorophenyl)-6-oxo-4,5-dihydro-H-pyrimidin-2-yl]-3-methoxy benzamide (−9.39596 kcal/mol), and Carpaine (−9.23601 kcal/mol). Hence, the given phytocompounds with minimum binding energy can be further used in a wet lab for the production of drugs to cure or inhibit the action of sickle cell protein.

The current review has highlighted the principle and methods by which molecular docking has been applied in the identification of novel phytocompounds yet challenges still remain, particularly with exactness and scoring capacity governed by quantum mechanics. Most docking program effectively predicts the binding methods of small particles ligands with receptor binding destinations. The present algorithm does not measure the complete energy related with intermolecular interactions with adequate precision.
However, in the present panorama of medication revelation, where high steady loss rates are a noteworthy concern, properly structured virtual screening procedures are efficient, financially savvy, and beneficial options. With the help of molecular docking, we can identify promising phytochemicals that can be used as future medicine for curing diseases.

From the above the experiment, it is concluded that Sickle cell anemia has been a major health concern worldwide and has proven to claim the lives of many people yearly [51]. Till now, no cure has been successfully developed for this genetic disease. The current study was focused on an assessment of the inhibitory activity of leaf extract and ripe fruit compounds of C. papaya leaves against sickle cell protein. Among the identified 24 compounds, three compounds [Xanthoangelol D,N-[(4R)-4-(3-fluorophenyl)-6-oxo-4, 5-dihydro-H-pyrimidin-2-yl]-3-methoxy benzamide, and Carpaime] showed better binding affinity toward the sickle cell protein (2HBS). Different modes of interaction such as hydrogen bonding and other hydrophobic interactions were observed between the ligands and the 2HBS protein of sickle cell anemia. The information acquired through this study on the binding mode of phytocompounds from C. papaya and 2HBS protein will highly facilitate the synthesis and testing of these compounds as drugs for sickle cell disease. The study suggested that compounds from the C. papaya will be potent drug candidates against sickle cell anemia.

ACKNOWLEDGMENT
The authors DRD, DK, and BPD are grateful to the DBT-BIF (Govt. of India) for funding and the authorities of the Department of Biosciences and Biotechnology, Fakir Mohan University, Balasore for providing laboratory facilities and encouragement. DK acknowledges UGC DS Kothari post doctoral scheme for his fellowship.

AUTHOR’S CONTRIBUTION
DK and BPD developed the concept. DRD, DK, and PK wrote the manuscript.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

