

REVIEW ARTICLE

AN OVERVIEW OF MOLECULAR DOCKING: PRINCIPLE, AND ITS APPLICATION IN DRUG DISCOVERY

Heairy Ghaghara¹, Aayushi Agarwal Bansal^{1*}

¹School of Pharmaceutical Sciences, Atmiya University, Rajkot, India.

Corresponding Author: Aayushi Agarwal Bansal Orcid Id: 0000-0002-3721-6169

Email: aayushi0709@gmail.com

Received on: 20/08/2025, Accepted on: 26/09/2025, Published on: 29/12/2025

ABSTRACT

Structure-based drug discovery utilises molecular docking as a computational tool to forecast the optimal binding conformation and affinity of specific ligands within the active site. Virtual screening, lead optimisation, and rational design of novel therapeutic agents are facilitated by this technique through the use of molecular recognition processes. Molecular docking is based on the principle of searching for suitable ligand orientations and conformations within the binding pocket using search algorithms, and then scoring functions are used to evaluate these poses in order to estimate binding energies derived from intermolecular interactions. The use of molecular docking in drug discovery is highly beneficial, particularly in structure-activity relationship studies, polypharmacology assessments, and pharmacophore modelling, but it presents challenges such as restricted receptor flexibility modelling and simplified treatment of solvent effects. These limitations are being increasingly addressed by the continuous evolution of computational methodologies, scoring functions, and integration with machine learning. This article provides a comprehensive overview of the molecular docking process, its underlying principles, and its broad applications in modern drug discovery pipelines.

Keywords: Molecular Docking, Scoring function, Virtual Screening, Lead Optimisation, Applications, Softwares.

INTRODUCTION OF MOLECULAR DOCKING

Molecular docking is a computational modelling technique that anticipates the optimal arrangement of ligands, small molecules, when they are bound to corresponding target macromolecules (such as proteins, enzymes or nucleic acids) to form 'a stable complex' as per Fig.1 [1]. Understanding molecular recognition processes that support biological functions and therapeutic interventions requires the ability to predict interactions.. At its core, molecular docking simulates the interaction between two molecules by examining the binding

modes of a given ligand at defined receptor sites. Primarily, the goal is to determine the most energetically effective conformation and orientation (pose) of the ligand while computing the binding affinity of any complex produced. Structure-based drug design (SBDD) involves molecular docking, which provides information about how small molecules interact with biological targets at the molecular level. Docking's computational efficiency, which anticipates these interactions and streamlines the screening of large compound libraries for experimental drug discovery means that there is less time, cost, or available resources.

