Attachement 2b

Summary of professional achievements

1 Name

Piotr Setny

2 Diplomas and scientific degrees

- **PhD in Physics** – Physics Department, University of Warsaw, December 2008, thesis entitled *Studies of hydrophobic effects using computer simulations*, with honors
- **MSc in Physics** – Physics Department, University of Warsaw, September 2003, thesis entitled *Investigation of differences and similarities in CK2 kinase interactions with ATP and GTP using molecular dynamics methods*, Summa cum laude
- **physician** – Medical University of Warsaw, June 2001

3 Employment in research institutes

- **since 07.2013** Centre of New Technologies, University of Warsaw, adiunct,
- **08.2009 – 06.2013** Technische Universität München, Germany, postdoctoral researcher,
- **01.2009 – 07.2009** University of California, San Diego, USA, visiting scholar,
- **10.2003 – 12.2008** Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw, senior specialist.

4 Scientific acheivement

4.1 Title of scientific achievement

Modelling and analysis of hydration effects in biomolecular systems

4.2 List of publications documenting the scientific achievement.

H1 Riccardo Baron*, **Piotr Setny***, J. Andrew McCammon, (*equal contribution)
Water in Cavity-Ligand Recognition
J. Am. Chem. Soc., 2010, 132:12091,

H2 **Piotr Setny***, Riccardo Baron*, J. Andrew McCammon, (*równy udział w pracy)
How Can Hydrophobic Association Be Enthalpy Driven?
J. Chem. Theory and Comp., 2010, 6:2866,

H3 **Piotr Setny**, Peter Kekenes-Huskey, Riccardo Baron, J. Andrew McCammon, Joachim Dzubiella,
Solvent fluctuations in hydrophobic cavity-ligand binding kinetics
Proc. Natl. Acad. Sci. U.S.A., 2013, 110:1194,

H4 **Piotr Setny**, Anita Dudek,
Explicit Solvent Hydration Benchmark for Proteins With Application to the PBSA Method,
J. Chem. Theory and Comp., 2017, 13:2762,

H5 **Piotr Setny**, Martin Zacharias,
Hydration in Discrete Water. A Mean Field, Cellular Automata Based Approach to Calculating Hydration Free Energies,
J. Phys. Chem. B., 2010, 114:8667,

H6 **Piotr Setny**,
Hydration in Discrete Water (II): From Neutral to Charged Solutes,
J. Phys Chem. B, 2015, 119:5970,

H7 **Piotr Setny**,
Prediction of water binding to protein hydration sites with discrete, semi-explicit solvent model,
J. Chem. Theory and Comp., 2015, 11:5961,

H8 **Piotr Setny**, Marta D. Wiśniewska,
A water mediated conformational preselection mechanism in substrate binding cooperativity to Protein Kinase A
Proc. Natl. Acad. Sci. U.S.A., 2018, *in press*, https://doi.org/10.1073/pnas.1720024115.

4.3 Scientific goals, results and their potential applications

4.3.1 Introduction

Water in biomolecular systems. Liquid water seems to be indispensable for the life as we know it. Traditionally, however, aqueous environment has been perceived only as a

passive medium, simply embedding the molecules involved in cellular activities. With our growing awareness of physical forces driving life processes, water appears to be one of the key determinants of biomolecular structure, dynamics and, consequently, function [1].

The important role of water in biomolecular systems is apparent at multiple levels of spatiotemporal organisation, starting from mesoscopic liquid phase, to partially structured hydration layers, to specifically bound solitary water molecules. a) As a liquid phase made of small, strongly interacting molecules, water contributes to so called hydrophobic effects [2]. They are responsible for the assembly of supramolecular structures such as biological membranes, protein folding, or specific intermolecular interactions. b) Owing to the permanent dipole moment of water molecules aqueous environment is a medium with high dielectric constant [3]. The resulting screening of electrostatic interactions mitigates their dominant role, allowing biomolecular systems to function under the subtle control of diverse physical effects. c) Owing to mutual hydrogen bonds, water molecules form dynamic, local structures surrounding the solvated objects [4]. The presence of such hydration layers modulates intersolute interactions, in a way that depends on their local surface topography and physicochemical properties. Such effects are particularly important for the interaction of macromolecules with ions, selective conduction through transmembrane channels, conformational dynamics, or association kinetics. d) Solitary water molecules embedded in hydrogen bonds networks within macromolecular structures contribute to their stability, and to specific interactions with other binding partners [5]. Their presence is of particular practical importance in rational drug design, since accounting for water molecules specifically located within receptor binding sites is necessary for the proper prediction of ligand binding modes. Finally, e) water molecules participate in multiple enzymatic reactions, contribute to proton transfer (Grotthuss mechanism), or take part in molecular transport through transmembrane channels [1].

