Charles University in Prague

Faculty of Mathematics and Physics

DOCTORAL THESIS

Ing. Michaela Nekardová

Computational Studies of Interactions of Small Molecules with Their Biological Targets

Institute Organics Chemistry and Biochemistry,

Czech Academy of Sciences

Supervisor of the doctoral thesis:

Prof. Ing. Pavel Hobza, Dr. Sc., dr. h. c., FRSC

Study programme: Physics

Specialisation: Biophysics, Chemical and Macromolecular Physics

Prague 2020

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular the fact that the Charles University in Prague has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 paragraph 1 of the Copyright Act.

In Prague, 2th February 2020   Michaela Nekardová

Acknowledgement

First and foremost, I thank my family and close friends.

I am thankful to my adviser prof. Pavel Hobza for his support and advice.

I am indebted to my consultants dr. Jindřich Fanfrlík and dr. Jan Řezáč for their assistance and especially for their technical advice.

My great thanks belong to all my collaborators and co-authors as well.

Last but not least, I appreciate the colleagues from prof. Hobza's group for creating a friendly working atmosphere.

Title: Computational studies of interactions of small molecules with their biological targets.

Author: Ing. Michaela Nekardová

Department / Institute: Department of computational chemistry, Institute Organics Chemistry and Biochemistry, Czech Academy of Sciences

Supervisor of the doctoral thesis: Prof. Ing. Pavel Hobza, Dr. Sc., dr. h. c., FRSC,

Institute Organics Chemistry and Biochemistry, Czech Academy of Sciences

Abstract:

The thesis specializes in the computational description of pharmaceutically important compounds.

A substantial number of pharmaceutical drugs are small molecules that are bound to an active site of an enzyme by the "lock (binding site) and key (drug)" model through non-covalent interactions. The association of enzymes with drugs cause an increase or decrease in the activity of enzymes. The main topic is focused on the computational elucidation of the structural basis for the interactions of the purine-like compounds with the enzyme cyclin-dependent kinase 2 that belongs to the protein-kinase enzyme family. These enzymes play an important role in the cell cycle regulation; their increased activity significantly contributes to the loss of control over cell proliferation, which is one of the primary causes of cancer cell formation. The study describes the binding motifs of roscovitine, which shows an inhibitory effect on the function of cyclin-dependent kinases, and its analogues containing bioisosteric central heterocycles in the complex with cyclin-dependent kinase 2. The binding affinity between the cyclin-dependent kinase 2 enzyme and the inhibitors was quantified as calculated binding scores and evaluated in relation to the conformation of the optimized structures. The hybrid model combining the quantum mechanics – QM (DFT-D) and semiempirical quantum mechanics – SQM (PM6-D3H4X) method was used. The solvent effect was described by the continuum solvation model COSMO at the SQM level for the whole system.

The second topic is aimed at the computational estimate physicochemical properties of several groups of neuroactive compounds that modulate the activity of the *N*-methyl-*D*-aspartate receptor. This receptor belongs to the family of glutamate receptors that are present in nerve cells. It is thought that it is connected with a variety of neurological disorders such as epilepsy, Parkinson's and Alzheimer's diseases. The computational studies investigate the lipophilic qualities and solvation free energy of neuroactive compounds as these properties are inherent characteristics of the substances and influence their interactions with the *N*-methyl-D-aspartate receptor. The solvation free energy was calculated in the SMD continuum solvation model at the HF/6-31G* level. The calculated logP was estimated on the basis of the change in the molecular conformation related to the transfer between *n*-octanol and water. The studied compounds were optimized by the RI-DFT-D3/B-LYP/TZVPP//COSMO method and their SP energy was calculated at the same level of accuracy.

This thesis is based on five of the ten publications that have been published on these topics.

Table of Contents

Introduction

The process of discovering and developing new medicines plays a major part in the current scientific research. Progress in biomedical research has led to a greater understanding of the inner workings of human diseases at the molecular level. One of the expected upshots of this knowledge is the preparation of new drugs with an increased affinity, efficacy, selectivity, metabolic stability, and oral bioavailability. Simultaneously, new problems have appeared such as antimicrobial resistance that poses a global health problem. Microorganisms (bacteria, viruses, parasites, fungi) have become resistant to antimicrobial drugs (antibiotics, antivirals, antimalarials, anthelmintics, antifungals) which had previously been used to treat them. For instance, according to World Health Organization 480 000 people develop multi-drug-resistant tuberculosis each year. Furthermore, drug resistance makes the fight against HIV and malaria difficult as well. The antimicrobials are also standardly used as prevention against infections in medical procedures such as organ transplantation, cancer therapy, diabetes management and major surgeries.[1-3] The antimicrobial resistance complicates and prolongs the treatment and can lead to disability or death. Cancer and dementia also belong to serious diseases whose incidence increases each year. This is due to several factors, such as population growth and ageing, unhealthy life style, or environmental degradation.[4] Cancer is the second-leading cause of death in the world, Figure 0.1.[5] It is a broad category of diseases (more than 100 different types) involving abnormal cell growth and division. Based on where it occurs, four main categories of cancer are recognized: carcinomas (internal organs and glands, e. g. breast, lung), sarcomas (support tissues, e.g. muscles, nerves), leukemias (blood), and lymphomas (lymphatic system). There are many risk factors; some are preventable, such as smoking cigarettes, immoderate alcohol consumption, poor nutrition, or excess body weight, others such as age, inherited genetic defects, gender, skin type, or pollutants are not. Genetic factors can contribute to the development of cancer. However, survival rate rises due to improvements in cancer screening and cancer treatment. Nowadays, chemotherapy is one of the best cures in the fight against cancer.[6-8]

Estimated number of new cases in 2018, worldwide, all cancers, both sexes, all ages
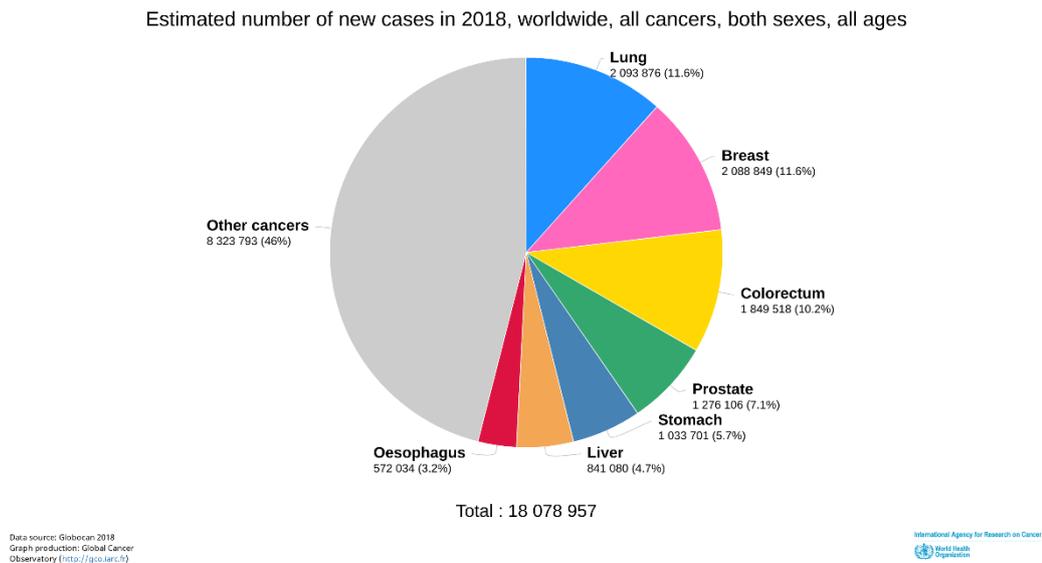
Figure 0.1: The global cancer statistics by The Global Cancer Observatory (GCO).

Dementia is a large group of brain diseases (e.g. Alzheimer disease, or Vascular dementia) with many causes that manifest themselves by deterioration in cognitive functions: memory, thinking, orientation, learning, judgement, behaviour, etc.[9] Worldwide, there are about 10 million new cases every year.[10] Although dementia mainly affects older people, it is not a normal part of ageing, but a neurological disorder result in loss of nerve cells and their connections in the brain. Alzheimer's disease is the most common cause of a dementia in older adults, but some types of this dementia affect young people as well.[10] Some causes of dementia or dementia-like symptoms can be reversed with treatment, e.g. vitamin B12 deficiency, infections and immune disorders, or subdural hematomas. To date, effective treatment of dementia has not been developed. Given the mentioned facts, it is evident that the need for new medicines is paramount.[9]

Drug discovery and development is a long-standing interdisciplinary effort by which new candidate medicines are researched. It includes lot of research fields that complement each other and often overlap to a certain extent.

In the past, compounds with therapeutics effect were identified from traditional (natural) medicaments (e. g. herbal drugs as quinine from the cinchona for treatment of malaria; morphine from the poppy used to suppress pain; or salicylic acid from white willow for the treatment of various skin diseases), or serendipitous discoveries (e. g. famed antibiotic

penicillin; anti-cancer drug cisplatin, anticoagulant warfarin, hormone insulin, or psychoactive drug chlordiazepoxide).[11-13]

Later classical pharmacology has found active substances in chemical libraries through phenotypic screening that is a time-consuming multi-step process. Synthetic molecules and natural compounds are screened in intact cells or whole animal organisms, which represent models of disease, for finding "lead compounds". They are biologically/pharmacologically active substances causing a desirable change in phenotype. The subsequent "target deconvolution" process identifies molecular targets (e.g. proteins) of the lead compounds. This crucial step elucidates mechanisms of action of the compounds and enables further optimization/modification of the active substances for better fitting of the biological target.

Recently, this process has become common practice; in contrast to classical pharmacology, it is called reverse pharmacology because the biological target is selected firstly. The approach works with a hypothesis that a modulation of the biological target is "disease modifying". This means that the alteration of specific activity of the biological target will cause beneficial curative effects. The drug target must also evince "druggability" which is the ability to bind a small molecule that modulates its activity. Potential drug compounds are searched in small compound libraries by widely used high-throughput screening (HTS). The technique is able to quickly carry out of millions of chemical, genetic and pharmacological *in vitro* tests of a biological or biochemical activity of a large number of drug-like compounds. HTS is useful for discovering ligands for various pharmacological targets such as enzymes, receptors, and ion-channels. The process is based on robotic conduct of experiments and automatic data processing when large number of compounds is tested against a chosen biological target. HTS enables to identify of numerous "hit" substances that undergo preliminary "hit to lead" optimization. Resultant lead compounds are usually optimized again for optimal fitting of the target.[13-17] Promising drug candidates are tested in pre-clinical studies whether they meet the preset criteria: pharmacodynamics and pharmacokinetic properties, ADME scheme (absorption, distribution, metabolism, excretion), and potential toxicity. After the validation of the required criteria, clinical trials can be initiated.[18]

A substantial number of pharmaceutical drugs are small molecules that are bound to an active site of an enzyme by the "lock (binding site) and key (drug)" model. The drugs cause an increase or decrease in the activity of enzymes, which elicit a physiological response (medical effect). The association of enzymes with drugs is driven by steric/spatial factors.

The interactions usually have non-covalent character that is fully reversible, and mostly include hydrogen bonds, dipol-dipol interactions, π-stacking, van der Waals interactions, and halogen bonds. Only small number of drugs are bound through chemical bond, i.e. covalently.[19-20]

There is a widespread consensus that the process of drug discovery and development requires many resources and it is very time-consuming. Powerful tools that enhance computational capacity have enabled extensive development of methods of computational chemistry that are able to minimize time and resource requirements. The computational approach significantly assists in solving chemical and biochemical problems. *In silico* simulations and modelling are based on the knowledge of three-dimensional structures, which enables us to understand the properties and the structural characteristics both of the small molecules and of the biological targets. It elucidates how shape and charge of a drug-like molecule influence its affinity to an active site of a biological target. It has been proved that computational chemistry is the effective way for the decription and prediction of interactions between small molecules and enzymes. The computational studies can serve to elucidate problems that cannot be satisfactory clarified in other way, generalize some phenomena, and provide a basis of further experimental work. Therefore, computer-aided drug design has become an integral part of a drug discovery and development. It can be mainly utilized for hit identification using virtual screening, hit-to-lead optimization and lead optimization.[21-23]

My work consists of two research topics specialized in pharmaceutically important compounds. The main topic is focused on the computational description of the structural basis for the interactions of small organic molecules with the enzyme cyclin-dependent kinase 2. Cyclin-dependent kinases belong to the serine/threonine protein-kinase enzyme family that play an important role in the cell-cycle regulation. Protein-kinases are enzymes which transfer the phosphate group from a high energy donor molecule (e. g. ATP) to a definite target molecule (substrate). The increased activity of cyclin-dependent kinases significantly contributes to the loss of control over cell proliferation, which is one of the primary causes of cancer cell formation. Thus, they represent a relevant target for cancer chemotherapy. Cyclin-dependent kinase 2 is essential for the cell cycle G1/S transition. It can be dispensable in the cell cycle of normally functioning cells, but it plays a critical role in the abnormal growth processes of cancer cells.[24-26] The discovery of roscovitine,[27] which is the ATP-competitive purine-based CDK inhibitor, has motivated further research leading

to the preparation of analogous compounds with an improved inhibitory efficiency. My work clarifies the binding motifs of the roscovitine and its analogues containing bioisosteric central heterocycles, which are non-covalently bound to the amino acids in the cyclin-dependent kinase 2 active site.[28] The binding affinity between the cyclin-dependent kinase 2 enzyme and the inhibitors was quantified as calculated binding scores and evaluated in relation to the conformation of the optimized structures. The hybrid quantum mechanics/semiempirical quantum mechanics (QM/SQM) method was used,[29] which employs the DFT-D method for QM and the PM6-D3H4X[30] method for SQM. The solvent effect was described by the continuum solvation model COSMO[31] at the SQM level for the whole system. The results have been successfully used for the preparation new efficient inhibitors.[32]

The second topic is aimed at the computational estimate of thermodynamic properties of several groups of neuroactive compounds that modulate the activity of N-methyl-$_D$-aspartate receptor. This receptor, which belongs to the family of glutamate receptors,[33] is glutamate-gated ion-channel protein present in nerve cells. It is involved in excitatory synaptic transmission, synaptic plasticity, and synapse formation underlying memory, and formation of neural networks during development in the central nervous system.[34-35] N-methyl-$_D$-aspartate receptor is associated with a variety of neurological disorders such as epilepsy, Parkinson's and Alzheimer's diseases.[36] The activity of the receptor can be influenced by allosteric modulators such as neurosteroids.[37] Although neuroactive compounds have a similar structure, they can inhibit as well as potentiate the receptor activity. My computational studies of the physicochemical properties of neuroactive compounds investigate the lipophilic qualities and solvation free energy, as these properties are inherent characteristics of the substances and influence their interactions with N-methyl-$_D$-aspartate receptor.[38-39] The solvation free energy of the compounds was calculated in the SMD continuum solvation model at the HF/6-31G* level.[40] The calculated logP was estimated on the basis of the change in the molecular conformation related to the transfer between n-octanol and water.[41-44] These computations complement and support the large experimental projects combining medicinal chemistry and cellular neurophysiology.

This thesis is based on five[28, 32, 45-47] of the ten publications that have been published on these topics. A publication list can be found at the end of the thesis together with the complete texts of the articles.

1. Non-covalent Interactions

Non-covalent interactions have an irreplaceable role in nature. Unique properties of both water and DNA are determined by hydrogen bonds that are the most important and frequent type of non-covalent interactions.[19] The use of non-covalent interactions is extensive in various disciplines such as: chemistry, physics, materials science, pharmacology, and particularly in biodisciplines.

A covalent bond is constituted if two subsystems with not fully occupied electron shells approach and overlap their orbitals, which involves the sharing of electron pairs (bonding pairs) between atoms. An effective distance between interacting atoms is shorter than 2Å. Unlike that, non-covalent interactions are formed between atoms that are more distant and even can have null orbital overlap. Electrical qualities enable these attractions. Electrically neutral molecules also form non-covalent interactions owing to asymmetric distribution of charge and polarizability. Non-covalent interactions can be expressed as

$$X + Y \leftrightarrow X \cdots Y.$$

Electrostatic interactions, inductive interactions, and London dispersion forces belong to long-range interactions. The exchange-repulsion term belongs to short-range interactions. All types of interactions act together in real systems.[20, 48]

1.1 Electrostatic Interactions

Electrostatic interactions are attractive or repulsive interactions between two molecules or ions and their electrostatic fields. Their strength is higher than that of the van der Waals forces. In vacuum or dielectric media, the electrostatic interactions are directed by Coulomb's Law.[48-49]

1.1.1 Hydrogen Bonding

A hydrogen bond (H-bond) is an electrostatic attractive force between a hydrogen atom, which is covalently bound to a more electronegative atom or group (donor), and an electronegative atom bearing a lone pair of electrons (acceptor): X-H⋯Y, where X and Y are electronegative atoms (X = F, O, N, also C) and Y has additionally lone pairs of electrons. Systems with delocalized electrons can be proton acceptors as well, e.g. H-bond between water and benzene (O-H⋯π). H-bond is the most frequent non-covalent composite interaction. H-bond is typical of the geometric specificity; it is the strongest if atoms X, H

a Y have the linear arrangement. H-bond is approximately tenfold weaker than a covalent bond but belongs among strong non-covalent interactions. Stabilization energy of a complex bound through H-bond is $3-5$ kcal.mol^{-1}. If a complex is formed through several H-bonds, stability evinces a synergy effect, and so the total sum is greater (as many as 40 kcal.mol^{-1}) than the simple sum of stabilization energy of single bonds. H-bond influences many chemical and physical properties of compounds such as melting and boiling point, solubility, and acid strength. This interaction is involved in stabilization of the secondary and tertiary structure of proteins. The secondary structures such as α-helix and β-sheet are stabilized strictly through H-bonds.[50]

1.1.1.1 Intermolecular Hydrogen Bond with Three-dimensional Network

The intermolecular H-bond with the three-dimensional network is characteristic of water. Each water molecule can create both two hydrogen bonds involving its own hydrogen atoms and two further hydrogen bonds utilizing the hydrogen atoms of neighbouring water molecules since oxygen bears two lone pairs of electrons. The hydrogen atoms are attracted towards the oxygen atom of neighbouring water molecules since the oxygen atom is partially negatively charged, and the hydrogen atom is partially positively charged. The network organisation of water molecules through H-bonds causes the collective ground state of liquid water to have an energy lower than the ground state of single gaseous molecules. The four hydrogen bonds around each water molecule tend to adopt the tetrahydral molecular geometry, which is the extensive ordering in the ice. In liquid water, the tetrahedral arrangement is detected only locally since thermal energy deforms or disturbs H-bonds in the clustering. Disorganization of linked water molecules is growing with increasing temperature. Despite this fact, H-bonds that connect water molecules in liquid to a network still keep large distances between molecules.[49-50]

1.1.1.2 Dihydrogen Bond

The dihydrogen bond is formed between two hydrogen atoms; one is covalently bound to a metal hydride and the other to a proton donor: M-H···H-Y, where M is an electropositive metal atom (boron, alkali metals, transition metals) and Y is an electronegative atom. Therefore, one hydrogen bears a positive and another a negative partial charge that enables the electrostatic attraction between both polarized hydrogen atoms.[50]

1.1.1.3 Blue-shifting Hydrogen Bond

The blue-shifting H-bond is a phenomenon showing the opposite characteristics to the classical H-bond. It means that during a formation of a complex, the X-H bond is shortened, the valence-vibration frequency is on the increase, and an intensity of a spectral line of the valence-vibration is on the decrease in wavelength.[50-51]

1.1.1.4 Intramolecular Hydrogen Bond

The intramolecular H-bond is formed inside a molecule if a hydrogen atom is bound to one atom covalently and to another through H-bond.[50]

1.1.1.5 Intraionic Hydrogen Bond

The intraionic H-bond is the special case of a hydrogen atom bound to two fluorine anions.[50]

1.1.2 Ionic Interactions

Ionic interactions are the electrostatic attractions of ions with the oppositely charged full permanent dipoles. It is typical in aqueous solutions containing ions that are surrounded by water molecules with oppositely charged dipoles. The interaction is also used in molecules with a localized charge on a particular atom.[50]

1.1.3 Non-covalent Interactions with σ-hole

Non-covalent interactions with σ-hole are formed between two electronegative atoms: Z-X⋯Y, X is a electronegative atom capable of creating an electrophilic area (halogen, chalcogen, pnictogen, tetrel, aerogen) that is bound to Y bearing a negative partial charge (lone pair of electrons e.g. at O or N, π-electrons). According to Coulomb's Law, interacting negative charged atoms usually repulse each other. However, in this particular case, the interaction is attractive because of the σ-hole, which is a positively charged area at X that can be associated with a negative partial charge at Y. This phenomenon was detected experimentally, but its clarification is owed to quantum chemistry computations. Electrostatic potential (ESP) calculations enable the visualization of charged regions of a molecule in three dimensions by colour-coded isodensity surfaces (ESP maps). It emerged that although X draws electrons closer owing to its higher electronegativity, it does not arrange them evenly, but as a specific ring with a positive charged area inside. This positive

ESP is called σ-hole and can interact with a negative charged Y. Electron density around a nucleus is anisotropic and grows with the atomic number.[50, 52]

1.2   Van der Waals Forces

1.2.1   Dipole-dipole Interactions

A dipole-dipole interaction, also called Keesom force, is electrostatic interaction between permanent dipoles in molecules. It is the strongest and the most important type of van der Waals interaction. The uneven sharing of electrons in heteroatom systems leads to a constitution of permanent dipoles, quadrupoles or even higher multipoles. Thus, interacting polar molecules that have oppositely charged poles are electrostatically attracted. The system tends towards reducing of potential energy through attractive interactions of aligned molecules. The interaction has a maximal effect if two dipoles are identically oriented and lie on one axis. In the opposite case, dipoles are repulsed. Superposition of contributions of interactions between individual multipoles of molecules determines the intensity of the resultant interaction. The intensity of the contributions is dependent on the intensity of the multipole moments of molecules and an intermolecular distance. The intensity of contributions decreases faster with the increasing distance of subsystems in the case of higher multipoles than in the case of lower multipoles.[50]

1.2.2   Dipole-induced Dipole Interactions

A dipole-induced dipole interaction, alternatively called Debye force or polarization interaction, is an electrostatic interaction between a polar molecule with a permanent dipole and a spherically symmetric non-polar molecule with no permanent dipole. The approach of these systems causes that a molecule with a permanent dipole induces (polarizes) a dipole moment in a non-polar molecule. This adjustment establishes an attractive interaction between a permanent and induced dipole. The intensity of an inductive contribution is dependent upon the intensity of the permanent multipole moment of a molecule, polarizability of a neutral molecule, and an intermolecular distance. The intensity of contributions decreases with the increasing distance of subsystems faster in the case of higher permanent multipoles than in the case of lower multipoles. If a connection between a molecule with a permanent dipole and a molecule with an induced dipole is discontinued, an induced dipole is revoked.[50]

1.2.3 London Dispersion Forces

All atoms and molecules exhibit London dispersion forces, also called instantaneous dipole-induced dipole interactions. They are the weakest type of non-covalent interaction but in great numbers can have significant contributions. Fluctuations of the electron distribution around an atom or molecule create instantaneous dipole moment that produces an induced dipole moment in a neighbouring atom or molecule. Interactions of short-term dipoles lead to synchronization of their oscillations, which is the basis of attractive dispersive interactions. The strength of the dispersive interactions is proportional to the polarizability of the interacting groups. It depends on the total number of electrons and the size of the occupied area since polarizability increases with the increasing volume of electrons. Dipole moments nullify each other in long-term average; hence molecules appear non-polar.[50]

1.2.4 Pauli Repulsion

The exchange-repulsion term for fermions (particles with half-integer spin: electrons, protons, neutrons, etc.) is called Pauli repulsion and is related to the Pauli exclusion principle saying that no two identical fermions can occupy the same quantum state in the same quantum system simultaneously. In the context of non-covalent interactions, the effect is elicited by the overlapping of electron densities of two electrons with the same spin quantum number in the same electron shell. The exchange-repulsion term acts against the approaching of the systems and its intensity increases with the decreasing distance of the systems.[50]

1.3 π-interactions

π-interactions are a type of non-covalent interactions that integrate π-systems of conjugated molecules (e.g. benzene, ethylene, acetylene) incorporating both covalent chemical σ- and π-bonds. Whereas σ-bond is created by head-on overlapping between atomic orbitals, which means that electron density of each of these atomic orbitals is located at a shared nodal plane passing through the two bonded nuclei, π-bond is formed by the overlapping between two lobes of atomic orbitals above and below the nodal plane of the molecule. The engagement of p orbitals is common, but d orbitals can be involved as well. π-bonds are usually weaker than σ-bonds. π-interactions can be divided into three main subsets.

π-π interactions (alternatively called π-π stacking) are non-covalent interactions between the π-orbitals of aromatic rings. Dispersive interactions represent main contribution of the stacking effects due to high polarizability of aromatic rings. There are several possible configurations such as 'edge-to-face ', 'T-shaped ', 'sandwich ', or 'displaced '.

Cation-π interactions are non-covalent interactions between a cation such as Li^+ or Na^+ (monopole) and π system (negatively charged region of the quadrupole). An electron-rich π system above and below the benzene ring bears a partial negative charge that interacts with a positively charged ion. These interactions are strong as H-bond or even stronger. Anion-π interactions are organised reversely. An anion is placed over an electron-poor π-system that is generated by electronegative substituents on the conjugated molecule.

Polar-π interactions are non-covalent interactions between molecules with permanent dipoles (typically water) and the quadrupole moment of a π-system (e.g. benzene).[50, 53]

1.4 Hydrophobic Interactions

The hydrophobic interactions are not a classical force, but the activity leads to elimination of disadvantageous interactions of molecules in water. Hydrophobic molecules suspended in water increase the energy of a system since non-polar molecular parts cannot create energetically advantageous interactions with polar water molecules and, in addition, they tie water molecules down to form H-bonds each other. Hence, hydrophobic substances tend to cluster. The formation consists of non-polar parts oriented inside and polar parts outside (i.e. into water) the cluster, which effectively reduces the disadvantageous contacts of non-polar molecular parts with water molecules.[19, 50]

2. Theoretical Chemistry

Theoretical chemistry is considered to define the framework for branches as quantum chemistry, computational chemistry, molecular dynamics, molecular modelling, molecular mechanics, statistical thermodynamics, chemoinformatics, and other. The subjects are closely interconnected. The terms theoretical and quantum chemistry are often perceived as synonymous. Theoretical chemistry evolves theoretical basis of chemistry and generalizations of chemical phenomena. The main effort is focused on clarification of the relation between the structure and properties of molecular systems. Also, a scrutiny of dynamics, and thermodynamic and kinetic properties of chemical systems is essential. For investigation of chemical problems, mathematical and physical methods, and computer simulations are utilized. Quantum chemistry, alternatively called molecular quantum mechanics, is an application of quantum mechanics to the study of chemical systems. It enables to understand chemical problems at the fundamental level. The behaviours of valence electrons constitute the principal phenomenon in quantum chemistry. That is aimed at description of crucial issues such as molecular orbitals and orbital interactions, chemical bonding, the surface of potential energy, the ground state of individual atoms and molecules, the excited states, and chemical reactions including transition states. Quantum chemical analyses employ semi-empirical methods based on quantum mechanics as well. The precision of quantum computational studies is conditional on accuracy both used techniques and models, and a size of molecular systems. Since chemical problems of only simple atoms and molecules (i.e. H, H_2^+, H_2, and He) can be solved analytically, it is necessary to employ computers to assistance of solving the quantum many-body problem (systems consisting of a large number of interacting particles). Computational chemistry uses computer simulations based on methods of theoretical chemistry. Highly developed computer programs enable to calculate structures and properties of molecules effectively. Computations deputize or complete experiments, predict properties and interactions of compounds, and allow designing of new substances and materials. Rising the accuracy of the results is limited by scaling considerations since the computation time increases as a power of the number of atoms. Therefore, a development of mathematical approximations and iterative methods is the necessary part of theoretical chemistry, quantum chemistry in particular.[54-56]

2.1 Schrödinger Equation

A solution of a quantum chemical problem is based on the determination of the electronic structure of the molecule by solving the Schrödinger equation (in non-relativistic quantum chemistry) with the electronic molecular Hamiltonian. The electronic structure represents the state of motion of electrons in an electrostatic field induced by stationary nuclei. It includes both the wave functions of the electrons and the energies related to them. The wave function is a mathematical description of the quantum state of a quantum system. It is the function of the coordinate $\mathbf{r}$ and the time t expressed (in one dimension) by the notation $\Psi(\mathbf{r}, t)$. The wave function enables to obtain the energy of electrons, angular momentum, and orbital orientation in the shape of the quantum numbers n, l, and m_l. The square of the wave function, $|\Psi(\mathbf{r}, t)|^2$, is the probability density assessing the probability of finding the particle at a position $\mathbf{r}$ and time t in the space around the nucleus. The wave function is not a real, but a complex function, therefore, the Schrödinger equation does not have real, but complex solutions. The Schrödinger equation is linear partial differential equation that estimates the state of the quantum system. It is the equation of motion, which is analogous to Newton's second law of motion in classical mechanics, describing the time and spatial course of the wave function of a particle.

$$i\hbar \frac{\partial \Psi}{\partial t} = \widehat{H}\Psi \tag{2.1}$$

where Ψ is the N-electron wave function, $\widehat{H}$ is the operator representing the energy of the electrons and nuclei in a molecule that is called Hamiltonian, i is the imaginary unit, $\hbar$ is the reduced Planck constant, and t is time. Hamiltonian is the sum of the kinetic energies and the potential energies for all the particles in the system. For one particle, it is expressed as

$$\widehat{H} = \widehat{T} + \widehat{V} = -\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r}, t) \tag{2.2}$$

and for many particles

$$\widehat{H} = \sum_{n=1}^{N} \widehat{T}_n + V = -\frac{\hbar^2}{2}\sum_{n=1}^{N}\frac{1}{m_n}\nabla_n^2 + V(\mathbf{r}_1, \mathbf{r}_2 \cdots \mathbf{r}_N, t) \tag{2.3}$$

where $\hat{T} = \frac{\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}}{2m} = -\frac{\hbar^2}{2m} \nabla^2$ is the kinetic energy operator of particle, in which m is the mass of the particle, and $V = V(\mathbf{r}_1, \mathbf{r}_2 \cdots \mathbf{r}_N, t)$ is the potential energy function (in one particle case, $\hat{V} = V(\mathbf{r}, t)$ is the potential energy operator).

