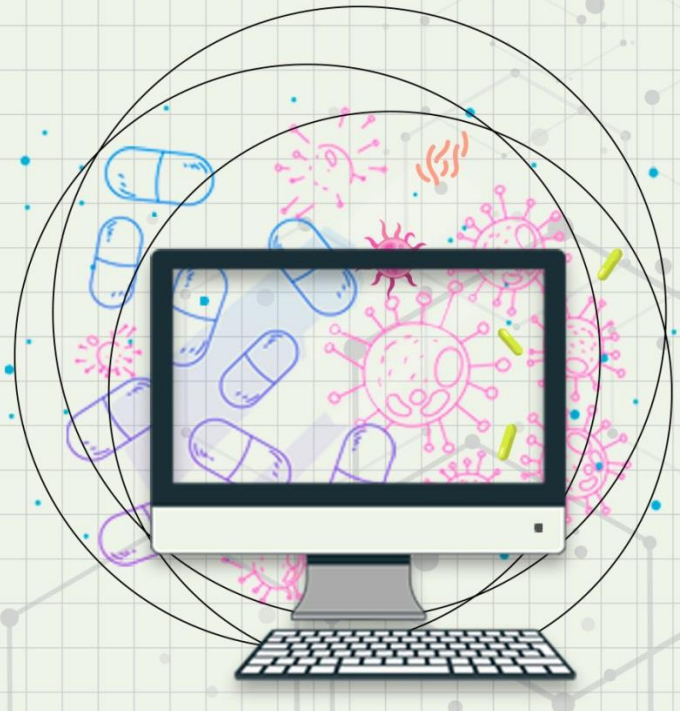
300 years
Russian Academy of Sciences

80 years
IBMC
Institute of Biomedical Chemistry

XXX Symposium on Bioinformatics and Computer-Aided Drug Discovery

PROCEEDINGS BOOK



Institute of Biomedical Chemistry

Moscow, Russia (Virtual), September 16-18 2024

DOI: 10.18097/BCADD2024

ISBN: 978-5-900760-20-9

PROCEEDINGS BOOK OF THE XXX SYMPOSIUM "BIOINFORMATICS AND COMPUTER-AIDED DRUG DISCOVERY" – Moscow: Institute of Biomedical Chemistry, 2024

The materials of the XXX International Symposium "Bioinformatics and Computer-Aided Drug Discovery" (Virtual, 16-18 September 2024) are presented. This Symposium is dedicated to the emerging challenges and opportunities for *in silico* drug discovery. Contemporary fields of biomedical science devoted to the analysis of normal and pathological states of the organism and revealing the pathological processes at the cellular and molecular levels are discussed.

The main topics include: development and practical application of computational methods for finding and validation of new pharmacological targets, *in silico* design of potent and safe pharmaceutical agents, optimization of the structure and properties of drug-like compounds, rational approaches to the utilization of pharmacotherapeutic remedies in medical practice.

This information will be useful for researchers whose investigations are dedicated to creating computational methods and their application to drug research and development using bio- and chemoinformatics methods based on post-genomic technologies. It can also be useful for undergraduate, graduate, and postgraduate students specializing in the relevant fields.

Responsible editors: Corr. Member of Rus. Acad. Sci. V.V. Poroikov, Prof. R.G. Efremov

PITFALLS OF SARS-COV2 MAIN PROTEASE COVALENT INHIBITION MODELING WITH THE COMBINED QUANTUM AND MOLECULAR MECHANICS APPROACHES

I.V. Polyakov[1,2]***, M.G. Khrenova***[1,2]

[1]*Lomonosov Moscow State University, Moscow, Russia*
[2]*Emanuel Institute of Biochemical Physics RAS, Moscow, Russia*

The protein-ligand interactions of target enzyme and potential drug molecules can be assessed with different models including the quantum mechanics-based approaches which are invaluable in studying covalent inhibitors [1]. Combined quantum and molecular mechanics (QM/MM) methods are routinely used to study mechanisms of chemical reactions in proteins and protein complexes [2]. The SARS-CoV-2 main protease (M^{pro}) interaction with the drug molecules was recently simulated with different QM/MM protocols [3, 4]. M^{pro} inhibition mechanism by PF-07321332 drug nirmatrelvir was uncovered; first, the ion pair intermediate is formed by Cys145 side chain deprotonated by the His41 side chain, second, the covalent C-S bond is formed, third, the resulting intermediate undergoes a proton transfer from the protonated but the resulting free energy profiles were in a striking disagreement of up to 11 kcal/mol in the reaction barrier height which translates in enormous discrepancies of the catalytic rate constant.