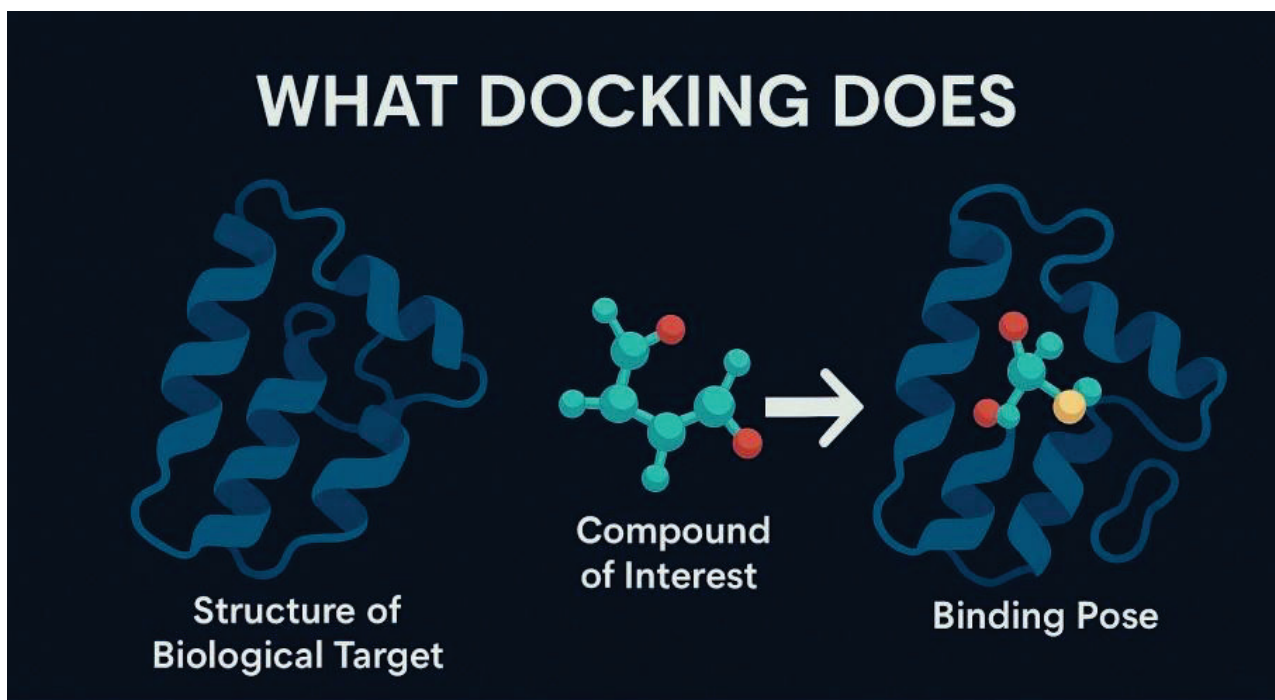


Figure 1: Molecular docking: predicting how a compound of interest fits into the structure of a biological target, determining its optimal binding pose for interaction.

Principle of Molecular Docking

The principle of molecular docking is based on the “lock and key” or “induced fit” hypothesis:

- The “lock and key” model suggests molecular interactions arise through exact geometric complementarity between receptor and ligand, similar to a key fitting within a particular lock, as per Fig.2 [2].

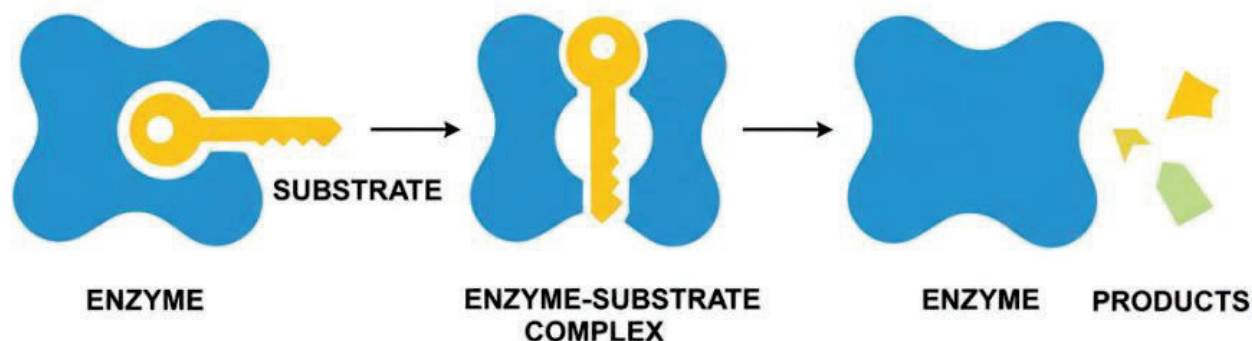


Figure 2: Enzyme action mechanism with the model of "lock and key" to depict how a particular substrate (key) corresponds with the active site (lock) of the enzyme to produce an enzyme-substrate complex, upon which the enzyme facilitates the reaction and releases the products formed without any change, ready for reuse.

- The “induced fit” model suggests that when ligands bind to proteins, conformational changes occur that result in the best possible complex formation, as per Fig. 3 [3].