$\hat{p} = -i\hbar\nabla$ is the momentum operator, $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$ the is the nabla/del operator, the dot means the scalar/dot product of vectors (the algebraic operation), $\nabla \cdot \nabla = \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ is the Laplace operator (Laplacian), i is the imaginary unit, and $\hbar$ is the reduced Planck constant.

The time-dependent Schrödinger equation for one single particle is then described as

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r}, t) = \left[\frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{r}, t)\right]\Psi(\mathbf{r}, t) \tag{2.4}$$

In quantum chemistry, the finding out of the electronic structure of atoms and molecules is standardly accomplished by the time-independent Schrödinger equation

$$\hat{H}\Psi = E\Psi \tag{2.5}$$

where Ψ is the wave function describing the stationary state of the system, $\hat{H}$ is the time-independent Hamiltonian, and E is the energy 'eigenvalue', which is a constant corresponding with the energy level of the system. The time-independent Schrödinger equation for one single particle is expressed as (with the legends as above).

$$\left[\frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\Psi(\mathbf{r}) = E\Psi(\mathbf{r}) \tag{2.6}$$

Although solving of the Schrödinger equation provides wave functions for electrons in atoms and molecules, an exact solution can be obtained only for the hydrogen atom. For other atomic or molecular systems, approximate solutions must be employed.[54-57]

2.2   Born-Oppenheimer Approximation

The Born-Oppenheimer approximation is considered to be one of the basic concepts in quantum chemistry. It enables to decouple quantum mechanical motion of the electrons from the motion of the nuclei, so into electronic and nuclear problems. Since the mass of the atomic nucleus is much heavier than the mass of the electron (a proton is 1836 times heavier than an electron), the nuclei move much more slowly than the electrons, which

means that the nuclei can be considered as nearly fixed in relation to electron motion. It is supposed that electrons respond to the nuclei motion instantly. The molecular wave function has this factorized form in which Ψ_e is parametrically dependent on the positions of the nuclei

$$\Psi(\mathbf{r}, \mathbf{R}) = \Psi_e(\mathbf{r}, \mathbf{R}) \cdot \Psi_n(\mathbf{R}) \tag{2.7}$$

where $\mathbf{r}$ is the set of electronic (e) and $\mathbf{R}$ of nuclear (n) coordinates.

A molecular coordinate system representing the general problem may be written in the atomic units as

$$\hat{H} = -\underbrace{\sum_{i=1}^{N} \frac{1}{2} \nabla_i^2}_{\hat{T}_e} - \underbrace{\sum_{A=1}^{M} \frac{1}{2M_A} \nabla_A^2}_{\hat{T}_n} - \underbrace{\sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_A}{r_{iA}}}_{\hat{V}_{en}} + \underbrace{\sum_{i=1}^{N-1} \sum_{j>i}^{N} \frac{1}{r_{ij}}}_{\hat{V}_{ee}} + \underbrace{\sum_{A=1}^{M-1} \sum_{B>A}^{M} \frac{Z_A Z_B}{R_{AB}}}_{\hat{V}_{nn}} \tag{2.8}$$

where i, j are electrons (N), A, B are nuclei (M), $r_{iA} = |\vec{r}_{iA}| = |\vec{r}_i - \vec{R}_A|$, $r_{ij} = |\vec{r}_{ij}| = |\vec{r}_i - \vec{r}_j|$, $R_{AB} = |\vec{R}_{AB}| = |\vec{R}_A - \vec{R}_B|$, M_A is the ratio of the mass of nucleus A to the mass of an electron, Z_A is the atomic number of nucleus A, T_e is the operator for the kinetic energy of the electrons, T_N is the operator for the kinetic energy of the nuclei, V_{eN} is the operator for the Coulomb attraction between electrons and nuclei, V_{ee} is the operator for the repulsion between electrons, and V_{NN} is the operator for the repulsion between nuclei.

This compact notation is the more common expression

$$\hat{H} = \hat{T}_e(\mathbf{r}) + \hat{T}_n(\mathbf{R}) + \hat{V}_{en}(\mathbf{r}, \mathbf{R}) + \hat{V}_{ee}(\mathbf{r}) + \hat{V}_{nn}(\mathbf{R}) \tag{2.9}$$

The term $\hat{V}_{eN}(\mathbf{r}, \mathbf{R})$ does not allow unequivocally to separate $\hat{H}$ into the nuclear and electronic parts since its contribution is large and cannot be neglected. Hence the $\mathbf{R}$ dependence is taken as the parameter, $\hat{T}_N(\mathbf{R})$ is neglected because it is significantly smaller than $\hat{T}_e(\mathbf{r})$, and $\hat{V}_{NN}(\mathbf{R})$ is made as a constant. Thus, it is thought that the nuclear and electronic separation is approximately correct. The electronic wave function $\Psi_e = \Psi_e(\mathbf{r}; \mathbf{R})$ directly depends on the electronic coordinates and parametrically on the nuclear coordinates. It describes the motion of the electrons and electronic states for fixed nuclear coordinates. An electronic wave function is obtained for each nuclear configuration and for each value of the nuclear positions, and the electronic system is in the electronic ground state corresponding to the lowest energy $E_e(\mathbf{R})$. The electronic Hamiltonian describing the motion of N electrons in the field of M fixed nuclei is then written as

$$\hat{H}_e(\mathbf{R}) = \hat{T}_e(\mathbf{r}) + \hat{V}_{eN}(\mathbf{r};\mathbf{R}) + \hat{V}_{ee}(\mathbf{r}) \tag{2.10}$$

and the Schrödinger equation is expressed as

$$\hat{H}_e(\mathbf{R})\Psi_e(\mathbf{r};\mathbf{R}) = E_e(\mathbf{R})\Psi_e(\mathbf{r};\mathbf{R}) \tag{2.11}$$

The geometry dependent electronic energy $E_e = E_e(\mathbf{R})$ puts the potential energy.

$$E_{tot}^{pot}(\mathbf{R}) = E_e + \hat{V}_{nn}(\mathbf{R}) \tag{2.12}$$

into effect in the Schrödinger equation for the nuclear motion. It is termed the potential energy surface (PES) and describes the relationship between the energy of an atomic or molecular system and its geometry. PES is considered to be one of the basic concepts in computational chemistry. The nuclear hamiltonian describing a motion of nuclei in the average field of electrons is expressed as.

$$\hat{H}_N = -\sum_{A=1}^{M}\frac{1}{2M_A}\nabla_A^2 + \underbrace{\langle -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \underbrace{\sum_{i=1}^{N-1}\sum_{j>i}^{N}\frac{1}{r_{ij}}\rangle + \sum_{A=1}^{M-1}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}}}_{E_e(\mathbf{R})}}_{E_{tot}^{pot}(\mathbf{R})} \tag{2.13}$$

The wave function Ψ_n in the nuclear Schrödinger equation

$$\hat{H}_n\Psi_n = E\Psi_n \tag{2.14}$$

describes vibration, rotation, and translation of a molecule. The total energy E of the molecule is comprised of electronic, vibrational, rotational, and translational energy.[54-57]

2.3 Slater Determinant

Slater determinant is the approximate expression enabling to obtain an accurate solution of a wave function of a multi-electron system. A wave function must meet anti-symmetry requirement and so complies with Pauli exclusion principle, which means changing a sign upon exchange of spacial and spin coordinates. The wave function is expressed as a linear combination of Slater determinants Φ_I:

$$\Psi(1,2,\dots,n) = \sum_I c_I \, \Phi_I \tag{2.15}$$

The Slater determinant constitutes a product wave function for an assemblage of electrons in which each electron is described as a wave function termed the spin-orbital $\phi(x)$ that x describes the position and spin of a single electron:

$$\Phi_I(1,2,\dots,n) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \phi_1(x_1) & \phi_1(x_2) & \dots & \phi_1(x_n) \\ \phi_2(x_1) & \phi_2(x_2) & \dots & \phi_2(x_n) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_n(x_1) & \phi_n(x_2) & \dots & \phi_n(x_n) \end{vmatrix} \tag{2.16}$$

where the coefficient $\frac{1}{\sqrt{n!}}$ is the normalization factor and ϕ_i is a single-particle function. Slater determinant has mathematical characteristics ensuring the validity of the Pauli principle, so antisymmetrized function, and it provides linear combinations of all possible permutations of the system.

The expansion of a wave function of atomic or molecular orbitals ϕ_i into a finite series of single-particle functions χ_v is utilized as the basic approximation:

$$\phi_i = \sum_i c_{iv} \chi_v \tag{2.17}$$

where the base function χ_v can be Slater determinants, Gauss functions, or other functions. A collection of base functions is also called a base. If the wave function ϕ_i represents a function of molecular orbitals, this expansion is called Molecular Orbitals–Linear Combination of Atomic Orbitals (MO-LCAO), although base functions not have to be connected only with atom orbitals. The expansion coefficients c_{iv} are determined as a solution of the SCF problem. A single Slater determinant is used as an approximation to the electronic wavefunction in Hartree–Fock theory.[54-56]

3. Methods

Lennard-Jones potential is a mathematically simple model enabling approximations of the interaction between a pair of neutral atoms or molecules. Non-covalently interacting systems with no permanent dipoles (neutral atoms or molecules) can be described by the rough mathematical model Lennard-Jones potential:

$$V(r) = 4 \in \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right] \tag{3.1}$$

where the parameter ε is the depth of the potential well, σ is the finite distance at which the inter-particle potential is zero: $V(r) = 0$ (it can be also $r = \infty$), r is the distance between the particles. $\frac{1}{r^{12}}$ is the repulsive term that describes Pauli repulsion at short ranges due to overlapping electron orbitals, and $\frac{1}{r^6}$ is the attractive long-range term describing attraction at long ranges.[54-56]

Although methods that are more accurate exist, Lennard-Jones potential is still used in some computer simulations. Of course, in the presented computational studies, more recent and significantly accurate methods were utilized.

3.1   Hartree–Fock Method

The Hartree–Fock (HF) method, also called the self-consistent field method (SCF), is the fundamental approximation used to solve the time-independent Schrödinger equation for the determination of the wave function and the energy of a many-electron system in a stationary state. It is a variational method that provides the wave function of a many-body system assumed to be in the form of a Slater determinant for fermions and of a product wave function for bosons. It correctly treats the statistics of the many-body system, antisymmetry for fermions and symmetry for bosons under the exchange of particles. The variational parameters of the method are the single-particle wave functions composing the many-body wave function. The HF method is based on the Born–Oppenheimer and orbital approximations. Under the Born–Oppenheimer approximation the nuclear and electronic degrees of freedom of a molecule are decoupled, and the nuclei are held fixed while the electronic contribution to the energy is calculated. In the orbital approximation, the electrons occupy individual spin-orbitals, and as a consequence the N-electron Schrödinger equation is transformed into N one-electron equations. Both approximations facilitate

computation, and the HF method proceeds by selecting a trial wave function (a molecular orbital formed as a linear combination of atomic orbitals, LCAO-MO) containing adjustable parameters, and subsequently solving a set of N coupled integro-differential equations through an iterative (self-consistent field) procedure.

The HF method solves the electronic Schrödinger equation that is expressed as

$$\left[-\frac{1}{2}\sum_i \nabla_i^2 - \sum_{A,i}\frac{Z_A}{r_{Ai}} + \sum_{A>B}\frac{Z_A Z_B}{R_{AB}} + \sum_{i>j}\frac{1}{r_{ij}}\right]\Psi(r;R) = E_{el}\Psi(r;R) \tag{3.2}$$

It is equivalent to this notation

$$\left[\hat{T}_e(r) + \hat{V}_{eN}(r;R) + \hat{V}_{NN}(R) + \hat{V}_{ee}(r)\right]\Psi(r;R) = E_{el}\Psi(r;R) \tag{3.3}$$

Since the HF method is the simplest wavefunction-based method, it works as a starting point for most methods describing a multi-electron atom or molecule more accurately.[54-56]

3.2 Density Functional Theory

Density functional theory (DFT) is a method, in which the fundamental role is played by the electron density ψ2. It is a computational quantum mechanical method employed functionals of the electron density to investigate the electronic structure (or nuclear structure) of many-body systems (e. g. atoms, molecules). It is based on the Hohenberg–Kohn theorem. It state that the electron density uniquely determines the external potential and the number of electrons of an atomic or molecular system. There are many corrections for improvement of speed and robustness, in particular adding of empirical dispersion correction to functionals. DFT is versatile, one of the most widely used computational methods. It enables calculations of both structure prediction and thermodynamics and kinetics. DFT calculation of total energies and forces are under very general conditions, it scales with N3 or better.[54-56]

3.3 Semiempirical Quantum Mechanics

Semiempirical quantum mechanical methods are quantum chemical electronic structure calculation techniques based on a formalism for ab initio quantum mechanics (e.g. molecular orbital) or density functional theory. They apply further approximations and use parameters from empirical data. These methods are less accurate than higher level calculations, but they are usually much faster. A commonly used approximation is called

Modified Neglect of Diatomic Overlap (MNDO). It is employed in the AM1 and PM3 methods that are suitable for QM/MM computations on proteins. Semiempirical methods usually include only the valence electrons, in which some integrals are ignored or replaced by empirically based parameters. Another popular method was developed by J. J. Stewart is PM6. The PM6-D3H4X method with corrections for both halogen and hydrogens and dispersion has become the most accurate SQM method for computations of interactions of biomolecules. [54-56, 58] It allows rapid optimization of geometries of large systems as three modifications were made to the conventional semiempirical procedure: the matrix algebra method for solving the self-consistent field (SCF) equations was replaced with a localized molecular orbital method (MOZYME), Baker's Eigen following technique for geometry optimization was replaced with the L-BFGS function minimizer, and some of the integrals used in the NDDO set of approximations were replaced with point-charge and polarization functions.[59]

3.4   Molecular Dynamics

Molecular dynamics (MD) simulations generate information about atomic positions and velocities. The energy surface is explored by solving Newton's laws of motion for the system. It enables studying both thermodynamic properties and time dependent (kinetic) phenomenon. The thermodynamic state of a system is usually defined by a small set of parameters, e.g. the temperature, T, the pressure, P, and the number of particles, N. Other thermodynamic properties may be obtained from the equations of state and other fundamental thermodynamic equations. The mechanical or microscopic state of a system is defined by the atomic positions, q, and momenta, p; these can also be considered as coordinates in a multidimensional space called phase space. For a system of N particles, this space has 6N dimensions. A single point in phase space, denoted by G, describes the state of the system. An ensemble is a collection of points in phase space satisfying the conditions of a particular thermodynamic state. A molecular dynamics simulation generates a sequence of points in phase space as a function of time; these points belong to the same ensemble, and they correspond to the different conformations of the system and their respective momenta. Molecular dynamics simulations are used to describe e.g. interactions of a protein with drug-like molecules, solvation of molecules, the conformational changes, or behaviour of biomolecules.[54]

3.5 Solvent Models

A solvent model is a computational method that estimate behaviour of solvated condensed phases. Quantitative prediction of the physical properties of liquids is important for many applications. Explicit solvent models describe water clusters, liquid water, and aqueous solutions. These models are usually used in the application of molecular mechanics (MM) and dynamics (MD) or Monte Carlo (MC) simulations. Implicit solvent models, also called continuum solvents, are a method describing the solvated phase without the presence of real water molecules. Continuum solvents employ the dielectric constant (ε) as the main parameter that is responsible for defining the degree of the polarizability of the solvent. This parameter can be supplemented with further parameters. Depending on the type of computations, different types of implicit solvent models are used. The Generalized Born (GB) implicit solvent model is widely used for molecular dynamics simulations of proteins and nucleic acids. This approach makes a model of hydration effects and provides solvent-dependent forces with efficiencies comparable to molecular-mechanics calculations on the solute alone. The Conductor-like screening model (COSMO) is a calculation method for determining the electrostatic interaction of a molecule with a solvent. These models approximate solvent by a dielectric continuum surrounding the solute molecules outside of a molecular cavity. The quantum mechanical-based Solvation Model Density (SMD) model has significantly better agreement with experimental data than other models.[54-56]

4. Interactions between CDK2 and Inhibitors

4.1   Introduction

4.1.1   Cell Cycle

In a simplified way, the cell cycle is the life cycle of a cell. It is a highly regulated sequential process occurring in a cell that results in duplication of its DNA and division of cytoplasm and organelles to produce two daughter cells. The cell cycle is a four-stage process comprising of interphase: the first gap (G1) phase, the synthesis (S) phase, the second gap (G2) phase, and the mitotic phase (M). This ordered sequence of events progresses from quiescence, i.e. the zero gap (G0) phase, to proliferation (G1, S, G2, and M), and back to G0, Figure 4.1.
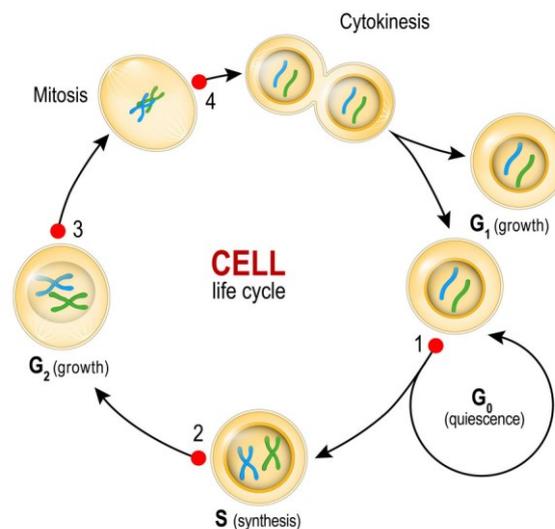
Figure 4.1: The schema of the cell cycle.

Interphase consists of these three steps. Throughout G1, the cell increases in size, duplicates organelles and forms the molecular building blocks. In the course of S, the cell synthesizes a complete copy of the DNA (each of the 46 chromosomes) in its nucleus. It also duplicates the centrosome (an organelle serving as a microtubule-organizing centre and a regulator of cell-cycle progression) helping separate DNA during M. During G2, the cell grows larger, creates proteins and organelles, and begins to reorganize its contents in preparation for mitosis. During M, the cell divides its copied DNA and cytoplasm to create two new cells.

M phase involves two different division-related processes: mitosis in which the replicated chromosomes are separated into two new nuclei; and cytokinesis in which the cytoplasm, organelles, and cell membrane are divided into two new cells. In G0, or resting phase, cells exist in a quiescent state outside of the replicative cell cycle. DNA replication and mitosis are controlled with the regulation of the activity of cyclin-dependent protein kinase enzymes, which are heterodimers consisting of a catalytic kinase subunit with a cyclin subunit.[60]

The cell cycle is controlled at three checkpoints, which are mechanisms ensuring proper division of the cell. The G1 checkpoint assesses whether all conditions are positive for cell division to proceed and inspects the cell about DNA damage. If a cell meets the requirements such as an appropriate size or sufficient energy reserves, it will enter S phase and begin the DNA replication. The G2 checkpoint ensures accurate replication of all of the chromosomes. If the DNA has been correctly replicated, cyclin dependent kinases signal the outset of mitotic cell division. In case of mistakes or damage of the DNA, the cell cycle is stopped. The M checkpoint determines whether all the sister chromatids are correctly fastened to the spindle microtubules. The cycle will not proceed until the kinetochores of each pair of sister chromatids are connected to at least two spindle fibers arising from opposite poles of the cell.[61] All of the main checkpoint transitions in the cell cycle are signalled by cyclins and cyclin dependent kinases.[62]

Although cell proliferation is essential for growth, development, and regeneration of eukaryotic organisms, it also causes cancer that belongs to some of the most serious diseases.[63-64]

4.1.2 Cyclin-dependent Kinases

Kinases belong to the family of phosphotransferases, which are enzymes catalysing the transfer of the phosphate group from the high-energy phosphate-donating molecule (e.g. ATP) to a specific target molecule (substrate). This transesterification is termed phosphorylation and ordinarily results in a functional change of substrate. It makes a phosphorylated substrate and ADP. The converse process is called dephosphorylation. Phosphorylation of molecules usually leads to enhancement or inhibition of their activity and influences their ability to associate with other molecules. Kinases play a fundamental role in signal transduction and regulation of complex processes in cells.

Protein kinases are the enzymes from the kinase family adjusting other molecules by phosphorylation. They mainly catalyse binding of the phosphate group to proteins on their serine, threonine, tyrosine, or histidine residues. Protein kinases are divided into these groups: AGC, CaM, CK1, CMGC, STE, TK, TKL.

Cyclin dependent kinases (CDKs) are the subgroup of the CMGC group. The individual CDKs are numbered CDK1, CDK2, etc. They phosphorylate other proteins on their serine or threonine residues. CDKs are mainly involved in regulation of the cell cycle and transcription, mRNA processing, the differentiation of nerve cells, and cellular metabolism. To be active, CDKs must first associate with a cyclin protein. The phosphorylation state of CDKs is also essential for their activity since they are subject to regulation by other kinases and phosphatases. Active CDKs phosphorylate other proteins to change their activity, which arouse actions influencing the next stage of the cell cycle.

Cyclins create a family of proteins that controls the cell progression within the cell cycle. They are marked by capitals, e.g. cyclin A, cyclin B, etc. Cyclins are regulatory subunits of heterodimer kinases that response for the start of the process running in particular phases of the cell. Cyclins have no enzymatic activity; their role consist in activation of CDK enzymes. They aim CDKs to specific subcellular locations. A single cyclin molecule interacts with a CDK molecules to create a complex. The formation results in activation of the CDK active site and it is completed by phosphorylation. In the course of the cell cycle, their concentration is cyclically variated, which corresponds to the increase or the decrease of activity of relevant kinase subunits. In the cells, the production of cyclines and their actions are realized in a strict time sequence. The formation of CDK complexes launches the cascade of actions carrying out duplication of DNA and mitosis. Cyclins can be divided into four classes based on their behaviour in the cell cycle: G1 cyclins, G1/S cyclins, S cyclins, and M cyclins. Specific combinations of CDKs and cyclins are employed in define stages of the cell cycle. CDKs are expressed in cells whereas cyclins are produced at specific stages of the cell cycle, in response to various molecular signals, Figure 4.2.[65]

CDK-activating kinase (CAK) belongs to the CDK family and activates the cyclin-CDK complex by phosphorylating threonine residue 160 in the CDK activation loop. It works as positive regulators of CDK1, CDK2, CDK4, and CDK6.[24, 26]
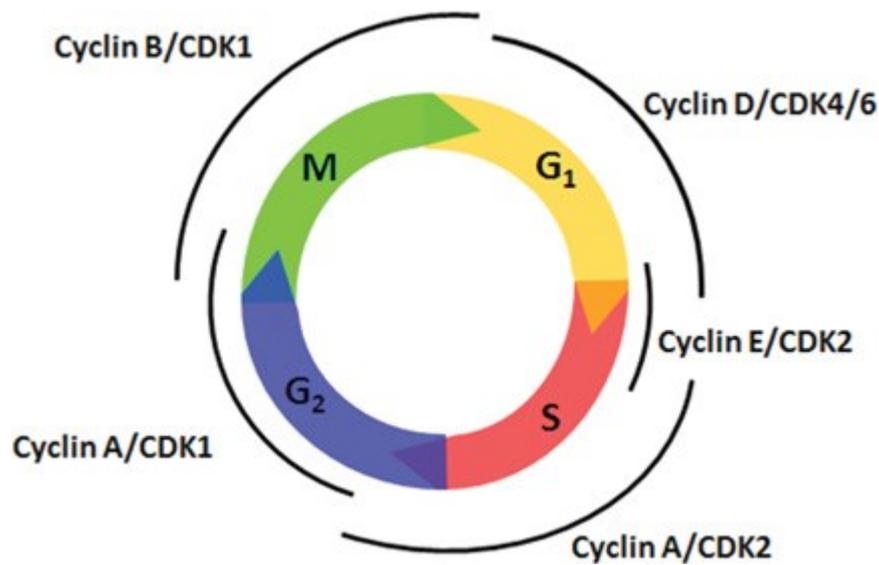
Figure 4.2: Control of the cell cycle progression by phosphorylation of cyclin-dependent kinase substrates.

The association of cyclins with kinases represents the pivotal mechanism of the cell cycle regulation. For the discoveries of CDKs and cyclins as key regulators of the cell cycle, Leland H. Hartwell, Tim Hunt, and Paul M. Nurse won the Nobel Prize in Physiology or Medicine in 2001.[66]

4.1.2.1 Cyclin-dependent Kinase 2

CDK2 is employed during G_1 and S phase of the cell cycle as a G_1-S phase checkpoint control. Although CDK2 is largely dispensable in the cell cycle of normally functioning cells, it is critically associated with the abnormal growth processes of cancer cells in multiple cancer types. It is a key modulator of various oncogenic signalling pathways; its abnormal activity leads to loss of proliferative control during oncogenesis. Furthermore, the overexpression of cyclin A and/or E, which are the regulatory subunits of CDK2 is a crucial oncogenic process in several cancers.[24, 67] Therefore, CDK2 and its cyclin partners represent eligible targets for cancer therapeutics.[68] Because of extreme similarity between the active sites of CDK2 and other CDKs, especially CDK1, finding highly selective small molecules with inhibitory effect on CDK2 is difficult. In addition, CDKs evince redundant functions. CDK1 is the only essential cyclin dependent kinase in the cell cycle, thus its inhibition could lead to undesirable side effects.[69]

The CDK2 bilobal architecture consists of the amino-terminal domain (N-terminal lobe; residues 1-82) and the carboxy-terminal domain (C-terminal lobe; residues 83-297). The N-terminal lobe is largely constituted by β-sheets (five anti-parallel β-strands) with one αC-helix that is crucial for cyclin binding. The C-terminal lobe is mainly made up of α-helices and contains the activation segment named the T-loop (residues 145-172) and the activating phosphorylation site Thr160. The T-loop enables binding of the Ser/Thr (phosphor-acceptor) region of substrates for phosphorylation. The N-terminal and C-terminal lobes are connected by the flexible hinge region (residues 81-84) that is the important part of the active site (ATP-binding site).The active site is a cleft between the N-terminal and C-terminal lobes (Figure 4.3[70]). The active site can be influenced by cyclin binding. In CDK without associated cyclin, T-loop blocks the cleft, and the position of several key amino acid residues is not convenient for ATP-binding. If the CDK2 forms the complex with the cyclin, two α-helices change position to allow ATP binding: the L12 helix becomes a beta strand and helps rearrange the T-loop; the PSTAIRE helix enables change the position of the key amino acid residues in the active site. Phosphorylation by CAK at Thr 161 on the T-loop increases the CDK2-cyclin complex activity. Other CDKs have very similar structural characteristics.[71]
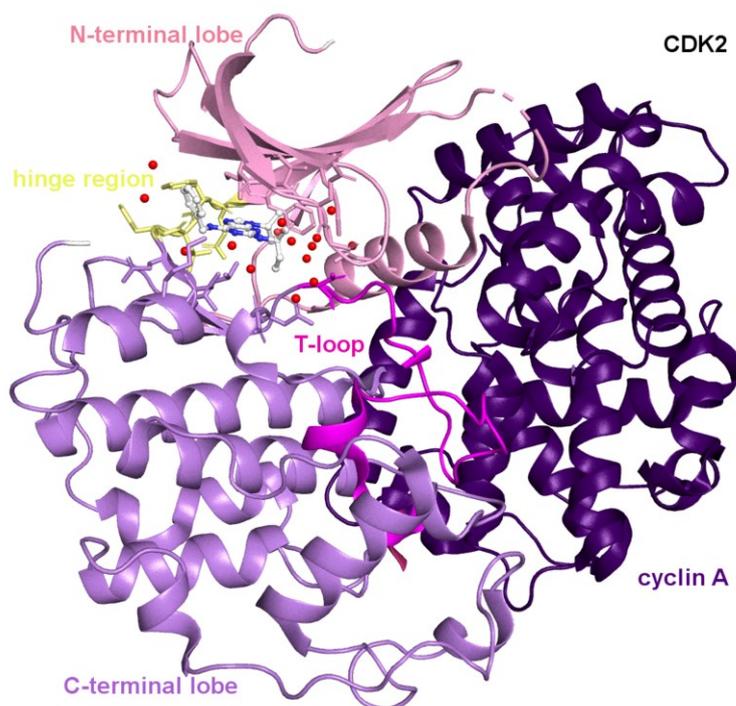
Figure 4.3: The non-optimized X-ray crystal structure of CDK2 with the inhibitor roscovitine bound in the active site (PDB ID: 3DDQ). The kinase enzyme is depicted by pink (N-terminal lobe) and violet (C-terminal lobe) colour, the cyclin A by dark violet one, the T-loop by magenta one, and the hinge region by yellow colour.