Even though we start to appreciate important, multifaceted role of water in biomolecular systems, in many cases its thorough investigation, understanding, and, ultimately, description with a quantitative, predictive model is still beyond our capabilities. Aside from complexity and diversity of water-mediated effects, one of important reasons is limited spectrum of experimental methods providing sufficient spatio-temporal resolution necessary to trace the structure and dynamics of specific water molecules [6]. On the one hand, method based on crystallography, and utilising X-rays or neutron diffraction allow detection of individual, strongly bound water molecules within macromolecular structures, but can't provide insights into more dynamic, partially disordered water regions. On the other hand, spectroscopic methods, for instance in terahertz range of electromagnetic field spectrum give information about local dynamics of hydration layers around macromolecules, but can't probe their structure [7]. A separate problem is related to the interpretation of water's role in biomolecular processes based on thermodynamic measurements. Typically they characterise the properties of a given system as a whole, without the possibility to dissect contributions arising due to its particular elements.

Modelling of hydration effects. Direct insights into hydration effects can be gained by using computer simulations in which, aside of the molecular system of interest, water is

modelled explicitly at atomic resolution [8]. Owing to steady increase of computer power, such explicit solvent simulations can now be used to efficiently sample the configuration space of a system consisting of tens of thousands of atoms [9]. Such system size corresponds to a medium-sized protein surrounded by a water shell of several Å thickness. One particularly useful feature of computer simulations is the ability to consider idealised model systems, exposing only selected aspects of studied phenomena to subsequent theoretical analysis. Interpreted through statistical physics methods, computer simulations allow the estimation of free energy changes, as well as their enthalpy, and entropic contributions [10]. With properly sampled statistical ensemble they provide unique possibility to interpret experimental findings based on fully atomistic view of the system of interest. Simulations of molecular dynamics (MD) give also insights into dynamical parameters such as water relaxation times at selected locations or reaction kinetics.

Typical computer simulations based on classical molecular mechanics do not take into account quantum properties of matter, in particular electron polarisability. Nevertheless, given practically available computer resources, they seem to provide adequate balance between the accuracy of system representation and the possibility to sample its states. As such, they remain the method of choice to study hydration effects in biomolecular systems at the atomistic level.

In some applications, however, such as computer aided drug design, explicit solvent simulations are still too time consuming. Such situation calls for alternative, simplified methods, which for a given configuration of molecules under study allow the assessment of hydration free energy, thus quantifying the thermodynamic contribution of aqueous environment [11]. Such implicit solvent methods are typically based on the assumption that water can be described as a continuous medium of high dielectric constant, separated from low dielectric solute interior by some kind of boundary surface. Hydration free energy of such system is than expressed as a sum of a) electrostatic term, accounting for the free energy of solute paritial charges in the solvent reaction field, and b) non-polar term, corresponding to the work needed to create the boundary surface around now neutralised solute. A popular way of evaluating the electrostatic term in applications to biomolecular systems is based on the solution of Poisson-Boltzmann (PB) equation [12] or its analytical approximation given by the generalised Born (GB) model [13]. The non-polar term is typically assumed as proportional to the area of the boundary surface (surface area – SA), with certain proportionality constant playing the role of a microscopic surface tension [14].

The accuracy of implicit solvent models in predictions of hydration free energies for small solutes is satisfactory [15]. Little is known, however, about the actual quality of their predictions for large, macromolecular solutes such as proteins, since there are no experimental data allowing for direct, quantitative comparisons. Unfortunately, aside from reports documenting successful applications of implicit solvent methods, there are numerous cases in which the results were not in agreement with experimental findings or the results obtained with fully atomistic models [16, 17, 18]. Aside from conceptual problem arising from attempts to evaluate hydration free energy of a system based only on its few fixed conformations, it is likely due to several fundamental shortcomings of standard implicit solvent models:

a) The boundary surface separating solute interior from the solvent region is typically a geometric construct based on the positions of solute atoms and their predefined radii. Its topography is thus insensitive to physicochemical properties of local environment and does not account for shifts resulting from high electric field, or, so called, drying effects responsible for the expulsion of water molecules from sterically accessible but hydrophobic regions [19].

b) Water polarisation in the first hydration layer is different from bulk polarisation due to packing effects and peculiar patterns of hydrogen bonds enforced by the presence of solute atoms [20].

c) Standard versions of implicit solvent models do not take into account the difference in hydration of opposite charges that arises due to asymmetry of charge distribution in water molecule [21].

d) Implicit solvent can neither predict nor account for the presence of isolated water molecules present within macromolecular structures.

e) The work of cavity creation (the non-polar term) is not linearly proportional to its surface area in molecular scales. The actual proportionality depends on local surface curvature (for small solutes of several Å diameters, the work scales with volume, being an entropic effect, different from the pressure-volume work), as well as on packing and type of solute atoms [2].

Overcoming the above limitations is a matter of intense research. It is focussed on the optimisation of atomic parameters (radii, partial charges), the definition of boundary surface [22], the problem of charge asymmetry[23], or better approximation of nonpolar contributions [24]. In parallel, alternative approaches are being developed, such as the model of polarisable dipoles [25], methods based on the density functional theory [26], or strategies utilising water density distribution derived from explicit solvent simulations [27]. A separate problem is related to the prediction of buried water molecules and the assessment of their thermodynamic effect. Most available approaches are either too costly or too inaccurate, to be routinely used for practically oriented applications such as the evaluation of receptor-ligand interactions.

