Project Closing Report

OTKA K111862

Development of a mixed quantum mechanical/molecular mechanical method and its application to biochemical systems

QM/MM methodological developments

The MRCC quantum chemical program was selected to build a QM/MM package in which the QM and MM subsystems are separated by frozen localized orbitals. The MRCC program was selected as it includes highly efficient implementations of local electron correlation methods and it is developed in the group of Mihály Kállay at the Budapest University of Technology and this ensures effective communication and flexibility, both crucial in the development of complex computer codes.

The Amber program package was selected for treating the MM subsystem and to drive the calculations. The Amber package is well suited for being part of our QM/MM program as it has already been interfaced with other QM codes, its source code is available, and it is a free package that is widely used.

The QM/MM package building included the following principal steps:

1) The orbital equations are built in the MRCC code.

2) The MRCC and Amber (MM) codes interfaced so that the resulting program is suitable for QM/MM calculations with both the link atom and the frozen orbital methods.

3) Equations required for calculating analytical gradients were derived. They include the derivatives with respect to the coordinates of atoms with fixed orbitals at the subsystem boundary. Forces were implemented.

4) An algorithm was developed and implemented to automatically generate the fixed orbitals connecting the MM and QM subsystems. This is a significant advance, since, for the first time, it allows for parameter free frozen orbital QM/MM calculations. The algorithm analyses the topology of the molecules, identifies atom and bond types and finds optimal positions for the QM/MM subsystem boundary.

5) The Huzinaga-equation based separation was implemented in WFT-in-DFT and DFT-in-DFT embedding schemes.

Calculations were performed for model systems with both the link atom and the frozen orbital approaches. The quality of the results obtained with the two methods is similar and is comparable with that of other QM/MM computational schemes [TCA2015]. Although the frozen orbital approach is theoretically well founded, while the link atom approach is highly intuitive, the similar quality results obtained with the two approaches render the link atom approach more appealing owing to its facile implementation and application. Therefore, the link atom approach was used in the applications described below.

An analysis of the various DFT-in-DFT embedding schemes using the Huzinaga-equation based separation revealed that QM/MM is competitive with DFT-in-DFT methods as far as the system is partitioned across single bonds [JCP2016].

Application of QM, MM and QM/MM methods to various problems

D-Amino acid oxidase (DAAO) is a flavoenzyme whose inhibition is expected to have therapeutic potential in schizophrenia. In a project aiming at identifying new D-amino acid oxidase (DAAO) inhibitors the binding mode of benzo[d]-isoxazols was to be determined. Based on the X-ray structure of related complexes two binding modes appeared to be feasible for benzo[d]-isoxazols. In order to determine the binding mode realized by these compounds, the binding free energy of several benzo[d]-oxazole derivatives was calculated with free-energy perturbation (FEP) in both binding modes and the calculated values were compared to experimental data. The calculated free energies in one of the binding mode exhibit more favourable correlation with experimental data than those in the other binding mode. We also performed high level QM/MM calculations of the interaction energies to confirm the validity of the predicted binding mode. Ligand–protein interaction energies were computed with QM/MM methods implemented in the Amber-Gaussian and Amber-MRCC programs. Representative complex structures were selected from molecular dynamic simulations performed for complexes of both binding modes. Ligand–protein interaction energies were calculated by QM/MM using dRPA method with DFT PBE functional and aug'-cc-pVTZ basis set and applying BSSE correction. Significantly different interaction energies were obtained for the two binding modes; the mode predicted by FEP calculations exhibited more favourable interaction energies, thus confirming the correctness of this bonding mode. It is worth noting that QM/MM calculations feature high level interaction energy evaluations with limited sampling, while FEP calculations feature a purely classical force field and enhanced sampling. The consistent results obtained by these orthogonal methods increase the confidence in the validity of the proposed binding mode [JCAMD2018].

The detailed mechanism of the oxidative half-reaction of DAAO was investigated. DAAO catalyses hydride transfer from the substrate to the flavin in the reductive half-reaction, and the flavin is reoxidized by O_2 in the oxidative half-reaction. The mechanism of the oxidative half-reaction was elucidated by combining the results of DFT level QM/MM calculations with experimental data from the literature.