4.1.3   Small-molecule Inhibitors of Cyclin-dependent Kinases

As CDKs play a fundamental role in the controlling of cell division, their mutations are often found in cancerous cells. The mutations induce uncontrolled growth of the cells, block the checkpoints of the cell cycle, and result in abnormal proliferation. They mainly occur in lymphomas, melanoma, breast cancer, pancreatic tumours, and lung cancer.[63-64] The development of effective drugs worked on the assumption that inhibition of CDKs can stop the proliferation of cancerous cells. The CDK inhibition has proved as the valid target of cancer therapy.[67, 72] To date, a large number of inhibitors that are able both to stop proliferation and to cause apoptosis (a form of programmed cell death) have been synthesized. The inhibitors are bound to CDKs inside the ATP binding pocket (Figure 4.3), thereby they competitively inhibit the CDK activity. The occupation of the inside of the ATP binding pocket impedes the phosphorylation of other proteins.[73-74]

The first-generation of CDKs inhibitors had both nonselective effect and an unacceptable toxicity.[75-76] Therefore, considerable effort has been invested into the identification of

potent substances with both high selectivity and optimal pharmacokinetic/pharmacodynamic, and toxicological properties. For instance, the recent approval of palbociclib for the treatment of ER-positive and HER2-negative breast cancer represents the culmination of such efforts.[77] Nevertheless, CDKs inhibitors rather evince panselective inhibitory effect, but they differ in effect on individual CDKs.[78-79]

Flavopiridol, also named alvocidib, was the first CDK inhibitor to subject to clinical trials.[80-81] It was investigated as an inhibitor for the treatment of acute myeloid leukaemia. Flavopiridol is synthetic analog derived from the chromone alkaloid extracted from natural product rohitukine produced by endophytic fungi isolated from Amoora rohituka and Dysoxylum binectariferum (both belong to the mahogany family Meliaceae, which is a flowering plant family of mostly trees and shrubs).[82]

Olomoucine was also one of the first-generation CDK inhibitors and the first one with the purine as the central core.[83] It is a cytokinin analogue showing ATP-competitive inhibitory effect. Cytokinins are a class of plant growth substances, phytohormones, which are involved primarily in cell growth and differentiation. Natural cytokinins have an intact adenine and some of them contain an aromatic ring substituent at C6.[84] It was discovered that C2, C6 and N9 substituents on adenine are crucial for inhibition of CDK. Cytokinin analogues with C2-, C6-, N9-substituted purines (2,6,9-trisubstituted purines) truly evince a significant inhibitory effect.[83, 85] Crystallization analysis of the complex CDK2 with olomoucine showed specificity in the orientation of the molecule bound into the active site of the enzyme since the adenine group of olomoucine had different orientation from ATP. Olomoucine had the five-membered ring turned towards the hinge region of the enzyme whereas ATP was in the opposite direction. Moreover, the substituent at the position 2 made hydrogen bonds with amino acids situated outside of the binding pocket; ATP was unable to achieve these contacts.[86-87] Olomoucine evinced inhibitory effect with relative selectivity on some of CDKs, which set off studies of its biochemical activity and development of more effective substances. [83, 88] Results indicated that adding proper substituents at the 2, 6, 9 positions can enhance binding affinity and selectivity.[89-90]

The best known structural modification of olomoucine is roscovitine (seliciclib, CYC202).[27, 85, 91] It is a panselective CDK inibitor more potent than olomoucine in inhibiting CDK2 activity, the difference between their experimentally determined IC_{50} values (the concentration of a drug that is required for 50% inhibition in vitro) is one order of magnitude (olomoucine $IC_{50} = 7$ μM; roscovitine $IC_{50} = 0.2$ μM). The structural

information from the CDK2-roscovitine complex confirmed that the purine moiety of the inhibitor holds an opposite position to ATP. It is thought that this important difference is caused by the presence of the benzyl group at the position 6. Substituents at the positions 2 and 9 are also important for the inhibitory activity; specifically, branched alkyls with polar groups (e. g. hydroxybutyl) are suitable for the position 2 and alkyl groups (e. g. isopropyl) for the position 9.[83, 92-93] Roscovitine is bound to the hinge region of CDK2 by two hydrogen bonds between Leu 83 and N7 of the purine core and the NH motif attached to position 6. Another hydrogen bond is present between the hydroxybutyl group at C2 and Glu 12 at the edge of the active site. The dispersion interaction between the isopropyl group on N9 and Phe 80 also importantly contributes to the binding affinity.[27] In this work, roscovitine is employed as „reference inhibitor". Development followed on from the structural characteristics of olomoucine and roscovitine with the aim of increasing selectivity and overall potency showed that a modification of substituents at the 2, 6, 9 positions of the purine moiety can significantly improve inhibitory effect. For example, purvalanol A (IC_{50} = 0.07 μM) is about one order of magnitude better inhibitor then roscovitine, and purvalanol B (IC_{50} = 0.007 μM) about two order of magnitude for the CDK2-cyclin A complex.[78, 92, 94]

In order to design inhibitors whose shape and charge distribution are complementary to the binding site of the target protein, it is also used so called "scaffold hopping".[95-97] It is the replacement of the central core (scaffold) of known active compounds by suitable isosteres. This technique has enabled the identification of several series of purine bioisosteres, some of which contain inhibitors with biochemical and biological properties superior to those of the analogous purines.[96, 98-99] Despite numerous potent CDK inhibitors, only a few progressed past Phase I clinical trials. One of the most potent analogs, pyrazolo[1,5-a]pyrimidine-based dinaciclib (SCH 727965),[100] is currently being profiled in several clinical trials both as monotherapy and in combination with diverse agents.[76]

4.2   Structural Basis of the Interaction of CDK2 with Roscovitine and Its Analogues Having Bioisosteric Central Heterocycles

As has been said above, the drug development of selective inhibitors is complicated due to the similarity of CDKs active sites. Therefore, understanding of binding interactions and binding motifs between CDKs and inhibitors is essential. The main aim of this work is to elucidate the structural basis for the interactions between roscovitine and its analogs

(containing 13 different bioisosteric central heterocycles) and CDK2 that plays a critical role in many malignancies. Although many direct roscovitine bioisosteres have a central scaffold of nearly identical shape to that of roscovitine, their inhibitory activities can be dramatically different – the experimentally determined IC_{50} values of the inhibitors span three orders of magnitude, Figure 4.4.[70, 98, 101] To clarify this non-trivial structure-activity relationship, we quantified the protein-inhibitor affinities of known purine bioisosteres[98] as calculated binding scores and evaluated them in relation to the conformation of the optimized structures. We also considered the contributions of the separated scaffolds and substituents in a complex with the whole protein to the binding affinity. We used the hybrid quantum mechanics/semi-empirical quantum mechanics (QM/SQM) method[29] which employs the DFT method with an empirical dispersion for QM and the PM6[30] method for SQM part. The solvent effect was described by the continuum solvation model COSMO[31] at the SQM level for the whole system. To confirm the calculations experimentally eight inhibitors, including two novel compounds, were profiled in biochemical assays.
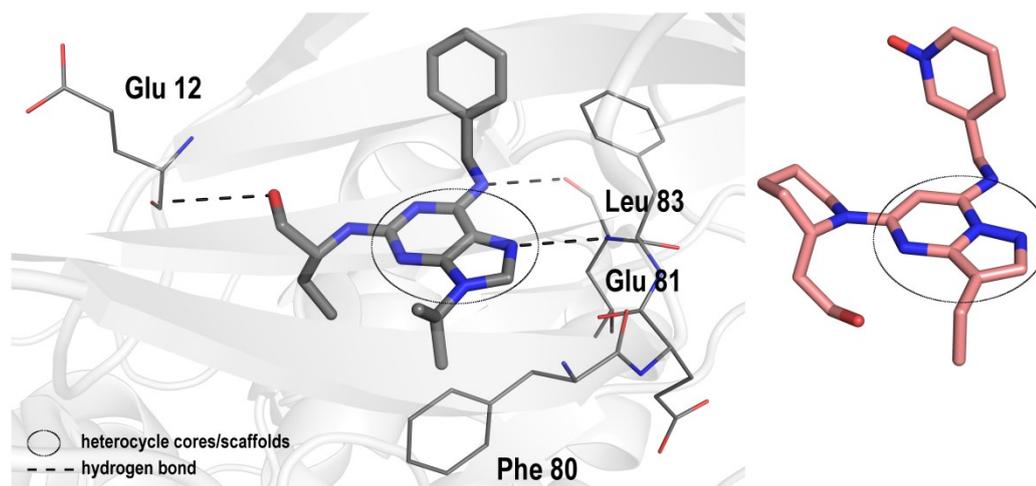
Figure 4.4: The non-optimized X-ray crystal structure of CDK2 with the inhibitor roscovitine bound in the active site (gray colour; IC$_{50}$=0.18 µM PDB ID: 3DDQ; only the key residues are depicted), and dinaciclib (pink colour; IC$_{50}$=0.001 µM; the structure of the molecule is taken from the crystal structure of inactive CDK2; PDB ID: 4KD1).

4.2.1 Computations of Binding Affinity of Inhibitors to CDK2

4.2.1.1 Adjustment of Crystal Structure

For the computational study, we used the active form of CDK2 (phospho-CDK2-cyclin A) in a complex with roscovitine. Generally, an inactive catalytic subunit of a cyclin-dependent kinase converts to an active form by association with a regulatory subunit cyclin. The formation of this protein complex (CDK-cyclin) is accompanied by a change of the position of two alpha helices, which enables ATP binding. For computational studies of CDK2, the important difference between the active and inactive form is the position of the residue Glu 51. In the active form, the residue Glu 51 occupies the space approximately 8 Å from the inhibitor where the carboxyl group of Glu 51 is oriented to the binding site and its negative charge contributes to the charge distribution around the inhibitor. In addition, Lys 33 is bound between Glu 51 and Asp 145, therefore it is not sufficiently flexible for direct interactions with the inhibitor[102-103]. Furthermore, the active form of CDK2 is standardly used for biochemical assays as well. In the inactive form, Glu 51 is located at the edge of the protein, considerably distant from the binding site.

The starting geometry taken from the CDK2-roscovitine complex (PDB ID: 3DDQ[70]) was prepared in the following way. To begin with, all the crystal waters and cyclin A were removed. Secondly, hydrogens were added to the protein by the Reduce program[104] and to the inhibitor by the Chimera program (version 1.9).[105] Next, the complex was adjusted by

TLEaP in the AMBER program (version 14).[106] We considered all the aspartates, glutamates, lysines and arginines as charged. All other amino acids including histidines (HIE residues) were taken as neutral. Then, all hydrogens were relaxed by the annealing in AMBER (20 ps, the initial temperature was 2500 K, cooled down to 0 K). We used the parameters from the FF03 force field for the protein and from the GAFF force field for the inhibitor. The partial charges for the inhibitors were calculated using the RESP procedure at the HF/6-31G* level.[107]

Subsequently, the geometries of the inhibitors were manually built in PyMOL (version 2.3.3)[108] using the geometry of roscovitine taken from the crystal structure[70] and the molecules were optimized by the RI-DFT-D/B-LYP/SVP method with Turbomole program (version 6.1).[109] The resolution of identity (RI) approximation was employed. Then, the empirical dispersion correction (D) by Jurecka et al.[110] (the employed damping function is described in Ref.[110]), and the COSMO continuum solvation model[31] were applied for the gradient optimization. After that, conformers of each compound were generated by the molecular dynamics with AMBER (30 ns, 400 K, one step was 2 fs). Following this, the planar scaffolds of the molecules were frozen in order to preserve the coordinates of the purine core, as defined in the crystal structure of the CDK2-roscovitine complex (Figure 4.4). The snapshots were saved each 10 ps and optimized by RI-DFT-D/B-LYP/SVP. At this point, the energetically favourable conformers of each inhibitor were inserted into the CDK2 active site, where only the substituents at the scaffolds were relaxed by the molecular dynamics with AMBER (20 ps, 310 K) for clashes/close contacts removing and adjusting for the active site. It should be noted that the range of the single point energies between the most stable conformer (with the lowest/most negative energy) and the other conformers was taken ~5 kcal/mol. Finally, the resulting complexes were optimized using the QM/SQM method.[29] It should be also noted that we tried to generate the poses of the inhibitors by the standard docking (AutoDock, AutoDock Vina)[111-112], but the corresponding calculated binding free energies did not correlate with the experimental data (not shown).

4.2.1.2 QM/SQM Setup

A sufficiently large region within the protein is required for the most accurate computations to be attained.[113] Nevertheless, the computational demand increases with the number of atoms, which is a standard problem of quantum mechanical methods. Therefore, we used

the QM/SQM method based on the ONIOM computational approach which enables the inclusion of the whole protein-inhibitor complex.[114-115] This less demanding technique describes the protein by defining two structure layers which are described at different levels of accuracy.

In our set-up, the small QM region included the inhibitor and the surrounding crucial amino acids (Figure 4.5). We included the whole residues Glu 8, Lys 9, Ile 10, Gly 11, Glu 12, Lys 20, Lys 33, Phe 80 (without atom N), Phe 82, Leu 83, Asp 86, Asn 132, Asp 145, and the atoms of Val 7 (C, O); Gly 13 (N, CA); Tyr 19 (C, O); Ala 21 (N, CA); Ala 31 (CA, CB); Leu 32 (C, O); Lys 34 (N, CA); Glu 51 (CB, CG, CD, OE1, OE2); Glu 81 (N, CA, C, O); His 84 (N, CA, C, O); Gln 85 (N, CA, C, O); Leu 87 (N, CA); Pro 130 (C, O); Gln 131 (N, CA, C, O); Leu 133 (N, CA); Leu 134 (CB, CG, CD1, CD2); Ala 144 (C, O); Phe 146 (N, CA). We used the RI-DFT-D/B-LYP/SVP method for the optimization and the more accurate RI-DFT-D/TPSS/TZVP method for the calculation of the interaction energies.

For the optimization, the large SQM region was comprised of the residues within 12 Å of the inhibitor. In order to keep the residues on the surface of the SQM region in the geometry that corresponded to the X-ray conformation of the protein, the residues more distant than 10 Å from the inhibitor were frozen during the optimization. The SQM layer around the QM region was described by the PM6-D3H4X method combined with the linear scaling procedure (MOZYME).[30] The modified semi-empirical quantum mechanics PM6-D3H4X method included the corrections for dispersion (D3) energy and hydrogen (H4) and halogen binding (X).[58, 116-117] The dispersion energy was estimated by the two-body interatomic model.[58] The resulting stabilization energies agreed well with the benchmark data.[58, 116-117] In the case of the single point calculations, the large region covered the whole protein-inhibitor complex. The solvent effect was consistently described by the COSMO implicit solvation model[31, 118] at the SQM level for the whole protein-inhibitor complex. The coupling between the QM and the SQM parts was done with the Cuby framework (version 3)[119] developed in our laboratory. It employs Turbomole[109] for the QM and Mopac (version 2012)[30] for the SQM region. In general, the principle of the ONIOM technique is the following:

$$E_{ONIOM} = E_{LR}^{SQM} - E_{SR}^{SQM} + E_{SR}^{QM} \qquad (4.1)$$

where E_{LR}^{SQM} is the energy of the large region (LR) calculated at the SQM level, E_{SR}^{SQM} is the energy of the small region (SR) calculated at the SQM level and E_{SR}^{QM} is the energy of the small region calculated at the QM level.
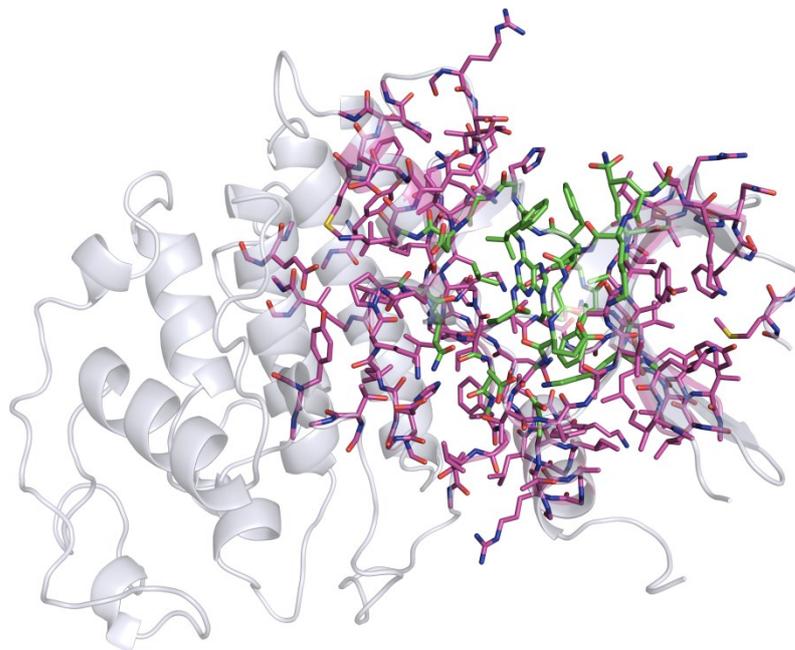
Figure 4.5: The computational set-up includes ~400 atoms in the QM region (green colour) and ~1900 atoms in the SQM region (magenta colour). The QM region was described by the DFT-D method. The SQM region was described by the PM6-D3H4X method combined with the linear scaling procedure (MOZYME).

4.2.1.3 Scoring Function

The binding affinity (score) is the sum of several terms.[29] The scoring function is standardly constructed by the equation

$$Score = \Delta E_{int} + \Delta\Delta G_{solv}^{COSMO} + \Delta G'_{conf}^{COSMO}(inhib.) - T\Delta S_{int}(inhib., prot.) \qquad (4.2)$$

The gas phase interaction energy (ΔE_{int}) was calculated as the difference between the energy of the protein-inhibitor complex and the energy of the isolated protein and inhibitor subsystems. This term was described at the QM/SQM level (see above).

The change of the solvation free energy ($\Delta\Delta G_{solv}^{COSMO}$) of the protein-inhibitor complex was calculated at the SQM level in COSMO.

The interaction 'free' energy, which is fundamental for the description of the protein-inhibitor binding affinities, was calculated as the sum of the gas phase interaction energy and the change of the solvation free energy:

$$\Delta G'_{int} = \Delta E_{int} + \Delta\Delta G_{solv}^{COSMO} \qquad (4.3)$$

The change of the conformational 'free' energy of the inhibitor $\left(\Delta G'_{conf}^{COSMO}(inhibitor)\right)$ was calculated as the energy difference between the conformation of the isolated inhibitor taken from the protein-inhibitor complex and the free inhibitor optimized in COSMO. The conformation 'free' energy was calculated as the sum of the gas phase RI-DFT-D/TPSS/TZVP energy and the PM6 COSMO solvation free energy.

The 'free' energy change of the protein $\left(\Delta G'_{conf}^{COSMO}(protein)\right)$ is commonly calculated as the energy difference between the protein in its optimal solution structure and the conformation adopted in the complex. It is usually described by annealing approach at the molecular mechanics level, which is significantly less accurate than the QM/SQM method. This fact can make the accuracy of the score considerably worse. However, since the geometries of the studied compounds are very similar, the term was omitted in this study.

The interaction entropy ($-\Delta TS_{int}(inhibitor, protein)$) term was determined under a rigid rotor – harmonic oscillator approximation from the vibrational analysis at the molecular mechanics level using the parm03 AMBER potential[120] combined with the generalized Born solvent model. For the inhibitors, $-\Delta TS_{int}$ was calculated as the difference between the vibrational entropy of the unbound inhibitors and the inhibitors bound to the protein. $-\Delta TS_{int}$ of the protein was calculated using selected key amino acids

(Ile 10, Gly 11, Glu 12, Gly 13, Phe 80, Glu 81, Phe 82, Leu 83, Asp 86, Gln 131, Asn 132, Asp 145) as the difference between the vibrational entropy of the geometry in the protein-inhibitor complex and in the protein without the inhibitor. To calculate this term, we used the Partial Hessian vibrational analysis[121] implemented in Cuby.[119]

The calculated terms were correlated with the experimental binding free energies $\left(\Delta G_{exp}^{0}\right)$ expressed from the IC_{50} values via the equation

$$\Delta G^{0}_{exp} = RTln\frac{IC_{50}}{1 + \dfrac{[S]}{K_{M}}} \tag{4.4}$$

where [S] = 15 μM is the concentration of substrate and KM = 200 μM is the Michaelis constant. The computational model was evaluated by using two statistic parameters: the coefficient of determination (R^2) and the predictive index (PI).[122] R^2, the square of the Pearson correlation coefficient between the experimental and predicted data values, quantifies how well the regression line approximates the real data points. It acquires the values 0 to 1, where 0 corresponds to impossibility and 1 to certainty. PI measures the reliability of the prediction, where +1.0 is perfectly correct prediction, −1.0 is perfectly incorrect prediction, and 0.0 is random.

4.2.1.4 Fragmentation of Inhibitors

The fragmentation of the inhibitors was used for the description of $\Delta G'_{int}$ of the contributions of the inhibitor parts (molecular fragments) in complex with the whole protein. The inhibitors taken from the optimized protein-inhibitor complex were divided into four parts (scaffold and three substituents; Figure 4.6) by using PyMOL[108] and the $\Delta G'_{int}$ values were described at the QM/SQM level (see above).
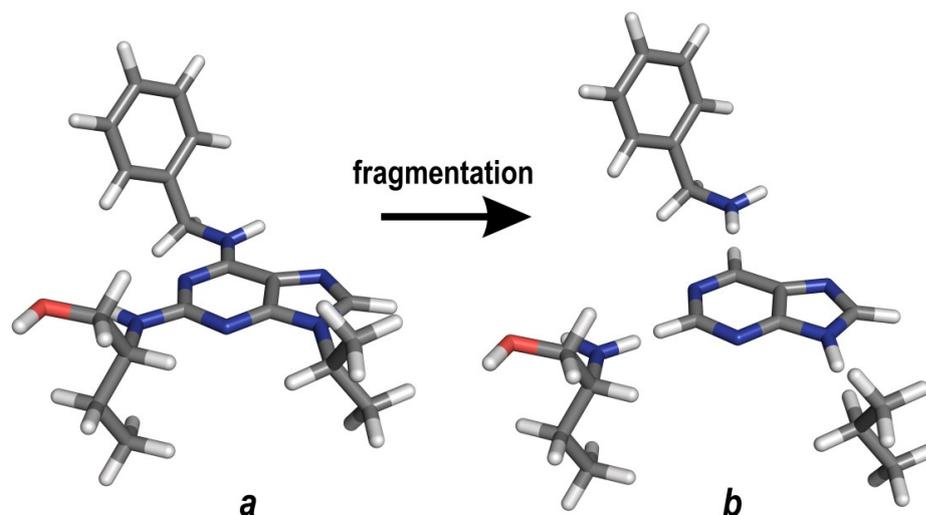
Figure 4.6: The contributions of the molecular fragments to the inhibitory effect. The inhibitors taken from the optimized protein-inhibitor complexes (Figure a) were divided into four parts: the scaffold and three substituents (Figure b).

4.2.1.5 Electrostatic Potential

The molecular electrostatic potentials (ESP) were calculated at the HF/cc-pVTZ level with the Gaussian 09 program[123] using the structures taken from the optimized protein-inhibitor complexes. The ESP were fitted at the number of the charge points around the molecules by using the Merz-Singh-Kollman scheme[124]. The isodensity surface was plotted on the 0.001 a.u. The images were created by Molekel program (version 4.3).[125] The ESP visualizes charged regions of the molecule in the three dimensions by colour-coded maps. Generally, the ESP is important for qualitative prediction of the manners in which the inhibitors fit into the protein active site.

4.2.2 Results of Computational Analysis

We investigated 17 compounds presented in Figure 4.7. The IC_{50} values were determined for eight of them, including two novel inhibitors A2 and F2. To the best of our knowledge, the syntheses of the other substances have not been reported. The suitability of the used methodology was tested on the dataset comprising 29 CDK2 inhibitors with known IC_{50} values. The dataset includes roscovitine and its 22 derivatives, which were published in Ref.[93], the compounds A2, A3, B1, E3, F2, and dinaciclib. We applied the single conformation approach, which means that for each inhibitor, the most energetically favourable protein-inhibitor complex (with the lowest $\Delta G'_{int}$) was used for the correlation

with the experimental data. Moreover, the root-mean-square deviation (RMSD) value, which was calculated for the position of roscovitine in the X-ray and the optimized protein-inhibitor complex, was 0.43 Å.

The calculated scores were compared with the experimental ΔG_{exp}^0 values. The correlation between the interaction 'free' energies $\Delta G'_{int}$, which are the sums of the fundamental terms $\Delta E_{int} + \Delta \Delta G_{solv}^{COSMO}$, and ΔG_{exp}^0 is very close (PI = 0.94; R^2 = 0.84). By the addition of the $\Delta G'^{COSMO}_{conf}(inhibitor)$ and $-TS_{int}(inhibitor, \; protein)$ terms, the correlation between the scores and the ΔG_{exp}^0 values decreased (PI = 0.82; R^2 = 0.69), but the values of the scores simultaneously converged to the experimental data, as shown in Figure 4.8.

The departure of the computationally predicted values from the experimentally measured ones is typical for the used methodology. Mean absolute deviation (MAD) was calculated by the formula

$$MAD = \frac{\sum_i^n |\Delta G_{calc} - \Delta G_{exp}^0|}{n} \tag{4.5}$$

MAD is 25.51 kcal/mol for the $\Delta G'_{int}$ values and 15.72 kcal/mol for the scores. These aspects are addressed in greater detail in Discussion (see below).

Furthermore, the precision/sensitivity of the used methodology was also verified by the comparison of the $\Delta G'_{int}$ values to the sums of $\Delta G'_{int}$ of the contributions of the individual molecular fragments in the complex with the whole protein for each investigated system. We observed the close correlation (PI = 0.86; R^2 = 0.80).