In this work we report the results of QM/MM modeling of the nirmatrelvir reaction with the M^{pro}. The potential energy QM/MM surface (PES) reaction profiles were obtained with the combination of ChemShell [5] and Turbomole. The PBE0-D3/6-31G**//CHARMM36 level of theory was employed, total atom count was over 10.5 thousand and the quantum part comprised 132 atoms. Two different starting points from the relaxed QM/MM molecular dynamics trajectories were chosen such as to mimic the enzyme-substrate complex (ES) and the reaction intermediate (INT). The resulting energy profiles were found to be in striking disagreement. One of the profiles was characterized by several low energy barriers (2-7 kcal/mol), while the other contained an 18 kcal/mol energy barrier.

The source of these disagreements was found to be in the structure of the active site. The INT model enzyme-substrate structure was characterized by closer S-C attack distance and better solvation of the nitrogen atom of the nitrile group of nirmatrelvir. Basically, the static PES profiles imply that there is a non-obvious reaction coordinate associated with the substrate position on the surface of the M^{pro} active site which is not probed during the conventional PES scans but can be grasped by the QM/MM molecular dynamics simulation if the sampling coordinate and trajectory length were chosen correctly. Thus, one needs to be extremely cautious doing computational design of covalent inhibitors if the active site is on the surface of the protein.

This research was funded by the Russian Science Foundation (grant number 19-73-20032). The research was carried out using the equipment of the shared research facilities of high-performance computing resources at Lomonosov Moscow State University.

1. Sadybekov, A.V., Katritch, V. (2023) *Nature* 616, 673–685.
2. Khrenova, M.G., Mulashkina, T.I., Kulakova, A.M., et al. (2024) *Moscow Univ. Chem. Bull.* 79, 86–92.
3. Grigorenko, B.L., Polyakov I.V., Khrenova, M.G., et al. (2023) *J. Am. Chem. Soc.* 145(24), 13204-13214.
4. Ramos-Guzmán, C.A., Ruiz-Pernia, J.J., Tunon, I. (2021) *Chem. Commun.* **57**, 9096.
5. Lu, Y., Sen, K., Yong, C., et al. (2023) *Phys. Chem. Chem. Phys.*, **25**, 21816-21835.

MOLECULAR MODELING OF BACTERIAL RESISTANCE: THE ROLE OF DYNAMIC BEHAVIOR OF PROTEIN COMPLEXES WITH SUBSTRATES OR INHIBITORS

***M.G. Khrenova*[1,2], *A.V. Krivitskaya*[2], *Ivan V. Smirnov*[1,3], *A.M. Egorov*[1]**

[1]*Lomonosov Moscow State University, Moscow, Russia*
[2] *Fundamental Research Centre of Biotechnology of the Russian Academy of Sciences, Moscow, Russia*
[3] *Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry of the Russian Academy of Sciences, Moscow, Russia*

Among the mechanisms of antibiotic resistance, the important role plays those associated with the interaction of bacterial enzymes with beta-lactam antibiotics, since it is this group of compounds that occupies approximately two thirds of the market for antibacterial therapeutics. These proteins include antibiotic targets – penicillin-binding proteins, as well as beta-lactamases, which are responsible for inactivation of antibiotics. From the point of view of enzymology, all these enzymes belong to the class of hydrolases. Therefore, for a comprehensive study of the mechanisms of reactions, including their comparison and the search for ways to control enzymatic activity, it is necessary to develop common approaches. A large amount of experimental material has been accumulated in the literature on the substrate specificity of proteases of various types, including serine and zinc-dependent ones.