The first step of the catalytic cycle of DAAO, namely the reductive half-reaction, starts with a hydride transfer from the bound substrate to the FAD cofactor (**Error! Reference source not found.**). As the details of this reaction step are well established, it provides a good basis for validating the model system and the method applied. B3LYP/6-31+G** QM/MM relaxed coordinate scan gave a 15.4 kcal mol^{-1} barrier for the reductive half-reaction of DAAO with D-alanine as a substrate. This is in reasonable agreement with the experimental value of 13–15 kcal mol^{-1} ($1.8 \times 103–2.4 \times 105$ M^{-1} min^{-1}) calculated from stopped flow rate constant measurements at 25 °C [J Biol Chem. 1993 Jul 5;268(19):13850-7; J Biol Chem. 1977 Jul 10;252(13):4464-73.], and also with ab initio QM MD simulations performed for a smaller model system [Biochemistry. 2002 Dec 3;41(48):14111-21].

Several reaction mechanisms for the oxidative half-reaction have been investigated. In addition to the relaxed coordinate scans of the investigated steps, we also searched for the minimum energy crossing point (MECP) as the system is in the triplet state at the beginning of the oxidative half-reaction and it is singlet in the product state. This was performed with a MECP searching QM computer code that was transformed into a QM/MM MECP searching program. The combination of the computational results

with available experimental information led to the following proposal for the reaction mechanism. The oxidative half-reaction starts with a single electron transfer from the reduced FAD to the O_2 molecule. The resulting FAD semiquinone-superoxide anion system is a stable intermediate. The system is in the triplet state at the start of the oxidative half-reaction and the triplet–singlet transition occurs in the semiquinone-superoxide anion system with a low barrier. The next step is a proton-coupled electron transfer between the FAD and the oxygen species, and this is accompanied by a change in the electronic configuration from an open shell to closed shell singlet. The reaction completes with proton abstraction from the oxidized substrate via a chain of water molecules. The experimentally observed reaction barrier for the oxidative half-reaction is ∼11 kcal mol^{-1} (106 M^{-1} min^{-1} range) and it is assigned to the first electron transfer as no intermediates of the oxidative half-reaction were experimentally observed [Biochemistry. 1966 Oct;5(10):3181-9; J Biol Chem. 1993 Jul 5;268(19):13850-7]. This is consistent with the calculated energy surface of the subsequent low barrier steps. The oxidized substrate contributes to O_2 activation in the rate-determining single-electron transfer as its presence lowers the barrier of the oxidative half-reaction. Our calculations confirm its additional role of providing a proton in the last step of H_2O_2 formation. The proposed detailed mechanism of the oxidative half-reaction of DAAO does not agree with the mechanism of other oxidases and provides another example for the versatility of oxidation in flavin-dependent enzymes [OBC2019].

In a study of the covalent inhibition of several enzymes, the reaction of small compounds equipped with a reactive group (warhead) with cysteine residues were investigated computationally and experimentally. Electrostatic potential minima around cysteine sulphur atoms as obtained by QM/MM calculations together with other descriptors were used to interpret i) the variation of cysteine reactivities in different enzymes against a set of compounds, and ii) the variation of cysteine reactivities within a single enzyme.

The same set of fragment-like compounds were experimentally tested against MurA and Cathepsine B (CatB) enzymes. It was found that MurA and CatB were labelled by 15 and 3 compounds, respectively. With the objective of interpreting the apparent reactivity difference between MurA Cys115 and CatB Cys29 we performed computational characterization of the relevant cysteines and their environment. We used three tools to obtain reactivity and accessibility descriptors. QSite by Schrödinger was used to perform mixed QM/MM calculations to inspect the reactive centre in the protein. It provided information about the electrostatic potential minima (ESP_{min}) on the sulphur atom of the cysteines, thus indicating the relative nucleophilicity of the targeted residues. By using the POPS algorithm, we could retrieve information on the cysteine accessibilities, in terms of solvent accessible surface areas (SASA) of both the whole residue and its side-chain sulphur (SG). Finally, the web-based platform Cpipe was used for reactivity and pK_a predictions for the analysed cysteines. By investigating the calculated properties, the lower ESP_{min} and pK_a values of CatB's Cys29 compared to MurA's Cys115 suggest a more pronounced nucleophilicity of the former, which is accompanied, however, by a lower accessibility. Overall, the lower accessibility of the catalytic cysteine in CatB could provide an explanation of the lower number of experimental actives found against this target.

