Univerzita Karlova v Praze
Farmaceutická fakulta v Hradci Králové
Katedra farmaceutické chemie a kontoly léčiv

# *In silico* screening inhibitorů SIRT6

## Diplomová práce

Autor: Tomáš Kučera
Vedoucí práce: prof. PharmDr. Martin Doležal, Ph.D.
Školitel-specialista: Maija Lahtela-Kakkonen, Ph.D.

Hradec Králové 2015

Charles University in Prague
Faculty of Pharmacy in Hradec Králové
Department of Pharmaceutical Chemistry and Drug Control

# *In silico* screening of SIRT6 inhibitors

## Master thesis

Author: Tomáš Kučera
Supervisor: prof. PharmDr. Martin Doležal, Ph.D.
Specialized supervisor: Maija Lahtela-Kakkonen, Ph.D.

Hradec Králové 2015

# Acknowledgment

I would like to convey my warmest gratitude to my specialized supervisor Maija Lahtela-Kakkonen, Ph.D. for her guidance, generous contribution of knowledge and experience, valuable comments as well as the support on the way. Furthermore I would like to thank the other members of Pharmaceutical and medicinal chemistry research group on Eastern Finland University who helped me during my work.

My acknowledgment goes also to prof. PharmDr. Martin Doležal, Ph.D. who helped me with my thesis on my home university, assoc. prof. doc. PharmDr. Kamil Musílek, Ph.D. for his mentoring during my student research work in my home country and Petr Nekorjak for the language corrections of my master thesis paper.

I would like to give special thanks to my colleague and friend Zuzana Maděryčová for her support right from the beginning of our study and membership in the Army of the Czech Republic.

Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum, provided under the programme "Projects of Large Infrastructure for Research, Development, and Innovations" (LM2010005), is greatly appreciated.

# Declaration

I declare that this master thesis is my original authorial work. All literature and other sources, which I used for the elaboration of this thesis, are listed on the reference page and properly cited in the text. This thesis was not used to earn any other or the same degree.

In Hradec Králové date signature of the author

# Abstrakt

**Název práce:** *In silico* screening inhibitorů SIRT6

**Autor:** Tomáš Kučera

**Katedra:** Katedra farmaceutické chemie a kontroly léčiv

**Vedoucí diplomové práce:** prof. PharmDr. Martin Doležal, Ph.D.

**Školitel-specialista:** Maija Lahtela-Kakkonen, Ph.D.

**Abstrakt:** SIRT6 je nazýván NAD-dependentní protein deacetylasa sirtuin-6 a je členem proteinové rodiny sirtuinů. Moduluje acetylaci histonu H3 (klinicky důležitých Lys9 a Lys56). SIRT6 je zajímavým cílem léčiv vzhledem k jeho roli při replikaci DNA, glykolýze a zánětu – proto je vývoj inhibitorů SIRT6 významný v souvislosti s diabetes mellitus, artritidou a rakovinou.

Cílem práce bylo najít nové molekuly inhibující deacetylační aktivitu SIRT6 za využití metod výpočetní chemie a molekulového modelování. Snažili jsme se najít zejména nové struktury, které by bylo možné optimalizovat v dalších fázích cesty za léčivem.

Jako vstupní data bylo použito 9 inhibitorů a krystalová struktura SIRT6 (PDB kód 3K35). Ze skupiny metod založených na ligandech byly vybrány farmakofor a chemická podobnost, z metod založených na struktuře to pak byl molekulový docking. Farmakofor byl definován po strukturálním srovnání čtyř známých ligandů a testován na souboru ligandů a neligandů. Jako vzor pro chemickou podobnost (BIT_MACCS fingerprint) byly použity známé ligandy a jejich fragmenty. Molekulový docking probíhal převážně v softwaru MOE.

Celkem bylo vybráno a pro *in vitro* testování doporučeno 44 molekul. Dosud jich bylo otestováno 11 a čtyři z nich vykazují signifikantní inhibiční aktivitu na SIRT6.