Herein, we present the results of molecular modeling, which includes calculations of molecular dynamic trajectories with classical and combined QM/MM potentials, analysis of geometric parameters and characteristics of electron density, as well as data processing using artificial intelligence methods to determine differences in stationary catalytic parameters observed in the experiment. Also, we demonstrate the importance of dynamic behaviour of complexes and its relation to the binding efficiency. The following examples will be considered: (1) mechanisms of interaction of metallo-beta-lactamases L1 and NDM-1 with cephalosporins, carbapenems and inhibitors, organic boric acids and unithiol; (2)interpretation of the effect of amino acid substitutions in penicillin-binding protein PBP-2 on interaction with ceftriaxone.

The work was supported by the Ministry of Science and Higher Education of the Russian Federation, agreement no. 075-15-2024-536. The research was carried out using the equipment of the shared research facilities of high-performance computing resources at Lomonosov Moscow State University.

MOLECULAR MODELING OF HUMAN LINE-1 ORF2 PROTEIN

***A.M. Kulakova*[1,2], *T.I. Mulashkina*[1], *M.G. Khrenova*[1,3]**

[1]*Lomonosov Moscow State University, Moscow, Russia*
[2]*Institute of Biomedical Chemistry, Moscow, Russia*
[3]*Bach Institute of Biochemistry, Federal Research Centre "Fundamentals of Biotechnology" of the Russian Academy of Sciences, Moscow, Russia*

The long interspersed element-1 (LINE-1) is an ancient genetic parasite that has the ability to move to different genome locations through a 'copy and paste' mechanism. LINE-1 retrotransposition can lead to a variety of genetic disorders and may play a role in the development of cancer, autoimmune diseases and the aging process. Human LINE-1 encodes two proteins: open reading frame 1 protein and open reading frame 2 protein (ORF2p). ORF2p consists of several domains, two of them being a reverse transcriptase (RT) domain and endonuclease domain. Inhibiting their activities could be a promising approach for therapy.

Herein, we present the results of ORF2p full-atom molecular modelling. Firstly, we study the dynamic behaviour of different form of ORF2p: apo-form, complex with RNA template, hybrid complex with RNA template and DNA primer, and complex with RNA template, DNA primer and deoxythymidine triphosphate. All studied system has an 'open' or 'thumb up' conformation of ORF2p reverse transcriptase core, which corresponds to the active form of ORF2p. All these calculations were the molecular dynamics simulation in NPT (p=1 atm., T=300 K) ensemble with classical CHARMM force field.

Also, we demonstrate the dynamic behavior of enzyme-substrate complex of ORF2p with thymidine triphosphate which was calculate using molecular dynamic simulation with combined quantum mechanics/molecular mechanics (QM/MM) potentials. The quantum part was calculated using Kohn-Sham method of the density functional theory with the PBE0 functional with D3 dispersion correction and the 6-31G** basis. The molecular mechanical part was described using the CHARMM force field.

The work was supported by the Ministry of Science and Higher Education of the Russian Federation, agreement no. 075-15-2024-643. The research was carried out using the equipment of the shared research facilities of high-performance computing resources at Lomonosov Moscow State University.

ON THE EFFICIENCY OF CARBAPENEM ACTIVATION BY BETA-LACTAMSES

***D.D. Kholmanskikh*[1,2], *M.E. Kuznetsov*[1,2], *D.A. Tatarnikova*[1], *A.V. Krivitskaya*[2,3], *M.G. Khrenova*[1,2]**

[1]*Lomonosov Moscow State University, Moscow, Russia*
[2]*Emanuel Institute of Biochemical Physics of the Russian Academy of Sciences, Moscow, Russia*
[3]*Fundamental Research Centre of Biotechnology of the Russian Academy of Sciences, Moscow, Russia*

Antibiotic resistance is a constant problem of bacterial infection treatment. Beta-lactams represent the main group of antibiotics that disrupt cell wall building due to the inhibition of penicillin-binding proteins. Beta-lactamases are enzymes that appeared evolutionary to inactivate beta-lactam compounds. Now, four classes of beta-lactamases are known. Three of them, A, C and D, are serine hydrolases and B class are zinc-dependent enzymes, called metallo-beta-lactamases. Beta-lactam compounds are also different. Among them, the most known are penicillins, cephalosporins, monobactams and carbapenems. Many experimental studies are conducted to determine the rate steady state kinetic parameters of antibiotic inactivation by beta-lactamases. Theoretical studies can serve rational background of variations in catalytic parameters.