Covalent inhibition of MAO-A was also investigated in the same study. Although several known MAO-A inhibitors bind covalently to the FAD cofactor, to the best of our knowledge, no validated cysteine-binding covalent inhibitor for MAO-A has been reported yet. The same set of cysteine selective reactive compounds investigated against MurA and CatB resulted in 12 MAO-A active compounds. Therefore, we performed a comparative analysis of the cysteine residues near to the active site of MAO-A. We identified Cys323 in a position that its labelling is likely to block the access to the active site. Moreover, since two additional residues, Cys201 and Cys321, are found near the active site, the reactivity and

accessibility of these three cysteines were characterized as explained above. By comparing the values obtained for these three residues in MAO-A, we observed that Cys323 not only has the most negative ESP_{min}, but also the largest SASA considering both the whole residue and the side-chain sulphur. These data suggest that Cys323 is the cysteine residue having the highest nucleophilic character and accessibility among the ones analysed in MAO-A. Furthermore, Cpipe predicted it to be reactive, together with Cys321, despite their high pK_a values. Altogether, these data suggest that Cys323 is potentially targetable and support that Cys323 labelling may lead to MAO-A inhibition with a novel covalent mechanism of action. Indeed, the labelling of Cys323 was proved by subsequent MS/MS measurements [JCIM2020a].

The catalytic mechanism and covalent inhibition of UDP-N-Acetylglucosamine enolpyruvyl transferase (MurA) was investigated. The MurA enzyme catalyses the first committed step in the biosynthesis of peptidoglycan involving the transfer of an enolpyruvil group from phosphoenolpyruvate (PEP) to UDPN-acetylglucosamine (UNAG) forming UDP-N-acetylglucosamine enolpyruvate (UNAGEP). This pathway is essential for the growth of bacteria but missing in mammals, that nominates MurA as an attractive antibacterial target. The aim of our study was to understand the details of the covalent reaction between Cys115 of MurA and its natural substrate, PEP, on one hand, and several covalent inhibitors, on the other hand.

We showed that the UNAG- or UNAM-induced conformational change at the flexible loop region of MurA facilitates the proton transfer between Cys115 and His394. This process yields the highly nucleophilic thiolate form of Cys115 that is able to participate in various reactions. The role of His394 as a deprotonating agent is confirmed by i) molecular dynamics simulations showing the proximity of Cys115 and His394 in the loop closed conformation, ii) sequence alignment of MurA enzymes from various species showing the conservation of His394 and iii) QM/MM MD simulations showing low barrier exothermic proton transfer from Cys115 to His394. Cys115 is involved in the catalytic mechanism and in the covalent inhibition of MurA, and its activation by deprotonation is a prerequisite for these reactions. It was shown that the natural substrate PEP is able to form a covalent complex with Cys115 by first abstracting a proton from His394 and then binding to Cys115. The mechanism investigated assumed prior proton transfer from Cys115 to His394 and is in line with the formation of a reversible covalent complex as proposed in Zhu et al. J Biol Chem. 2012 Apr 13;287(16):12657-67. The free energy profile for the binding of several covalent inhibitors was calculated. Three groups of compounds, namely, oxirane derivatives, haloketones, and Michael acceptors, were studied. It was found that QM/MM MD simulations with the DFTB3 approximate density functional tight binding method and the AMBER force field produced sensible free energy profiles with lower barriers for the inhibitors compared to the inactive compounds. The calculated barrier for fosfomycin and terreic acid, both are known oxirane-based inhibitors of MurA, was found to be lower than that for an inactive oxirane derivative. Similarly, the active compounds with haloketone warheads were separated from a similar inactive compound based on the free energy barriers of their covalent binding, and analogous separation was also reproduced for Michael acceptors. The finding that calculated free energy barriers separate actives from inactives within a set of compounds with similar reactive warheads may find use in supporting covalent ligand design [JCIM2019].

The reaction mechanism of uridine-pseudouridine transformation catalysed by the box H/ACA pseudouridine synthase was investigated. Pseudouridylation is an abundant posttranslational modification that concerns almost all type of RNAs and the malfunctioning of pseudouridine synthase

enzymes are linked to severe diseases. We first performed QM calculations on model systems. Our results excluded some formerly proposed mechanisms. The Michael addition scheme was found to be unlikely since no stable adduct is formed between the uridine and the catalytic aspartate. The nucleophilic substitution scheme is ruled out owing to the unfavourable steric arrangement of the reactants. Our results are in favour of the glycal scheme and provide details for the mechanism that is likely to start with the glycosidic bond cleavage between the ribose and uracil. This step is potentially followed by or coupled to the deprotonation of the C2′-atom of the sugar by the conserved catalytic aspartate. It was found that the rebinding of the basis to the sugar via a C-C bond occurs readily only after the reprotonation of C2′-atom [TCA2018].