All the results of the verification of the used methodology implied that our computational approach could be used for the accurate description of the binding affinity of the studied compounds.
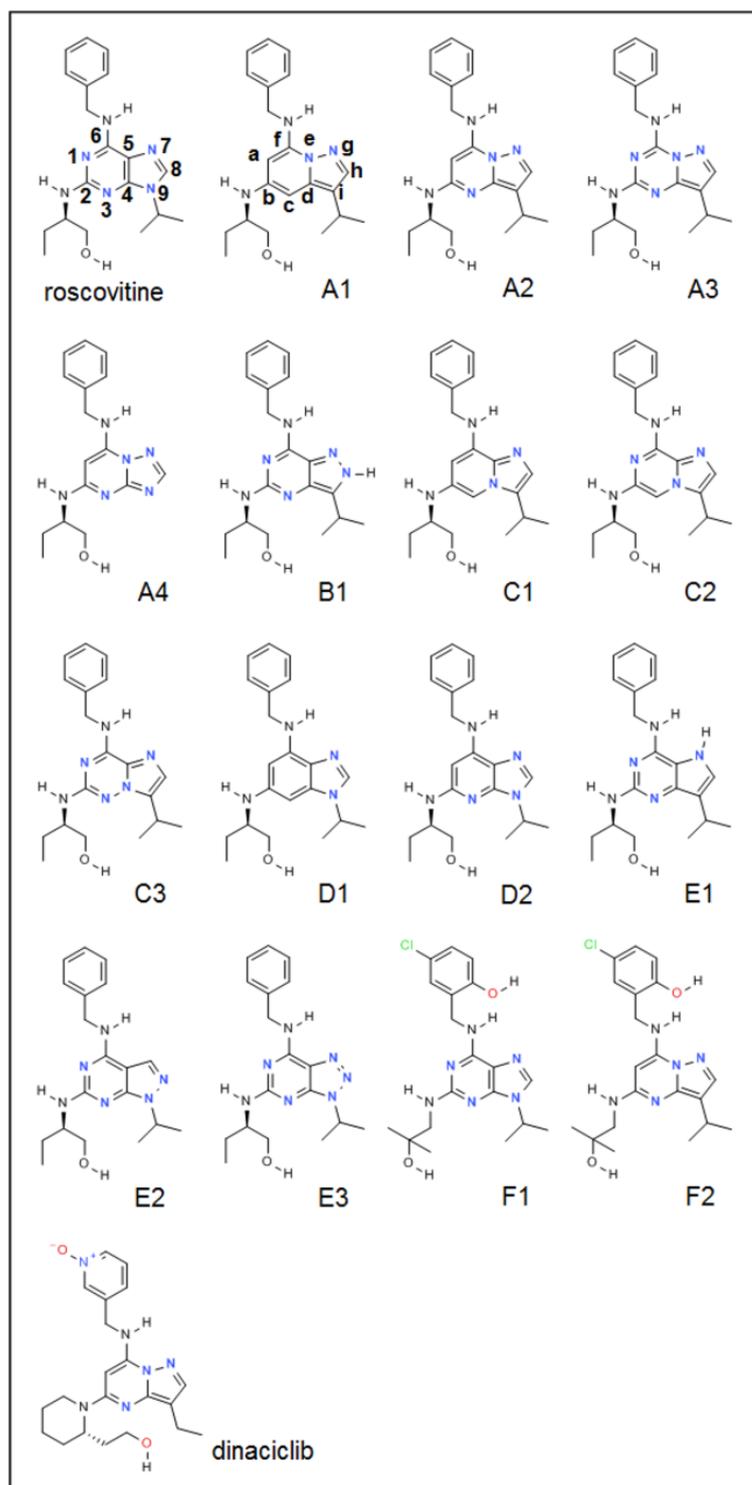
Figure 4.7: The structures of roscovitine, its direct bioisosteres A1-E3, compounds F1 and F2, and dinaciclib. The letter labeling used for compound A1 is used for the corresponding positions in the other compounds. The compounds are clustered into groups A-E according to the positions of nitrogens in the five-membered ring of the central heterocyclic core. Group F includes inhibitors with the substituted benzyl group.
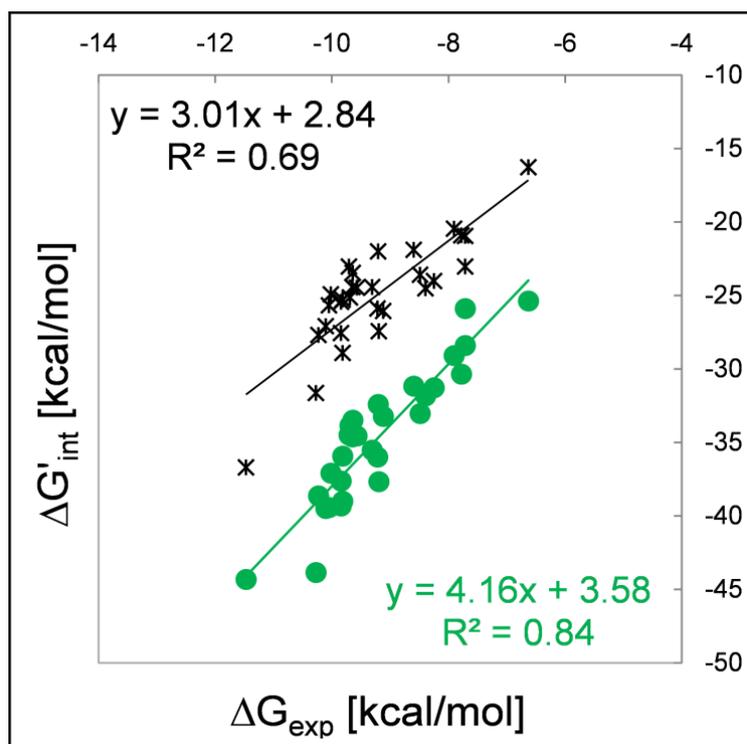
Figure 4.8: The plot of the correlations between the computationally predicted values (score as black points; interaction 'free' energies as green points) and the experimentally measured values.

The $\Delta G'_{int}$ values of the inhibitors in the complex with CDK2 and the $\Delta G'_{int}$ of the contributions of their scaffolds and substituents are summarized in Table 4.1. The terms of the binding scores of all the compounds are presented in Table 4.2. The ESP maps of roscovitine and its bioisosteres are visualized in Figure 4.9 and the active site of CDK2 is in Figure 4.10.

					$\Delta G'_{int}$ of contributions of protein-fragment complex				
									[kcal. mol^{-1}]
inhibitor	**IC$_{50}$ [μM]**	**ΔG^0_{exp}**	**score**	**$\Delta G'_{int}$**	**scaffold**	**benzyl-amine**	**hydroxyl-butyl**	**isopropyl**	**sum of $\Delta G'_{int}$ of contributions**
roscovitine	0.180 ± 0.095	-8.60	-21.91	-31.18	-12.84	-7.81	-6.93	-1.51	-29.09
A1	---	---	-21.80	-31.73	-12.77	-10.22	-6.87	-1.31	-31.18
A2	0.027 ± 0.005	-9.64	-24.45	-34.62	-13.13	-9.99	-7.80	-2.32	-33.25
A3	0.014 ± 0.002	-10.02	-23.47	-33.48	-12.69	-7.61	-7.32	-1.55	-29.18
A4	---	---	-22.94	-30.14	-10.97	-8.47	-8.97	---	-28.41
B1	0.027 ± 0.007	-9.64	-24.93	-37.09	-16.75	-7.79	-7.63	-1.24	-33.40
C1	---	---	-23.03	-31.71	-13.73	-9.60	-5.48	-2.01	-30.81
C2	---	---	-21.74	-31.40	-14.05	-8.12	-6.91	-1.35	-30.44
C3	---	---	-23.60	-31.92	-13.29	-7.97	-6.58	-1.17	-29.02
D1	---	---	-20.02	-30.28	-11.22	-10.35	-7.29	-1.30	-30.15
D2	---	---	-24.64	-33.16	-12.74	-9.62	-7.90	-1.31	-31.57
E1	---	---	-19.20	-28.51	-9.37	-8.07	-7.67	-4.72	-29.82
E2	---	---	-15.55	-23.33	-6.49	-7.57	-5.81	-4.72	-24.59
E3	6.280 ± 0.160	-6.63	-16.27	-25.38	-7.19	-7.54	-7.31	-1.98	-24.02
						hydroxy chlorobenzylamine	**hydroxyl-isobutyl**		
F1 *a*	0.019 ± 0.003	-9.84	-25.44	-37.63	-10.02	-16.96	-4.84	-4.35	-36.17
F1 *b*	---	---	-25.44	-33.07	-10.17	-9.87	-5.57	-2.49	-28.10
F1 *c*	---	---	-23.38	-34.00	-11.39	-11.01	-1.15	-3.79	-27.33
F1 *d*	---	---	-21.91	-30.23	-12.81	-6.56	-0.08	-3.84	-23.29
F2 *a*	---	---	-26.24	-39.55	-5.44	-17.40	-6.44	-4.68	-33.95
F2 *b*	0.009 ± 0.001	-10.27	-31.65	-43.84	-6.64	-22.54	-4.61	-4.95	-38.74
F2 *c*	---	---	-26.02	-37.90	-7.05	-19.38	-3.84	-5.88	-36.16
F2 *d*	---	---	-21.91	-30.63	-10.90	-10.88	-1.67	-5.10	-28.55
						pyridine N-oxide	**hydroxy amine**	**ethyl**	
dinaciclib	0.001 ± 0.001	-10.86	-36.71	-44.31	-14.65	-18.42	-6.67	-1.12	-40.87

Table 4.1: Summary of the results. The half-maximal inhibitory concentrations and their standard deviations; the experimental binding free energies expressed via the equation (4.4); the computational scores and the interaction 'free' energies of the protein-inhibitor complexes; and the interaction 'free' energies of contributions of the individual molecular fragments (the scaffolds and substituents) in the complex with the whole protein. The labels a, b, c and d by the compounds F1 and F2 mean four most energetically favourable poses of F1 and F2 in the complex with CDK2.

	I	II	III	IV	V	VI	VII	VIII = III + IV	IX = III + IV + V + VI +VII
						kcal. mol^{-1}				
compound	IC$_{50}$ [μM]	ΔG_{exp}	ΔE_{int}	$\Delta\Delta G_{solv}^{COSMO}$	$\Delta G_{conf}'^{COSMO}$	$-T\Delta S_{int}(in.)$	$-T\Delta S_{int}(pr.)$	$\Delta G_{int}'$	score
3a	0.020	-9.81	-82.98	43.97	4.24	3.98	1.86	-39.00	-28.92
3b	0.020	-9.81	-83.53	47.60	4.56	4.47	1.59	-35.92	-25.30
3c	0.790	-7.78	-80.45	50.09	3.18	3.91	2.36	-30.36	-20.91
3d	0.070	-9.12	-88.02	54.80	2.48	3.18	1.49	-33.22	-26.07
3e	0.260	-8.39	-75.51	43.68	2.52	2.72	2.05	-31.83	-24.53
3f	0.220	-8.48	-74.91	41.87	4.80	3.00	1.64	-33.04	-23.61
3g	0.340	-8.24	-71.90	40.63	4.46	1.44	1.35	-31.27	-24.02
3h	0.890	-7.71	-69.09	43.19	1.30	2.36	1.28	-25.90	-20.95
3i	0.890	-7.71	-73.87	45.47	2.06	1.93	1.37	-28.40	-23.04
3j	0.031	-9.57	-77.82	43.26	4.49	4.26	1.34	-34.56	-24.48
3k	0.060	-9.20	-82.14	49.72	3.62	5.06	1.73	-32.42	-22.01
3l	0.059	-9.21	-93.07	57.08	4.63	3.89	1.57	-35.99	-25.90
3m	0.630	-7.90	-78.35	49.27	5.13	2.18	1.30	-29.08	-20.47
3n	0.061	-9.19	-87.62	49.96	4.61	3.81	1.80	-37.66	-27.44
3o	0.025	-9.69	-83.52	49.68	3.48	3.75	1.47	-33.84	-25.14
3p	0.009	-10.23	-82.07	43.44	4.93	4.41	1.60	-38.63	-27.69
3r	0.013	-10.05	-80.19	40.77	7.86	4.10	1.79	-39.42	-25.67
3s	0.024	-9.70	-81.67	47.17	4.82	4.60	2.05	-34.50	-23.03
3t	0.012	-10.10	-79.76	40.26	7.77	2.69	1.93	-39.50	-27.11
3u	0.019	-9.84	-81.94	42.61	6.19	3.78	1.78	-39.33	-27.57
roscovitine	0.180	-8.60	-73.56	42.38	3.99	3.77	1.50	-31.18	-21.91
olomoucine II	0.050	-9.30	-81.59	46.07	5.42	3.84	1.83	-35.52	-24.43
A1	---	---	-73.51	41.78	4.72	3.43	1.77	-31.73	-21.80
A2	0.027	-9.64	-76.84	42.22	4.85	3.74	1.58	-34.62	-24.45
A3	0.014	-10.02	-75.98	42.49	4.97	3.52	1.53	-33.48	-23.47
A4	---	---	-71.90	41.76	2.45	3.32	1.42	-30.14	-22.94
B1	0.027	-9.64	-80.61	43.52	6.10	4.02	2.03	-37.09	-24.93
C1	---	---	-73.96	42.25	3.30	3.56	1.83	-31.71	-23.03
C2	---	---	-73.51	42.11	5.21	2.84	1.61	-31.40	-21.74
C3	---	---	-74.86	42.94	3.33	3.48	1.52	-31.92	-23.60
D1	---	---	-73.46	43.18	4.70	3.87	1.70	-30.28	-20.02
D2	---	---	-75.75	42.59	3.07	3.87	1.57	-33.16	-24.64
E1	---	---	-71.99	43.48	4.37	3.23	1.71	-28.51	-19.20
E2	---	---	-65.48	42.15	3.70	2.76	1.32	-23.33	-15.55
E3	6.280	-6.63	-69.04	43.65	3.89	3.76	1.46	-25.38	-16.27
F1 *a* /3q	0.019	-9.84	-87.44	49.81	6.46	4.40	1.33	-37.63	-25.44
F1 *b*	---	---	-80.39	47.33	2.38	3.80	1.44	-33.07	-25.44
F1 *c*	---	---	-77.43	43.43	4.73	4.10	1.79	-34.00	-23.38
F1 *d*	---	---	-78.17	47.94	2.58	4.00	1.74	-30.23	-21.91
F2 *a*	---	---	-82.79	43.24	7.40	3.87	2.04	-39.55	-26.24
F2 *b*	0.009	-10.27	-104.25	60.41	6.17	4.06	1.96	-43.84	-31.65
F2 *c*	---	---	-85.34	47.44	6.76	3.63	1.49	-37.90	-26.02
F2 *d*	---	---	-81.32	50.69	3.07	3.72	1.88	-30.63	-21.91
dinaciclib	0.001	-11.47	-112.15	67.84	2.31	3.92	1.38	-44.31	-36.71

Table 4.2: Summary of the computational terms. I: the IC$_{50}$ values. II: the experimental binding free energy expressed via the equation (4.4). III: the gas phase interaction energy. IV: the change of the solvation free energy of the protein-inhibitor complex. V: the change of the conformational 'free' energy of the inhibitor. VI, VII: the sum of the interaction entropy of inhibitor/protein. VIII: the interaction 'free' energy as the sum of the terms fundamental for the description of the protein-inhibitor binding affinities. IX: the score describing the protein-inhibitor binding affinities as the sum of all relevant computational terms. The labels a, b, c and d by the F1 and F2 compounds refer to the different poses of the compounds in the complex with CDK2.
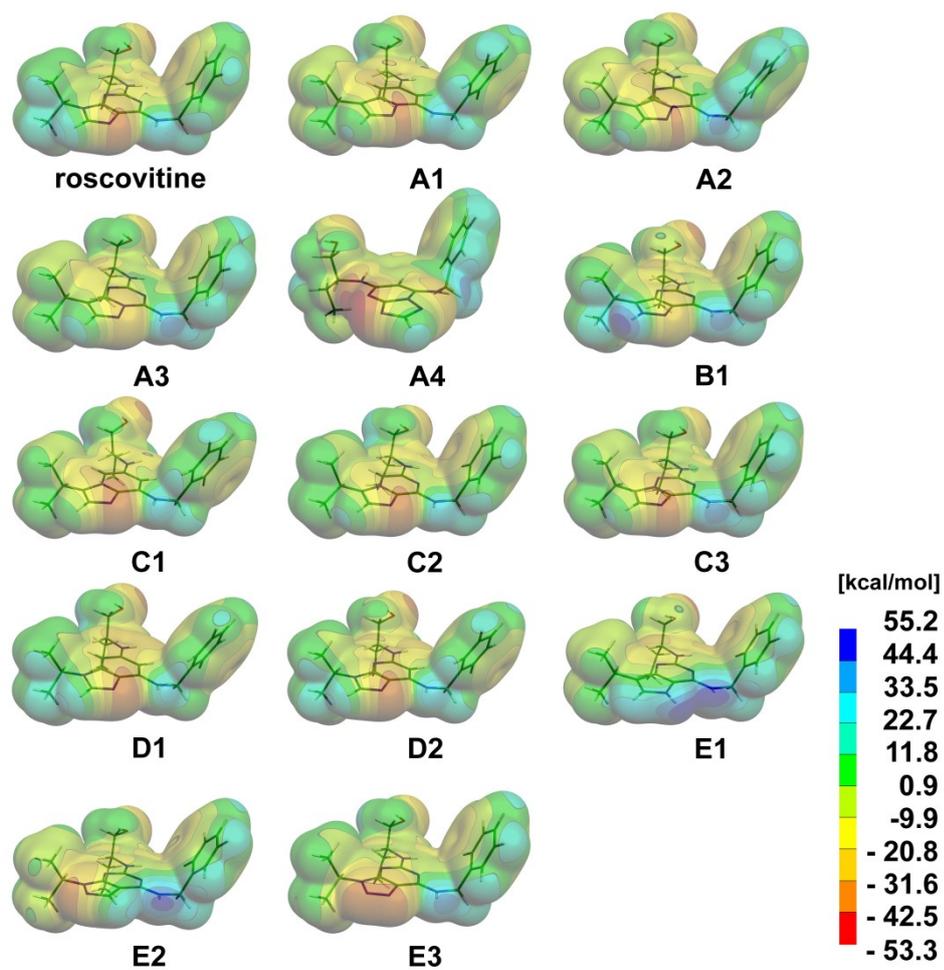
Figure 4.9: The electrostatic potentials of roscovitine and its bioisosteres were computed on the geometries taken from the optimized CDK2-inhibitor complexes. The charge distribution is indicated by the colour-coded maps. The depicted front parts of the molecules are oriented to the hinge region in the active site, where the basic binding motifs of the inhibitors are located.
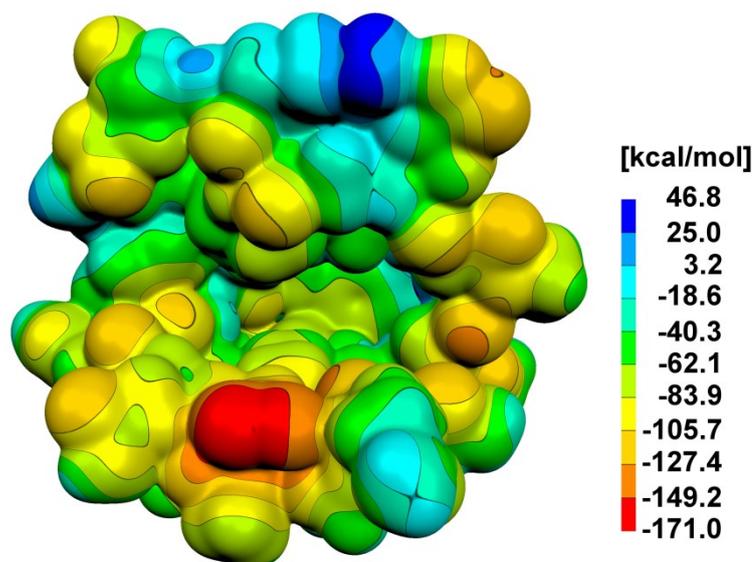
Figure 4.10: The ESP surface displays the charge distribution of the CDK2 active site in the three dimensions by the colour-coded map. The inner part of the cleft shows the hydrophobic character of the binding site.

To compare $\Delta G'_{int}$ of roscovitine with the bioisosteres and assess the contributions of their molecular fragments to the overall binding affinity, the compounds were divided into five groups (A-E) according to the positions of the nitrogen atoms in the five-membered ring, which is crucial for the interaction with the hinge region of CDK2 (Figure 4.11). Group F includes two inhibitors possessing substituted benzyl groups (Figure 4.12). For the evaluation of the binding modes, we used the generally accepted values: the distance between the donor and acceptor atom 1.5-2.6 Å for the hydrogen bond (H-bond) and 2.7-3.5 Å[126-127] with the directionality 151-180° for the halogen bond (X-bond).[128-129]

Group A (compounds A1, A2, A3 and A4 in Figure 4.11) includes the inhibitors with nitrogen atoms at positions e and g in the five-membered ring. These compounds have the same calculated binding mode as roscovitine, but A2 and A3 are significantly more efficient inhibitors (Table 4.1), which is likely caused by the combination of several factors. For A2, A3, A4, the calculated charge distribution at the benzylamine nitrogen at position f is more positive compared to roscovitine, which suggests that the H-bond between the nitrogen and Leu 83 is stronger (Figure 4.9). A1, A2, and A3 also carry less positive charge at positions h and i. A4 lacks the propyl group at i and thus cannot form the dispersion interaction with Phe 80, which likely decreases the binding affinity. The presence of nitrogen at c

(compounds A2-A4) makes the geometry of the molecule more compact and thus conformationally more stable while the carbon at c (compound A1) is less favourable because it causes hydrogen repulsion interaction. The benzyl groups of A1, A2 and A4 are turned to the protein due to the hydrogen repulsion between this group and carbon at a and therefore contributes significantly to $\Delta G'_{int}$. In A3, the benzyl group is rotated towards the nitrogen at a, which stabilizes the molecule geometry but simultaneously reduces the interactions of the benzyl group with the protein. Since these two factors are essentially compensatory, both conformations of the benzyl group have similar influence on the total binding affinity.
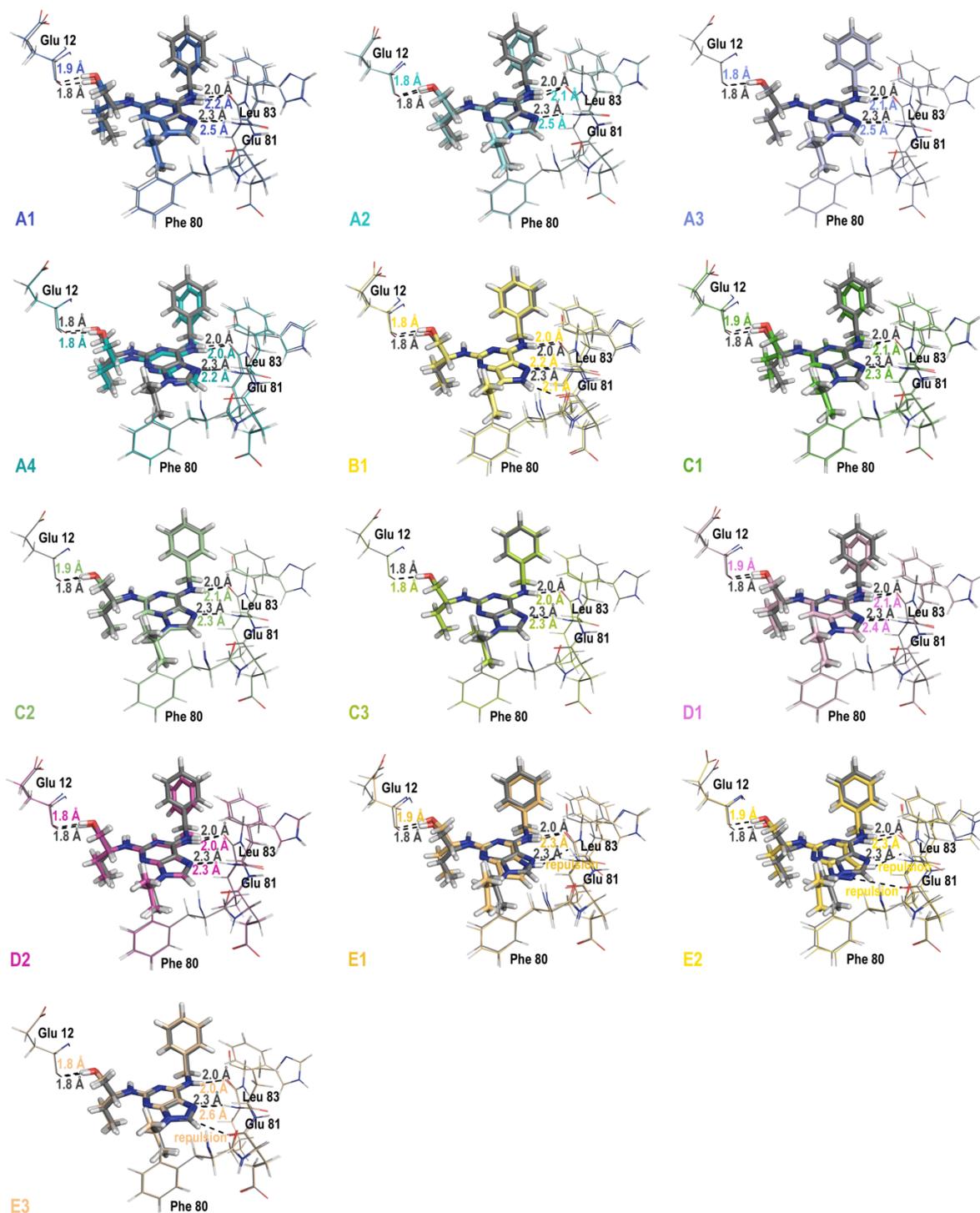
Figure 4.11: Comparison of the calculated binding modes of roscovitine and its bioisosteres (A-E) with different central heterocycles in the active site of CDK2 (only key residues are depicted). The binding modes of the inhibitors are the same, but their binding affinities are different; for example, the IC$_{50}$ value of A2 is 0.027 μM while IC$_{50}$ of E3 is 6.280 μM.

Group B consists of the inhibitor B1, which has nitrogen atoms at positions g and h in the five-membered ring (Figure 4.11). The inhibitory activity of the compound is similar to A2 and A3 (Table 4.1). The ESP map shows a significantly strong positively charged area surrounding h (Figure 4.9) which can be utilized for the H-bond with Glu 81; in contrast to roscovitine, which utilizes only a weak H-bond at this position. Correspondingly, $\Delta G'_{int}$ of the contribution of the scaffold is the highest of all the studied roscovitine bioisosteres.

Group C (compounds C1, C2, and C3 in Figure 4.11) includes the inhibitors with nitrogen atoms at positions g and d of the five-membered ring. These compounds have the same binding mode as roscovitine and their $\Delta G'_{int}$ values are also close (Table 4.1). Although $\Delta G'_{int}$ of the contributions of their scaffolds are higher in comparison to roscovitine, the results suggest that the presence of nitrogen at d has only a negligible effect on the total binding affinity. The ESP areas surrounding h and i have less positive character than in the case of roscovitine (Figure 4.9). Nevertheless, $\Delta G'_{int}$ of the contributions of the isopropyl groups are similar, which suggests that this term is (at least within the groups A-D) relatively independent of the scaffold.

Group D (compounds D1 and D2 in Figure 4.11) includes the inhibitors with nitrogen atoms at positions g and i. Both compounds have the same binding mode as roscovitine; D1 with calculated affinity lower and D2 higher than roscovitine ($\Delta G'_{int}$ in Table 4.1), which (as in the case of A1 vs. A2) is likely influenced by the presence of nitrogen at position c. $\Delta G'_{int}$ of the contribution of the D2 scaffold is higher compared to D1, but slightly lower than in the case of roscovitine. The ESP maps are very similar to that of roscovitine (Figure 4.9).

Group E (compounds E1, E2, and E3 in Figure 4.11) includes the inhibitors with the worst calculated binding affinities ($\Delta G'_{int}$ in Table 4.1). They are likely to be caused by the repulsion interactions with the hinge region, which is also in accordance with the ESP maps: E1 and E2 have positive charged areas at g, and E2 and E3 carry negative charge at h while roscovitine has these areas charged inversely (Figure 4.9). In the case of E1 and E2, the repulsion is located between position g and Leu 83; E2 and E3 also show additional repulsion between h and Glu 81. The overall calculated affinities of the compounds (E1 > E3 > E2) reflect the contributions calculated for their separated individual scaffolds. Interestingly, these compounds have the highest $\Delta G'_{int}$ of the contributions of the isopropyl group at position i among all the studied compounds, which is likely related to the fact that this group is differently oriented towards Phe 80.