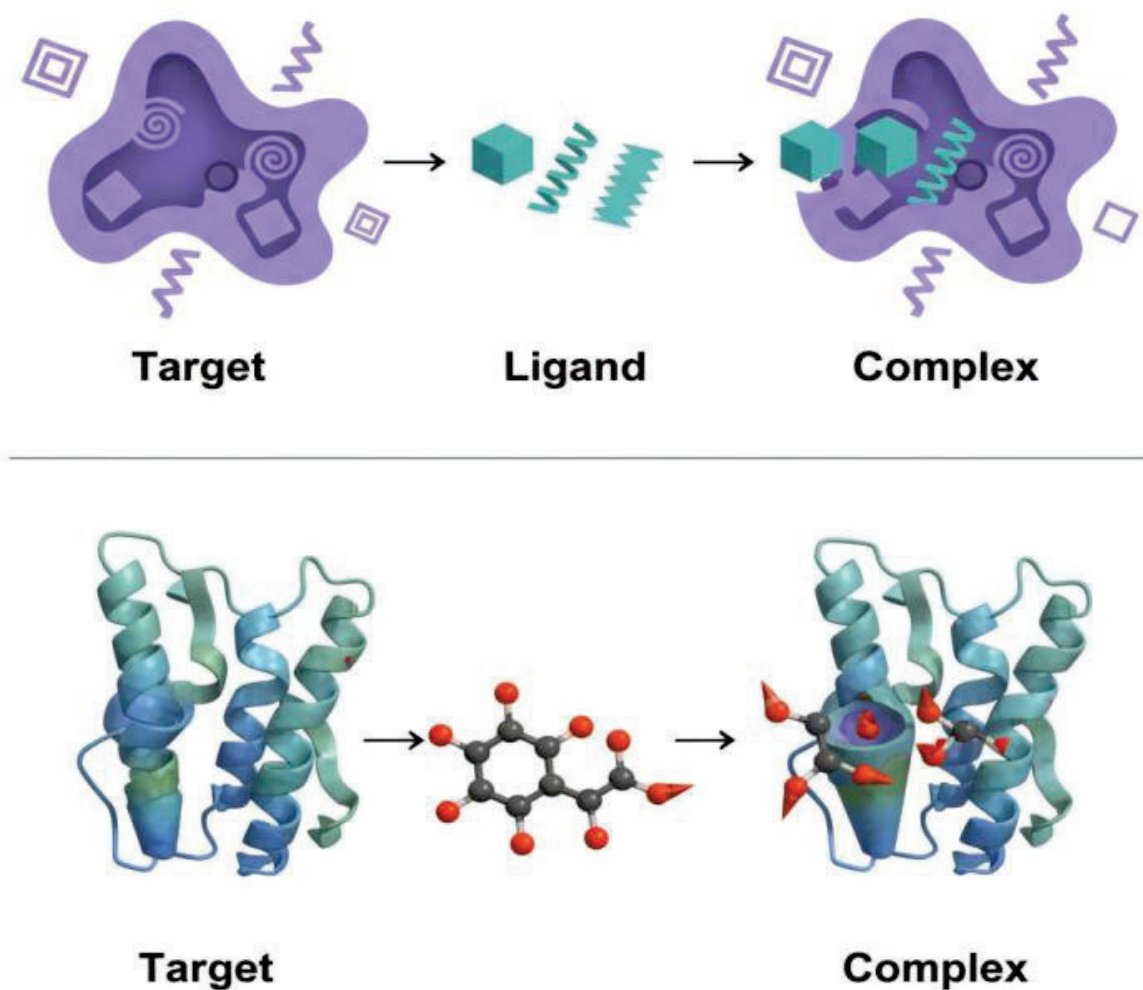


Figure 3: Process of molecular docking, in which a ligand (e.g., a drug molecule) binds specifically to a biological target (e.g., a protein). The top figure illustrates this process schematically, demonstrating the ligand inserting into the target site to create a target-ligand complex. The lower half illustrates a more realistic image: a ligand (represented as a ball-and-stick model) binds to and approaches a protein target (illustrated as a ribbon structure), forming a stable protein-ligand complex that is necessary for molecular recognition in drug discovery and inhibition of enzymes.