We found that significant conformational changes accompany the detachment of the basis from the sugar. Therefore, both QM/MM metadynamics and molecular dynamics with umbrella sampling calculations were initiated in order to better account for the dynamics of the system. These calculations confirmed that the basis-sugar detachment occurs with a reasonable free-energy barrier of 20-25 kcal/mol. We have not seen, however, the C2' deprotonation before the reattachment of the basis to the sugar via a C-C bond. We have also performed calculations on the subsequent reaction steps involving proton movement, bond reorganization and conformational changes to form pseudouridine, the final product of the catalysed reaction. These studies resulted in free-energy profiles that support the assumed reaction mechanism and confirm the significant role of the flexibility of the protein-RNA complex. The preparation of a manuscript reporting these results has been recently started [MS2020a].

According to the "amyloid hypothesis", amyloid β-peptide Aβ42 is a central pathological agent in Alzheimer's disease. Investigation performed for a series of mimetic peptides possessing a potent inhibitory effect on Aβ42 aggregate formations. One of the compounds with highest inhibitory effect was selected for an extensive physico-chemical evaluation including a Thioflavin-T assay monitoring fluorescence intensity at different times and transmission electron microscopy studies, as well as a cell viability assay to obtain further insights into the prevention of Aβ42-mediated cell toxicity. Various computational modelling techniques were also employed to identify the interactions of the selected compound with early stage Aβ42 monomer and look for possible explanations of how this mimetic peptide affects the conformational transition of the peptide. These simulations enabled us to obtain a picture for the experimentally hard to attain, very early stage interaction with the ligand at atomic level. Simulations showed that the compound diminished the formation of aggregation-prone structures by maintaining helical content which were reproduced by ligand-free control simulations. It was also found that the β-strand structure was significantly diminished at CHC region (residues K16LVFFA21). Regarding helical structures, the α-helical distribution was higher at N27KGAIIG33 residues and 310-helix propensities were increased at N-terminal and CHC regions in the presence of the compound. These secondary structure results are in good agreement with experimental [FEBS J. 2006 Feb;273(3):658-68] and theoretical [J Phys Chem B. 2011 Nov 10;115(44):12978-83] studies where it had been indicated that the loss of helical content can be observed during the early stage of toxic species formation. Molecular docking studies revealed that the compound interacted with the Aβ42 monomer through π-stacking interactions between its Fmoc group and the aromatic residues His14 or Tyr10. The compound buries its hydrophobic Fmoc group in the CHC leading to the dominance of helical conformation in this region. This result corroborates the simulations where the contact distribution of the ligand is also enhanced at the CHC region of the Aβ42 monomer. In conclusion, the experimental and theoretical work demonstrated that the selected compound modulates the early steps of the amyloidogenic process also inhibiting the cytotoxicity linked to Aβ42 aggregation [BC2018].

The effect of substituents on peptide halogen-bond (XB) acceptors were investigated with the objective of giving the basis for more accurate ligand–protein interaction models. Halogen bonds are highly important in medicinal chemistry as halogenation of drugs often improves both selectivity and efficacy towards protein active sites. However, accurate modelling of halogen bond interactions remains a challenge, since a thorough theoretical investigation of the bonding mechanism, focusing on the realistic complexity of drug-receptor systems, is lacking. The systematic studies on ligand/peptide-like systems revealed that peptide methyl and amino building blocks improve the stability of XB complexes by electrostatics, dispersion and charge transfer from the Lewis base to the halogenated Lewis acid. Conversely, the inclusion of a peptide carbonyl, adjacent to the XB pair, decreases the stability of the XBs. These results point out the great influence of the protein backbone environment in both the complex geometries and energies. The stability and geometry of XBs to hydrogen-bonds (HBs), as a function of the atom-pair distance were also compared. It was revealed that XBs and HBs share the same interaction mechanism, based on the subtle interplay of electrostatics, donor–acceptor orbital interactions, and dispersion, consistent with previous comparative studies. However, it was found that XBs are less stable than HBs, with bond lengths longer than HBs, because of a stronger Pauli repulsion. These findings demonstrate the need to incorporate quantum effects in molecular modelling approaches for drug design [JCIM2020b].