Group F consists of purine F1 and (heretofore unknown) pyrazolo[1,5-a]pyrimidine F2, which differ from the other compounds by the substituents at positions b (hydroxy-isobutyl) and f (2-hydroxy-4-chlorobenzylamine) (Figure 4.7). Both compounds are potent inhibitors of CDK2 (Table 4.1). The four most energetically favourable poses of F1 and F2 are depicted in Figure 4.12; the best calculated affinities are associated with poses a for F1 and poses b for F2. F1a forms four H-bonds: two with the hinge region (as roscovitine), one between the substituted benzylamine motif at f and the hinge region, and one H-bond located between the aminoalcohol substituent at b and Asp 145. F2b forms three H-bonds and one X-bond. However, only one H-bond is located between the scaffold and the hinge region of CDK2 because the scaffold is shifted significantly to the edge of the active site where the inhibitor forms two strong H-bonds between the substituted benzylamine at f and Asp 86; and between the aminoalcohol at b and Asp 145. Therefore, the $\Delta G'_{int}$ contribution of the F2 scaffold is lower than that of F1. In F2, $\Delta G'_{int}$ of the contribution of the substituted benzylamine group is significant both due to the weak X-bond between the chlorine atom of this group and Glu 8, and to the hydroxyl with Asp 86. However, in F1 the X-bond is absent due to the different orientation of the benzylamine motif; this, on the other hand, enables H-bond interaction with Leu 83.
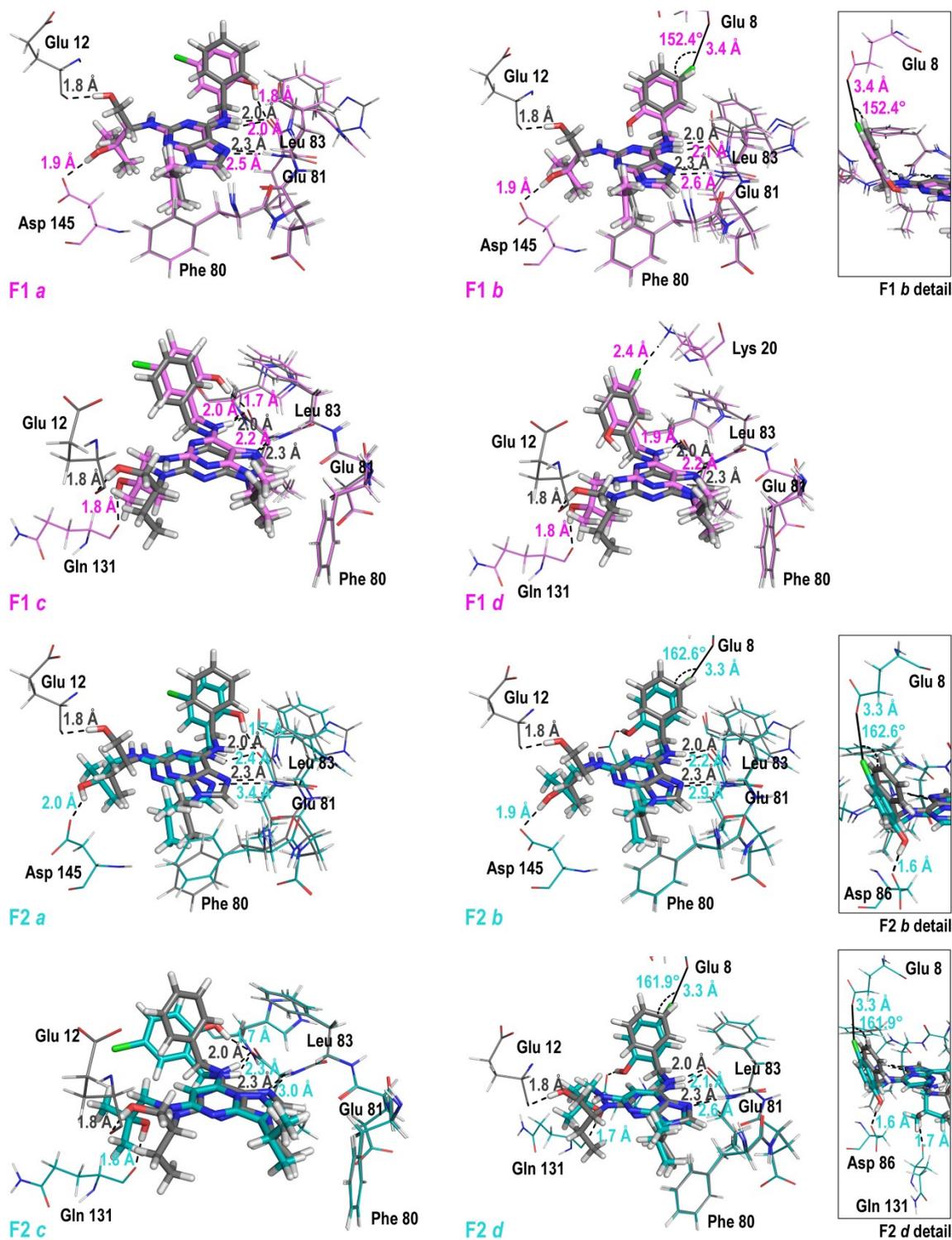
Figure 4.12: Comparison of the calculated binding modes of roscovitine and the four most energetically favourable modes of F1 and F2 in the active site of CDK2 (only key residues are depicted).

Figure 4.13 demonstrates the character of the contributions of the scaffolds and the substituents to the binding affinities of the protein-inhibitor complexes, including the newly prepared compounds A2 and F2 and the potent inhibitor dinaciclib, which is currently being profiled in clinical trials. Dinaciclib is bound through four H-bonds: two are constituted between the pyrazolo[1,5-a]pyrimidine scaffold and the hinge region, one is located between the aminoalcohol substituent at b and Asn 132, and one is formed by the interaction of the pyridine N-oxide at f and Lys 89. $\Delta G'_{int}$ of the contributions of the scaffold as well as the pyridine N-oxide motif are significant (Table 4.1). The comparison between F1 (IC_{50} = 0.019 μM) and F2 (IC_{50} = 0.009 μM), which have identical substituents and different scaffolds, shows the improvement of the binding affinity due to the modification of the atom arrangement of the scaffold, but this effect is not substantial. The compounds A2 (IC_{50} = 0.027 μM) and F2 with the pyrazolo[1,5-a]pyrimidine scaffolds and the different substituents demonstrate the increase of the inhibitory effect through the change of the substituents. On the other hand, comparison of dinaciclib (IC_{50} = 0.001 μM) and F2, which have also the same scaffold and different substituents, illustrates that the additional substituents do not always lead to a considerably better binding affinity. These results suggest non-additive contributions of the scaffolds and substituents to the inhibitory efficiency.
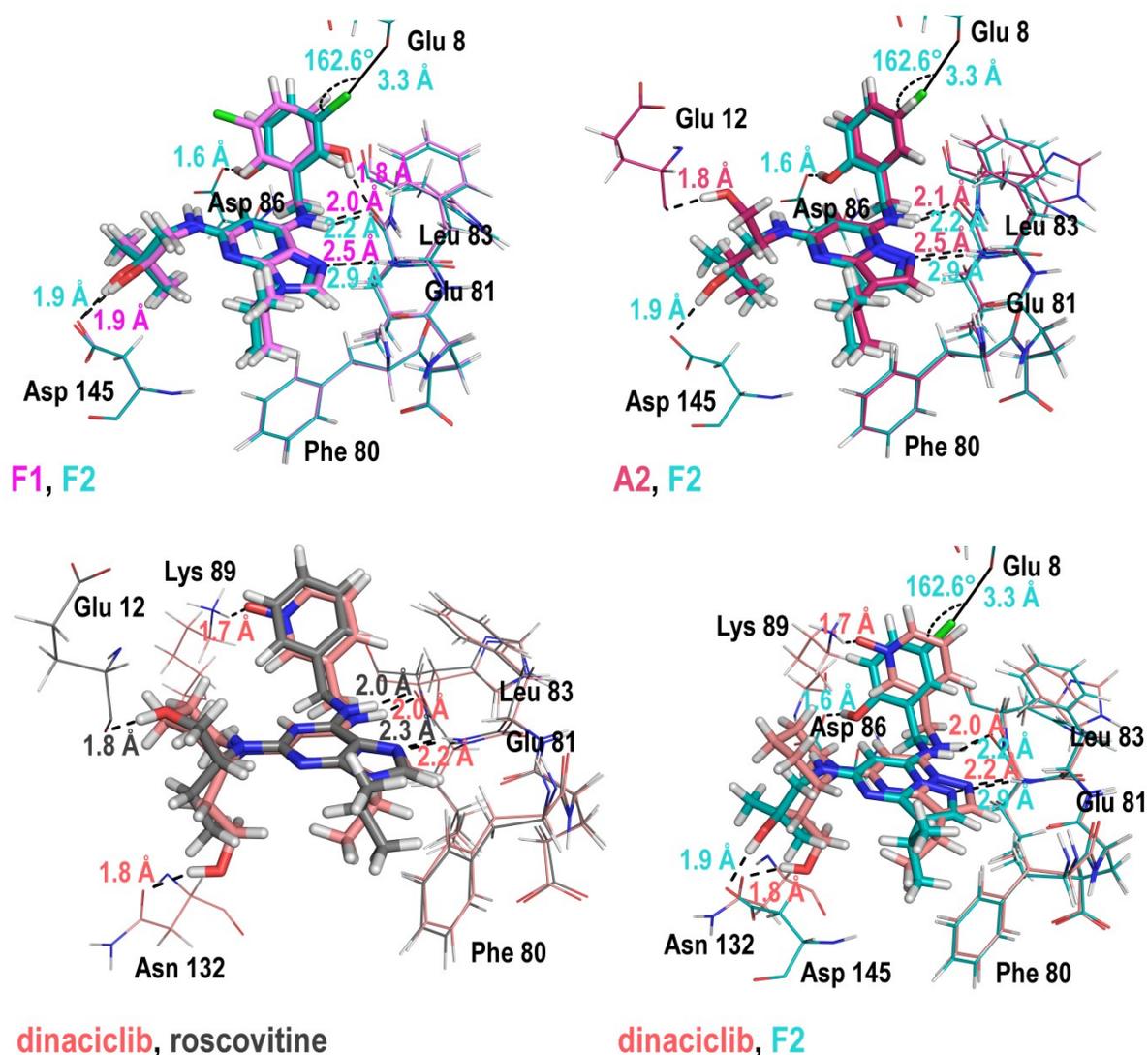
Figure 4.13: Comparison of the calculated binding modes of F1 and F2, A2 and F2, dinaciclib and roscovitine, dinaciclib and F2 in the active site of CDK2 (only key residues are depicted).

4.2.3 Discussion

The discovery of the first purine-based CDK inhibitors has spurred the design of novel compounds with improved biochemical and biological properties. Systematic efforts focused on bioisosteric replacement of the purine core have revealed several groups of inhibitors (e. g., pyrazolo[1,5-a]-1,3,5-triazines, pyrazolo[1,5-a]pyrimidines, pyrazolo[1,5-a]pyridines and pyrazolo[4,3-d]pyrimidines) with inhibitory activities superior to that of roscovitine, while other direct analogs with central pharmacophores of very similar shape

and polarity have proven to be significantly less active.[98] In this study, we aimed to explain these non-trivial trends in greater detail by theoretical and computational chemistry.

In order to calculate the binding affinities of the selected roscovitine analogs with high accuracy, we chose the computational approach based on the QM/SQM setup, which employs the RI-DFT-D method for QM part and the PM6-D3H4X method for SQM part. The solvent effect was described by the COSMO implicit solvation model.[31, 118] The Root mean square error determined for the L7 benchmark data set[130] by the RI-DFT-D/TPSS/TZVP method is 2.95 kcal/mol and 3.92 kcal/mol for the PM6-D3H4/SMB method. Our computationally demanding (and time-consuming) set-up included ~400 atoms in the QM region (which essentially corresponds to the limit of this method) and ~1900 atoms in the SQM region. One cycle of the gradient optimization took approximately 1 hour and the whole process ca. 2 weeks on 8 processors (Intel Xeon E5630 2.53 GHz). The single point calculations were done with the whole protein-inhibitor complex containing ~5000 atoms. The error bars of the computationally predicted values are not reported because the statistical error, which is straightforward to calculate in molecular dynamics or ensembles-using methods, cannot be evaluated in our single-conformation methodology.

Since the preferred binding mode of a small-molecule inhibitor in complex with a protein is defined by many factors that are impossible to address with complete accuracy[29], we applied the single conformation approach (sometimes called end-point approach), which takes into account the most energetically favourable protein-inhibitor complex for each compound. We used the standard computational model without explicit water molecules in the active site.[29] Our results are in accordance with the findings published in Ref.[131], which suggests a relatively minor influence of the waters on the binding affinity of structurally close CDK2 inhibitors (roscovitine and its analogs).

The achieved correlation between the $\Delta G'_{int}$ and ΔG^0_{exp} values is very good, but the addition of the other terms causes the decrease of the correlation between the scores and ΔG^0_{exp}. Similar results related to the CDK2 protein-inhibitor complexes were published in Ref.[132] This effect can be caused by several factors. The $-\Delta TS_{int}$ values were calculated at lower level of accuracy (because of the high computational demand), which in general results in less precise values. The terms can compensate each other to a certain extent, and it is impossible to perfectly quantify particular contributions of the terms. Although the addition

of the $\Delta G'^{COSMO}_{conf}(inhibitor)$ and $-TS_{int}(inhibitor, protein)$ terms decrease the correlation between the scores and ΔG^0_{exp}, it simultaneously converges the values of the calculated scores to the experimental data. We can expect that the addition of other terms as $\Delta G'^{COSMO}_{conf}(protein)$ should further decrease the deviation of the computationally predicted values from the experimentally measured ones. To consider the other terms is however beyond the scope of this study. Real protein-inhibitor interactions represent a complex and sensitive balance between the free energy gained through the formation of the protein-inhibitor complex and the free energy required for desolvation and change of the conformation of both the inhibitor and the active site of the protein. The protein-inhibitor binding is also accompanied by changes of entropy. These important terms are not possible to calculate at the same level of accuracy and the computed values are by one order of magnitude larger than the experimental binding free energy. In addition, the energy difference between a tight-binding inhibitor (with a nanomolar IC_{50} value) and a weak-binding inhibitor (with IC_{50} value of ca. 100 μM) is only about 5 kcal/mol.[133] Due to the approximate character of the scoring function, the agreement of the computationally predicted and the experimentally measured values is therefore often limited.

Our computational results indicate that the nature of the central heterocyclic core can be fundamental for the binding affinity of purine bioisosteres to CDK2, which is demonstrated by the comparison of roscovitine with the more potent inhibitors A2, A3 and B1 possessing bioisosteric cores. In addition, the scaffold influences the interaction of the attached substituents with the protein. Depending on the scaffold, the same substituents can be associated with different $\Delta G'_{int}$ of the contributions, which is illustrated e. g. by comparison of compounds F1 (purine scaffold) and F2 (pyrazolo[1,5-a]pyrimidine scaffold).

To conclude, we have demonstrated that the effects of the scaffold and the substituents are largely interdependent, and that their contribution to the binding affinity is not additive.

4.2.4 Conclusion

Computational methods have become an integral part of modern rational drug design. Proper in silico approaches can effectively describe and predict interactions between small-molecule inhibitors and their biomolecular targets. Our work demonstrates that detailed theoretical studies, based on accurate computations, can make the computer-aided drug design more efficient. In this project, we calculated relative binding affinities of several known and two newly prepared analogs of roscovitine with different bioisosteric central

cores to CDK2. The activities of the majority of the studied inhibitors were reported, but the structural basis of these experimental observations has not been explained to date. The results of our calculations are in accordance with the experimental biochemical data and suggest that the approach can be used more generally for prediction of other organic compound-biomolecule interactions.

4.3 3,5,7-Substituted Pyrazolo[4,3-*d*]pyrimidine Inhibitors of Cyclin-Dependent Kinases and Their Evaluation in Lymphoma Models

4.3.1 Introduction

This work is aimed at describing of synthesis and biological activity of new pyrazolo[4,3-d]pyrimidines as CDK2 inhibitors with potential therapeutic effect on Non-Hodgkin lymphomas (NHLs). These hematologic malignancies are most frequent in the western hemisphere, comprising approximately 30% of all hematologic cancers.[134] The presented study includes crystallographic and computational analysis as well.[32]

The recently described 5-alkylamino-3-isopropyl-7-[4-(2-pyridyl)benzyl]amino-1(2)H-pyrazolo[4,3-d]pyrimidines share some of the same substitutions as the inhibitors roscovitine, but showed substantially higher potency against CDKs and cancer cell lines than roscovitine did.[135-136] Because of the laborious synthesis of these derivatives, we sought to improve the synthetic accessibility of these compounds while retaining or even improving their biochemical potency. Herein we focus on modification of the substituted position 5 on the pyrazolo[4,3-d]pyrimidine scaffold and present a new series of 3,5,7-trisubstituted pyrazolo[4,3-d]pyrimidines in which the alkylamino group at the 5-position of the heterocycle is replaced by a 5-alkylthio group.

To verify the effect of this modification, we employed molecular modelling. We applied the computational procedure delineated in our previous work, in which we identified the pyrazolo[4,3-d]pyrimidine core as the most favourable central heterocycle of the purine bioisosteres of CDK2 inhibitors.[28] Importantly, we have shown that the contributions of the central heterocyclic cores and the individual substituents, quantified and evaluated in relation to conformations of the optimized protein-inhibitor complexes, are not simply additive. Depending on the scaffold, the same substituents can be associated with different interaction "free" energies $\Delta G'_{int}$ (in kcal.mol^{-1}). Our preliminary calculations suggested

that the replacement of the alkylamino group at the 5-position with the 5-alkylthio group can increase the binding affinity to CDK2.

4.3.2 Results

The theoretical study analyses the binding mode of compound **4.35** described in this work, and its nitrogen-containing pyrazolo[4,3-*d*]pyrimidine analog **5i** which was described previously.[136] We computationally investigated two conformers of each inhibitor. The conformers **4.35**calc1 and **5i**calc1 were identified by the computational procedure[28] as the most stable geometries. For this purpose, we utilized the **CR8** (PDB ID: 3DDP)[70] molecule as the template. The **4.35**calc2 conformer is identical to the **4.35** geometry published here and **5i**calc2 was designed according to this conformation. As CDK2 model, we used the crystal structure 3DDQ[70] which was computationally adjusted in our previously study.[28]
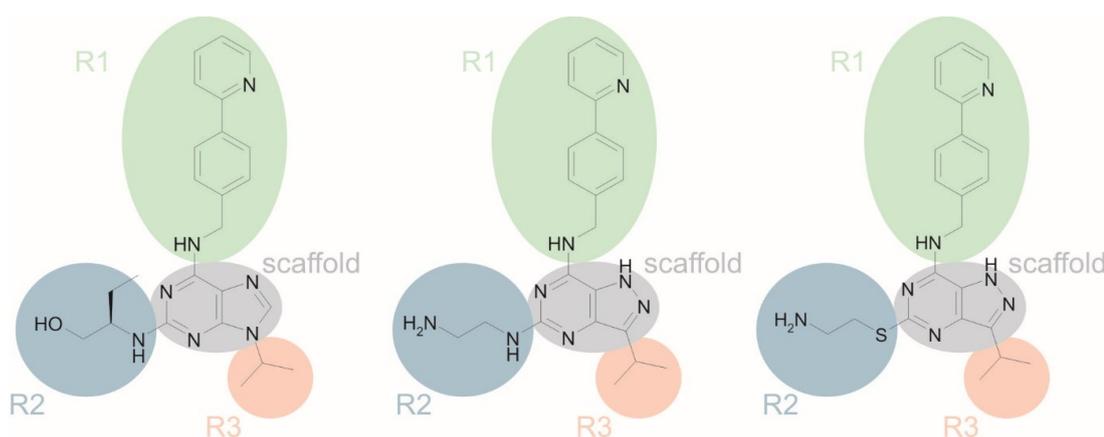
*Figure 4.14: The structures of CR8, **5i** (middle) and **4.35** (right). To assess the $\Delta G'_{int}$ contributions of the inhibitor parts, the inhibitors were fragmented into four parts (the scaffold and three substituents)*

We used the computational methods and methodology detailly described above in the foregoing work.[28]

The computationally optimized protein-inhibitor complexes have the following binding modes. The inhibitors are bound by three hydrogen bonds between the scaffold and the protein hinge region (carbonyl of Leu83, backbone NH of Leu83, carbonyl of Glu81), similar to the binding modes of **CR8**. The compounds also create a hydrogen bond between Lys89 and the 4-(2-pyridyl)benzylamine moiety at position 7. The compounds are stabilized by dispersion interactions between the 3-isopropyl group and Phe80 as well. The

inhibitors make other hydrogen bonds between the substituents at position 5 and the residues situated at the edge of the active site. The 2-aminoethylthio moiety of **4.35**calc1 is bound to carbonyl of Gln131 and the terminal amide of Asn132. The 2-aminoethylamino moiety of **5i**calc1 forms the hydrogen bonds with the same residues. An analogous 4-hydroxybutylamine group of **CR8** is bound by two hydrogen bonds with the backbone NH and Glu12 carbonyl (Figure 4.15 and Figure 4.16).

The evaluation of the computational data (Table 4.3) shows that the interaction between the protonated amino group of Lys89 and the free electron pair of the pyridine nitrogen substantially contributes to the total $\Delta G'_{int}$ values of all the complexes. The $\Delta G'_{int}$ of **4.35**calc1 indicates that this inhibitor is the most efficient. The $\Delta G'_{int}$ contributions of the scaffolds denote the larger values of the pyrazolo[4,3-d]pyrimidine core (**4.35**calc1 and **5i**calc1) compared with the purine core (**CR8**), in line with our previous report.[28] Importantly, the $\Delta G'_{int}$ contribution of the 2-aminoethylthio moiety of **4.35**calc1 is significantly higher than that of the 2-aminoethylamino of **5i**calc1. In addition, the total $\Delta G'_{int}$ of **4.35**calc1 (as well as the sum of the $\Delta G'_{int}$ of its fragments) indicates higher binging affinity than **5i**calc1. Thus, the computational results indicate that the replacement of the alkylamino group at the 5-position of the heterocycle by the 5-alkylthio group is completely relevant.

compound	IC$_{50}$ [µM]	ΔG_{exp}	$\Delta G'_{int}$ [$kcal.mol^{-1}$]					
			fragments of inhibitors				suma of fragments	CDK2-inhibitor complex
			scaffold	R1	R2	R3		
CR8	0.062	-9.87	-11.69	-17.92	-8.24	-3.81	-41.66	-44.37
4.35calc1	0.002	-11.92	-15.25	-18.28	-7.41	-3.20	-44.15	-47.15
4.35calc2	---	---	-13.74	-17.43	-5.34	-3.55	-40.07	-44.19
5icalc1	0.018	-10.61	-14.62	-17.99	-4.52	-3.11	-40.24	-44.62
5icalc2	---	---	-18.40	-17.94	0.20	-4.55	-40.70	-43.57

Table 4.3: The $\Delta G'_{int}$ values of the protein-inhibitor complexes and $\Delta G'_{int}$ of the protein-fragment complexes that describe the contributions of the molecular parts (the scaffold and the substituents) to the binding affinity.
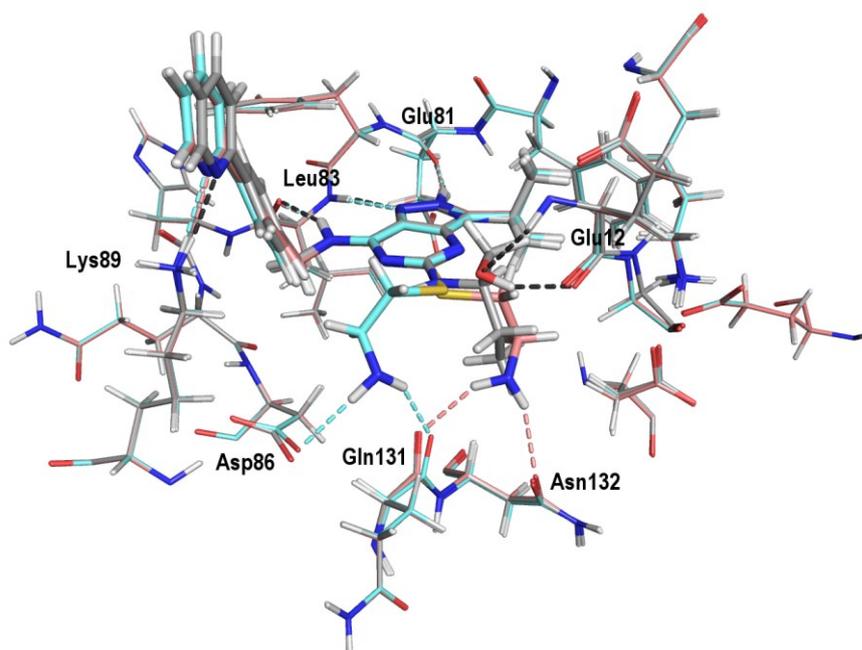
*Figure 4.15: The computationally optimized complexes of CDK2 with the inhibitors **4.35** and **CR8**. Only the key residues are depicted. The geometries of the **4.35** conformers (**4.35**$_{calc1}$ – salmon colour, **4.35**$_{calc2}$ – aquamarine colour; IC$_{50}$ = 0.002) in comparison with **CR8** (gray colour; IC$_{50}$ = 0.062) are illustrated.*
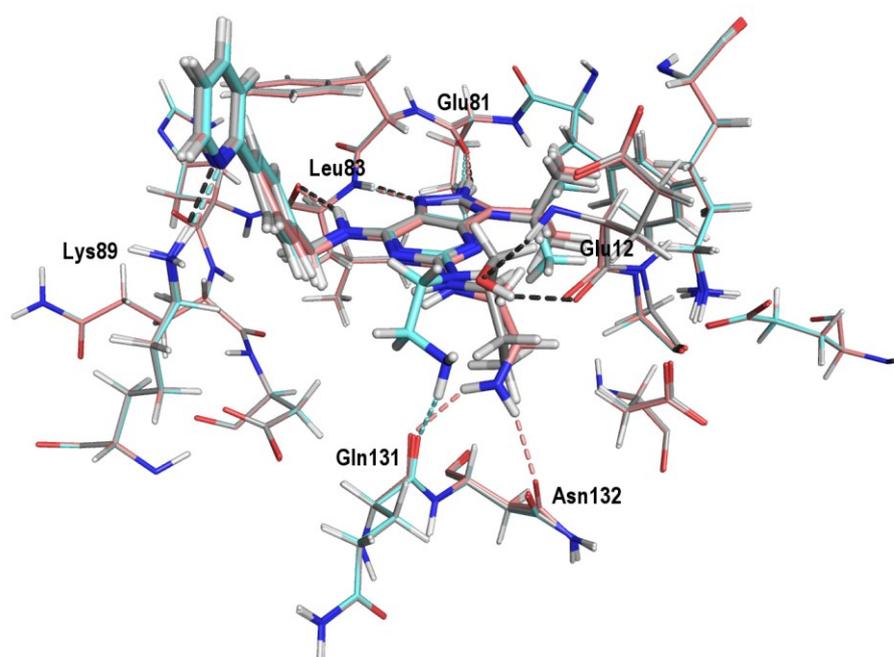
*Figure 4.16: The computationally optimized complexes of CDK2 with the inhibitors **5i** and **CR8**. Only the key residues are depicted. The geometries of the **5i** conformers (**5i**$_{calc1}$ – salmon colour, **5i**$_{calc2}$ – aquamarine colour; IC$_{50}$ = 0.018) in comparison with **CR8** (gray colour; IC$_{50}$ = 0.062) are illustrated.*
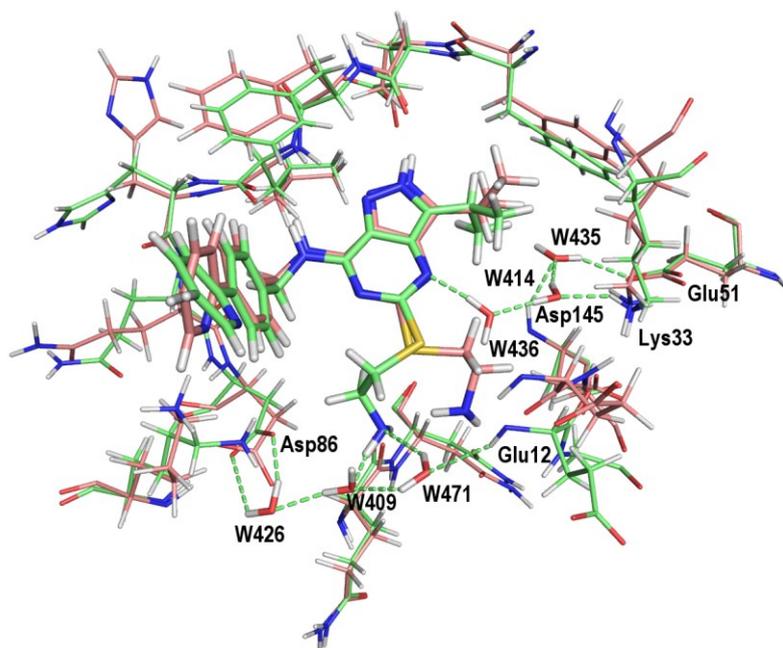
Figure 4.17: The nonoptimized complex of CDK2 with 4.35 (lime colour) compared with the computed complex of CDK2 with the conformer 4.35calc1 (salmon colour). Only the key residues are depicted. This demonstrates how the less stable 4.35 conformer is substantially stabilized by the hydrogen bonds with two crystal water chains. The geometry of hydrogens of the water molecules was modeled in PyMol.
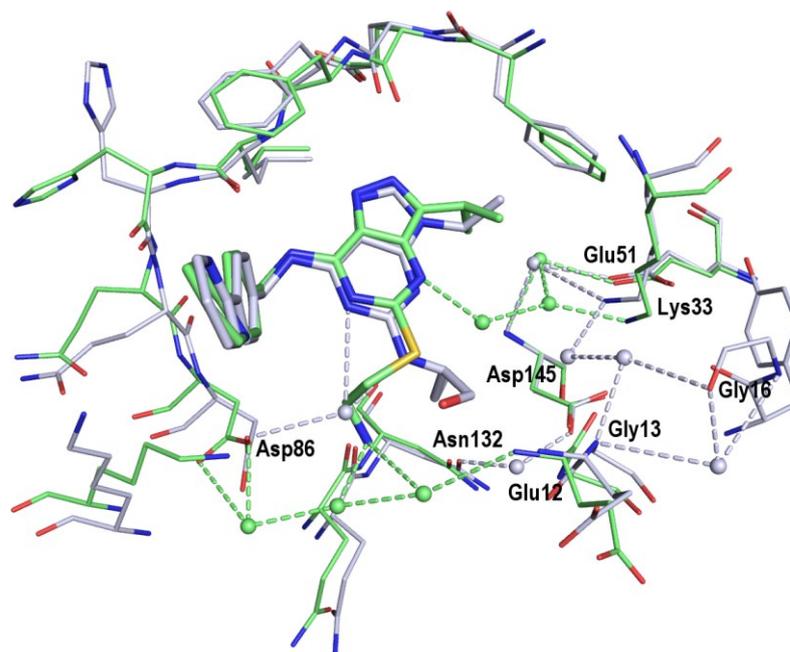
*Figure 4.18: The nonoptimized complex of CDK2 with **4.35** (lime colour) compared with the nonoptimized complex of CDK2 with **CR8** (blue-white colour). Only the key residues are depicted. The inhibitors stabilized by hydrogen bonds with the crystal water chains are shown.*

4.3.3 Discussion

The computationally predicted conformer **4.35**$_{calc1}$ differs in the geometry of the 2-aminoethylthio moiety at position 5 from the conformer detected in the crystal structure. This substituent in **4.35**$_{calc1}$ and **5i**$_{calc1}$ is similarly oriented to the analogous moiety of **CR8** at position 5 (Figure 4.15, Figure 4.16, and Figure 4.18). In contrast, in the crystal structure, the substituent is oriented in the opposite direction. We evaluated the effect of this orientation in the **4.35**$_{calc2}$ and **5i**$_{calc2}$ conformers. **4.35**$_{calc2}$ corresponds to the crystal structure and **5i**$_{calc2}$ was modelled accordingly. The amino group of the moiety at position 5 of **4.35**$_{calc2}$ is bound to the carboxylate of Asp86 and the backbone carbonyl of Gln131, and the amino group of the moiety at position 5 of **5i**$_{calc1}$ forms the hydrogen bond with the backbone carbonyl of Gln131 (Figure 4.15 and Figure 4.16). The comparison of the total $\Delta G'_{int}$ values of these conformers shows lower binding affinity than those of **4.35**$_{calc1}$ and **5i**$_{calc1}$. The $\Delta G'_{int}$ contributions of the moiety at position 5 are also lower than **4.35**$_{calc1}$ and **5i**$_{calc1}$; the **5i**$_{calc2}$ moiety even results in repulsion (Table 4.3). We conclude that although the **4.35**$_{calc1}$ conformer is more stable than the **4.35**$_{calc2}$ conformer, the **4.35**$_{calc2}$ conformer

(in complex with the protein) is favourable because the protein-inhibitor complex can be substantially stabilized by two crystal water molecule chains (Figure 4.16, Figure 4.17, and Figure 4.18).[137]

4.3.4 Conclusion

Herein, we describe the synthesis and biological activity of new pyrazolo[4,3-d]pyrimidines that inhibit CDK2 as demonstrated by biochemical assays and crystallographic analysis. Our findings confirmed the in vitro and in vivo sensitivity of aggressive non-NHLs to CDK inhibitors and provide a rationale for their future clinical evaluation. In addition, the combination of CDK inhibitors selected based on the biological mechanism (targeting predominantly MCL-1) and the BCL-2 targeting agent venetoclax proven to exert synergistic effects exhibit synthetic lethality in vivo and should be considered for combinatorial treatment approaches in patients. The implementation of suitable substitutions at position 5 of the pyrazolo[4,3-*d*]pyrimidine scaffold can improve the inhibitory effect. The substituent at position 5 is mostly in contact with the flexible parts of the enzyme surface or with the solvent (water molecules) and therefore can be more variable, whereas the side chains at positions 3 and 7 do not allow such variability. Despite low differences in the CDK2 inhibition measured with the studied compounds, we found some relationships that are associated with the size/length of the substituent at position 5 or its polarity. Compound 3, which contains an acidic sulfanyl functional group in position 5, was the weakest in the series, with nearly three-fold lower potency than an analogous 3,7-disubstituted compound described earlier.[136] The methylation of the 5-sulfanyl functional group results in short lipophilic substituents in **4.14**, increasing the potency more than ten-fold.