**Klíčová slova:** SIRT6, sirtuin, inhibitor, virtuální screening, histon

# Abstract

**Title:** *In silico* screening of SIRT6 inhibitors

**Author:** Tomáš Kučera

**Department:** Department of Pharmaceutical Chemistry and Drug Control

**Supervisor:** prof. PharmDr. Martin Doležal, Ph.D.

**Specialized supervisor:** Maija Lahtela-Kakkonen, Ph.D.

**Abstract:** SIRT6 is called NAD-dependent protein deacetylase sirtuin-6 and it is a member of sirtuin protein family. It modulates acetylation of histone H3 (clinically important Lys9 and Lys56). The SIRT6 enzyme is an interesting drug target because of its role in DNA replication, glycolysis and inflammation – that is why the design of SIRT6 inhibitors is relevant in context of diabetes mellitus, arthritis and cancer.

The aim of the work was to identify small molecules to inhibit deacetylase activity of SIRT6 using methods of computational chemistry and molecular modeling. We tried to find new lead structures with possibility to be optimized in next phases of the drug discovery process.

The 9 known inhibitors and crystal structure of SIRT6 (PDB code 3K35) were used as input data during the modeling. Pharmacophoric and chemical similarity searches were selected from the group of ligand-based methods and molecular docking from the group of structure-based methods. The pharmacophore was defined after structural alignment of four known ligands and tested on set of ligands and non-ligands. As pattern molecules for chemical similarity search (BIT_MACCS fingerprint), known ligands and their fragments were used. Docking was done mainly using software MOE.

Together 44 molecules were selected and recommended for *in vitro* testing. 11 compounds have been tested so far and four of them show significant inhibition activity on SIRT6.

**Keywords:** SIRT6, sirtuins, inhibitor, virtual screening, histone

# Contents


**1 Introduction** **3**

1.1 Histones . . . 3

1.2 SIRT6 and its inhibitors . . . 3

1.3 Aims of the work . . . 5

**2 Theoretical part** **7**

2.1 Virtual screening . . . 7

2.2 Ligand-based methods . . . 7

2.2.1 Quantitative structure-activity relationship . . . 7

2.2.2 Chemical fingerprint . . . 11

2.2.3 Pharmacophore . . . 11

2.3 Structure-based methods . . . 12

2.3.1 Docking . . . 12

2.3.2 Scoring . . . 16

**3 Methods** **19**

3.1 Protein preparation . . . 19

3.2 Small molecules preparation . . . 19

3.3 Docking in software MOE . . . 20

3.3.1 Software verification . . . 20

3.3.2 Docking protocol development . . . 21

3.3.3 Screening using docking . . . 21

3.4 Pharmacophore . . . 22

3.5 Molecular similarity . . . 23

3.6 Molecular similarity of fragments . . . 24

3.7 Docking in software Schrödinger Glide . . . 24

3.8 Final selection . . . 25

3.9 *In vitro* testing . . . 25

**4 Results** **27**

4.1 Virtual screening . . . 27

4.2 *In vitro* testing . . . . . 41

**5 Discussion** **43**

5.1 Methods . . . . . 43

5.1.1 Docking . . . . . 43

5.1.2 Pharmacophore . . . . . 44

5.1.3 Fingerprint and fingerprint of fragments . . . . . 44

5.2 *In vitro* results . . . . . 44

**6 Conclusions** **47**

**Bibliography** **59**

**List of abbreviations** **60**

**List of tables** **61**

**List of figures** **64**

# 1 Introduction

## 1.1 Histones

Nucleosome is a basic building unit of chromatin. It is formed from double-chain DNA and protein core. The core includes eight subunits – co called histone proteins. There are four histone subtypes H2A, H2B, H3 and H4, each in two copies (fig. 1.1). Histone proteins in eukaryotic cells are evolutionary very old and conservative. (For example, histone H4 from sheep and pea differ only in two residues). They contain high level of basic amino acids as arginine and lysine. Positive-charged residues in histones are important for interaction with and binding to negative-charged sugar-phosphate backbone of DNA. [1, p. 211] Positive charge on the histone tails can be removed by acetylation and thus causing interruption of the protein-DNA interaction and releasing of chromatin structure. Liberated DNA sequences can be then transcribed. Histone tail acetylation is controlled by histone acetyltransferases and histone daecetylases. Histone acetyltransferases transfer the acetyl group to lysine residue and histone deacetylases remove it. [2, 3]