The binding efficiency might be related to the conformational diversity as well as flexibility of the entire protein. This can be studied using classical molecular dynamic (MD) simulations with the subsequent analysis of trajectories. Further hydrolysis rate depends on the interactions of the enzyme with the substrate in the active site. For hydrolases there are two most important participants of the reaction. Those are a nucleophilic particle and a so-called oxyanion hole. The nucleophile can by either an oxygen atom of the catalytic serine residue or an oxygen atom of the hydroxide anion for serine and metallo-beta-lactamases, respectively. The oxyanion hole is formed by either NH groups or zinc cations depending on the particular type of beta-lactamase. The proper way to deepen the understanding of these interaction is molecular dynamics with combined quantum mechanics / molecular mechanics (QM/MM) potential. Herein, we demonstrate results on QM/MM MD simulations of several carbapenems with a metallo-beta-lactamase NDM-1 to explain differences in their observed kinetic parameters.

The authors acknowledge financial support from the Russian Science Foundation, project 19-73-20032. The research is carried out using the equipment of the shared research facilities of HPC computing resources at Lomonosov Moscow State University.

MODELING OF THE PHOSPHORYL TRANSFER MECHANISM IN THE ACTIVE SITE OF PROTEIN KINASE A

***M.S. Leonova*[1], *A.V. Krivitskaya*[2], *T.I. Mulashkina*[1,2]**

[1]*Moscow State University, Moscow, Russia*
[2]*Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, Moscow, Russia*

Protein phosphorylation is one of the most important mechanisms in cell signaling and gene regulation. Consequently, this reaction is treated as a therapeutic target for cancerous, immune, inflammatory and neurodegenerative diseases. Enzymes called protein kinases catalyze the reaction of the γ-phosphate transfer from the ATP to specific residues such as serine, threonine or tyrosine. Among the superfamily of protein kinases, cAMP-dependent protein kinase (protein kinase A, PKA) was the first to be characterized. Moreover, PKA has been analyzed most thoroughly ever since becoming a model of all kinases because of the highly conserved core of the enzyme.

Although, PKA is one of the best studied kinases, there are still questions about nature of the phosphoryl transfer mechanism. Phosphorylation reactions are followed by the cleavage of the P-O bond. Considering transition state, the nucleophile and leaving group may be bonded to phosphorus to varying degrees, so associative and dissociative mechanisms may be recognized. In the case of associative mechanism, the bond formation with the nucleophile starts when the ADP and the *γ-phosphate group are still* rather close. Conversely, dissociative mechanism describes transition state nature where the new bond is yet to be formed properly but the bond between the ADP and the leaving group is already almost broken.

Thus, the aim of this study was to determine the type of mechanism in the reaction of serine phosphorylation of the substrate SP20 in the active site of PKA. The geometric and electron-density criteria of structures corresponding to possible conformations of the enzyme-substrate complex were analyzed. In addition, the Gibbs energy profile of the serine phosphorylation reaction in the active site of PKA was predicted using molecular modeling methods.

