GENERAL ARTICLE

International Journal of Science, Technology and Society
Vol 3, No. 1 (2017) : 01-06

Computational techniques for designing new lead molecules in the process of drug discovery

Suresh Kumar, Samiyara Begum, Hemant Kumar Srivastava

Department of Chemistry, Indian Institute of Technology Guwahati, Guwahati 781 039, INDIA

Publication Info

Article history:
Received : 30.06.2017
Accepted : 25.08.2017
DOI: https://doi.org/10.18091/ijsts.v3i01.10953

Key words:
Drug design, Molecular Docking, QSAR and Pharmacophore Modeling

Corresponding author:
Suresh Kumar

Email:
sureshkr@iitg.ernet.in

ABSTRACT

Computational techniques are important in the field of drug discovery. These techniques are generally categorized in two methods namely 'structure-based' and 'ligand-based' methods. The present review discusses the theory of the most important methods, recent successful applications, pharmacophore modeling and quantitative structure-activity relationship (QSAR) studies. A brief introduction of molecular docking methods and their development and applications in drug discovery process is also included. Basic theories and fundamental techniques including sampling algorithms and scoring functions are discussed.

INTRODUCTION

The design and development of a new drug is a challenging task as timely delivery and managing the expenditures is extremely difficult [1-5]. Although there is a noticeable increase in funding for research and development in the process of drug discovery [6-9], only a small part of the produced drugs (less than 10 percent drugs) pass the clinical trials and only a few drugs reach to market [10-12]. This generates the demand of all pharmaceutical industries and research fields to find alternative ways to increase the efficiency and productivity in the process of drug discovery [13-16]. The computational techniques like docking, molecular dynamics (MD) simulations, QSAR, virtual screening, pharmacophore modeling etc. are helpful for pharmaceutical and biotechnology based companies in designing some new drug like molecules [17-21]. In this mini review, we focus on a brief discussion on various computational techniques to understand their basic details and usefulness in the process of drug discovery [22-25]. This review also reports the impact of computational approaches in drug designing and a few available publications on docking, mapping, homology modeling etc. [26-29].

STRUCTURE-BASED DRUG DESIGN (SBDD)

The molecular modeling of SBDD is a powerful tool to study the structure-activity relationships. It is the design and optimization of a chemical structure with the aim to identify a suitable compound for chemical testing to find some new medications [30]. SBDD is based on the knowledge of 3D structure of the lead molecule. Generally, shape, size and charge of the lead molecule affects its ability to interaction with the biological target and understanding this affect is necessary to for the design and development of new drugs. If the chemical structure of the receptor is not available, it can be predicted by homology modeling techniques [31]. Homology modeling is the modeling of a protein on the basis of known amino acid sequences of a protein and the modeled protein is comparable with similar homologous protein [32].

LIGAND-BASED DRUG DESIGN (LBDD)

The method of LBDD relies on the knowledge of small molecules that bind to the target of interest. Here, the known molecular pattern that bind to a target is used to find molecules with similar patterns [33]. Some popular LBDD approaches are pharmacophore modeling, molecular

similarity approaches, quantitative structure–activity relationship etc. [34]. LBDD is generally used in cases where chemical structure of the receptor cannot be defined [35]. However, this is very useful technique and various efforts have been made to strengthen the LBDD process in last two decades.

QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIP (QSAR)

The relationship between the features of the chemical structure of ligands and their biological activity is called QSAR [36-38]. It is among the most practical and widely used tool in analogue/ligand-based drug design process. The principal objective of QSAR is to find a relationship between the biological activity and physiochemical properties for a set of ligands. Various molecular descriptors are used in QSAR to understand the representative molecular properties. 2D descriptors, independent of 3D orientation or conformation of the molecules, have some common property to reflect the structure of molecules [39, 40]. Constitutional descriptors are one of the most important 2D descriptors that reveals the molecular composition of a compound without using molecular geometry [41].