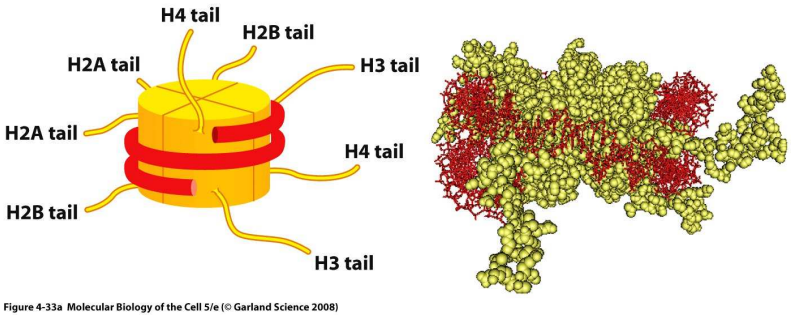



**Figure 1.1:** Nucleosome, from Albrets [1, p. 218]. Scheme of nucleosome shows its octameric structure. Eight tail chains on the N-terminus of histones include modifiable residues.

## 1.2 SIRT6 and its inhibitors

There are 18 known human histone daecetylases. Classes I, IIA, IIB and IV are zinc-dependent. NAD-dependent histone deacetylases are categorized to the class III and they are called sirtuins. [4] The family of sirtuins has got seven members

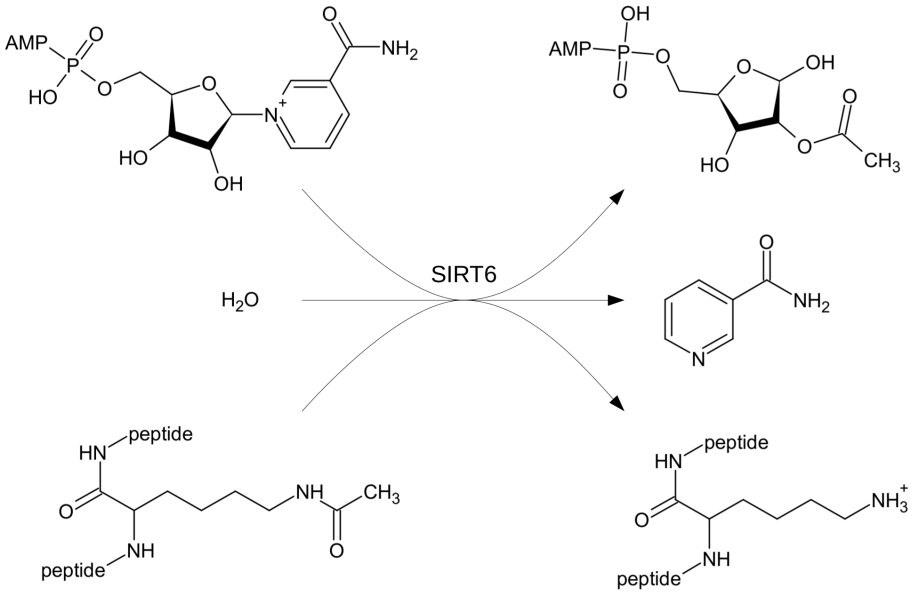



**Figure 1.2:** Mechanism of lysine deacetylation catalyzed by sirtuin, according to Sauve [5]

and in our work, we focus to the sirtuin 6 (SIRT6). The figure 1.2 shows chemical mechanism of deacetylation catalyzed by SIRT6.

Target of SIRT6 deacetylation activity are clinically important Lys9 and Lys56 on histone H3 (H3K9 and H3K56). [6] Deacetylation of H3K9 is connected with damage of DNA and activation of repair mechanisms. The acetylation of H3K56 is related to chromatin stability. [2, 7, 8] SIRT6 has got complex role in regulation of cell signaling pathways. As a short example of influenced cell processes, we can note following: [9, 10]

- DNA repair and genome stability
- cell proliferation
- glucose metabolism
- lipid metabolism
- inflammation.

In context of previous, SIRT6 is an interesting therapeutic target with connection to diabetes mellitus, arthritis and cancer and design of its inhibitors is a point of interest of current biomedicine.

## 1.3 Aims of the work

The main aim of the work was to identify small molecules to inhibit deacetylase activity of SIRT6. We preferred to find new lead structures with the possibility of being optimized in next phases of the drug design process. The methods of *in silico* screening were used for this work.

The partial aim was to optimize selected virtual screening methods for this work.