Molecular Docking involves two major components:

Search Algorithm

A search algorithm explores the possible orientations, conformations, and positions of a ligand within a defined binding pocket. The flexibility of the ligand (and sometimes the receptor) allows multiple possible docking poses to be generated [4].

Common search algorithms include systematic search to identify exhaustively evaluate all possible positions and conformations. Random Search (Monte Carlo) is to identify random orientations and conformations through sampling. The evolution of genetic algorithms involves mimicking natural selection processes to evolve the most suitable binding positions. Simulated Annealing enables the gradual cooling of a system to reach optimum global energy efficiency. The method of particle swarm optimisation employs stochastic estimation with a population model. Depending on the system's complexity and precision, these algorithms are often chosen for their computational efficiency and accuracy.

Scoring Function

Once possible poses are generated, a scoring function ranks them by estimating the strength and stability of the ligand-target interaction. These scoring functions approximate the binding free energy based on physicochemical interactions [5].

Types of scoring functions:

Utilise molecular mechanics force fields from disciplines like electrostatics and van der Waals to compute interaction energies using field-based techniques. Utilise weighted terms based on experimental data for scoring purposes. Obtain statistical data from known protein-ligand complexes to gain knowledge. The use of consensus scoring combines various scoring functions to improve precision.

Docking tools, including Auto Dock and Auto dock Vina algorithms as well as GOLD/GLIDE and MOE derivatives use different combinations of these algorithms and scoring functions [6-7].

The Role of Molecular Docking in Drug Discovery

Molecular docking has widespread applications throughout the drug discovery and development pipeline, also briefly summarized in (Table 1) [8].

Virtual Screening (VS)

Molecular docking is extensively used in virtual high-throughput screening (vHTS) to rapidly evaluate large compound libraries against a target. It helps identify 'hit' molecules with potential binding affinity and desirable interaction profile [9].

Lead Compound Optimisation

Docking is a crucial tool in SAR studies, where chemical modifications of lead compounds are guided by docking results to improve their binding affinity, selectivity, and pharmacokinetic properties [10].

Structure-Based Drug Design (SBDD)

The rational design of novel inhibitors or agonists can be made easier by molecular docking, which takes into account the 3D structure of the target. It predicts how small modifications in ligands affect binding modes and energies, facilitating the design of optimised drug candidates [11, 12].

Prediction of Binding Mechanism

Docking studies can predict the molecular mechanism of ligand binding by identifying key interactions such as hydrogen bonds, π - π stacking, salt bridges, and hydrophobic contacts within the binding site [13].

Drug Repurposing and Polypharmacology

Docking helps assess the ability of existing drugs to bind new targets (drug repurposing) or evaluate a drug's interaction with multiple targets (polypharmacology). This is valuable in identifying secondary activities and potential off-target effects [14, 15].

Pharmacophore Modelling

Docking results inform the development of pharmacophore models, which represent the spatial arrangement of features critical for biological activity, aiding in further screening and design of new drug candidate [16].

Protein-Protein Interaction Modulation

The challenge in discovering drugs is the ability to inhibit protein-protein interactions (PPIs). The model created by docking can help identify potential therapeutic strategies that could be used to disrupt or stabilise specific PPIs [17].