MOLECULAR DOCKING

Molecular docking is the attempt to find the "best" matching between the ligand and the receptor to form a stable complex. It can evaluate different possible conformations of ligand receptor binding, which makes it reliable for understanding the mechanism of interaction [42, 43]. Molecular docking is mainly divided into three types: (i) docking program to identify potential ligands from a library of chemical compounds, (ii) binding mode of potential ligands and (iii) prediction of binding poses [44]. In rigid molecular docking (lock and key model), the internal geometry of the receptor and ligand is fixed during docking. On the other hand, receptor is rigid and ligand is flexible in induced-fit or flexible docking. Docking algorithms can be used to find ligands, multiple proteins and their binding conformations that are close to experimentally determined structures [45]. Different types of algorithms such as molecular dynamics, monte carlo methods, genetic algorithms, fragment-based methods, point complementary methods, distance geometry methods, systematic searches

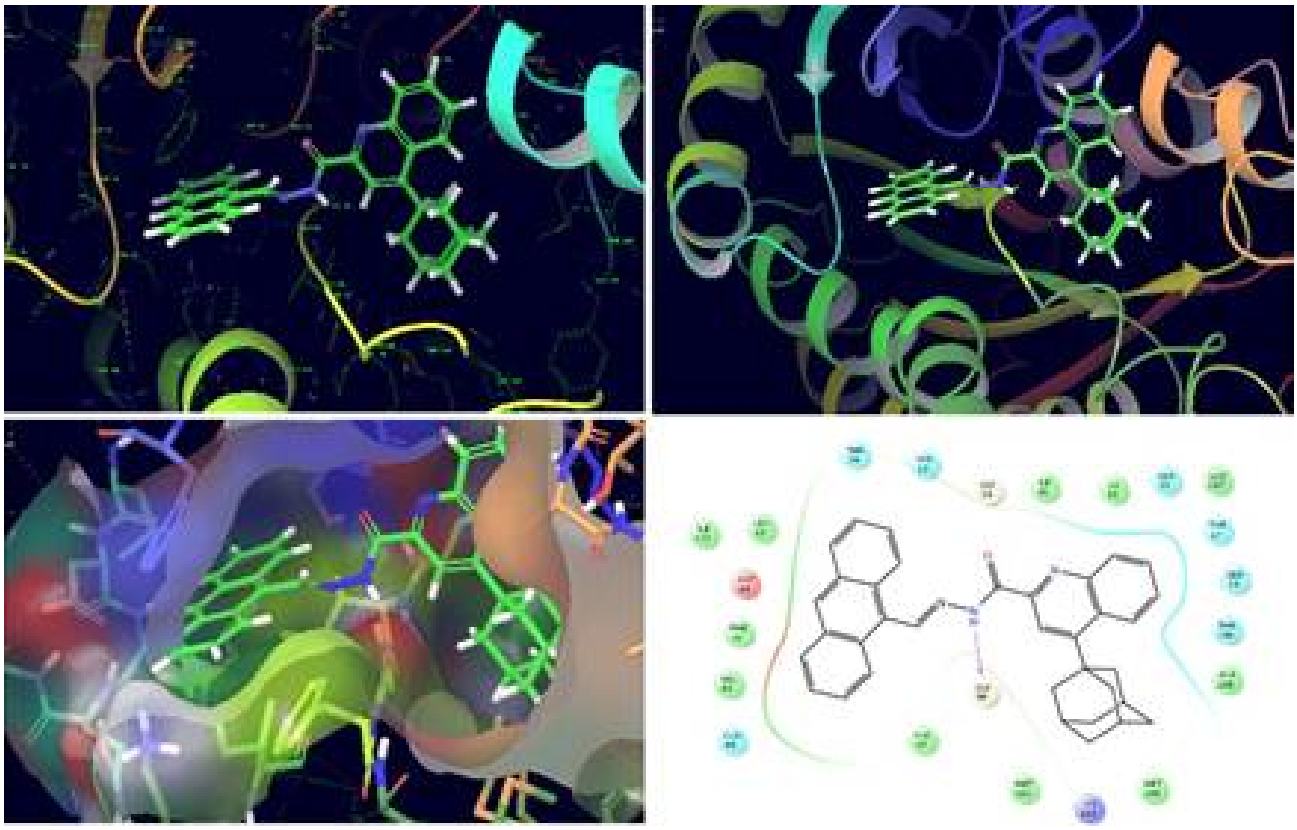

Fig. 1. Images of molecular docking studies

are frequently used for docking calculations. Figure 1 depicts the docked poses of an anti-TB drug molecule with a receptor (PDB-ID 1ZID). The ligand-receptor interaction diagram, showing the important interactions, are also included in this figure. Molecular docking consists of multiple steps as discussed below:

(a) Receptor preparation: The X-ray crystallographic 3D structure of receptor as available on Protein Data Bank (PDB) or any other reliable source is modified with stabilizing charges, filling missing residues, removal of water molecules from the cavity etc.