Ion channels formed by quadruplexes built either from guanine or from modified nucleobases were investigated. Molecular dynamics calculations were performed for several model systems to investigate the behaviour of monovalent cations within the channel. Previous studies distinguished container or channel like character for the central ion channel when different nucleobases were applied in the formation of the structure. We extended these investigations by QM/MM calculations to clarify the chemical background of the different character, as well as to determine the reaction barrier for channel crossing. It was observed that the outcomes of the calculations were highly sensitive to the selection of the snapshots from the dynamical trajectory. Therefore, molecular dynamics simulations with varying parameters and sampling were performed to generate representative geometries for subsequent QM/MM calculations to obtain reliable barriers of channel crossing [MS2020b].

Publications produced by the support of OTKA 111862

BC2018 Gera, János ; Szögi, Titanilla ; Bozsó, Zsolt ; Fülöp, Livia ; Barrera, Exequiel E ; Rodriguez, Ana M ; Méndez, Luciana ; Delpiccolo, Carina ML ; Mata, Ernesto G ; Cioffi, Federica et al. Searching for improved mimetic peptides inhibitors preventing conformational transition of amyloid-β42 monomer BIOORGANIC CHEMISTRY 81 pp. 211-221. , 11 p. (2018)

JCAMD2018 . Orgován, Zoltán ; Ferenczy, György G ; Steinbrecher, Thomas ; Szilágyi, Bence ; Bajusz, Dávid ; Keserű, György M. Validation of tautomeric and protomeric binding modes by free energy calculations. A case study for the structure based optimization of d-amino acid oxidase inhibitors JOURNAL OF COMPUTER-AIDED MOLECULAR DESIGN 32 : 2 pp. 331-345. , 15 p. (2018)

JCIM2019 Mihalovits, Levente M. ; Ferenczy, György G. ; Keserű, György M. Catalytic Mechanism and Covalent Inhibition of UDP- N -Acetylglucosamine Enolpyruvyl Transferase (MurA): Implications to the Design of Novel Antibacterials JOURNAL OF CHEMICAL INFORMATION AND MODELING 59 : 12 pp. 5161-5173. , 13 p. (2019)

JCIM2020a Andrea Scarpino, László Petri, Damijan Knez, Tímea Imre, Péter Ábrányi-Balogh, György G. Ferenczy, Stanislav Gobec, György M. Keserű WIDOCK: A Warhead-Independent

Virtual Screening Protocol for Cysteine Targeted Covalent Inhibitors JOURNAL OF CHEMICAL INFORMATION AND MODELING in revision

JCIM2020b Enrico Margiotta, Stephanie C. C. van der Lubbe, Lucas de Azevedo Santos, Gabor Paragi, Stefano Moro, F. Matthias Bickelhaupt and Célia Fonseca Guerra Halogen Bonds in Ligand-Protein Systems: Molecular Orbital Theory for Drug Design JOURNAL OF CHEMICAL INFORMATION AND MODELING in revision

JCP2016 Bence, Hégely ; Péter, R Nagy ; György, G Ferenczy ; Mihály, Kállay Exact density functional and wave function embedding schemes based on orbital localization JOURNAL OF CHEMICAL PHYSICS 145 : 6 Paper: 064107 , 12 p. (2016)

MS2020a Kiss, Dóra Judit, Menyhárd, Dóra.K., Oláh, J., Tóth, G., Stirling, A., Ferenczy, György G. The reaction pathway of pseudouridylation by box H/ACA pseudouridine synthase: Insight from quantum mechanical/molecular mechanical free energy calculations MANUSCRIPT IN PREPARATION.

MS2020b J. Gera, G. Ferencz, Z. Kupihar, L. Kovacs and G. Paragi Ionmobility in natural and artificial quadruplex structures MANUSCRIPT IN PREPARATION.

OBC2019 Kiss, Dóra Judit ; Ferenczy, György G. A detailed mechanism of the oxidative half-reaction of d-amino acid oxidase: another route for flavin oxidation ORGANIC & BIOMOLECULAR CHEMISTRY 17 : 34 pp. 7973-7984. , 12 p. (2019)

TCA2015 Hégely, Bence ; Bogár, Ferenc ; Ferenczy, György G ; Kállay, Mihály A QM/MM program using frozen localized orbitals and the Huzinaga equation THEORETICAL CHEMISTRY ACCOUNTS 134 : 11 Paper: 132 , 12 p. (2015)

TCA2018 Kiss, D.J.; Oláh, J.; Tóth, G.; Menyhárd, D.K.; Ferenczy, G.G. Quantum chemical calculations support pseudouridine synthase reaction through a glycal intermediate and provide details of the mechanism THEORETICAL CHEMISTRY ACCOUNTS 137 : 12 Paper: 162 (2018)