Table 1: Role of Molecular Docking in Drug Discovery

Application Area	Purpose	Application Area	Examples
Virtual screening(VS)	Rapidly screens large compound libraries against protein targets	Identifies 'hit' molecular with strong affinity and interaction potential	Discovery of inhibition for HIV-1 protease using docking based high throughput screening
Lead Compound Optimisation	Guide Structure Activity Relationship (SAR) studies	Improves selectivity, potency, and ADMET properties	Optimization of Imatinib(Gleevec) analogs to enhance binding with BCR-ABL kinase
Structure Based Drug Design(SDDS)	Uses 3D structure to design/reinforce ligands	Facilities rational design of novel inhibition/agonists	Development of Oseltamivir against influenza neuraminidase.
Prediction of Binding Mechanism	Identifies crucial ligand-protein interaction(H-Bond, salt bridge etc)	Explain molecular mechanism and guide subsequent modification	Molecular Docking predicted binding pocket interaction for Sildenafil (Viagra) with PDE5 enzyme.
Pharmacophore Modelling	Extract spatial/chemicals features necessary for activity	Improve the design of Pharmacophore - based screening librates	Docking-derived pharmacophore models helped in identifying novel COX-2 inhibitors.
Protein-Based Interaction(PPI) Modulation	Help design molecules to inhibit or stabilize protein-protein interfaces	Enable new strategies against challenging "undruggable" target	Docking used in discovering small molecules that disrupt MDM2-p53 interaction in cancer therapy.

MOLECULAR DOCKING SOFTWARE AND TOOLS

Molecular docking software development has successfully reached great heights, being bounded only by the computational power and algorithmic sophistication available. Docking systems of today consider varying theoretical and practical aspects, giving research techniques to choose from for the application [18].

AutoDock

AutoDock is one of the most popular open-source docking programs, with development from work at The Scripps Research Institute. The software uses genetic algorithms and Lamarckian genetic algorithms for conformational searching. AutoDock's scoring function combines molecular mechanics force fields with empirically weighted terms, providing good predictions of binding energies for diverse molecular systems [19].

Dock

DOCK, developed at the University of California, San Francisco, pioneered shape-fitting algorithms in molecular docking. The program uses sphere sets to represent molecular surface complementarity and includes various scoring functions, such as ChemScore and GB/SA solvation scoring. Its modular structure allows for the integration of new algorithms and scoring methods, making it especially useful for academic research.

FlexX

FlexX uses an incremental construction algorithm for ligand docking. It breaks ligands into fragments and rebuilds them within the binding site. The program includes multiple scoring functions, including FlexX-Score, PLP, and Drug Score, providing versatility in predicting binding modes. Its speed and accuracy make it well-suited for virtual screening application [20].

GOLD

GOLD (Genetic Optimization for Ligand Docking) uses advanced genetic algorithms for conformational searching. The software offers multiple scoring functions and provides many options for customizing docking protocols. GOLD excels at handling protein flexibility and generates reliable predictions for various protein-ligand complexes.

Glide

Glide implements a hierarchical series of filters to search for possible locations of ligands in the active site region of receptors. The program employs Monte Carlo sampling methods and includes specialized scoring functions optimized for different applications. Glide's systematic search algorithms provide particularly accurate results for pharmaceutical application [21].

APPLICATION OF MOLECULAR DOCKING

- **Enhanced Algorithms:** SwissDock 2024 employs Attracting Cavities 2.0 and AutoDock Vina, providing users with an option between an extremely accurate algorithm (Attracting Cavities) and a rapid algorithm (AutoDock Vina). The algorithms more accurately describe ligand-target interaction energies, such as non-covalent and covalent bindings, solvation, and conformational change, and yield more trustworthy predictions.
- **Flexible Input and Automated Preparation:** Ligands may be provided in various formats like SMILES, Mol2, or PDBQT, while targets are accepted as PDB or PDBQT files or retrieved from the primary databases (Protein Data Bank, AlphaFold). The program prepares the input molecules automatically for docking, involving protonation state modification and addition of absent atoms or side chain [22].
- **Improved User Interface and Accessibility:** SwissDock 2024 has a newly designed web interface and also offers command-line access, making it available to different user levels as well as offering batch processing or pipeline integration. Users can visualize docking results in real time in 3D and examine several ligand poses with extended molecular interactions.
- **Integration with Other Computational Tools:** Docking results can be transferred directly to allied SwissDrugDesign tools for subsequent analysis, like ADME prediction, target prediction, or similarity searches, to enable end-to-end drug discovery workflows [23].
- **Covalent Docking:** SwissDock 2024 now includes docking of covalent ligands, which establish lasting chemical bonds with target proteins, a growing mechanism of drug action, enabling the exploration of irreversible inhibitors or targeted covalent drugs.
- **Wider Target Profile:** In addition to proteins, docking is also applied to other macromolecules such as nucleic acids and polysaccharides, facilitating the exploration of a wider variety of biological targets that are of interest in antivirals, gene therapy, and glycomics.
- **Sophisticated Scoring and Search Space Definition:** Users may define very specific search boxes for docking sites graphically, change sampling exhaustivity, and give preference to binding cavities by depth and accessibility, enhancing efficiency and relevance of docking outcomes.
- **Computational Efficiency and High Throughput:** By transferring computations to server-side and algorithm optimization, SwissDock and analogous platforms minimize the demand for high local computational