(b) Ligand preparation: Different conformations of ligands are generated and minimized stepwise for increased activity and selectivity, as well as drug-like properties and used for docking against target biomolecules.

(c) Receptor grid generation: Setting position constraint of the grids used for grid-based scoring in docking.

(d) Docking and scoring: Here ligand is docked with the receptor and the defined grid of receptor interactions are checked. The scoring function generates scores depending on the selected ligand with best fit.

PHARMACOPHORE MODELING

Pharmacophore modeling, as explained by Ehrlich, is a technique that analyzes the important structural features (chemical groups) responsible for the biological activity of a drug molecule. There is a considerable expansion of the concept and application range of a pharmacophore model. According to IUPAC [46], pharmacophore model defines the ensemble of steric and electronic features that are needed to ensure the optimal supramolecular interactions with a specific biological target and trigger the biological response. A pharmacophore model can be either ligand-based or structure-based manner. In ligand-based manner, a set of active molecules are superposed and extract the common chemical features that are responsible for the biological activity [47]. In structure-based manner, the possible interacting points between the biological target and the ligands are summarized. The various steps of pharmacophore modeling are:

(a) Select a training set of ligands: Selection of a set of structurally different molecules having both active and inactive compounds.

(b) Conformational analysis: Production of a set of low energy conformations to contain the bioactive conformation of the selected molecules.

(c) Molecular superimposition: Superimposition of all suitable combinations of low-energy conformations of the molecules.

(d) Abstraction: Transformation of the superimposed molecules into an abstract representation.

(e) Validation: Validation of the model to account the differences in biological activity for a range of molecules.

MOLECULAR DYNAMICS (MD) SIMULATION

Understanding of molecular motions is undoubtedly essential in any successful drug discovery process. MD simulation calculates the accurate binding affinities of ligand-receptor complex, analyzes the binding position and the stability with respect to time [48]. Recently Cavalli and coworkers summarized the role of molecular dynamics and related methods in drug discovery process [49]. This topic is of interest for various researchers in the field of drug discovery [50-53].

QUANTUM MECHANICS/MOLECULAR MECHANICS (QM/MM)

QM/MM is used to study the structural changes and the binding strength upon the formation of ligand-receptor complex. It is an advanced computational technique to perform the geometry optimization, single point energy calculation, Hessian and gradient calculations. It is useful in treating atomic complex system and has the potential to be used as an essential part in drug discovery for the identification of the lead and understand of the ligand receptor interaction in detail [54-59].

Some important terminology:

S. No.	Term	Explanation
1	Ligand	Ligand is an organic substance that forms complex with biomolecule to serve biological purpose. Ligand usually produces a signal for binding to a site on the target biomolecule.
2	Receptor	Target biomolecules act as receptors that bind to signaling ligand and initiate a physiological response. Cell membrane receptors permit signaling ligands to effect the cell function without entering the cell.

S. No.	Term	Explanation
3	Descriptor	Molecular descriptors are set of numerical or binary values representing various molecular properties of a compound. They are classified into two types i.e. 2D and 3D.
4	Active site	Active site is the specific region of an enzyme where a substrate binds and proceed to chemical reaction.
5	Pharmacophore	A pharmacophore modeling refers to the geometrical description of the chemical functionalities of a ligand to interact with the receptor.
6	Affinity	Activity is the relationship between the structure of a ligand and the receptor which suggests the strength of their binding.
7	Dataset	It is the set of biologically active molecules with ascending or descending order of activity values.
8	Training set	It is a set of data to assess the potentially predictive relationships.
9	Test set	Test set is the data set used to evaluate the strength and efficacy of a predictive relationship.

ACKNOWLEDGMENT

The authors gratefully acknowledge the financial support from SERB (DST, New Delhi India). Authors also acknowledge the research facilities at Indian Institute of Technology Guwahati, Assam.

REFERENCES

Pierce AC, Jacobs M and Stuver-Moody C (2008). Docking study yields four novel inhibitors of the protooncogene Pim-1 kinase. J Med Chem., 51: 1972-1975.