Scientific goals of the research under consideration were to further investigate the (thermo)dynamics of hydration effects in biomolecular systems and to develop new modelling methods. Specific results include: a) the analysis of hydration influence on thermodynamic and kinetic parameters of association in a model receptor-ligand system based on explicit solvent MD simulations (publications, **H1** – **H3**), b) the comparison of hydration free energy estimates for proteins based on explicit and implicit solvent models (publication **H4**), c) a novel approach to the modelling of biomolecular hydration, attempting to address the shortcomings of standard implicit solvent methods while maintaining their high computational efficiency (publications **H5** – **H7**), d) the analysis of the role played by specifically bound water molecules in the catalytic subunit of protein kinase A for substrate binding cooperativity (publication **H8**).

4.3.2 Specific aspects of hydration effects in biomolecular systems

The first group of publications was devoted to the role of hydration effects for the association process in a model receptor-ligand system. A simplified model of a biomolecular receptor binding site was represented by a hemispherical cavity of ~ 8 Å radius, embedded in paraffin-like material. Its more or less hydrophobic character was tuned by the adjustment of a point charge localised at the bottom of the cavity. Ligands were modelled as a methane-like spherical particles with various total charges. Extensive explicit solvent MD simulations were carried out in order to obtain thermodynamic parameters of the association process, as well as, to investigate the influence of solvent fluctuations within the binding cavity on association kinetics. According to my best knowledge, those were the first studies dealing with detailed analysis of hydration effects in a host-guest system involving concave binding sites of various electrostatic and hydrophobic characteristics. Most research to date was focussed on the interactions between generally convex binding partners, thus bringing only limited insights into more realistic receptor-ligand association.

The publication **H1** is devoted to the analysis of potentials of mean force and their entropic and enthalpy contributions for the association in seven different versions of the model receptor-ligand system with various charge configurations. The results demonstrate that contributions to the total binding free energy that originate from hydration effects are of the same magnitude as direct receptor-ligand interactions. Moreover, it was shown that the role of aqueous solvent is not merely limited to scaling of interactions observed *in vacuo*, but it can qualitatively change the overall thermodynamic effect of the binding process. It is well exemplified by less favourable association in the system with oppositely charged binding partners than in the neutral one. It is due to the fact that strong electrostatic attraction in the first case was almost completely balanced by charge dehydration penalty, while relatively weak dispersion interactions in the second case, were substantially augmented by hydrophobic effect. The second interesting observation was related to systems in which the charge was present in only one of the binding partners. Depending wether it was localised within the cavity or in the ligand, the association changed from favourable to unfavourable, respectively, even though direct host-guest interactions were the same. This discrepancy resulted from different dehydration penalties associated with different charge localisation. Interesting insights were also obtained for enthalpy and entropic contributions to free energy, which turned out to depend much on the sign of solutes charge due to different structure of their hydration shells.

Particularly interesting thermodynamic signature was observed in the hydrophobic system, described in the publication **H2**. In agreement with previous results, the association was highly favourable. Unexpectedly, however, the change in free energy was dominated by enthalpy contribution, contrary to the standard view that hydrophobic interactions are driven by entropy. The analysis of thermodynamic effects during the association process indicated that unfavourable loss of entropy was concurrent with the onset of so called drying transition, that is abrupt solvent expulsion from the binding cavity once the approaching ligand reached certain critical distance. The entropic penalty was associated with quenching of pronounced wetting-dewetting solvent fluctuations that were observed in the binding site exposed to the bulk solvent. Importantly, the drying transition was not

associated with a change in free energy due to perfect enthalpy-entropy compensation. It determined, however, the dominant, enhalpy-driven character for the entire process since ligand dehydration alone that was responsible for the overall change in free energy was not enough to modify the signs of enthalpy and entropy contributions.

Although the effects described in the publication **H2** were based on a model system, their characterisation is of high practical importance. Following constant advances in computational and experimental methods, large hydrophobic enclosures filled with disordered, dynamic solvent of potentially high entropy are observed in a growing number of biomolecular systems. Dehydration of such regions, for instance during nonpolar ligand binding, indeed seems to involve peculiar hydrophobic effects with atypical thermodynamic signature.

The aforementioned spontaneous wetting-dewetting transitions in the hydrophobic binding cavity, turned out to have a nontrivial impact on association kinetics, whose analysis is the subject of the publication **H3**. The mean first passage time for ligand binding in the purely hydrophobic system evaluated using explicit solvent MD simulations was about 100 ps longer than estimated based on brownian dynamics diffusion in the respective mean force potential, under the assumption of bulk-like diffusion constant for the ligand. Such effect was unexpected, since the drying transition, being a small scale hydrophobic collapse, was assumed to speed up the association with respect to standard diffusive process. A detailed analysis led to the conclusion that water fluctuations contributed to the occurrence of forces acting on the ligand in a longer time scale than bulk-like random collisions with solvent molecules. Their maximum influence was observed as the ligand was at the entrance to the binding cavity, and resulted in ~ 6 fold increase in the effective translational friction coefficient, leading to slowing down of the association process. This new and relatively unexplored aspect of hydration effects may be of potential importance for modulation of binding kinetics in bimolecular systems. Its further investigation has been pursued in collaboration with prof. Joachim Dzubiella from Helmholtz Zentrum Berlin, resulting in a series of publications (**A4** – **A5**, below).