5. Physicochemical Properties of Neuroactive Compounds

5.1 Introduction

The projects presented in this part are aimed at the computational estimate of the thermodynamic properties of several groups of neuroactive compounds that modulate the activity of the N-methyl-D-aspartate receptor (NMDAR). The NMDA is one of three types of glutamate receptors and ligand-gated ion channel protein present at nerve cells ,[33] Figure 5.1.[138] The NMDAR plays key physiological roles in controlling synaptic plasticity and memory function.[34-35, 139] Therefore, it is associated with the pathophysiology of several central nervous system (CNS) disorders and recently has been identified as a locus for disease-associated genomic variation.[36] The receptor is named after the N-methyl-D-aspartat agonist molecule showing selective binding to this type of receptor. Activation of NMDAR by glutamate and glycine binding results in the opening of the ion channel that is nonselective to cations and enables their transport through the cell membrane.[140] The opening and closing of the ion channel are mainly gated by ligand binding, the current flow through the ion channel is voltage dependent. Extracellular ions Mg^{2+} and Zn^{2+} can be bound to specific sites on the receptor and block the passage of other cations through the open ion channel. Depolarization of the cell dislodges and repels the Mg^{2+} and Zn^{2+} ions from the pore, thus allowing a voltage-dependent flow of Na^+ and small amounts of calcium Ca^{2+} ions into the cell and potassium K^+ out of the cell.[141-144] Ca^{2+} flux through NMDARs is thought to be essential in synaptic plasticity, which is a cellular mechanism for learning and memory, and its disorder is connected with dementia such as Alzheimer's disease.[145]

NMDAR is comprised of two glycine-binding GluN1 and two glutamate-binding GluN2 subunits that form the central cation-permeable channel pore. The crystal structure of the intact heterotetrameric GluN1/GluN2B NMDAR shows that the receptor is arranged as a dimer of GluN1-GluN2B heterodimers with the two-fold symmetry axis running through the entire molecule composed of an amino-terminal domain (ATD), a ligand-binding domain (LBD), and a transmembrane domain (TMD). The ATD and LBD are much more highly packed in the NMDA receptors than non-NMDA receptors, which may explain why ATD regulates ion channel activity in NMDAR but not in non-NMDAR.[138]

The activity of NMDAR can be influenced by allosteric modulators such as neurosteroids, or other neuroactive compounds.[37]Although neuroactive compounds have a similar structure, they can inhibit as well as to potentiate the receptor activity.
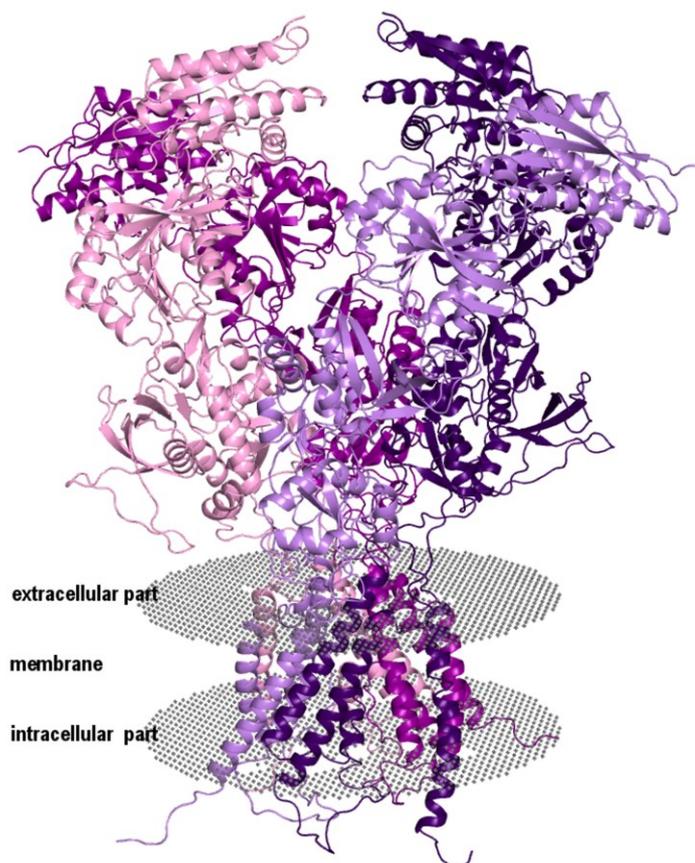
Figure 5.1: The crystal structure of GluN1a/GluN2B N-methyl-D-aspartate receptor ion channel (PDB ID: 4PE5).

5.2 Neurosteroid-like Inhibitors of *N*-Methyl-D-Aspartate Receptor: Substituted 2-Sulfates and 2-Hemisuccinates of Perhydrophenanthrene

5.2.1 Introduction

This study is comprised of the design, synthesis, and assessment of the biological activity of a library of perhydrophenanthrene 2-sulfates and 2-hemisuccinates (**1-10**). Their ability to modulate NMDAR-induced currents were tested on recombinant GluN1/GluN2B receptors. The design of new NMDAR modulators could lead to the development of pharmaceutically attractive, drug-like compounds, as a broad variety of central nervous system diseases (neurodegeneration, ischemia, traumatic brain injury, etc.)[146-148] have been associated with glutamate induced excitotoxicity under pathological conditions, a specific form of neuronal cell death caused by overactivation of NMDARs.

We focused on clarification of the critical role of the steroidal D-ring. Therefore, we have synthesized a series of fully saturated phenanthrenes that have not been previously reported. The two major steroidal structural features necessary for maintaining inhibitory biological activity on NMDARs are retention of the 5β-stereochemistry and axial 3α-configuration.[149] In addition, the stereochemistry at steroidal positions C-8 and C-9 was maintained. As such, these new perhydrophenanthrene analogues (**1−10**) closely mimic the steroidal ABC ring arrangement. The inhibitory effect of neurosteroids on NMDARs is also associated with the presence of a positively or negatively charged substituent at the C-3 position.[149-150] Hence, the sulfate moiety was selected as the primary choice. To evaluate these new perhydrophenanthrene analogues as druglike candidates, we drew inspiration from the published data for pregnanolone C-3 hemisuccinate ester, a compound with neuroprotective activity both in vitro and in vivo models of neurodegeneration.[151-152] Thus, the hemisuccinate moiety was chosen to extend the structure−activity relationship (SAR) study.

5.2.2 Computational Estimate of Thermodynamic Properties

5.2.2.1 Preparation of Structures

The studied compounds were manually built in PyMOL[108] using the geometry of the molecule taken from the crystal structure (PDB ID: 3CAV),[153] and were relaxed by the RI-DFT/B-LYP/SVP method with the program Turbomole (version 6.1).[109] The empirical dispersion correction (D)[110] and COSMO continuum solvation model[31] were applied on the

gradient optimization. The most stable local minima of the compounds were generated by the molecular dynamics simulation with a general AMBER empirical force field (the simulation was run for 30 ns; the constant temperature was 400 K).[120] The package AMBER 14 MD was used.[104] The partial charges of the molecules were calculated using the RESP procedure[107] at the HF/6-31G* level. The resulting geometries were minimized by the RI-DFT-D/B-LYP/SVP//COSMO method and their single-point energies (SP) were calculated at the RI-DFT-D3/B-LYP/TZVPP//COSMO level.[154] The chosen structures were re-optimized by the RI-DFT-D3/B-LYP/TZVPP//COSMO method and their SP were calculated at the same level of accuracy.

5.2.2.2 Computational Methods

The solvation free energy (ΔG_{solv}) of the compounds was calculated in the SMD continuum solvation model[40] (the transfer from vacuum to water and from *n*-octanol to water) at the HF/6-31G* level with the program Gaussian (version 09).[123] In this method, the single-point energies were computed with an identical molecular geometry for both the transfer from vacuum to water and from *n*-octanol to water. ΔG_{solv} estimates the behaviour of the single-molecules in water and *n*-octanol that represents the membrane environment. The negative values of ΔG_{solv} signify the free energy gained, and the positive values signify free energy required during the transfer from the first phase to the second phase.

The partition coefficient (P) is defined as the ratio of concentrations of a neutral solute in *n*-octanol and water, and it represents the solute lipophilicity. It is usually reported as the common logarithm:

$$logP = log \frac{c_{neutral,octanol}}{c_{neutral,water}} \tag{5.1}$$

The calculated logP was obtained via the equation

$$logP = \frac{\Delta G_{ow}}{-RT \ln(10)} \tag{5.2}$$

where ΔG_{ow} is the transfer free energy, R is the molar gas constant and T is temperature (298.15 K).[43]

The ΔG_{ow} was calculated on the basis of the change in the molecular conformation related to the transfer between *n*-octanol and water. The compounds were optimized at the M06-2X/6-31G* level in SMD with Gaussian, and ΔG_{ow} was expressed as the difference

between the total energies in water and in n-octanol, because this energy includes the internal energy of the molecule.[43]

The distribution-coefficient (D), which is also presented as a common logarithm, takes into account both the neutral and ionized form of the solute in both phases and is used for estimating the lipophilicity of ionizable species.[155]

$$logD = log \frac{c_{ionized,octanol} + c_{neutral,octanol}}{c_{ionized,water} + c_{neutral,water}} \tag{5.3}$$

The logD values were predicted at pH = 7.4, which is the physiological pH of blood serum, using the program MarvinSketch.[156]

5.2.3 Results

In the computational study, we have investigated the lipophilic qualities of the studied compounds in the context of their inhibitory effect since the lipophilicity belongs to the basic characteristics of the neurosteroids, which influences their interactions with NMDAR.[47] The relevant physicochemical properties[38] of compounds **1-10** were estimated by quantum mechanics computational methods and by physicochemical properties predictor. The computational results are summarized in Table 5.1.

compound	$IC_{50}[\mu M]$	$\Delta G_{exp}[kcal.mol^{-1}]$	$\Delta G_{solv}[kcal.mol^{-1}]$: transfer from				logP	logD
			vacuum to water		n-octanol to water			
			neutral	charged	neutral	charged		
1	74.6	-5.7	-20.70	-75.89	0.98	-6.58	2.10	1.47
2	63.4	-5.8	-23.96	-84.84	2.35	-6.03	4.08	0.97
3	33.0	-6.1	-16.21	-71.51	2.46	-5.02	3.07	1.49
4	29.2	-6.2	-19.55	-80.19	3.89	-4.40	5.25	0.99
5	224.0	-5.0	-20.77	-75.69	0.68	-6.84	1.71	1.18
6	87.6	-6.1	-23.99	-84.64	2.13	-6.26	3.92	0.68
7	77.0	-5.6	-17.18	-72.67	1.94	-5.55	2.74	1.21
8	33.3	-6.1	-20.58	-81.33	3.34	-5.02	4.98	0.70
9	15.6	-6.6	-14.15	-69.60	2.61	-4.95	3.11	1.55
10	23.2	-6.3	-17.46	-78.33	4.03	-4.33	5.10	1.04
Correlation for sulfate derivatives (1,3,5,7,9)			0.78	0.78	0.80	0.77	0.81	0.71
Correlation for hemiester derivatives (2,4,6,8,10)			0.93	0.93	0.97	0.96	0.93	0.28
Correlation for compounds 1-10			0.36	0.05	0.63	0.77	0.36	0.03

Table 5.1: Summary of the computational values logP, logD, and ΔG_{solv} $[kcal.mol^{-1}]$ of the studied compounds. The negative values of ΔG_{solv} signify the free energy gained, and the positive values signify free energy required, during the transfer from the first phase to the second phase.

The interactions of neurosteroid/neurosteroid-like compounds with NMDAR constitute a complicated process. First, the micelles occurring in the extracellular liquid fuse with the membrane; thereafter the single-molecules leave the membrane and finally enter the channel vestibule, which is the hydrophobic site of action.[157] The ΔG_{solv} values show that the charged molecules prefer water while the neutral ones n-octanol. These results suggest that the transfer between n-octanol and water is possible and it can be accompanied by the transition between the neutral and the ionized state. The lipophilic character of compounds **1-10** was estimated by the logP and logD coefficients. The logP is defined as the logarithm of the ratio of neutral solute concentrations in n-octanol and water. On the other hand, the logD is the logarithm of the ratio of the sum of the concentrations of the neutral and the ionized form of the solute in each of the two phases and thus is depending on pH. As compounds **1-10** can be both ionized and neutral, the description of the lipophilicity by the logD is more appropriate.[155] The correlations between the calculated data and the experimental ΔG_{exp} data show two different groups of the compounds evincing similar behaviour only for ΔG_{solv} during the transfer from n-octanol to water. The correlations indicate only a trend because the number of the evaluated compounds is not representative but just a minimum. The sulfated analogues (**1**, **3**, **5**, **7**, **9**) show R^2 = 0.80/0.77 (neutral/charged form) for ΔG_{solv}, and R^2 = 0.81/0.71 for logP/logD. The dependence is

lower than for the hemisuccinate analogues (**2**, **4**, **6**, **8**, **10**), where $R^2 = 0.97/0.96$ (neutral/charged form) for ΔG_{solv}, and $R^2 = 0.93/0.28$ for logP/logD. The correlation between logD values of the hemisuccinates and the ΔG_{exp} values has not found. This can indicate that the real ratio of the neutral and ionized form of the solute of the hemisuccinates is different from the calculated, which is most likely caused by the different character of their functional hemisuccinate group in the aqueous and *n*-octanol environment (e.g. pKa) as compared with sulfate moiety. The hemisuccinate analogues with slightly better inhibitory effect than the sulfated ones (except for the compound **10**) also evince the logD values lower and logP values higher than the sulfates. The sulfate and hemisuccinate analogues, which bear ether of ester moiety on the skeleton, have a similar inhibitory effect except for the significantly different values of the compound **5** (sulfate) and the compound **10** (hemisuccinate). We hypothesize such a phenomenon can be explained by conformation/deformation energy of the molecule during its motion inside NMDAR to the site of the action.[157] The hemisuccinate derivative is considerably more flexible as it has more rotational C-C bonds and so more conformers. On the other hand, the sulfate moiety is rather rigid. The deformation energy is expressed as the difference of the single point energies between the global minimum (extended form) and the local minimum (compact conformer) in aqueous phase. The deformation energy is 5.08 kcal/mol for compound **5** and 13.47 kcal/mol for compound **6**. The results suggest that compound **6** inclines to compose the compact form through the intramolecular interaction of the aliphatic chain with the carbon skeleton and the ester moiety (Figure 5.2 A). Although this intramolecular interaction has an adverse energy effect, it facilitates the passage of the molecule to the site of the inhibitory action.[157] On the contrary, the conformation of the compound **5** remains nearly extended (Figure 5.2 B). Therefore, the conformers can have different ΔG_{solv} as well. The methyl ether of methyl ester substituents can also contribute to both conformation change of the particular molecule, e.g. compound **4** tends to form the compact conformer, probably just as all hemisuccinate analogue, but the methyl group at C-8 fastens the ether moiety at C-7.

Therefore, the conformer is less clenched (Figure 5.2 C) and the deformation energy is 11.82 kcal/mol. The comparison of pregnanolone sulfate (IC$_{50}$ = 24.6 μM, logP = 2.93 and logD = 1.67)[47] with compounds **9** (IC$_{50}$ = 15.6 μM, logP = 3.11 and logD = 1.55) and **10** (IC$_{50}$ = 23.2 μM, logP = 5.10 and logD = 1.04) shows better inhibitory effect of both compounds with logD values lower and logP values higher than pregnanolone sulfate. Also,

5β-Androstan-3α-yl sulfate (IC_{50} = 1.2 μM, logP = 4.07 and logD = 2.11)[47] evinces significantly better inhibitory activity than compounds **9** and **10**, while its logD value is significantly higher than in the case of compounds **9** and **10** and it has lower logP value than compound **10**. Compound **9** (sulfate) shows the best inhibitory activity from the series of compounds **1-10** and it has the highest logD among the studied compounds and the highest logP among the sulfated analogues (Table 5.1). Compound **10** (hemisuccinate) which is the second-best studied inhibitor, has the highest logD among the hemisuccinate analogues but lower log D value than the sulfated analogues; its logP is the second-best among all the compounds.
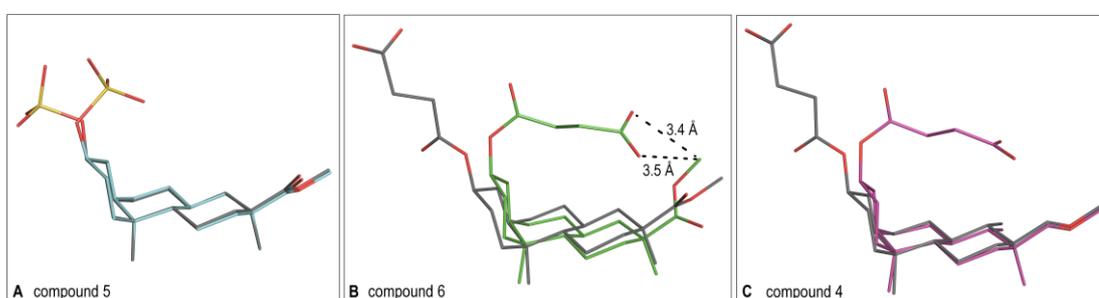
Figure 5.2: The conformers of compounds 4-6. A. The global (gray) and local minimum (blue) of compound 5. B. The global (gray) and local minimum (green) of compound 6. C. The global (gray) and local minimum (magenta) of compound 4.

5.2.4   Conclusion

In this study, we have examined new compounds (**1-10**), with fully saturated phenanthrene skeletons that mimic the steroidal ABC ring system, in order to evaluate their ability to modulate NMDAR-induced current. We have shown, that the steroidal D-ring can be fully or partially degraded with a variety of polar substituents while maintaining inhibitory activity. We have also shown that the potency of these new inhibitors is comparable with the endogenous ligand, pregnanolone sulfate. Also, we have shown that the degradation of the steroidal D-ring did not lead to NMDA subtype selectivity and that there is a higher potency to inhibit NMDAR responses than AMPA responses. Moreover, the IC_{50} values determined for GABAR responses were different from those determined for NMDAR responses for compound **9** that was evaluated as the most potent modulator of a focused library of perhydrophenanthrene analogues. This indicates that the steroidal D-ring plays a significant role for NMDAR/GABAR selectivity and gives a perspective for further design

of receptor selective neurosteroid-like based drugs. This new structural motif offers new prospects for the further modification and optimization of the pharmacological and pharmacokinetic properties of these neuroactive compounds. Finally, the computational analysis has defined the structure-function relationship of the steroid-like compounds in connection with their lipophilicity, solvation free energy, and conformation change which have a significant effect on the inhibition of NMDAR responses. Our results denote the importance of the lipophilicity and the solvation free energy for the action of the steroids at the NMDAR, which is in an agreement with our previous results for sulfated steroids modified on the D-ring.[47] The results also indicate that a decrease of the lipophilicity related to the structural modification need not lead to the decrease of the inhibitory activity. Besides, the results point out the weighty effect of the conformation change of the steroids. Therefore, we conclude that the results are more likely to suggest a different character of the inhibitory action of both inhibitory groups.

5.3 Positive Modulators of the *N*-Methyl-D-Aspartate Receptor: Structure-Activity Relationship Study on Steroidal 3-Hemiesters

5.3.1 Introduction

Herein, we report the synthesis and structure-activity relationship study of pregn-5-ene and androst-5-ene dicarboxylic acid esters (**2-24**) on positive modulation of *N*-methyl-D-aspartate receptor function. Next-generation genome sequencing allowed identification of mutations in GRIN genes encoding for human NMDAR subunits that have been associated with various neurodevelopmental disorders.[158] The analysis of mutated receptors revealed various forms of trafficking and functional defects.[159] The inhibitory effect of neurosteroids on NMDARs is dependent upon a 3α negatively charged substituent and 5βstereochemistry.[160] This class of neurosteroids is represented by endogenous pregnanolone sulfate (20-oxo-5β-pregnan-3α-yl sulfate).[161] In contrast, a 3β-negatively charged moiety in combination with Δ5-stereochemistry (a double bond between C-5 and C-6) favors potentiation of NMDARs.[162] This class of neurosteroids is represented by endogenous pregnenolone sulfate (20-oxo-pregn-5-en-3β-yl sulfate, compound 1, PES).[150] PES is an abundantly occurring neurosteroid synthesized de novo in the central nervous system, which exhibits different modulatory effects on several types of receptors, specifically potentiating the responses elicited by NMDARs[163] while inhibiting currents mediated by γ-aminobutyric acid type A receptors (GABAARs),[164] glycine receptors,[165]

and α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptors (AMPARs).[163] A memory enhancement effect after the administration of PES in vivo has been reported.

5.3.2 Results

Our results for steroids with inhibitory action at NMDARs showed the plasma membrane as the route for the steroid to reach its binding site on the receptor.[150] We could speculate that steroids with a potentiating effect, which are structurally similar to the inhibitory ones, may employ the same route. Therefore, we have analyzed the role of lipophilicity in the potency of compounds at NMDARs.

The computational analysis was used to evaluate the lipophilic qualities of compounds **1−24**. The relevant physicochemical properties (ΔG_{solv}, logP, and logD values), which are commonly used for describing the compound lipophilicity,[38] were calculated by quantum mechanics computational methods and with a physicochemical properties predictor.[41] The used methods and methodology was described above. The computational results are summarized in Table 5.2. Compounds **5** and **14** show similar values of log D (1.29 and 1.28, respectively). However, the EC_{50} value for compound **5** (19.0 μM) is 4 times higher than for compound **14** (4.7 μM). Similarly, more lipophilic compounds **11** and **22** (logD values of 2.47 and 2.42, respectively) show an approximately 5-fold difference in EC_{50} values. These results suggest that compound lipophilicity and presumed solubility in the membrane play only a minor role in determining the compound potency and indirectly indicate that potency is more dependent on the specific (structural/binding) interactions between the compound and NMDAR. Our recent results for steroids with inhibitory action at NMDARs showed the plasma membrane as the route for the steroid to reach its binding site on the receptor.[150] We could speculate that steroids with a potentiating effect, which are structurally similar to the inhibitory ones, may employ the same route. Therefore, we have analysed the role of lipophilicity in the potency of compounds at NMDARs.

| | ΔG_{solv} $[kcal.mol^{-1}]$: transfer from | | | | | |
| | vacuum to water | | n-octanol to water | | | |
compound	neutral	charged	neutral	charged	logP	logD
1 (PES)	-21.10	-75.74	1.76	-5.67	2.02	1.26
2	-17.62	-79.36	3.40	-4.88	3.49	0.71
3	-20.10	-81.91	3.79	-5.33	3.81	0.84
4	-20.30	-85.84	4.17	-5.56	4.27	0.73
5	-21.68	-88.61	4.42	-5.21	4.41	1.29
6	-21.10	-88.10	5.18	-4.35	4.97	1.75
7	-21.09	-88.77	5.87	-4.02	5.46	2.09
8	-20.70	-88.52	6.54	-3.24	6.00	2.64
9	-12.15	-75.10	6.43	-2.78	5.29	2.02
10	-12.51	-79.10	6.82	-3.03	5.74	1.91
11	-13.86	-81.37	7.06	-2.62	6.14	2.47
12	-13.25	-80.81	7.87	-1.76	6.44	2.92
13	-13.23	-81.47	8.51	-1.44	6.76	3.27
14	-12.73	-75.49	5.45	-3.60	4.64	1.28
15	-12.92	-79.73	5.95	-3.96	4.91	1.18
16	-14.35	-81.93	6.15	-3.57	5.22	1.74
17	-13.66	-81.30	6.95	-2.70	5.81	2.19
18	-13.69	-81.88	7.62	-2.26	6.29	2.54
19	-22.65	-89.39	3.52	-6.08	3.73	1.07
20	-22.03	-88.88	4.32	-5.21	4.32	1.53
21	-22.02	-89.78	5.01	-4.91	4.81	1.87
22	-21.57	-89.57	5.72	-4.23	5.25	2.42
23	-21.45	-89.76	6.31	-3.59	5.26	3.00
24	-26.87	-89.56	5.77	-2.82	5.40	3.94

Table 5.2: Summary of the computational values logP, logD, and ΔG_{solv} $[kcal.mol^{-1}]$ of the studied compounds. The negative values of ΔG_{solv} signify the free energy gained and the positive values signify free energy required, during the transfer from the first phase to the second phase.

5.3.3 Conclusion

In this study, we examined a library of compounds **2−24**, bearing a C-3 hemiester moiety, Δ5- double bond, and various modifications at position C-17, in order to evaluate their ability to modulate the activity of NMDARs. The results of our experiments indicate that the C-17 substituent of the D-ring can be structurally modified or fully degraded while maintaining a positive modulatory effect of the steroid. We have also shown that the potentiating effect of these compounds exhibits a dependency on the length of the C-3 substituent for each D-ring modification. The most efficacious and potent modulators, respectively, from all tested compounds (**2−24**) were compound **15**, exhibiting the E_{max} value of 452%, and compound **18**, exhibiting the EC_{50} value of 1.8 μM. In addition, we

have shown that the selected compound **6** has a subunit-independent effect at recombinant NMDARs, which is similar to that at native NMDARs and has only a minor inhibitory effect at AMPARs. Our data further indicate that compound **6** is an inhibitor of native GABAARs.