resources, enabling large-scale virtual screening projects in drug discovery [24].

LIMITATIONS AND CHALLENGES

Despite its widespread use, molecular docking has certain limitations:

- While biological targets are dynamic, the protein receptor is typically viewed as rigid or semi-flexible in most docking studies.
- Many docking protocols fail to include explicit solvent molecules or simplify the process of solvent effects, which can impact ligand binding.
- Some existing scoring functions are not accurate in accurately predicting binding free energies of specific ligand classes or interactions.
- The computational cost of large-scale docking campaigns with large libraries or highly flexible ligands can be significant [25].

Emerging solutions:

- Using Molecular dynamics (MD) simulations for receptor flexibility.
- Using large datasets, scoring functions can be trained using machine learning techniques.
- To optimise energy calculations, one can employ docking and Free Energy Perturbation (FEP) or Thermodynamic Integration (TI) techniques [26].

CONCLUSION

Molecular docking has become a key part of modern drug discovery. It offers a computational way to predict how small molecules bind to biological targets like proteins. By simulating these interactions, docking helps find promising drug candidates from large chemical libraries. This process significantly reduces the time and cost of experimental screening. Docking also plays an important role in drug design, virtual screening, and lead optimisation. It provides insights into binding modes, affinities, and structure-activity relationships. These features make docking a useful tool in the early stages of developing therapies. They allow researchers to prioritise compounds for further testing and refinement.

Although it is widely used, the accuracy of molecular docking is still improving. Traditional scoring functions often do not fully consider complex molecular interactions and dynamic biological environments. Recent advances in algorithms, the inclusion of molecular flexibility, and the use of machine learning techniques are boosting docking performance. Machine learning models, trained on large datasets, help refine scoring functions and predict more accurate binding poses. As these technologies keep evolving, molecular docking is becoming more reliable and predictive. This reinforces its critical role in designing and optimising new therapeutic agents.