4.3.3 Representation of hydration effects in computer simulations

The results described in publications **H1** – **H2** seem to confirm important, and non-trivial role of aqueous environment in shaping bimomolecular free energy landscape. It motivates the question about the accuracy and consistency with which hydration effects are represented by the available computational approaches. Experimental hydration free energies are known for many small, druglike solutes. Their reproduction by calculations based on explicit solvent simulations or by means of implicit solvent models, has been evaluated in a number of studies, bringing knowledge regarding the expected accuracy and awareness of possible limitations. In the case of large biomolecules, such as proteins or nucleic acids, there is no experimental data allowing for direct, quantitative comparison with theoretical results. Accordingly, our belief that they reproduce hydration effects for large solutes with sufficient accuracy is based on indirect measurements and on the assumption that models and their parameters validated for small compounds can be also applied for macromolecules.

An attempt to validate such assumptions is the subject of publication **H4**. It is based

on calculations of differences in hydration free energy that arise due to conformational changes of five diverse protein structures. The calculations were based on explicit solvent MD simulations, independently for two popular water models: TIP3P and SPC/E, as well as using implicit solvent approach, utilising Poisson equation and surface area or volume based non-polar terms. The results indicate that discrepancies between the two explicit water models are much smaller (roughly 5 times) than between them and the considered implicit solvent model, where the root mean square error reaches 10 kcal/mol. Notably, in all cases discrepancies scaled with solute size rather than absolute hydration free energy. It indicates that potential problems are related to the representation of local hydration peculiarities (the size and complication of the first hydration layer increase with the solute size), rather than the description of global electrostatic effects due to bulk water polarisation (for smaller proteins, even changes in hydration free energy in the range of 100 kcal/mol were predicted consistently by all models).

Hydration free energy changes evaluated based on explicit solvent simulations were treated as reference, and used to investigate the transferability of implicit solvent model parameters such as protein dielectric constant, boundary surface definition, or the effective surface tension between different kinds of solutes. The results indicate that there is no single, universally optimal parameter set. Moreover, some assumptions valid for small molecules (represented in this work by isolated amino acids) are no longer justified in the case of large protein structures, since their complex surface topography and large solvent inaccessible core limit the range of usable surface definitions and corresponding values of internal dielectric constant. In particular, it was demonstrated that frequently used van der Waals surface definition is not adequate for macromolecular.

Significant problems were demonstrated for quantitative representation of non-polar effects by implicit solvent model. According to explicit solvent simulations, non-polar contributions to hydration free energy differences between distinct protein conformations can be in the order of tens kcal/mol. In some cases their incorrect reproduction by implicit models was the main source of overall differences with respect to simulation results. This observation highlights the lack of simplified, yet accurate treatment of the non-polar effects. While this problem is not well recognised in the context of drug-like solutes, whose non-polar contributions are small, it may become an accuracy limiting factor for macromolecules.

In order to obtain possibly accurate reference results for hydration free energy changes based on explicit solvent simulations, a new reestimation method of thermodynamic parameters was introduced. It increases the accuracy of results in the case of state functions, whose values are recorded for a closed thermodynamic cycle. The reestimation is based on the introduction of a constraint, that forces the summation of subsequent parameter changes to zero, subject to minimisation of differences from their originally estimated values, weighted by respective uncertainties. The procedure leads to a set of results that automatically fulfil the requirement of zero sum over a closed cycle, and whose uncertainties are smaller than the original ones. It is worth emphasising, that the method is applicable not only to hydration free energies, but to any set of thermodynamic data that fulfils initial conditions.

Another non-standard methodological aspect of the work **H4** was the analysis of corrections to explicit solvent hydration free energies accounting for that use periodic

boundary conditions in explicit solvent simulations. It was demonstrated that such corrections, though neglected in most other studies, can reach several kcal/mol. Accordingly, their evaluation is essential for proper comparison of explicit solvent results with predictions of implicit solvent models that are typically obtained under assumption of non-periodicity.

4.3.4 Semi-explicit solvent model

Inherent limitations of standard implicit solvent models prompt efforts to develop alternative approaches. One such concept is described in publications **H5** – **H7**. It is based on discrete solvent representation, defined on a body centred cubic lattice (BCC lattice). The model assumes that lattice points can be either occupied by a water molecule or empty. The occupancy is governed by local water excess chemical potential, which depends on effective interaction with the solute and the rest of aqueous environment. The solute-solvent term is evaluated based on orientation-averaged energy of a solvent probe in force-field like electrostatic and Lennard-Jones potentials generated by solute atoms. The solvent-solvent term is described by hydrogen bond interactions, whose presence depends on the occupancy of surrounding lattice. The definition of this term benefits from the geometry of BCC lattice, in which the arrangement of nearest neighbours reflects tetrahedral distribution of up to four hydrogen bonds possible for a single water molecule.