REFERENCES
1. Gohike H, Klege G. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. Angew Chem Int Ed 2002;41(15):2644–76.
2. Kitchen DB, Decorneze H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. Nat Rev Drug Discov 2004;3(11):935.
3. Warren GL, Andrews CW, Capelli AM, Clarke B, LaLonde J, Lambert MH, et al. A critical assessment of docking programs and scoring functions. J Med Chem 2006;49(20):5912–31.
4. Shun TY, Lazo JS, Sharlow ER, Johnston PA. Identifying active from HTS data sets: practical approaches for the selection of an appropriate HTS data-processing method and quality control review. J Biomol Screen 2011;16(1):1–4.
5. Bordogna A, Pandini A, Bonati L. Predicting the accuracy of protein-ligand docking on homology models. J Comput Chem 2011;32(1):81–98.
6. Pujadas G, Vaque M, Ardevol A, Blade C, Salvado MJ, Blay M, et al. Protein-ligand docking: a review of recent advances and future perspectives. Curr Pharm Anal 2008;4(1):1–9.
7. Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE. A geometric approach to macromolecular-ligand interactions. J Mol Biol 1982;161(2):269–88.
8. Hermann JC, Marti-Arbona R, Fedorov AA, Fedorov E, Almo SC, Shoichet BK, et al. Structure-based activity prediction for an enzyme of unknown function. Nature 2007;448(7155):775.
9. Kalyanaraman C, Imker HJ, Fedorov AA, Fedorov EV, Glasner ME, Babbitt PC, et al. Discovery of a dipeptide epimerase enzymatic function guided by homology modeling and virtual screening. Structure 2008;16(11):1668–77.
10. Sousa SF, Fernandes PA, Ramos MJ. Protein–ligand docking: current status and future challenges. Proteins Struct Funct Bioinformatics 2006;65(1):15–26.
11. Jones G, Willlett P, Glen RC, Leach AR, Taylor R. Development and validation of a genetic algorithm for flexible docking. J Mol Biol 1997;267(3):727–48.
12. Huang SY, Grinter SZ, Zou X. Scoring functions and their evaluation methods for protein–ligand docking: recent advances and future directions. Phys Chem Chem Phys 2010;12(40):12899–908.
13. Englebienne P, Moitessier N. Docking ligands into flexible and solvated macromolecules. 5. Force-field-based prediction of binding affinities of ligands to proteins. J Chem Inf Model 2009;49(11):2564–71.
14. Murray CW, Auton TR, Eldridge MD. Empirical scoring functions. II. The testing of an empirical scoring function for the prediction of ligand-receptor binding affinities and the utility of Bayesian regression to improve the quality of the model. J Comput Aided Mol Des 1998;12(5):503–19.
15. Huang SY, Zou X. An iterative knowledge-based scoring function to predict protein–ligand interactions: II. Derivation of interaction potentials. J Comput Chem 2006;27(15):1866–75.
16. Desjarlais RL, Sheridan RP, Seibel DL, Dixon JS, Kuntz ID, Venkataraghavan R. Using shape complementarity as an initial screen in designing ligands for a receptor binding site of known three-dimensional structure. J Med Chem 1998;31(4):722–9.
17. Shoichet BK, Kuntz ID. Protein docking and complementarity. J Mol Biol 1991;221(1):327–46.
18. Goodford PJ. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. J Med Chem 1985;28(7):849–57.
19. Ewing TJ, Makino S, Skillman AG, Kuntz ID. DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. J Comput Aided Mol Des 2001;15(5):411–28.
20. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. J Comput Chem 2004;25(9):1157–74.
21. Allen WJ, Fochtman BC, Balus BC, Rizzo RC. Customizable de novo design strategies for DOCK: application to HIV gp41 and other therapeutic targets. J Comput Chem 2017;38(30):2641–63.
22. Morris GM, Goodsell DS, Huey R, Olson AJ. Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4. J Comput Aided Mol Des 1996;10(4):293–304.
23. Ravindranath PA, Forli S, Goodsell DS, Olson AJ, Sanner MF. AutoDockFR: advances in protein-ligand docking with explicitly specified binding site flexibility. PLoS Comput Biol 2015;11(12):1004586.
24. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 2010;31(2):455–61.
25. Thomsen R, Christensen MH. MolDock: a new technique for high-accuracy molecular docking. J Med Chem 2006;49(11):3135–21.
26. Dhanik AR, McMurray JS, Kavraki LE. DINC: a new AutoDock-based protocol for docking large ligands. BMC Struct Biol 2013;13(1):S11.
27. Ferreira L, dos Santos R, Oliva G, Andricopulo A. Molecular docking and structure-based drug design strategies. Molecules 2015;20(7):13384–421.

28. Bunn HF. Pathogenesis and treatment of sickle cell disease. N Engl J Med 1997;337(11):762–9.

29. Bownas J. Genetic profile: sickle cell anemia. National Institutes of Health Publication, Maryland, pp 96–4057, 2000.

30. Buchanan GR, DeBaun MR, Quinn CT, Steinberg MH. Sickle cell disease. ASH Educ Program Book 2004;2004(1):35–47.

31. Pauling L, Itano HA, Singer SJ, Wells IC. Human beta-globin messenger RNA. III. Nucleotide sequences derived from complementary DNA. J Biol Chem 1977;252(14):5040–53.

32. Perutz MF, Lehmann H. Molecular pathology of human haemoglobin. Nature 1968;219(5157):902.

33. Kurjogi M, Satapute P, Jogaiah S, Abdelrahman M, Daddam J, Ramu V, et al. Computational modeling of the staphylococcal enterotoxins and their interaction with natural antitoxin compounds. Int J Mol Sci 2018;19(1):133–45.

34. Fall AB, Vanhaelen-Fastré R, Vanhaelen M, Lo I, Toppet M, Denfer A, et al. In vitro antitoxic activity of a rearranged limonoid isolated from Khaya senegalensis. Planta Med 1999;65(03):209–12.

35. Das DR, Kumar D, Kumar P, Dash BP. Molecular docking and its application in search of antitoxic agent from Carica papaya. J Appl Biol Biotech 2020;8(01):105–116. DOI: 10.7324/JABB.2020.80117