In this study, molecular dynamic trajectories were clustered, and three stable conformations of the enzyme-substrate complex were found, differing by the mutual arrangement of the substrate's serine and the phosphate tail of ATP. Representative frames were selected from these clusters, and molecular dynamics calculations were performed with the potentials of the combined quantum mechanics/molecular mechanics (QM/MM) method: for analyzing the state of the enzyme-substrate complex — without adding a bias potential, and for plotting the Gibbs energy profile — with the addition of a bias potential using the umbrella sampling method. The quantum subsystem included side chains of residues Lys72, Asp166, Lys168, serine of substrate SP20 and 7 molecules of water, as well as two magnesium cations and their coordination spheres represented by side chains of residues Asn171 and Asp184, phosphate groups of the ATP and 3 molecules of water. This subsystem was described at the DFT level with the PBE0 hybrid functional and D3 dispersion correction, using the 6-31G** basis set. The CHARMM force field was used to describe the MM subsystem. The difference between the distances of the breaking bond ($P\text{-}O_{ATP}$) and nucleophilic attack ($P\text{-}O_{Ser}$) was chosen as the reaction coordinate: $\Delta = d(P\text{-}O_{ATP}) - d(P\text{-}O_{Ser})$.

The mean value of the breaking bond length and the value of the Laplacian of electron density along the line of the breaking P-O bond can be used as criteria for determining the type of mechanism. Applying these criteria to the cleaved $P\text{-}O_{ATP}$ bond showed that, regardless of the conformation of the enzyme-substrate complex, the phosphorylation reaction in the active site of PKA occurs by a dissociative mechanism. Analysis of the Gibbs energy profiles computed for the phosphorylation reaction of serine in the active site of PKA also showed that the reaction proceeds by a dissociative mechanism. In addition, the Gibbs energy profile for the conformational transition of the detected conformations was obtained, and it was determined which of the conformations is the most reactive.

This work was supported by the Russian Science Foundation (project no. 23-13-00011). The research is carried out using the equipment of the shared research facilities of HPC computing resources at Lomonosov Moscow State University.

DETERMINATION OF SUBSTRATE ACTIVATION IN ACTIVE SITES OF HYDROLASES USING NEURAL NETWORKS

***T.I. Mulashkina*[1,2], *M.A. Krinitskiy*[1,3,4], *M.G. Khrenova*[1,2,5]**

[1]*Moscow State University, Moscow, Russia*
[2]*Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, Moscow, Russia*
[3]*Moscow Institute of Physics and Technology, Moscow, Russia*
[4]*Shirshov Institute of Oceanology of the Russian Academy of Sciences, Moscow, Russia*
[5]*Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry of the Russian Academy of Sciences, Moscow, Russia*

For enzyme-catalyzed chemical reactions to occur, the substrate must be activated at the enzymes active site. This can be determined by analyzing Laplacian electron density maps in the reaction region. These maps show deconcentration of electrons near the electrophile in the direction of nucleophile attack for reactive molecules and concentration of electrons for nonreactive molecules. By studying molecular dynamics trajectories calculated using quantum mechanics and molecular mechanics methods (QM/MM), the efficiency of substrate activation can be estimated quantitatively from the ratio between reactive and nonreactive states. A typical QM/MM trajectory is tens of picoseconds longcorresponds to tens of thousands of frames in which Laplacian maps. Manually processing such a volume of information is difficult. So, in this work, we developed a convolutional neural network based on the ResNet50 architecture. This network is capable of distinguishing between reactive and non-reactive enzyme-substrate complexes.

The neural network was trained on images of Laplacian electron density maps calculated for enzyme-substrate complexes of the main protease of SARS-CoV-2 with three different substrates. During training and validation, artificial data augmentation was used to increase invariance to rotations and translations of the neural network. To evaluate its performance, sets of Laplace electron density maps were randomly selected from QM/MM MD trajectories of bacterial metallo-beta-lactamase NDM-1 complexes and antibiotic imipenem as well as capralactam-lipase CALB complexes. The training and test data were balanced by the target feature, the reaction state or its absence.

Based on the training results, the proposed neural network distinguishes between reactive and non-reactive states in all data sets with a correct answer rate exceeding 97%. The difference between the training and test samples does not exceed one percentage point.

The attention field of the neural network was also analyzed. This analysis showed that the key factor in determining the reactivity of enzyme-substrate complexes is indeed deconcentration of electron density near electrophilic atoms.

This work was supported by the Scientific and Educational School of Moscow State University (project no. 23-Sh03-04).