Solvent distribution hydration free energy for a given solute conformation is obtained in an iterative manner, starting from lattice that is uniformly occupied by solvent. Solvent evolution is propagated by emptying lattice points whose instantneous excess chemical potential is higher (less favourable) than a given threshold. It is iterated till a stationary distribution is obtained.

Theoretical formulation of the model and its preliminary parametrisation, based on experimental hydration free energies for 156 drug-like molecules, were described in the publication **H5**. It was demonstrated that, using only 6 adjustable parameters, the model reproduces hydration free energies of 16 test molecules with accuracy similar to free energy calculations based on explicit solvent simulations. Notably, the proposed approach avoids most limitations of standard implicit solvent models mentioned in the section 4.3.1. In particular, it does not depend on any geometrically defined solute-solvent boundary, but instead, solvent distribution in the presence of a solute is obtained as a result of calculations. This gives a unique possibility of capturing non-trivial hydration phenomena such as drying effects. Of particular importance in this context is the problem of discriminating between isolated hydration sites within macromolecular structures and sterically accessible to water, but void cavities (see publication **H7** below). The proposed model does not rely on any arbitrary chosen surface tension, typically needed to evaluate non-polar contributions. Interestingly, even though the scaling of non-polar hydration free energies with solute size was not taken into account during model parametrisation, it turned out to quantitatively reproduce the dependence of microscopic water surface tension on interface curvature in agreement with the scaled particle theory (SPT).

Further development of the model is described in the publication **H6**. It includes: a) improvements in the description of the first hydration shell surrounding a solute, allowing to exclude one of six adjustable parameters, b) the introduction of long range analytic

corrections for electrostatic and dispersion interactions, c) model reparametrisation based on a new training set of 153 small solutes (including 33 atomic and molecular ions) with experimentally determined hydration free energies. Calculations for a test set of 533 solutes (with 11 ions) demonstrated that, with similar computational load as a standard PBSA method, the model delivers better hydration free energy predictions, in particular for charged solutes.

Studies described in the publication **H7** were devoted to further explore the capability of the model to deliver spatial solvent distribution in the presence of a solute. A method was introduced to convert an array of occupied BCC lattice points into most probable locations of distinct water molecules and to estimate their excess free energies. It is of particular practical importance for modelling of macromolecular solutes and their complexes, whose structures trap multiple isolated water molecules. Their correct positioning and the knowledge of affinity to the host system is important for proper theoretical description. Owing to practical applications, for instance in computer aided drug design, algorithmic developments in this are are in high demand.

The results derived by the semi-explicit hydration model were compared with reference water binding free energies to protein structures, based on explicit solvent free energy perturbation calculations employing the double decoupling method. It was demonstrated that the accuracy of occupancy predictions for putative hydration sites (that is the probability of correct discrimination between an actually hydrated site and a cavity that should remain empty) was larger than 90 %. Also, the numerical estimates of water binding free energies were in good agreement with reference data (root mean square error for 48 considered sites was ~ 2 kcal/mol).

4.3.5 The effects of specifically bound water molecules

Macromolecular structures contain multiple isolated water molecules, bound in specific hydration sites. The conservation of hydration sites distribution observed in a number of evolutionary related proteins raises a question concerning their putative functional role. The aforementioned sem-explicit hydration model (publicatins **H5** – **H7**) was applied for systematic analysis of hydration sites locations in catalytic subunits of protein kinases. 13 hydration sites that were universally present in structures representing diverse subfamilies of the kinase family were selected for detailed studies. Two of them, located in the region suggesting possible role in substrate binding turned out to be particularly interesting. Based on long MD simulations involving the protein kinase A (PKA) it was demonstrated (publication **H8**) that water molecules within those two hydration sites contribute to experimentally observed binding cooperativity of ATP and peptide substrate in PKA. A single water molecule within the first site (site A) was shown to participate in a specific hydrogen bond network. It was involved in propagation of the effect of local conformational change triggered by ATP binding to peptide binding site, promoting its conformation that is favourable for substrate binding. In turn, several water molecules located in the second site (site B) turned out to adopt highly dynamic configurations, whose readjustments allowed maintaining the flexibility of the peptide positioning segment, enabling its fine-tuning by the aforementioned allosteric signal. The simulations also shed light on the atomistic mechanism

of tyrosine 204 to alanine (Y204A) mutation in PKA, whose effect is experimentally known to suppress substrate binding cooperativity. The results suggest that the mutation affects water fluctuations within the site B, thus impairing the flexibility of the peptide positioning region and making it insensitive to allosteric modulation. The gathered insights, seem to provide an example of important functional role of buried water molecule, indicating an interesting, yet poorly investigated research topic.

4.3.6 Application of the achieved results

A practical application of the achieved results in envisioned for the semi-explicit solvent model. At the moment, the model is implemented in a computer program written in C programming language. The program, available for external users, allows the calculation of hydration free energies for molecular solutes described by standard atomistic force field, as well as the prediction of water distribution and affinity to distinct hydration sites. Owing to code parallelisation the program efficiently uses contemporary multicore processors, delivering predictions for mid-sized proteins ($\sim$ 250 amino acids) within several minutes.