REFERENCES

- [1] Agarwal, S., & Mehrotra, R. J. J. C. (2016). An overview of molecular docking. *JSM chem*, 4(2), 1024-1028.
- [2] Fan, J., Fu, A., & Zhang, L. (2019). Progress in molecular docking. *Quantitative Biology*, 7(2), 83-89.
- [3] Raval, K., & Ganatra, T. (2022). Basics, types and applications of molecular docking: A review. *IP International Journal of Comprehensive and Advanced Pharmacology*, 7(1), 12-16.
- [4] Meng, X. Y., Zhang, H. X., Mezei, M., & Cui, M. (2011). Molecular docking: a powerful approach for structure-based drug discovery. *Current Computer-Aided Drug Design*, 7(2), 146–157.
- [5] Morris, G. M., Huey, R., Lindstrom, W., et al. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry*, 30(16), 2785–2791.
- [6] Ferreira, L. G., dos Santos, R. N., Oliva, G., & Andricopulo, A. D. (2015). Molecular docking and structure-based drug design strategies. *Molecules*, 20(7), 13384–13421.
- [7] Kitchen, D. B., Decornez, H., Furr, J. R., & Bajorath, J. (2004). Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature Reviews Drug Discovery*, 3(11), 935–949.
- [8] Pagadala, N. S., Syed, K., & Tuszynski, J. (2017). Software for molecular docking: a review. *Biophysical Reviews*, 9(2), 91–102.
- [9] Åqvist, J., Medina, C. and Samuelsson, J.-E. (1994). A new method for predicting binding affinity in computer-aided drug desing. *Protein Eng.* 7, 385–391.
- [10] Baker, E. N. and Hubbard, R. E. (1984). Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol.* 44: 97–179.
- [11] Bartlett, M. S. and Smith, J. W. (1991). *Penemocystis carinii*, an opportunist in immuno-compromised patients. *Clin. Microbiol. Rev.* 4, 137–149.
- [12] Edman, J. C., Edman, U., Cao, M., Lundgren, B., Kovacs, J. A. and Santi, D. V. (1989). Isoaltion and expression of the *Pneumocystis carinii* dihydrofolate reductase gene. *Proc. Natl. Acad. Sci. USA* 86: 8625–8629.
- [13] Fischer, D., Lin, S. L., Wolfson, H. L. and Nussinov, R. (1995). A geometry-based suite of molecular docking processes. *J. Mol. Biol.* 248: 459–477.
- [14] Gallant, J. E., Moore, R. D. and Chaisson, R. E. (1994). Prophylaxis for opportunistic infections in patients

- with HIV infection. *Ann. Intern. Med.* 120: 932–944.
- [15] Ghose, A. K. and Crippen, G. M. (1985). Geometrically feasible binding modes of a flexible ligand molecule at the receptorsite. *J. Comp. Chem.* 6: 350–359.
- [16] Good, A. C. and Mason, J. S. (1995). Three-dimensional structure database searches. *Rev. Comp. Chem.*, 7, 67–117.
- [17] Goodsell, D. S. and Olson, A. J. (1990). Automated docking of substrates to proteins by simulated annealing. *Proteins* 8: 195–202.
- [18] Gordin, F. M., Simon, G. L., Wofsy, C. B. and Mills, J. (1984). Adverse reactions to trimethorim-sulfamethoxazole in patients with the acquired immunodeficiency syndrome. *Ann. Intern. Med.* 100: 495–499.
- [19] Greer, J. and Bush, B. L. (1978). Macromolecular shape and surface maps by solvent exclusion. *Proc. Natl. Acad. Sci. USA* 75: 303–307.
- [20] Greer, J., Erickson, J. W., Baldwin, J. J. and Varney, M. D. (1994). Application of the three-dimensional structures of protein target molecules in structure-based drug desig. *J. Med. Chem.* 37: 1035–1054.
- [21] Gschwend, D. A. and Kuntz, I. D. (1996). Orientational sampling and rigid-body minimization in molecular docking, revisited; On-the-fly optimization and degeneracy removal. *J. Comput.-Aided Mol. Design.*
- [22] Gschwend, D. A., Sirawaraporn, W., Santi, D. V. and Kuntz, I. D. (1996). Specificity in structure-based drug design: Identification of a novel, selective inhibitor of *Pneumoystis carinii* dihydrofolate reductase.
- [23] Guida, W. C. (1994). Software for structure-basedx drug design. *Curr. Opin. Struc. Biol.* 4: 777–781.
- [24] Hart, T. N. and Read, R. J. (1992). A multiple-start Monte Carlo docking method. *Proteins* 13: 206–222.
- [25] Helmer-Citterich, M. and Tramontano, A. (1994). PUZZLE: A new method for automated protein docking based on surface shape complementarity. *J. Mol. Biol.* 235: 1021–1031.
- [26] Holloway, M. K., Wai, J. M., Halgren, T. A., et al. (1995). *A priori* prediction of activity for HIV-1 protease inhibitors employing energy minimization in the active site. *J. Med. Chem.* 38: 305–317.