Hao WS, Hu YB, Niu CS, Huang XY, Chang CPB, Gibbons J and Xu J (2008). Discovery of the catechol structural moiety as a Stat3 SH2 domain inhibitor by virtual screening. *Bioorg* Med Chem Lett., 18: 4988-4992.

Salam NK, Huang THW, Kota BP, Kim MS, Li YH and Hibbs DE (2008). Novel PPAR-gamma agonists identified from a natural product library: A virtual screening, induced-fit docking and biological assay study. Chem Biol Drug Des., 71: 57-70.

Kenyon V, Chorny I, Carvajal WJ, Holman TR and Jacobson MP (2006). Novel Human Lipoxygenase Inhibitors Discovered Using Virtual Screening with Homology Models. J. Med. Chem., 49: 1356-1363.

Li HL, Huang J, Chen LL, Liu XF, Chen T, Zhu J, Lu WQ, Shen X, Li J, Hilgenfeld R and Jiang HL (2009). Identification of Novel Falcipain-2 Inhibitors as Potential Antimalarial Agents through Structure-Based Virtual Screening. J Med Chem., 52: 4936-4940.

Shah F, Gut J, Legac J, Shivakumar D, Sherman W, Rosenthal PJ and Avery MA (2012). Computer-aided drug design of falcipain inhibitors: virtual screening, structure–activity relationships, hydration site thermodynamics, and reactivity analysis. J. Chem. Inf. Model., 52: 696-710

Kitchen DB, Stahura FL and Bajorath J (2004). Computational techniques for diversity analysis and compound classification. Mini-Rev Med Chem., 4: 1029-1039.

Perola E, Walters WP and Charifson PS (2004). A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance. Proteins., 56: 235-249.

Stote RH, Kellenberger E, Muller H, Bombarda E, Roques BP, Kieffer B, and Mely Y (2004). Structure of the His44 '! Ala single point mutant of the distal finger motif of HIV-1 nucleocapsid protein: A combined NMR, molecular dynamics simulation, and fluorescence study. Biochemistry., 43: 7687–7697.

Sutherland JJ, Nandigam RK, Erickson JA and Vieth M (2007). Lessons in molecular recognition. 2. Assessing and improving cross-docking accuracy. J Chem Inf Model., 47: 2293-2302.

McGaughey GB, Sheridan RP, Bayly CI, Culberson JC, Kreatsoulas C, Lindsley S, Maiorov V, Truchon JF and Cornell WD (2007). Comparison of topological, shape, and docking methods in virtual screening. J Chem Inf Model., 47: 1504-1519.

Cornell WD (2007). Virtual screening methods for the identification of lead compounds for drug discovery. Abstr Pap Am Chem S., **234**.

Warren GL, Andrews CW, Capelli AM, Clarke B, LaLonde J, Lambert MH, Lindvall M, Nevins N, Semus SF, Senger S, Tedesco G, Wall ID, Woolven JM, Peishoff CE and Head MS (2006). A critical assessment of docking programs and scoring functions. J Med Chem., 49: 5912-5931.

Srivastava HK and Sastry GN (2012). Molecular Dynamics Investigation on a Series of HIV Protease Inhibitors: Assessing the Performance of MM-PBSA and MM-GBSA Approaches. J Chem Inf Model., 52: 3088-3098.

Cross JB, Thompson DC, Rai BK, Baber JC, Fan KY, Hu YB and Humblet C (2009). Comparison of Several Molecular Docking Programs: Pose Prediction and Virtual Screening Accuracy. J Chem Inf Model., 49: 1455-1474.

Stahl M and Rarey M (2001). Detailed analysis of scoring functions for virtual screening, J Med Chem., 44: 1035-1042.

Bissantz C, Folkers G and Rognan D (2000). Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. J Med Chem. 43: 4759-4767.

Cummings MD, DesJarlais RL, Gibbs AC, Mohan V and Jaeger EP (2005). Comparison of automated docking programs as virtual screening tools. J Med Chem., 48: 962-976.

Chen YZ and Zhi DG (2001). Ligand-protein inverse docking and its potential use in the computer search of protein targets of a small molecule. Proteins-Structure Function and Genetics., 43: 217-226.