The program serves also as an engine for actively developed Internet service (available under: https://gsolvate.biomod.cent.uw.edu.pl). It offers the users a possibility to submit their structure of interest, for instance in Protein Databank file format, and execute all tasks remotely. In the near future, the service will be extended to provide a descriptor of local protein surface hydration propensity based on semi-explicit solvent calculations. Such descriptors may be used for the detection of regions on protein surface involved in complex formation with other binding partners. The completion of Internet service is scheduled for the end of 2018.

5 Other scientific acheivements

5.1 Hydrophobic effects involving drying transitions

My PhD thesis, completed under the supervision of Dr. Maciej Geller, was devoted to studies of hydrophobic effects. I introduced a model of a concave, hydrophobic receptor binding site in a form of a hemispherical cavity of $\sim$ 8 Å radius that was embedded in paraffin-like material. Using explicit solvent simulations I demonstrated (publication **A1**) that water does not form distinct hydration layers next to the concave cavity walls, but rather remains disorganised, resembling a liquid-vapour interface. Furthermore, collective behaviour of water molecules turned out to lead to pronounced density fluctuations within the pocket, adopting the form of wetting-dewetting transitions.

Subsequent work was devoted to the analysis of hydration effects accompanying association of the considered pocket with a model hydrophobic ligand, represented by a methane-like particle. It was shown that at some critical cavity-ligand separation, cavity interior becomes permanently dewetted (publication **A2**). Solvent expulsion occurred even though volume between the binding partners was still sterically accessible to water molecules, thus constituting a model of a hydrophobic collapse. The analysis of potentials

of mean force derived based on MD simulations (publications **A2** and **A3**) indicated that association in a system that undergoes a drying transition can occur without free energy barrier. This is in contrast to hydrophobic association between previously studied solutes whose size and surface topography did not trigger dewetting, thus implying the need of additional work for crossing through hydration layers.

The obtained results raised intriguing questions concerning the influence of barrie-rless potential of mean force in the considered system on association kinetics. Following initial interesting findings (see publication **H3** above), this research topic is further pursued in the group of prof. Joachim Dzubiella from Helmholtz Zentrum Berlin with me as an external collaborator (publications **A4** – **A5**).

A1 P. Setny, M. Geller, *Water properties inside nanoscopic hydrophobic pocket studied by computer simulations.* J. Chem. Phys., 2006, 125:144717,

A2 P. Setny, *Water properties and potential of mean force for hydrophobic interactions of methane and nanoscopic pockets studied by computer simulations.* J. Chem. Phys., 2007, 127:54505,

A3 P. Setny, *Hydrophobic interactions between methane and a nanoscopic pocket: three dimensional distribution of potential of mean force revealed by computer simulations.* J. Chem. Phys., 2008, 128:125105,

A4 R. G. Weiss, P. Setny, J. Dzubiella, *Solvent Fluctuations Induce Non-Markovian Kinetics in Hydrophobic Pocket-Ligand Binding.* J. Phys. Chem. B, 2016, 120:8127,

A5 R. G. Weiss, P. Setny, J. Dzubiella, *Principles for Tuning Hydrophobic Ligand–Receptor Binding Kinetics.* J. Chem. Theory Comput., 2017, 13:3012.

5.2 Variational implicit solvent model

The results concerning drying effects in a model receptor-ligand system, obtained during my PhD studies, spurred collaboration with the research group of prof. J. Andrew McCammon from University of California, San Diego. The system served for parametrisation of the Variational Implicit Solvent Model (VISM), in which hydration free energy is expressed as a functional of the solute-solvent boundary. The actual position of solvent surface around solute atoms is obtained upon minimisation of the functional, which takes into account terms dependent on steric and dispersion solute-solvent interactions, as well as the area and local curvature of the boundary. The application of VISM to describe the association in the considered model system demonstrated that it is capable of reproducing drying effects (publications **B1** and **B2**). In agreement with the results of MD simulations VISM indicated two distinct hydration free energy minima, corresponding to and empty and solvent-occupied binding cavity, as well as complete drying transition for cavity-ligand separations smaller than a certain threshold. My role in this project was to conduct additional MD simulations in order to increase the accuracy of previous free energy calculations and to get new insights into the details of dewetting process, as well as to participate in VISM parametrisation and interpretation of its results.

B1 P. Setny, Z. Wang, L.-T. Cheng, B. Li, J. A. McCammon, J. Dzubiella, *Dewetting-Controlled Binding of Ligands to Hydrophobic Pockets.* Phys. Rev. Lett., 2009, 103:187801,

B2 L.-T. Cheng, Z. Wang, P. Setny, J. Dzubiella, B. Li, J. A. McCammon, *Interfaces and hydrophobic interactions in receptor-ligand systems: A level-set variational implicit solvent approach.* J. Chem. Phys., 2009, 131:144102.