The aim of the theoretical part was to describe some of ligand- and structure-based screening methods which can be used for this part of development.

# 2 Theoretical part

## 2.1 Virtual screening

Drug development is a costly and time-consuming process which has got several subsequent phases. In past, it was *in vitro*, preclinical and clinical testing. Nowadays, *in silico* the phase usually precedes the others. Virtual screening is one of *in silico* methods. These techniques became popular in 1980s and their popularity has increased with the development of computational technologies and growing computational performance. [11]

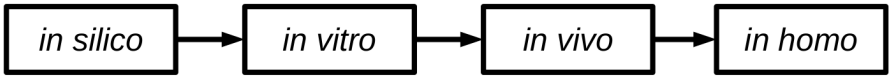



**Figure 2.1:** General scheme of current drug design

Virtual screening is *in silico* methods in which the drug target or known active compounds are used for identifying novel compounds that are likely to have the desired effects. Virtual screening can be divided into two strategies: Ligand-based virtual screening and structure-based virtual screening. Both approaches can be applied simultaneously. There are several advantages in virtual screening. It is a fast method compared to *in vitro* methods. Also the number of compounds screened with virtual screening is larger than with *in vitro* or *in vivo*. It is also much cheaper and safe compared to biological methods. [12]

## 2.2 Ligand-based methods

### 2.2.1 Quantitative structure-activity relationship

QSAR (quantitative structure-activity relationship) methods tries to find correlation between molecule structure and its biological activity. There is relatively easy way to quantify biological activity of molecule (parameters like $K_i$, EC50, IC50) but mathematical quantification of chemical structure is the real challenge. The molecular weight, lipophilicity, steric and electronic properties, acidity, presence of

some functional groups and other parameters can be used as chemical descriptors. [13, 14]

We recognize some types of QSAR based on level of used chemical descriptors. 1D QSAR is based on the correlation between biological activity and global physicochemical properties such as logP or $pK_a$ value. 2D QSAR takes into account the structural patterns and the connectivity and correlates them with biological activity. The 3D-QSAR is nowadays most popular QSAR method. It calculates with spatial structure of a molecule (usually with its steric properties). The 4D-QSAR computes with more conformations of the molecule. The 5D-QSAR adds induced-fit models to 4D-QSAR. The 6D-QSAR adds solvation models. [15, 13]

### 3D-QSAR

This type of QSAR works with three-dimensional representation of molecules. It focuses mainly on atoms' properties in their spatial localization. Studied properties usually have geometrical or electrostatic character.

**Descriptors** Conversion of three dimensional chemical data to linear or tabular data is not a trivial task. The problem is solved by using different descriptors which mostly describe physicochemical character of a molecule. There are many different approaches for defining these descriptors. Many descriptors can be calculated with simple empirical formulas based on the structure and connectivity of molecule. For example molecular weight is easy to calculate and log P can be estimated from atomic contributions. [15]

CoMFA (Comparative molecular field analysis) is one of most popular 3D QSAR method that is alignment-dependent. It was designed by Richard Cramer in 1988. [16] COMFA is based on the assumption that the non-covalent interactions between protein and ligand can be related to biological activity. In CoMFA, ligand-receptor interactions are represented by steric and electrostatic interactions between molecule and a probe atom. The success of CoMFA relies heavily on relative positioning of ligands in the fixed lattice. First, positions of ligands in the space are needed. Then the ligands are placed in three cubic grid and the steric (eqn 2.1 where $E_{vdW}$ is the van der Waals interaction energy $A_i$ and $B_i$ are van der Waals repulsion

and attraction parameters, r is the distance between atom and grid point) and electrostatic (eqn 2.2, where $E_{Coulomb}$ is the electrostatic interaction energy, $q_i$ is the point charge on the atom i of the molecule, $q_j$ is the charge of the probe atom, D is the dielectric constant and r is the distance between atom and grid point) properties are calculated in each node of the grid. Results of this calculation are input data for the multilinear regression analysis. [14, 16]

$$E_{vdW} = \sum(\frac{A_i}{r^{12}} - \frac{B_i}{r^6}) \tag{2.1}$$

$$E_{Coulomb} = \frac{q_i q_j}{Dr} \tag{2.2}$$

Another method for finding descriptors is the Comparative Molecular Similarity Indices Analysis (CoMSIA). It is similar to CoMFA but it uses Gaussian potential instead of classical CoMFA functions. An advantage of this method is introduction of hydrophobic properties.