The structure-activity relationship study including computational evaluation of physicochemical properties revealed interesting dependence between structure and modulatory effect. In summary, we conclude that PES analogues modified at the C-3 and/or at the D-ring offer new prospects for further optimization of pharmacological and pharmacokinetic properties of these neuroactive compounds.

5.4 A New Class of Potent N-Methyl-D-Aspartate Receptor Inhibitors: Sulfated Neuroactive Steroids with Lipophilic D-Ring Modifications

5.4.1 Introduction

20-Oxo-5β-pregnan-3α-yl sulfate (pregnanolone sulfate; 3α5βS) is an endogenous neurosteroid that inhibits responses of NMDARs.[161] Similarly to the open channel blockers, 3α5βS acts in a use-dependent manner (requiring receptor activation by agonists), but unlike them, its effect is voltage-independent.[161-162] Kinetic properties of steroid binding and inhibition are slow and not typical of a simple receptor−ligand interaction in an aqueous solution.[150] This indirectly suggests the importance of the plasma membrane as a compartment where the steroid accumulates to reach its binding site on NMDARs. The inhibitory effect of neurosteroids on NMDARs is dependent upon the bent steroid ring structure that is associated with a 5β-stereochemistry, whereas the more planar arrangement of 5α-pregnanes favors potentiation.[149] In addition, the stereochemistry at position C-3 is also crucial to the inhibitory effect. As such, steroids which have the combined 3α5β configuration were demonstrated as potent inhibitors of NMDA-induced currents.[149, 162] To elucidate the effect of charge and chain length on modulatory action, a series of sulfate, hemiesters, permethylated amines, and carboxylic acids were synthesized (it has been established that C-3 substituent must bear charge to maintain the biological activity, while uncharged derivatives have no significant effect on NMDA modulation).[149-150] Our previous study on the structure−activity relationship of 3α5βS analogues also demonstrated that analogues that have various substituents at the C-7 carbon may participate in steroid binding on the NMDA receptor[166] as well. The following paper describes the SAR study of

pregnanolone sulfate analogues with structural modifications on the steroidal D-ring that has not been previously reported. To elucidate the structure−activity relationship of the acetyl moiety of the pregnane skeleton and the relevance of the 20-oxo group, a series of nonpolar 3α5βS analogues were synthesized and their biological activity was evaluated on human embryonic kidney cells (HEK293) transfected with plasmids encoding GluN1-a/GluN2B/GFP genes. In addition, we have evaluated the action on recombinant (GluN1/Glu2A-D), native NMDA receptors expressed in hippocampal neurons, native AMPA/kainate, and native GABA receptors. A detailed understanding of the structure−activity relationships between neurosteroids and NMDA receptors is important for the development of drugs with potential therapeutic use. It has already been shown that a synthetic analogue of 3α5βS pregnanolone hemisuccinate ester has neuroprotective activity in both in vitro and in vivo models of neurodegeneration.

5.4.2 Results

The computational results of the lipophilic qualities of the studied compounds are summarized in the Table 5.3. The calculated data correlate well with the experimental ΔG_{exp} data. The correlation between ΔG_{exp} and the values estimated by the accurate QM computational methods (ΔG_{solv}, logP) is higher than between the values obtained by the physicochemical properties predictor (logP, logD), which is in agreement with the theory. The difference in the accuracy of both methods is plotted for logP values in Figure 5.3, where is $R^2 = 0.92$ for QM computations and $R^2 = 0.79$ for predictor. Compound **12**, the inhibitor with iodine moiety at C-17, is the outlier among the values, which were computed by QM methods and therefore was not included in the correlation. The description of iodine by QM computational methods is nontrivial and less reliable. As the results show, the lipophilic character of this iodine compound is significantly underestimated, its logP is even less than logP of 3α5βS, while it is a more potent inhibitor. The high correlation between ΔG_{exp} and ΔG_{solv} (in the case of water/n-octanol phase $R^2 = 0.94$ for neutral systems and $R^2 = 0.92$ for charged systems; in the case of vacuum/water phase $R^2 = 0.71$ for neutral systems and $R^2 = 0.74$ for charged systems) indicate that the (de)solvation free energy plays an important part in an inhibitory effect. The similar results for log P ($R^2 = 0.92$) and log D ($R^2 = 0.78$) show that the inhibition activity of this group of neuroactive steroids closely relates with their lipophilicity. These results are in accordance with expectations and can be used as valid parameters for further structural predictions: comparison of logP values (QM computation) vary from 3.62 to 5.83 for this structural group. However, the inhibitory

effect of neurosteroid is multiparameter nature and, therefore, for the purpose of design of new compounds, especially drug candidate, other aspects should be considered, e.g., enthalpy, entropy, solubility, permeability data, Lipinski rule of five, etc.

| compound | $IC_{50}[\mu M]$ | $\Delta G_{exp}[kcal.mol^{-1}]$ | $\Delta G_{solv}[kcal.mol^{-1}]$: transfer from | | | | logP | logD |
| | | | vacuum to water | | n-octanol to water | | | |
			neutral	charged	neutral	charged		
3	0.40	-8.77	-13.49	-68.48	4.46	-3.02	4.94	2.84
5	0.50	-8.63	-14.06	-68.71	4.23	-3.20	4.77	2.78
7	0.16	-9.31	-13.04	-67.98	4.87	-2.68	5.26	3.13
12	0.60	-8.52	-18.08	-71.33	1.80	-5.49	2.51	2.75
15	1.20	-8.11	-14.00	-68.80	3.53	-3.92	4.07	2.11
17	1.60	-7.94	-14.59	-69.40	3.52	-3.98	3.88	2.05
19	0.60	-8.52	-13.56	-68.48	4.00	-3.52	4.46	2.40
21	0.80	-8.35	-14.56	-69.54	4.10	-3.37	4.65	2.44
25	7.00	-7.06	-17.41	-71.52	3.06	-4.42	3.70	1.98
30	1.10	-8.16	-13.97	-68.92	3.96	-3.49	4.46	2.56
36a	0.90	-8.28	-13.86	-68.85	3.76	-3.77	4.19	2.10
36b	0.70	-8.43	-13.65	-68.53	3.93	-3.57	4.38	2.10
43	2.20	-7.75	-14.73	-69.05	3.53	-3.79	4.04	2.05
45	N/A	N/A	-15.05	-69.73	3.10	-4.25	3.62	1.75
48	5.40	-7.22	-14.24	-69.45	3.43	-3.95	3.90	1.81
53	0.08	-9.69	-12.91	-67.69	5.42	-2.13	5.83	3.58

Table 5.3: Summary of the computational values logP, logD, and ΔG_{solv} [$kcal.mol^{-1}$] of the studied compounds. The negative values of ΔG_{solv} signify the free energy gained, and the positive values signify free energy required, during the transfer from the first phase to the second phase.
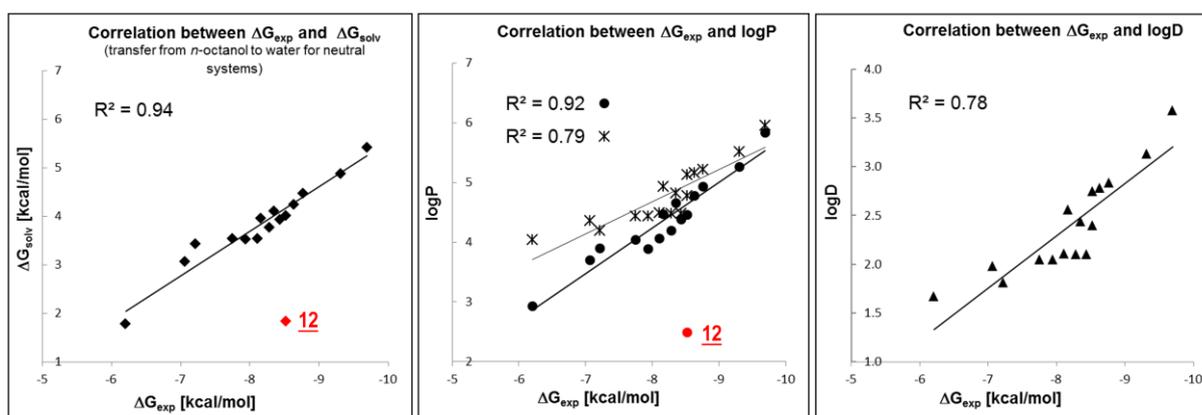
Figure 5.3: Correlation between the experimental ΔG_{exp} values and calculated logP, logD and ΔG_{solv} values for the synthetized neuroactive steroids (3, 5, 7, 15, 17, 19, 21, 25, 30, 36a, 36b, 43, 48 and 53).

5.4.3 Conclusion

In this study, we were able to show that the pregnane acetyl group can be substituted with a variety of nonpolar substituents while still maintaining biological activity. In turn, this discovery has led to the development of NMDA receptor inhibitors that are even more potent than the endogeneous ligand pregnanolone sulfate. In addition, we have shown that nonpolar substitution did not lead to NMDA subtype selectivity and as well as that there is higher potency to inhibit NMDA receptor responses than AMPA responses. Our computational data also revealed a correlation between the lipophilicity and the IC_{50} values, even throughout the relatively broad degree of structural variations. This supports our theory that lipophilic modification, and thus conveniently logP, could offer some guidance in the design of new neurosteroid inhibitors. The finding that the increase in the steroid potency was correlated with the lipophilicity also highlights the importance of the plasmatic membrane as a route of the steroid access to the receptor.[150, 157] The results of our previous experiments show that steroids have behavioral effects associated with altered brain function (neuroprotective effect)[167] and importantly, do not induce psychotomimetic symptoms, which can be the result of the combined effect of the steroid on NMDA, AMPA, and GABA receptors. However, the impact of the steroid lipophilicity to cross the blood−brain barrier has to be established. Furthermore, this structural flexibility offers new prospects for the further modification and optimization of the pharmacological and pharmacokinetic properties of these new neuroactive steroids. On the other hand, this data revealed a new avenue of investigation, steroid lipophilicity, and how it correlates with plasma membrane interaction. For instance, is the lower IC_{50} value caused by the steroid's high affinity toward the NMDA receptor or by its high concentration near the receptor?[168] Finally, as the sulfate moiety at position C-3 is susceptible to hydrolysis by sulfatases, the synthesis of a hemiester or carboxylic acid derivative with a nonpolar substituent on the steroidal D-ring may offer a more metabolically stable compound, one which could penetrate into the CNS without succumbing to metabolic degradation.

Summary

The aim of the thesis has consisted of the computational/theoretical analyses of the interactions between the small molecules and biomolecules. Both the cyclin-dependent kinase 2 enzyme and the *N*-methyl-D-aspartate receptor play a critical role in human biology, thus, they have become the valid pharmaceutical target. The presented computations complete and support the results of the experimental projects based on medicinal chemistry and cellular neurophysiology.

I described and elucidated the binding motifs and interactions between CDK2 and roscovitine and its analogs containing 13 different bioisosteric central heterocycles. The results were successfully utilized in the preparation of new pyrazolo[4,3-d]pyrimidines. Their biological activity was demonstrated by biochemical assays and crystallographic analysis. Our findings confirmed the in vitro and in vivo sensitivity of aggressive non-NHLs to CDK inhibitors and provide a rationale for their future clinical evaluation.

My computational analyses of the physicochemical properties of neuroactive compounds investigated the lipophilic qualities (logP and logD) and solvation free energy since these properties are inherent characteristics of the substances and influence their interactions with NMDAR and membrane. Our results confirm that the lipophilicity of neuroactive compounds has a strong influence on their interactions with NMDAR and membrane, but it is only one of several important characteristics. However, all the studied groups of neuroactive molecules have demonstrated the desired activity on NMDAR.

Bibliography

1.   WHO, *Antibiotic resistance: multi-country public awareness survey.* . 10665/26724 ed.; WHO Press, World Health Organization: Geneva, Switzerlan, 2015.

2.   Van Puyvelde, S.; Deborggraeve, S.; Jacobs, J., Why the antibiotic resistance crisis requires a One Health approach. *Lancet Infect Dis* **2018,** *18* (2), 132-134.

3.   Neu, H. C., The crisis in antibiotic resistance. *Science* **1992,** *257* (5073), 1064-73.

4.   Zhang, Y.-B.; Pan, X.-F.; Chen, J.; Cao, A.; Zhang, Y.-G.; Xia, L.; Wang, J.; Li, H.; Liu, G.; Pan, A., Combined lifestyle factors, incident cancer, and cancer mortality: a systematic review and meta-analysis of prospective cohort studies. *British Journal of Cancer* **2020**.

5.   Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R. L.; Torre, L. A.; Jemal, A., Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians* **2018,** *68* (6), 394-424.

6.   Siegel, R.; DeSantis, C.; Virgo, K.; Stein, K.; Mariotto, A.; Smith, T.; Cooper, D.; Gansler, T.; Lerro, C.; Fedewa, S.; Lin, C.; Leach, C.; Cannady, R. S.; Cho, H.; Scoppa, S.; Hachey, M.; Kirch, R.; Jemal, A.; Ward, E., Cancer treatment and survivorship statistics, 2012. *CA: A Cancer Journal for Clinicians* **2012,** *62* (4), 220-241.

7.   Global Cancer Observatory. (http://gco.iarc.fr).

8.   Sawyers, C., Targeted cancer therapy. *Nature* **2004,** *432* (7015), 294-297.

9.   Ames, D.; O'Brien, J. T.; Burns, A., *Dementia*. 5th ed.; Taylor & Francis Inc: 2017.

10.   Alzheimer's, A., 2017 Alzheimer's disease facts and figures. *Alzheimers Dement* **2017,** *13* (4).

11.   Pina, A. S.; Hussain, A.; Roque, A. C., An historical overview of drug discovery. *Methods Mol Biol* **2009,** *572*, 3-12.

12.   Dias, D. A.; Urban, S.; Roessner, U., A historical overview of natural products in drug discovery. *Metabolites* **2012,** *2* (2), 303-36.

13.   Warren, J., Drug discovery: lessons from evolution. *Br J Clin Pharmacol* **2011,** *71* (4), 497-503.

14.   Takenaka, T., Classical vs reverse pharmacology in drug discovery. *BJU Int* **2001,** *88 Suppl 2*, 7-10; discussion 49-50.

15.   Lee, J. A.; Uhlik, M. T.; Moxham, C. M.; Tomandl, D.; Sall, D. J., Modern phenotypic drug discovery is a viable, neoclassic pharma strategy. *J Med Chem* **2012,** *55* (10), 4527-38.

16.   Swinney, D. C.; Anthony, J., How were new medicines discovered? *Nat Rev Drug Discov* **2011,** *10* (7), 507-519.

17.   Keseru, G. M.; Makara, G. M., Hit discovery and hit-to-lead approaches. *Drug Discov Today* **2006,** *11* (15-16), 741-8.

18.   Ng, R., *Drugs: From Discovery to Approval*. 3rd Edition ed.; Wiley-Blackwell: 2015; p 560.

19.   Lodish, H.; Berk, A.; Zipursky, S. L.; Matsudaira, P.; Baltimore, D.; Darnell, J., Noncovalent Bonds. In *Molecular Cell Biology*, 4th edition ed.; New York: W. H. Freeman: 2000.

20.   Piela, L., Intermolecular interactions. In *Ideas of Quantum Chemistry*, Elsevier Science 2007; pp 681-761.

21.   Tropsha, A.; Bajorath, J., Computational Methods for Drug Discovery and Design. *J Med Chem* **2016,** *59* (1), 1.

22.   Sliwoski, G.; Kothiwale, S.; Meiler, J.; Lowe, E. W., Jr., Computational methods in drug discovery. *Pharmacol Rev* **2014,** *66* (1), 334-95.

23.   Materi, W.; Wishart, D. S., Computational systems biology in drug discovery and development: methods and applications. *Drug Discov Today* **2007,** *12* (7-8), 295-303.

24.   Malumbres, M., Cyclin-dependent kinases. *Genome Biol* **2014,** *15* (6), 122.

25. Tsai, L. H.; Harlow, E.; Meyerson, M., Isolation of the human cdk2 gene that encodes the cyclin A- and adenovirus E1A-associated p33 kinase. *Nature* **1991,** *353* (6340), 174-7.

26. Morgan, D. O., Principles of CDK regulation. *Nature* **1995,** *374* (6518), 131-4.

27. De Azevedo, W. F.; Leclerc, S.; Meijer, L.; Havlicek, L.; Strnad, M.; Kim, S. H., Inhibition of cyclin-dependent kinases by purine analogues: crystal structure of human cdk2 complexed with roscovitine. *Eur J Biochem* **1997,** *243* (1-2), 518-26.

28. Nekardova, M.; Vymetalova, L.; Khirsariya, P.; Kovacova, S.; Hylsova, M.; Jorda, R.; Krystof, V.; Fanfrlik, J.; Hobza, P.; Paruch, K., Structural Basis of the Interaction of Cyclin-Dependent Kinase 2 with Roscovitine and Its Analogues Having Bioisosteric Central Heterocycles. *Chemphyschem* **2017,** *18* (7), 785-795.

29. Fanfrlik, J.; Bronowska, A. K.; Rezac, J.; Prenosil, O.; Konvalinka, J.; Hobza, P., A reliable docking/scoring scheme based on the semiempirical quantum mechanical PM6-DH2 method accurately covering dispersion and H-bonding: HIV-1 protease with 22 ligands. *J Phys Chem B* **2010,** *114* (39), 12666-78.

30. Stewart, J. J. P., Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *J Mol Model* **2007,** *13* (12), 1173-1213.

31. Klamt, A.; Schuurmann, G., Cosmo - a New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient. *J Chem Soc Perk T 2* **1993,** (5), 799-805.

32. Jorda, R.; Havlicek, L.; Sturc, A.; Tuskova, D.; Daumova, L.; Alam, M.; Skerlova, J.; Nekardova, M.; Perina, M.; Pospisil, T.; Siroka, J.; Urbanek, L.; Pachl, P.; Rezacova, P.; Strnad, M.; Klener, P.; Krystof, V., 3,5,7-Substituted Pyrazolo[4,3- d]pyrimidine Inhibitors of Cyclin-Dependent Kinases and Their Evaluation in Lymphoma Models. *J Med Chem* **2019,** *62* (9), 4606-4623.

33. Strong, K. L.; Epplin, M. P.; Bacsa, J.; Butch, C. J.; Burger, P. B.; Menaldino, D. S.; Traynelis, S. F.; Liotta, D. C., The Structure-Activity Relationship of a Tetrahydroisoquinoline Class of N-Methyl-d-Aspartate Receptor Modulators that Potentiates GluN2B-Containing N-Methyl-d-Aspartate Receptors. *J Med Chem* **2017,** *60* (13), 5556-5585.

34. Traynelis, S. F.; Wollmuth, L. P.; McBain, C. J.; Menniti, F. S.; Vance, K. M.; Ogden, K. K.; Hansen, K. B.; Yuan, H.; Myers, S. J.; Dingledine, R., Glutamate receptor ion channels: structure, regulation, and function. *Pharmacol Rev* **2010,** *62* (3), 405-96.

35. Collingridge, G., Synaptic plasticity. The role of NMDA receptors in learning and memory. *Nature* **1987,** *330* (6149), 604-5.

36. Chen, H. S.; Lipton, S. A., The chemical biology of clinically tolerated NMDA receptor antagonists. *J Neurochem* **2006,** *97* (6), 1611-26.

37. Kemp, J. A.; McKernan, R. M., NMDA receptor pathways as drug targets. *Nat Neurosci* **2002,** *5 Suppl*, 1039-42.

38. Faassen, F.; Kelder, J.; Lenders, J.; Onderwater, R.; Vromans, H., Physicochemical properties and transport of steroids across Caco-2 cells. *Pharm Res* **2003,** *20* (2), 177-86.

39. Barton, P.; Davis, A. M.; McCarthy, D. J.; Webborn, P. J., Drug-phospholipid interactions. 2. Predicting the sites of drug distribution using n-octanol/water and membrane/water distribution coefficients. *J Pharm Sci* **1997,** *86* (9), 1034-9.

40. Marenich, A. V.; Cramer, C. J.; Truhlar, D. G., Universal Solvation Model Based on the Generalized Born Approximation with Asymmetric Descreening. *J Chem Theory Comput* **2009,** *5* (9), 2447-64.

41. Mannhold, R.; Poda, G. I.; Ostermann, C.; Tetko, I. V., Calculation of molecular lipophilicity: State-of-the-art and comparison of log P methods on more than 96,000 compounds. *J Pharm Sci* **2009,** *98* (3), 861-93.

42. Kibbey, C. E.; Poole, S. K.; Robinson, B.; Jackson, J. D.; Durham, D., An integrated process for measuring the physicochemical properties of drug candidates in a preclinical discovery environment. *J Pharm Sci* **2001,** *90* (8), 1164-75.

43. Kolar, M.; Fanfrlik, J.; Lepsik, M.; Forti, F.; Luque, F. J.; Hobza, P., Assessing the accuracy and performance of implicit solvent models for drug molecules: conformational ensemble approaches. *J Phys Chem B* **2013,** *117* (19), 5950-62.

44.   Bannan, C. C.; Calabro, G.; Kyu, D. Y.; Mobley, D. L., Calculating Partition Coefficients of Small Molecules in Octanol/Water and Cyclohexane/Water. *J Chem Theory Comput* **2016,** *12* (8), 4015-24.

45.   Krausova, B.; Slavikova, B.; Nekardova, M.; Hubalkova, P.; Vyklicky, V.; Chodounska, H.; Vyklicky, L.; Kudova, E., Positive Modulators of the N-Methyl-d-aspartate Receptor: Structure-Activity Relationship Study of Steroidal 3-Hemiesters. *J Med Chem* **2018,** *61* (10), 4505-4516.

46.   Slavikova, B.; Chodounska, H.; Nekardova, M.; Vyklicky, V.; Ladislav, M.; Hubalkova, P.; Krausova, B.; Vyklicky, L.; Kudova, E., Neurosteroid-like Inhibitors of N-Methyl-d-aspartate Receptor: Substituted 2-Sulfates and 2-Hemisuccinates of Perhydrophenanthrene. *J Med Chem* **2016**.

47.   Kudova, E.; Chodounska, H.; Slavikova, B.; Budesinsky, M.; Nekardova, M.; Vyklicky, V.; Krausova, B.; Svehla, P.; Vyklicky, L., A New Class of Potent N-Methyl-D-Aspartate Receptor Inhibitors: Sulfated Neuroactive Steroids with Lipophilic D-Ring Modifications. *J Med Chem* **2015,** *58* (15), 5950-66.

48.   Zhou, H. X.; Pang, X., Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation. *Chem Rev* **2018,** *118* (4), 1691-1741.

49.   Chapter 13 - INTERMOLECULAR INTERACTIONS. In *Ideas of Quantum Chemistry*, Piela, L., Ed. Elsevier: Amsterdam, 2007; pp 681-761.

50.   Hobza, P.; Muller-Dethlefs, K., *Non-Covalent Interactions: Theory and Experiment* 2009.

51.   Hobza, P.; Havlas, Z., Blue-Shifting Hydrogen Bonds. *Chem Rev* **2000,** *100* (11), 4253-4264.

52.   Kolar, M. H.; Hobza, P., Computer Modeling of Halogen Bonds and Other sigma-Hole Interactions. *Chem Rev* **2016,** *116* (9), 5155-87.

53.   Riley, K. E.; Hobza, P., On the importance and origin of aromatic interactions in chemistry and biodisciplines. *Acc Chem Res* **2013,** *46* (4), 927-36.

54.   Cramer, C. J., *Essentials of Computational Chemistry: Theories and Models*. 2nd ed.; 2004.

55.   Szabo, A.; Ostlund, N. S., *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory.* Dover: 1996.

56.   Levine, I. N., *Quantum Chemistry* Prentice Hall: 2008.

57.   Chapter 2 - THE SCHRÖDINGER EQUATION. In *Ideas of Quantum Chemistry*, Piela, L., Ed. Elsevier: Amsterdam, 2007; pp 55-89.

58.   Rezac, J.; Hobza, P., Advanced Corrections of Hydrogen Bonding and Dispersion for Semiempirical Quantum Mechanical Methods. *J Chem Theory Comput* **2012,** *8* (1), 141-151.

59.   Stewart, J. J., Application of the PM6 method to modeling proteins. *J Mol Model* **2009,** *15* (7), 765-805.

60.   Gotta, M.; Meraldi, P., *Cell Division Machinery and Disease*. 2017; Vol. 1002.

61.   Barnum, K. J.; O'Connell, M. J., Cell cycle regulation by checkpoints. *Methods Mol Biol* **2014,** *1170*, 29-40.

62.   Endicott, J. A.; Noble, M. E., Structural principles in cell-cycle control: beyond the CDKs. *Structure* **1998,** *6* (5), 535-41.

63.   Hartwell, L. H.; Kastan, M. B., Cell cycle control and cancer. *Science* **1994,** *266* (5192), 1821.

64.   Otto, T.; Sicinski, P., Cell cycle proteins as promising targets in cancer therapy. *Nat Rev Cancer* **2017,** *17* (2), 93-115.

65.   Suryadinata, R.; Sadowski, M.; Sarcevic, B., Control of cell cycle progression by phosphorylation of cyclin-dependent kinase (CDK) substrates. *Biosci Rep* **2010,** *30* (4), 243-55.

66.   Foundation, N. The Nobel Prize in Physiology or Medicine 2001. https://www.nobelprize.org/prizes/medicine/2001/summary/.

67.   Canavese, M.; Santo, L.; Raje, N., Cyclin dependent kinases in cancer. *Cancer Biology & Therapy* **2012,** *13* (7), 451-457.

68.   Wood, D. J.; Korolchuk, S.; Tatum, N. J.; Wang, L.-Z.; Endicott, J. A.; Noble, M. E. M.; Martin, M. P., Differences in the Conformational Energy Landscape of CDK1 and CDK2 Suggest a Mechanism for Achieving Selective CDK Inhibition. *Cell chemical biology* **2019,** *26* (1), 121-130.e5.

69.   Brown, N. R.; Korolchuk, S.; Martin, M. P.; Stanley, W. A.; Moukhametzianov, R.; Noble, M. E.; Endicott, J. A., CDK1 structures reveal conserved and unique features of the essential cell cycle CDK. *Nat Commun* **2015,** *6*, 6769.

70.   Bettayeb, K.; Oumata, N.; Echalier, A.; Ferandin, Y.; Endicott, J. A.; Galons, H.; Meijer, L., CR8, a potent and selective, roscovitine-derived inhibitor of cyclin-dependent kinases. *Oncogene* **2008,** *27* (44), 5797-807.

71.   Morgan, D. O., *Cell Cycle: Principles of Control*. New Science Press: 2006; p 297.

72.   Deshpande, A.; Sicinski, P.; Hinds, P. W., Cyclins and cdks in development and cancer: a perspective. *Oncogene* **2005,** *24* (17), 2909-2915.

73.   Whittaker, S. R.; Mallinger, A.; Workman, P.; Clarke, P. A., Inhibitors of cyclin-dependent kinases as cancer therapeutics. *Pharmacol Ther* **2017,** *173*, 83-105.

74.   Peyressatre, M.; Prevel, C.; Pellerano, M.; Morris, M. C., Targeting cyclin-dependent kinases in human cancers: from small molecules to Peptide inhibitors. *Cancers (Basel)* **2015,** *7* (1), 179-237.

75.   DiPippo, A. J.; Patel, N. K.; Barnett, C. M., Cyclin-Dependent Kinase Inhibitors for the Treatment of Breast Cancer: Past, Present, and Future. *Pharmacotherapy* **2016,** *36* (6), 652-67.

76.   Asghar, U.; Witkiewicz, A. K.; Turner, N. C.; Knudsen, E. S., The history and future of targeting cyclin-dependent kinases in cancer therapy. *Nature reviews. Drug discovery* **2015,** *14* (2), 130-146.

77.   Dhillon, S., Palbociclib: first global approval. *Drugs* **2015,** *75* (5), 543-51.

78.   Jorda, R.; Hendrychova, D.; Voller, J.; Reznickova, E.; Gucky, T.; Krystof, V., How Selective Are Pharmacological Inhibitors of Cell-Cycle-Regulating Cyclin-Dependent Kinases? *J Med Chem* **2018,** *61* (20), 9105-9120.

79.   Sridhar, J.; Akula, N.; Pattabiraman, N., Selectivity and potency of cyclin-dependent kinase inhibitors. *AAPS J* **2006,** *8* (1), E204-E221.

80.   Senderowicz, A. M., Flavopiridol: the first cyclin-dependent kinase inhibitor in human clinical trials. *Invest New Drugs* **1999,** *17* (3), 313-20.

81.   Shapiro, G. I., Preclinical and clinical development of the cyclin-dependent kinase inhibitor flavopiridol. *Clin Cancer Res* **2004,** *10* (12 Pt 2), 4270s-4275s.