5.3 Modelling of protein – nucleic acid interactions

The main topic of my postdoctoral research conducted in the group of prof. Martin Zacharias at the Technical University of Munich was the development of methods aimed at the prediction of protein-nucleic acid complexes. In order to extend the functionality of the ATTRACT program, initially developed in the Zacharias group for protein-protein docking, I designed and parametrised a coarse-grained potential describing protein – nucleic acid interactions. Its parametrisation and the results of application to protein-RNA and protein-DNA docking were described in publications **C1** and **C2**, respectively. With the aim to account for conformational changes of macromolecules during complex formation, I systematically evaluated the applicability of elastic network models to describe global deformations of nucleic acid structures. Such models are widely used for proteins, however, due to specific architecture of nucleic acids their optimal implementation and actual performance for this group of macromolecules remained practically unknown. Based on the analysis of multiple variants of elastic network models of RNA and DNA structures I proposed an optimal placement of interaction centres and force constant definition (publication **C3**). The developed methods were practically applied for modelling of polymerase RNA complex with 6S RNA, conducted in cooperation with an experimental group (publication **C4**).

C1 P. Setny, M. Zacharias, *A coarse-grained force field for Protein-RNA docking.* Nucleic Acids Res., 2011, 39:9118,

C2 P. Setny, R. P. Bahadur, M. Zacharias, *Protein-DNA docking with a coarse-grained force field.* BMC Bioinformatics, 2012, 13:228,

C3 P. Setny, M. Zacharias, *Elastic Network Models of Nucleic Acids Flexibility.* J. Chem. Theory Comput.,2013, 9:5460,

C4 B. Steuten, P. Setny, M. Zacharias, R. Wagner, *Mapping the spatial neighbourhood of the regulatory 6S RNA bound to Escherichia coli RNA polymerase holoenzyme.* J. Mol. Biol, 2013, 425:3649.

5.4 Receptor-ligand interactions and computer aided drug design

The application of computer simulations for the estimation of receptor-ligand binding free energies was an important aspect of my master thesis, developed under the supervision of Dr. Maciej Geller. Using free energy perturbation methods I investigated the mechanism of dual specificity of the CK2 kinase towards its phosphate-donating substrate. I demonstrated that

it relies on the presence of several water molecules buried within the binding site. Binding of two different substrate with equal affinity turned out to be possible owing to rearrangement of specific hydrogen bond network maintained by those water molecules (publication **D1**).

Since the early years of my studies at the Physics Department, University of Warsaw, computer aided drug design (CADD) has been one of my important areas of interest. Working at the Interdisciplinary Centre for Mathematical and Computational Modelling (ICM, UW) I was responsible for the coordination of Polish country-wide licences for CADD software and for scientific support to Polish academic users. Summarising the experience gained during that time I wrote a review of CADD methods, available in electronic form at the ICM UW web page (`https://kdm.icm.edu.pl/kdm/Projektowanie_leków`). I also took part in two research projects utilising CADD techniques. The first, conducted under the supervision of prof. Bogdan Lesyng, was devoted to the development of selective ligands for JAK kinases. The second, coordinated by prof. Joanna Trylska, aimed at the investigation of aminoglycoside binding to the ribosomal site A and subsequent design of novel antibiotic compounds (publications **D2** and **D3**).

D1 P. Setny, M. Geller, *Refinement of X-ray data on dual cosubstrate specificity of CK2 kinase by free energy calculations based on molecular dynamics simulation.* Proteins, 2005, 58:511,

D2 J. Romanowska, P. Setny, J. Trylska, *Molecular dynamics study of the ribosomal A-site.* J. Phys. Chem. B, 2008, 112:15227,

D3 P. Setny, J. Trylska, *Search for novel aminoglycosides by combining fragment-based virtual screening and 3D-QSAR scoring.* J. Chem. Inf. Model., 2009, 49:390.

5.5 Interaction of influenza virus fusion peptide with lipid bilayer

This project is conducted in collaboration with experimental group of dr. Remigiusz Worch from the Polish Academy of Sciences. My role is to use computational modelling in order to obtain atomistic insights into the functioning of influenza virus fusion peptide within lipid bilayer and to provide structural interpretation of experimental findings. The studies are based on MD simulations employing temperature replica exchange method (REMD). I investigated the influence of various peptide conformations and configurations within the membrane on local lipid disorder, and explained the mechanism by which three specific, conserved amino acids may be responsible for enhancing peptide propensity to induce membrane fusion (publication **E1**). Other results were related to the functional role of N-terminal peptide group (publication **E2**). Current work is devoted to the influence of cholesterole within the lipid membrane on peptide function.

E1 R. Worch, J. Krupa, A. Filipek, A. Szymaniec, P. Setny, *Three conserved C-terminal residues of influenza fusion peptide alter its behavior at the membrane interface.* Biochim. Biophys. Acta - Gen. Subj., 2017, 1861:97,

E2 R. Worch, A. Dudek, J. Krupa, A. Szymaniec, P. Setny, *Charged N-terminus of Influenza Fusion Peptide Facilitates Membrane Fusion.* Int. J. Mol. Sci. 2018, 19:578.

References

[1] P. Ball, Water is an active matrix of life for cell and molecular biology, *Proc. Natl. Acad. Sci.* **2017**, 201703781 (2017).

[2] D. Chandler, Interfaces and the driving force of hydrophobic assembly., *Nature* **437**(7059), 640 (2005).

[3] A. Warshel, J. Åqvist, Electrostatic energy and macromolecular function, *Annu. Rev. Biophys. ...* **20**, 267 (1991).