Santiago DN, Pevzner Y, Durand AA, Tran M, Scheerer RR, Daniel K, Sung SS, Woodcock HL, Guida WC and Brooks WH (2012). Virtual Target Screening: Validation Using Kinase Inhibitors. J Chem Inf Model., 52: 2192-2203.

Brooks WH, Daniel KG, Sung SS and Guida WC (2008). Computational validation of the importance of absolute stereochemistry in virtual screening. J Chem Inf Model., 48: 639-645.

Huggins DJ, Sherman W and Tidor B (2012). Rational Approaches to Improving Selectivity in Drug Design. J Med Chem., 55: 1424-1444.

Paul N, Kellenberger E, Bret G, Muller P and Rognan D (2004). Recovering the true targets of specific ligands by virtual screening of the Protein Data Bank. Proteins., 54: 671-680.

Feher M and Williams CI (2012). Numerical Errors and Chaotic Behavior in Docking Simulations. J Chem Inf Model., 52: 724-738.

Corbeil CR and Moitessier N (2009). Docking Ligands into Flexible and Solvated Macromolecules. 3. Impact of Input Ligand Conformation, Protein Flexibility, and Water Molecules on the Accuracy of Docking Programs. J Chem Inf Model, 49: 997-1009.

Santos R, da Costa G, Franco C, Gomes-Alves P, Flammang P and Coelho AV (2009). First Insights into the Biochemistry of Tube Foot Adhesive from the Sea Urchin Paracentrotus lividus (Echinoidea, Echinodermata). Mar Biotechnol., 11: 686-698.

Repasky MP, Murphy RB, Banks JL, Greenwood JR, Tubert-Brohman I, Bhat S and Friesner R A (2012). Docking performance of the glide program as evaluated on the Astex and DUD datasets: a complete set of glide SP results and selected results for a new scoring function integrating WaterMap and glide. J Comput Aid Mol Des., 26: 787-799.

Schulz-Gasch T and Stahl, M. (2003) Binding site characteristics in structure-based virtual screening: evaluation of current docking tools, *J Mol Model.* **9**, 47-57.

Mason JS (2008). COMP 33-Perspectives and learnings on in silico pharmacology and biological fingerprints. Abstr Pap Am Chem S., 236.

Wang Y, Shaikh SA and Tajkhorshid E (2010). Exploring Transmembrane Diffusion Pathways With Molecular Dynamics. Physiology., 25: 142-154.

Burger A (1978). Drug Design and Development- Realistic Appraisal. J Med Chem., 21: 1-4.

Hanson S M, Newstead S, Swartz KJ and Sansom MSP (2015). Capsaicin Interaction with TRPV1 Channels in a Lipid Bilayer: Molecular Dynamics Simulation. Biophys J., 108: 1425-1434.

Vogt M and Bajorath J (2011). Introduction of the Conditional Correlated Bernoulli Model of Similarity Value Distributions and its Application to the Prospective Prediction of Fingerprint Search Performance. J Chem Inf Model., 51: 2496-2506.

Acharya C, Coop A, Polli JE and MacKerell AD (2011). Recent Advances in Ligand-Based Drug Design: Relevance and Utility of the Conformationally Sampled Pharmacophore Approach. Curr Comput-Aid Drug., 7: 10-22.

Loew GH, Villar HO and Alkorta I (1993). Strategies for Indirect Computer-Aided Drug Design. Pharm Res-Dordr., 10: 475-486.

Srivastava HK, Choudhury C and Sastry GN (2012). The Efficacy of Conceptual DFT Descriptors and Docking Scores on the QSAR Models of HIV Protease Inhibitors. Med Chem., 8: 811-825.

Srivastava HK, Chourasia M, Kumar D and Sastry GN (2011). Comparison of Computational Methods to Model DNA Minor Groove Binders. J Chem Inf Model., 51: 558-571.

Ravindra GK, Achaiah G and Sastry GN (2008). Molecular modeling studies of phenoxypyrimidinyl imidazoles as p38 kinase inhibitors using QSAR and docking. Eur J Med Chem., 43: 830-838.

Neurath AR, Strick N, Li YY and Debnath AK (2001). Cellulose acetate phthalate, a common pharmaceutical excipient, inactivates HIV-1 and blocks the coreceptor binding site on the virus envelope glycoprotein gp120. Bmc Infect Dis. 1: art. no.-17.