82.   Kumara, P. M.; Soujanya, K. N.; Ravikanth, G.; Vasudeva, R.; Ganeshaiah, K. N.; Shaanker, R. U., Rohitukine, a chromone alkaloid and a precursor of flavopiridol, is produced by endophytic fungi isolated from Dysoxylum binectariferum Hook.f and Amoora rohituka (Roxb).Wight & Arn. *Phytomedicine* **2014,** *21* (4), 541-6.

83.   Havlíček, L.; Hanuš, J.; Veselý, J.; Leclerc, S.; Meijer, L.; Shaw, G.; Strnad, M., Cytokinin-Derived Cyclin-Dependent Kinase Inhibitors: Synthesis and cdc2 Inhibitory Activity of Olomoucine and Related Compounds. *Journal of Medicinal Chemistry* **1997,** *40* (4), 408-412.

84.   Schaller, G. E.; Street, I. H.; Kieber, J. J., Cytokinin and the cell cycle. *Curr Opin Plant Biol* **2014,** *21*, 7-15.

85.   Vermeulen, K.; Strnad, M.; Krystof, V.; Havlicek, L.; Van der Aa, A.; Lenjou, M.; Nijs, G.; Rodrigus, I.; Stockman, B.; van Onckelen, H.; Van Bockstaele, D. R.; Berneman, Z. N., Antiproliferative effect of plant cytokinin analogues with an inhibitory activity on cyclin-dependent kinases. *Leukemia* **2002,** *16* (3), 299-305.

86.   Schulze-Gahmen, U.; Brandsen, J.; Jones, H. D.; Morgan, D. O.; Meijer, L.; Vesely, J.; Kim, S. H., Multiple modes of ligand recognition: crystal structures of cyclin-dependent protein kinase 2 in complex with ATP and two inhibitors, olomoucine and isopentenyladenine. *Proteins* **1995,** *22* (4), 378-91.

87.   Schulze-Gahmen, U.; De Bondt, H. L.; Kim, S. H., High-resolution crystal structures of human cyclin-dependent kinase 2 with and without ATP: bound waters and natural ligand as guides for inhibitor design. *J Med Chem* **1996,** *39* (23), 4540-6.

88.   Abraham, R. T.; Acquarone, M.; Andersen, A.; Asensi, A.; Belle, R.; Berger, F.; Bergounioux, C.; Brunn, G.; Buquet-Fagot, C.; Fagot, D.; et al., Cellular effects of olomoucine, an inhibitor of cyclin-dependent kinases. *Biol Cell* **1995,** *83* (2-3), 105-20.

89.   Krystof, V.; McNae, I. W.; Walkinshaw, M. D.; Fischer, P. M.; Muller, P.; Vojtesek, B.; Orsag, M.; Havlicek, L.; Strnad, M., Antiproliferative activity of olomoucine II, a novel 2,6,9-trisubstituted purine cyclin-dependent kinase inhibitor. *Cell Mol Life Sci* **2005,** *62* (15), 1763-71.

90.   Krystof, V.; Uldrijan, S., Cyclin-dependent kinase inhibitors as anticancer drugs. *Curr Drug Targets* **2010,** *11* (3), 291-302.

91.   Cicenas, J.; Kalyan, K.; Sorokinas, A.; Stankunas, E.; Levy, J.; Meskinyte, I.; Stankevicius, V.; Kaupinis, A.; Valius, M., Roscovitine in cancer and other diseases. *Ann Transl Med* **2015,** *3* (10), 135-135.

92.   Gray, N. S.; Wodicka, L.; Thunnissen, A. M.; Norman, T. C.; Kwon, S.; Espinoza, F. H.; Morgan, D. O.; Barnes, G.; LeClerc, S.; Meijer, L.; Kim, S. H.; Lockhart, D. J.; Schultz, P. G., Exploiting chemical libraries, structure, and genomics in the search for kinase inhibitors. *Science* **1998,** *281* (5376), 533-8.

93.   Zatloukal, M.; Jorda, R.; Gucky, T.; Reznickova, E.; Voller, J.; Pospisil, T.; Malinkova, V.; Adamcova, H.; Krystof, V.; Strnad, M., Synthesis and in vitro biological evaluation of 2,6,9-trisubstituted purines targeting multiple cyclin-dependent kinases. *Eur J Med Chem* **2013,** *61*, 61-72.

94.   Villerbu, N.; Gaben, A. M.; Redeuilh, G.; Mester, J., Cellular effects of purvalanol A: a specific inhibitor of cyclin-dependent kinase activities. *Int J Cancer* **2002,** *97* (6), 761-9.

95.   Bohm, H. J.; Flohr, A.; Stahl, M., Scaffold hopping. *Drug Discov Today Technol* **2004,** *1* (3), 217-24.

96.   Blaquiere, N.; Castanedo, G. M.; Burch, J. D.; Berezhkovskiy, L. M.; Brightbill, H.; Brown, S.; Chan, C.; Chiang, P. C.; Crawford, J. J.; Dong, T.; Fan, P.; Feng, J.; Ghilardi, N.; Godemann, R.; Gogol, E.; Grabbe, A.; Hole, A. J.; Hu, B.; Hymowitz, S. G.; Alaoui Ismaili, M. H.; Le, H.; Lee, P.; Lee, W.; Lin, X.; Liu, N.; McEwan, P. A.; McKenzie, B.; Silvestre, H. L.; Suto, E.; Sujatha-Bhaskar, S.; Wu, G.; Wu, L. C.; Zhang, Y.; Zhong, Z.; Staben, S. T., Scaffold-Hopping Approach To Discover Potent, Selective, and Efficacious Inhibitors of NF-kappaB Inducing Kinase. *J Med Chem* **2018,** *61* (15), 6801-6813.

97.   Bajorath, J., Computational scaffold hopping: cornerstone for the future of drug design? *Future Med Chem* **2017,** *9* (7), 629-631.

98.   Jorda, R.; Paruch, K.; Krystof, V., Cyclin-dependent kinase inhibitors inspired by roscovitine: purine bioisosteres. *Curr Pharm Des* **2012,** *18* (20), 2974-80.

99.   Sun, H.; Tawa, G.; Wallqvist, A., Classification of scaffold-hopping approaches. *Drug Discov Today* **2012,** *17* (7-8), 310-324.

100.   Paruch, K.; Dwyer, M. P.; Alvarez, C.; Brown, C.; Chan, T. Y.; Doll, R. J.; Keertikar, K.; Knutson, C.; McKittrick, B.; Rivera, J.; Rossman, R.; Tucker, G.; Fischmann, T.; Hruza, A.; Madison, V.; Nomeir, A. A.; Wang, Y.; Kirschmeier, P.; Lees, E.; Parry, D.; Sgambellone, N.; Seghezzi, W.; Schultz, L.; Shanahan, F.; Wiswell, D.; Xu, X.; Zhou, Q.; James, R. A.; Paradkar, V. M.; Park, H.; Rokosz, L. R.; Stauffer, T. M.; Guzi, T. J., Discovery of Dinaciclib (SCH 727965): A Potent and Selective Inhibitor of Cyclin-Dependent Kinases. *ACS Med Chem Lett* **2010,** *1* (5), 204-8.

101.   Martin, M. P.; Olesen, S. H.; Georg, G. I.; Schonbrunn, E., Cyclin-Dependent Kinase Inhibitor Dinaciclib Interacts with the Acetyl-Lysine Recognition Site of Bromodomains. *Acs Chem Biol* **2013,** *8* (11), 2360-2365.

102.   Davies, T. G.; Tunnah, P.; Meijer, L.; Marko, D.; Eisenbrand, G.; Endicott, J. A.; Noble, M. E. M., Inhibitor binding to active and inactive CDK2: The crystal structure of CDK2-cyclin A/indirubin-5-sulphonate. *Structure* **2001,** *9* (5), 389-397.

103.   Kontopidis, G.; McInnes, C.; Pandalaneni, S. R.; McNae, I.; Gibson, D.; Mezna, M.; Thomas, M.; Wood, G.; Wang, S.; Walkinshaw, M. D.; Fischer, P. M., Differential binding of inhibitors to active and inactive CDK2 provides insights for drug design. *Chem Biol* **2006,** *13* (2), 201-11.

104.   Case, D. A.; Babin, V.; Berryman, J. T.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Cheatham, I., T. E.; Darden, T. A.; Duke, R. E.; Gohlke, H.; Goetz, A. W.; Gusarov, S.; Homeyer, N.; Janowski, P.; Kaus, J.; Kolossváry, I.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Luchko, T.; Luo, R.; Madej, B.; Merz, K. M.; Paesani, F.; Roe, D. R.; Roitberg, A.; Sagui, C.; Salomon-Ferrer, R.; Seabra, G.; Simmerling, C. L.; Smith, W.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Kollman, P. A., AMBER 14. *University of California, San Francisco* **2014**.

105.   Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF chimera - A visualization system for exploratory research and analysis. *J Comput Chem* **2004,** *25* (13), 1605-1612.

106.   Salomon-Ferrer, R.; Case, D. A.; Walker, R. C., An overview of the Amber biomolecular simulation package. *Wires Comput Mol Sci* **2013,** *3* (2), 198-210.

107.   Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A., A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges - the Resp Model. *J Phys Chem-Us* **1993,** *97* (40), 10269-10280.

108.   The PyMOL Molecular Graphics System, Version 2.0, Schrödinger, LLC.

109.   Ahlrichs, R.; Bar, M.; Haser, M.; Horn, H.; Kolmel, C., Electronic-Structure Calculations on Workstation Computers - the Program System Turbomole. *Chem Phys Lett* **1989,** *162* (3), 165-169.

110.   Jurecka, P.; Cerny, J.; Hobza, P.; Salahub, D. R., Density functional theory augmented with an empirical dispersion term. Interaction energies and geometries of 80 noncovalent complexes compared with ab initio quantum mechanics calculations. *J Comput Chem* **2007,** *28* (2), 555-569.

111.   Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J., AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J Comput Chem* **2009,** *30* (16), 2785-2791.

112.   Trott, O.; Olson, A. J., Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J Comput Chem* **2010,** *31* (2), 455-461.

113.   Sumner, S.; Soderhjelm, P.; Ryde, U., Effect of Geometry Optimizations on QM-Cluster and QM/MM Studies of Reaction Energies in Proteins. *J Chem Theory Comput* **2013,** *9* (9), 4205-4214.

114.   Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K., ONIOM: A multilayered integrated MO+MM method for geometry optimizations and single point energy predictions. A test for Diels-Alder reactions and Pt(P(t-Bu)(3))(2)+H-2 oxidative addition. *J Phys Chem-Us* **1996,** *100* (50), 19357-19363.

115.   Dapprich, S.; Komaromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J., A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives. *J Mol Struc-Theochem* **1999,** *461*, 1-21.

116.   Rezac, J.; Fanfrlik, J.; Salahub, D.; Hobza, P., Semiempirical Quantum Chemical PM6 Method Augmented by Dispersion and H-Bonding Correction Terms Reliably Describes Various Types of Noncovalent Complexes. *J Chem Theory Comput* **2009,** *5* (7), 1749-1760.

117.   Rezac, J.; Hobza, P., A halogen-bonding correction for the semiempirical PM6 method. *Chem Phys Lett* **2011,** *506* (4-6), 286-289.

118.   Wick, C. R.; Hennemann, M.; Stewart, J. J.; Clark, T., Self-consistent field convergence for proteins: a comparison of full and localized-molecular-orbital schemes. *J Mol Model* **2014,** *20* (3), 2159.

119.   Rezac, J., Cuby: An integrative framework for computational chemistry. *J Comput Chem* **2016,** *37* (13), 1230-7.

120.   Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and testing of a general amber force field. *J Comput Chem* **2004,** *25* (9), 1157-74.

121.   Li, H.; Jensen, J. H., Partial Hessian vibrational analysis: the localization of the molecular vibrational energy and entropy. *Theor Chem Acc* **2002,** *107* (4), 211-219.

122.   Pearlman, D. A.; Charifson, P. S., Are free energy calculations useful in practice? A comparison with rapid scoring functions for the p38 MAP kinase protein system. *Journal of Medicinal Chemistry* **2001,** *44* (21), 3417-3423.

123.   Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J., Gaussian 09, Revision E.01. **2009.**

124.   Singh, U. C.; Kollman, P. A., An Approach to Computing Electrostatic Charges for Molecules. *J Comput Chem* **1984,** *5* (2), 129-145.

125.   Portmann, S.; Luthi, H. P., MOLEKEL: An interactive molecular graphics tool. *Chimia* **2000,** *54* (12), 766-770.

126.   Cerny, J.; Hobza, P., Non-covalent interactions in biomacromolecules. *Phys Chem Chem Phys* **2007,** *9* (39), 5291-303.

127.   Muller-Dethlefs, K.; Hobza, P., Noncovalent interactions: a challenge for experiment and theory. *Chem Rev* **2000,** *100* (1), 143-68.

128.   Wilcken, R.; Zimmermann, M. O.; Lange, A.; Joerger, A. C.; Boeckler, F. M., Principles and applications of halogen bonding in medicinal chemistry and chemical biology. *J Med Chem* **2013,** *56* (4), 1363-88.

129.   Politzer, P.; Murray, J. S.; Clark, T., Halogen bonding: an electrostatically-driven highly directional noncovalent interaction. *Phys Chem Chem Phys* **2010,** *12* (28), 7748-57.

130.   Sedlak, R.; Janowski, T.; Pitonak, M.; Rezac, J.; Pulay, P.; Hobza, P., Accuracy of Quantum Chemical Methods for Large Noncovalent Complexes. *J Chem Theory Comput* **2013,** *9* (8), 3364-3374.

131.   Zhang, B.; Tan, V. B. C.; Lim, K. M.; Tay, T. E.; Zhuang, S. L., Study of the inhibition of cyclin-dependent kinases with roscovitine and indirubin-3 '-oxime from molecular dynamics simulations. *J Mol Model* **2007,** *13* (1), 79-89.

132.   Dobes, P.; Fanfrlik, J.; Rezac, J.; Otyepka, M.; Hobza, P., Transferable scoring function based on semiempirical quantum mechanical PM6-DH2 method: CDK2 with 15 structurally diverse inhibitors. *J Comput Aided Mol Des* **2011,** *25* (3), 223-35.

133.   Lepsik, M.; Rezac, J.; Kolar, M.; Pecina, A.; Hobza, P.; Fanfrlik, J., The Semiempirical Quantum Mechanical Scoring Function for In Silico Drug Design. *Chempluschem* **2013,** *78* (9), 921-931.

134.   Rosenthal, A.; Younes, A., High grade B-cell lymphoma with rearrangements of MYC and BCL2 and/or BCL6: Double hit and triple hit lymphomas and double expressing lymphoma. *Blood Rev* **2017,** *31* (2), 37-42.

135.   Jorda, R.; Havlicek, L.; McNae, I. W.; Walkinshaw, M. D.; Voller, J.; Sturc, A.; Navratilova, J.; Kuzma, M.; Mistrik, M.; Bartek, J.; Strnad, M.; Krystof, V., Pyrazolo[4,3-d]pyrimidine Bioisostere of Roscovitine: Evaluation of a Novel Selective Inhibitor of Cyclin-Dependent Kinases with Antiproliferative Activity. *Journal of Medicinal Chemistry* **2011,** *54* (8), 2980-2993.

136.   Vymetalova, L.; Havlicek, L.; Sturc, A.; Skraskova, Z.; Jorda, R.; Pospisil, T.; Strnad, M.; Krystof, V., 5-Substituted 3-isopropyl-7-[4-(2-pyridyl)benzyl]amino-1(2)H-pyrazolo[4,3-d]pyrimidines with anti-proliferative activity as potent and selective inhibitors of cyclin-dependent kinases. *Eur J Med Chem* **2016,** *110*, 291-301.

137.   Schiebel, J.; Gaspari, R.; Wulsdorf, T.; Ngo, K.; Sohn, C.; Schrader, T. E.; Cavalli, A.; Ostermann, A.; Heine, A.; Klebe, G., Intriguing role of water in protein-ligand binding studied by neutron crystallography on trypsin complexes. *Nat Commun* **2018,** *9* (1), 3559.

138.   Karakas, E.; Furukawa, H., Crystal structure of a heterotetrameric NMDA receptor ion channel. *Science* **2014,** *344* (6187), 992-7.

139.   Li, F.; Tsien, J. Z., Memory and the NMDA receptors. *N Engl J Med* **2009,** *361* (3), 302-303.

140.   Furukawa, H.; Singh, S. K.; Mancusso, R.; Gouaux, E., Subunit arrangement and function in NMDA receptors. *Nature* **2005,** *438* (7065), 185-192.

141.   Dingledine, R.; Borges, K.; Bowie, D.; Traynelis, S. F., The glutamate receptor ion channels. *Pharmacol Rev* **1999,** *51* (1), 7-61.

142.   Liu, Y.; Zhang, J., Recent development in NMDA receptors. *Chinese medical journal* **2000,** *113* (10), 948-56.

143.   Cull-Candy, S.; Brickley, S.; Farrant, M., NMDA receptor subunits: diversity, development and disease. *Current opinion in neurobiology* **2001,** *11* (3), 327-35.

144.   Paoletti, P.; Neyton, J., NMDA receptor subunits: function and pharmacology. *Curr Opin Pharmacol* **2007,** *7* (1), 39-47.

145.   Birnbaum, J. H.; Bali, J.; Rajendran, L.; Nitsch, R. M.; Tackenberg, C., Calcium flux-independent NMDA receptor activity is required for Aβ oligomer-induced synaptic loss. *Cell Death & Disease* **2015,** *6* (6), e1791-e1791.

146.   Hynd, M. R.; Scott, H. L.; Dodd, P. R., Glutamate-mediated excitotoxicity and neurodegeneration in Alzheimer's disease. *Neurochem Int* **2004,** *45* (5), 583-95.

147.   Dong, X. X.; Wang, Y.; Qin, Z. H., Molecular mechanisms of excitotoxicity and their relevance to pathogenesis of neurodegenerative diseases. *Acta Pharmacol Sin* **2009,** *30* (4), 379-387.

148.   Lai, T. W.; Zhang, S.; Wang, Y. T., Excitotoxicity and stroke: Identifying novel targets for neuroprotection. *Prog Neurobiol* **2014,** *115*, 157-188.

149.   Weaver, C. E.; Land, M. B.; Purdy, R. H.; Richards, K. G.; Gibbs, T. T.; Farb, D. H., Geometry and charge determine pharmacological effects of steroids on N-methyl-D-aspartate receptor-induced Ca(2+) accumulation and cell death. *The Journal of pharmacology and experimental therapeutics* **2000,** *293* (3), 747-54.

150.   Borovska, J.; Vyklicky, V.; Stastna, E.; Kapras, V.; Slavikova, B.; Horak, M.; Chodounska, H.; Vyklicky, L., Jr., Access of inhibitory neurosteroids to the NMDA receptor. *Br J Pharmacol* **2012,** *166* (3), 1069-83.

151.   Weaver, C. E., Jr.; Marek, P.; Park-Chung, M.; Tam, S. W.; Farb, D. H., Neuroprotective activity of a new class of steroidal inhibitors of the N-methyl-D-aspartate receptor. *Proc Natl Acad Sci U S A* **1997,** *94* (19), 10450-4.

152.   Lapchak, P. A., The neuroactive steroid 3-alpha-ol-5-beta-pregnan-20-one hemisuccinate, a selective NMDA receptor antagonist improves behavioral performance following spinal cord ischemia. *Brain Res* **2004,** *997* (2), 152-8.

153.   Faucher, F.; Cantin, L.; Luu-The, V.; Labrie, F.; Breton, R., The crystal structure of human Delta4-3-ketosteroid 5beta-reductase defines the functional role of the residues of the catalytic tetrad in the steroid double bond reduction mechanism. *Biochemistry* **2008,** *47* (32), 8261-70.

154.   Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H., A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J Chem Phys* **2010,** *132* (15), 154104.

155.   Kah, M.; Brown, C. D., LogD: lipophilicity for ionisable compounds. *Chemosphere* **2008,** *72* (10), 1401-8.

156.   ChemAxon, Marvin was used for drawing, displaying and characterizing chemical structures, substructures and reactions, Marvin 15.1.19. *(http://www.chemaxon.com)* **2015.**

157.   Vyklicky, V.; Krausova, B.; Cerny, J.; Balik, A.; Zapotocky, M.; Novotny, M.; Lichnerova, K.; Smejkalova, T.; Kaniakova, M.; Korinek, M.; Petrovic, M.; Kacer, P.; Horak, M.; Chodounska, H.; Vyklicky, L., Block of NMDA receptor channels by endogenous neurosteroids: implications for the agonist induced conformational states of the channel vestibule. *Sci Rep* **2015,** *5*, 10935.

158.   Mota, S. I.; Ferreira, I. L.; Rego, A. C., Dysfunctional synapse in Alzheimer's disease - A focus on NMDA receptors. *Neuropharmacology* **2014,** *76 Pt A*, 16-26.

159.   Rotaru, D. C.; Yoshino, H.; Lewis, D. A.; Ermentrout, G. B.; Gonzalez-Burgos, G., Glutamate receptor subtypes mediating synaptic activation of prefrontal cortex neurons: relevance for schizophrenia. *J Neurosci* **2011,** *31* (1), 142-56.

160.   Reddy, D. S., Neurosteroids: endogenous role in the human brain and therapeutic potentials. *Prog Brain Res* **2010,** *186*, 113-37.

161.   Petrovic, M.; Sedlacek, M.; Horak, M.; Chodounska, H.; Vyklicky, L., Jr., 20-oxo-5beta-pregnan-3alpha-yl sulfate is a use-dependent NMDA receptor inhibitor. *J Neurosci* **2005,** *25* (37), 8439-50.

162.   Park-Chung, M.; Wu, F. S.; Farb, D. H., 3 alpha-Hydroxy-5 beta-pregnan-20-one sulfate: a negative modulator of the NMDA-induced current in cultured neurons. *Mol Pharmacol* **1994,** *46* (1), 146-50.

163.   Wu, F. S.; Gibbs, T. T.; Farb, D. H., Pregnenolone sulfate: a positive allosteric modulator at the N-methyl-D-aspartate receptor. *Mol Pharmacol* **1991,** *40* (3), 333-6.

164.   Majewska, M. D.; Mienville, J. M.; Vicini, S., Neurosteroid pregnenolone sulfate antagonizes electrophysiological responses to GABA in neurons. *Neuroscience letters* **1988,** *90* (3), 279-84.

165.   Wu, F. S.; Chen, S. C.; Tsai, J. J., Competitive inhibition of the glycine-induced current by pregnenolone sulfate in cultured chick spinal cord neurons. *Brain Res* **1997,** *750* (1-2), 318-20.

166.   Stastna, E.; Chodounska, H.; Pouzar, V.; Kapras, V.; Borovska, J.; Cais, O.; Vyklicky, L., Jr., Synthesis of C3, C5, and C7 pregnane derivatives and their effect on NMDA receptor responses in cultured rat hippocampal neurons. *Steroids* **2009,** *74* (2), 256-63.

167.   Rambousek, L.; Bubenikova-Valesova, V.; Kacer, P.; Syslova, K.; Kenney, J.; Holubova, K.; Najmanova, V.; Zach, P.; Svoboda, J.; Stuchlik, A.; Chodounska, H.; Kapras, V.; Adamusova, E.; Borovska, J.; Vyklicky, L.; Vales, K., Cellular and behavioural effects of a new steroidal inhibitor of the N-methyl-d-aspartate receptor 3alpha5beta-pregnanolone glutamate. *Neuropharmacology* **2011,** *61* (1-2), 61-8.

168.   Horak, M.; Vlcek, K.; Chodounska, H.; Vyklicky, L., Jr., Subtype-dependence of N-methyl-D-aspartate receptor modulation by pregnenolone sulfate. *Neuroscience* **2006,** *137* (1), 93-102.

List of Figures

List of Tables

List of Abbreviations

ATP	Adenosine triphosphate
CDK	Cyclin dependent kinase
CNS	Central nervous system
COSMO	Conductor-like screening model
DFT	Density functional theory
DNA	Deoxyribonucleic acid
ESP	Electrostatic potential
GB	Generalized Born implicit solvent model
HF	Hartree–Fock method
NHLs	Non-Hodgkin lymphomas
HTS	High-throughput screening
MAD	Mean absolute deviation
MD	Molecular dynamics
NMDAR	N-methyl-$_D$-aspartate receptor
PDB	Protein Data Bank
QM	quantum mechanics
RESP	Restrained Electrostatic Potential
RMSD	Root-Mean-Square Deviations
SAR	Structure-Activity Relationships
SMD	Solvation Model Density model
SQM	Semiempirical Quantum Mechanical

List of Attached Publications

The doctoral thesis is based on following articles:

A. **Nekardova, M.**, Vymetalova, L., Khirsariya, P., Kovacova, S., Hylsova, M., Jorda, R., Krystof, V., Fanfrlik, J., Hobza, P., Paruch, K.: Structural Basis of the Interaction of Cyclin-Dependent Kinase 2 with Roscovitine and Its Analogues Having Bioisosteric Central Heterocycles. *Chemphyschem*. 2017; 18(7):785-795.

B. Jorda, R., Havlicek, L., Sturc, A., Tuskova, D., Daumova, L., Alam, M., Skerlova, J., **Nekardova, M.**, Perina, M., Pospisil,T., Siroka, J., Urbanek, L., Pachl, P., Rezacova, P., Strnad, M., Klener, P., Krystof, V.: 3,5,7-Substituted Pyrazolo[4,3- d]pyrimidine Inhibitors of Cyclin-Dependent Kinases and Their Evaluation in Lymphoma Models. *J Med Chem*. 2019; 62(9):4606-4623.

C. Krausova, B., Slavikova, B., **Nekardova, M.**, Hubalkova, P., Vyklicky, V., Chodounska, H., Vyklicky, L., Kudova, E.: Positive Modulators of the N-Methyl-d-aspartate Receptor: Structure-Activity Relationship Study of Steroidal 3-Hemiesters. *J Med Chem*. 2018; 24;61(10):4505-4516.

D. Slavikova, B., Chodounska, H., **Nekardova, M.**, Vyklicky, V., Ladislav, M., Hubalkova, P., Krausova, B., Vyklicky, L., Kudova, E.: Neurosteroid-like Inhibitors of N-Methyl-d-aspartate Receptor: Substituted 2-Sulfates and 2-Hemisuccinates of Perhydrophenanthrene. *J Med Chem*. 2016; 59(10):4724-4739.

E. Kudova, E., Chodounska, H., Slavikova, B., Budesinsky, M., **Nekardova, M.**, Vyklicky, V., Krausova, B., Svehla, P., Vyklicky, L.: A New Class of Potent N-Methyl-D-Aspartate Receptor Inhibitors: Sulfated Neuroactive Steroids with Lipophilic D-Ring Modifications. *J Med Chem*. 2015; 58(15):5950-66.

Other attached articles:

F. Riedlova, K., **Nekardova, M.**, Kacer, P., Syslova, K., Vazdar, M., Jungwirth, P., Kudova, E., Cwiklik, L.: Distributions of therapeutically promising neurosteroids in cellular membranes. *Chem Phys Lipids*. 2016; 203:78-86.

G. Adla, S. K., Slavikova, B., Smidkova, M., Tloustova, E., Svoboda, M., Vyklicky, V., Krausova, B., Hubalkova, P., **Nekardova, M.**, Holubova, K., Vales, K., Budesinsky, M., Vyklicky, L., Chodounska, H., Kudova, E.: Physicochemical and biological properties of novel amide-based steroidal inhibitors of NMDA receptors. *Steroids*. 2016; 117:52-61.

H. Vyklicky, V., Smejkalova, T., Krausova, B., Balik, A., Korinek, M., Borovska, J., Horak, M., Chvojkova, M., Kleteckova, L., Vales, K., Cerny, J., **Nekardova, M.**, Chodounska, H., Kudova, E., Vyklicky, L.: Preferential Inhibition of Tonically over Phasically Activated NMDA Receptors by Pregnane Derivatives. *J Neurosci*. 2016; 36(7):2161-75.

I. Adla, S. K., Slavikova, B., Chodounska, H., Vyklicky, V., Ladislav, M., Hubalkova, P., Krausova, B., Smejkalova, T., **Nekardova, M.**, Smidkova, M., Monincova, L., Soucek, R., Vyklicky, L., Kudova, E.: Strong Inhibitory Effect, Low Cytotoxicity and High Plasma Stability of Steroidal Inhibitors of N-Methyl-D-Aspartate Receptors With C-3 Amide Structural Motif. *Front Pharmacol*. 2018; 12;9:1299.

J. Sivcev, S., Slavikova, B., Rupert, M., Ivetic, M., **Nekardova, M.**, Kudova, E., Zemkova, H.: Synthetic testosterone derivatives modulate rat P2X2 and P2X4 receptor channel gating. *J Neurochem*. 2019; 150(1):28-43.

Attached Publications