[4] V. Helms, Protein dynamics tightly connected to the dynamics of surrounding and internal water molecules., *Chemphyschem* **8**(1), 23 (2007).

[5] S. Park, J. G. Saven, Statistical and molecular dynamics studies of buried waters in globular proteins., *Proteins* **60**(3), 450 (2005).

[6] M. Levitt, B. H. Park, Water: now you see it, now you don't., *Structure* **1**(4), 223 (1993).

[7] V. Conti Nibali, M. Havenith, New insights into the role of water in biological function: Terahertz absorption spectroscopy and molecular dynamics simulations studies of the solvation dynamics of biomolecules, *J. Am. Chem. Soc.* **136**, 12800 (2014).

[8] M. P. Allen, D. J. Tildesley, Computer Simulation of Liquids, *Liq. Oxford Univ. Press. New York* **18**(195), 385 (1987).

[9] A. C. Pan, T. M. Weinreich, S. Piana-Agostinetti, D. E. Shaw, Demonstrating an order-of-magnitude sampling enhancement in molecular dynamics simulations of complex protein systems, *J. Chem. Theory Comput.* p. acs.jctc.5b00913 (2016).

[10] A. Pohorille, C. Jarzynski, C. Chipot, Good practices in free-energy calculations., *J. Phys. Chem. B* **114**(32), 10235 (2010).

[11] B. Roux, T. Simonson, Implicit solvent models., *Biophys. Chem.* **78**(1-2), 1 (1999).

[12] K. A. Sharp, B. Honig, Calculating total electrostatic energies with the nonlinear Poisson-Boltzmann equation, *J. Phys. Chem.* **94**, 7684 (1990).

[13] W. C. Still, A. Tempczyk, R. C. Hawley, T. Hendrickson, Semianalytical treatment of solvation for molecular mechanics and dynamics, *J. Am. Chem. Soc.* **112**, 6127 (1990).

[14] R. M. Levy, L. Y. Zhang, E. Gallicchio, A. K. Felts, On the nonpolar hydration free energy of proteins: surface area and continuum solvent models for the solute-solvent interaction energy., *J. Am. Chem. Soc.* **125**(31), 9523 (2003).

[15] D. L. Mobley, K. L. Wymer, N. M. Lim, J. P. Guthrie, Blind prediction of solvation free energies from the SAMPL4 challenge., *J. Comput. Aided. Mol. Des.* **28**(3), 135 (2014).

[16] R. Zhou, B. J. Berne, Can a continuum solvent model reproduce the free energy landscape of a beta -hairpin folding in water?, *Proc. Natl. Acad. Sci. U. S. A.* **99**(20), 12777 (2002).

[17] D. A. Pearlman, Evaluating the molecular mechanics poisson-boltzmann surface area free energy method using a congeneric series of ligands to p38 MAP kinase., *J. Med. Chem.* **48**(24), 7796 (2005).

[18] A. Cumberworth, J. M. Bui, J. Gsponer, Free energies of solvation in the context of protein folding: Implications for implicit and explicit solvent models, *J. Comput. Chem.* **37**, 629 (2016).

[19] L. Wang, B. J. Berne, R. A. Friesner, Ligand binding to protein-binding pockets with wet and dry regions., *Proc. Natl. Acad. Sci. U. S. A.* **108**(4), 1326 (2011).

[20] H. S. Muddana, N. V. Sapra, A. T. Fenley, M. K. Gilson, The electrostatic response of water to neutral polar solutes: implications for continuum solvent modeling., *J. Chem. Phys.* **138**(22), 224504 (2013).

[21] D. L. Mobley, J. R. Baker, A. E. Barber, C. J. Fennell, K. A. Dill, Charge asymmetries in hydration of polar solutes., *J. Phys. Chem. B* **112**(8), 2405 (2008).

[22] Grant J. Andrew, A smooth permittivity function for Poisson-Boltzmann solvation methods, *J. Comput. Chem.* **22**(6), 608 (2001).

[23] E. O. Purisima, T. Sulea, Restoring charge asymmetry in continuum electrostatics calculations of hydration free energies., *J. Phys. Chem. B* **113**(24), 8206 (2009).

[24] J. A. Wagoner, N. A. Baker, Assessing implicit models for nonpolar mean solvation forces: the importance of dispersion and volume terms., *Proc. Natl. Acad. Sci. U. S. A.* **103**(22), 8331 (2006).

[25] J. Florián, A. Warshel, Langevin Dipoles Model for ab Initio Calculations of Chemical Processes in Solution: Parametrization and Application to Hydration Free Energies of Neutral and Ionic Solutes and Conformational Analysis in Aqueous Solution, *J. Phys. Chem. B* **101**(28), 5583 (1997).

[26] A. Kovalenko, F. Hirata, Three-dimensional density profiles of water in contact with a solute of arbitrary shape: a RISM approach, *Chem. Phys. Lett.* **290**(1-3), 237 (1998).

[27] T. Lazaridis, Inhomogeneous Fluid Approach to Solvation Thermodynamics. 1. Theory, *J. Phys. Chem. B* **102**, 3531 (1998).

28.03.2018