Karelson M, Lobanov VS and Katritzky AR (1996). Quantum-chemical descriptors in QSAR/QSPR studies. Chem Rev. 96: 1027-1043.

Pasha FA, Srivastava HK and Singh PP (2005). Semiempirical QSAR study and ligand receptor interaction of estrogens. Mol Divers., 9: 215-220.

Yuriev E, Agostino M and Ramsland PA (2011). Challenges and advances in computational docking: 2009 in review. J Mol Recognit., 24: 149-164.

Meng XY, Zhang HX, Mezei M and Cui M (2011). Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. Curr Comput-Aid Drug., 7: 146-157.

Lopez-Vallejo F, Caulfield T, Martinez-Mayorga K, Giulianotti MA, Nefzi A, Houghten RA and Medina-Franco JL (2011). Integrating Virtual Screening and Combinatorial Chemistry

for Accelerated Drug Discovery. Comb Chem High T Scr., 14: 475-487.

Huang SY and Zou XQ (2010). Advances and Challenges in Protein-Ligand Docking. Int J Mol Sci., 11: 3016-3034.

McNaught AD (1997). International Union of Pure and Applied Chemistry and International Union of Biochemistry and Molecular Biology - Joint Commission on Biochemical Nomenclature - Nomenclature of carbohydrates (Recommendations 1996) (Reprinted from Pure Appl Chem, vol 68, pg 1919-2008, 1996). Adv Carbohyd Chem Bi., 52: 43-177.

Mason JS, Good AC and Martin EJ (2001). 3-D Pharmacophores in drug discovery. Curr Pharm Design., 7: 567-597.

Mccammon JA, Gelin BR and Karplus M (1977). Dynamics of Folded Proteins. Nature., 267: 585-590.

De Vivo M, Masetti M, Bottegoni G and Cavalli A (2016). Role of Molecular Dynamics and Related Methods in Drug Discovery. J Med Chem., 59: 4035-4061.

Srivastava HK and Sastry GN (2013). Efficient estimation of MMGBSA-based BEs for DNA and aromatic furan amidino derivatives. J Biomol Struct Dyn., 31: 522-537.

Kamal A, Shankaraiah N, Reddy CR, Prabhakar S, Markandeya N, Srivastava HK and Sastry GN (2010). Synthesis of bis-1,2,3-triazolo-bridged unsymmetrical pyrrolobenzodiazepine trimers via 'click' chemistry and their DNA-binding studies. Tetrahedron., 66: 5498-5506.

Kamal A, Bharathi EV, Ramaiah MJ, Dastagiri D, Reddy JS, Viswanath A, Sultana F, Pushpavalli SNCVL, Pal-Bhadra M, Srivastava HK, Sastry GN, Juvekar A, Sen S and Zingde S (2010). Quinazolinone linked pyrrolo[2,1-c][1,4]benzodiazepine (PBD) conjugates: Design, synthesis and biological evaluation as potential anticancer agents. Bioorgan Med Chem., 18: 526-542.

Durrant JD and McCammon JA (2011). Molecular dynamics simulations and drug discovery. Bmc Biol., 9:

Gleeson MP and Gleeson D (2009). QM/MM Calculations in Drug Discovery: A Useful Method for Studying Binding Phenomena?. J Chem Inf Model., 49: 670-677.

Lodola A and De Vivo M (2012). The Increasing Role of Qm/Mm in Drug Discovery. Adv Protein Chem Str., 87: 337-362.

Zhou T, Huang DZ and Caflisch A (2010). Quantum Mechanical Methods for Drug Design. Curr Top Med Chem., 10: 33-45.

Shaik S, Cohen S, Wang Y, Chen H, Kumar D and Thiel W (2010). P450 Enzymes: Their Structure, Reactivity, and Selectivity-Modeled by QM/MM Calculations. Chem Rev., 110: 949-1017.

Altarsha M, Wang DQ, Benighaus T, Kumar D and Thiel W (2009). QM/MM Study of the Second Proton Transfer in the Catalytic Cycle of the D251N Mutant of Cytochrome P450cam. J Phys Chem B., 113: 9577-9588.

Altarsha M, Benighaus T, Kumar D and Thiel W (2009). How is the Reactivity of Cytochrome P450cam Affected by Thr252X Mutation? A QM/MM Study for X = Serine, Valine, Alanine, Glycine. J Am Chem Soc., 131: 4755-4763.