Other alignment-dependent methods are for example genetically evolved receptor modeling, comparative binding energy analysis, adaptation of the fields for molecular comparison, hint interaction field analysis and comparative residue interaction analysis.

The second group of approaches for finding descriptors are so called alignment-independent methods. They eliminate some limitations of alignment-dependent methods (like computing time demands, possibility of user bias). They can be based on similar approaches like CoMFA (e. g. GRID – grid independent descriptors) or they can use different properties of compounds. For example Holo-QSAR uses fingerprint description of fragments which are made from the molecule, or CoSA (Comparative Spectral Analysis) uses mass spectrometry, nuclear magnetic resonance and infrared spectral data. [15]

**Statistical quantification** After numerical description of chemical structure, there is a big amount of data for each molecule. To evaluate the relationships between a large amount of calculated descriptors and the biological activity multivariate statistical analysis like partial least squares method (PLS) is needed.

PLS is the most commonly used method. It transforms the data set to smaller matrix and computes with this matrix of numbers in the next phases of the process. It is a difference from classical multilinear regression analysis.

**Quality of the model** The most usual approach for evaluation of QSAR model is external test. A part (10 to 20 %) of starting compounds is randomly selected and excluded from the model development. Activity of these compounds is then calculated using the done model. Computationally predicted activity is then compared with the measured one. This method reduces the set size (and the associated model quality) but it enables to estimate the predictive power. Only models, which demonstrates significant predictive power in external test, should be used as a *in silico* screening method. [17]

**3D-QSAR problems and limitations** The biggest problems of 3D-QSAR are connected with matter of the active conformation. The alignment procedure can dramatically affect the quality of the alignment-dependent 3D-QSAR models. There are some methods for the discovering the most probable spacial conformation like X-ray crystallography or docking. But there is no sureness that the used conformation is the active one in biological environment. Additionally, a small change in the ligand structure can bring totally different spatial position and interactions with the receptor. It of course changes the binding affinity.

The second problem is the amount of data and the demands on the computing time during their processing. Because of this aspect, CoMFA and other 3D-QSAR techniques are inapplicable in high-throughput virtual screening.

Also the selection of compounds has a great influence on the 3D QSAR models. The predictivity of QSAR model should be tested with external test set.

Even though some of the methods try to work with the receptor (e. g. the genetically evolved receptor modeling), most used 3D-QSAR approaches do not neglect the receptor, its properties and its interactions with the ligand. [18, 15]

### Multidimensional QSAR

The 4D-, 5D- and 6D-QSAR methods are usually called multidimensional QSAR. Of course, they compute with the three-dimensional space and the signs 4D-, 5D- and 6D do not mean next dimensions but taking more degrees of freedom into consideration.

### 2.2.2 Chemical fingerprint

Fingerprint descriptor is typically a binary bit string characterization of molecule structure. It serves for comparison of molecules and for determination of their similarity. [19, 20] This method is based on the assumption of 'neighborhood behavior' – the theory that chemically similar molecules exhibit similar biological activity. [21, 22] Different fingerprints are available: 2D fingerprints like MACCS or Unity and 3D fingerprint like Unity 3D. Usually a single fingerprint is compared with molecules in a database for the determination of their similarity. Similarity between fingerprints is estimated with the similarity coefficient: Tanimoto is the most widely used. The Tanimoto coefficient is between 0 (completely different) and 1 (completely similar). [23]

General advantage of fingerprint is that it demands only a little input data. This method can be used when the target is unknown and when there are only few known active ligands.

Descriptor similarity methods are relatively computationally inexpensive. It is the second argument for their using. The computation has in principle two phases. The first is obtaining the descriptors and in the second one, the descriptors are compared. The phases can be separated and the descriptor data can be computed only ones and used for many projects.

The fundamental problem is how to transform chemical data into linear numerical format. There are many different ways how to get fingerprint description of a given molecule. Very representative example for explanation was published by Xue [24]. This method is shown in figure 2.2.

### 2.2.3 Pharmacophore

Pharmacophore is an abstract description of molecular features that describes the interaction capability of either one, or a group of compounds towards a target.

Pharmacophore based methods are also based on the assumption that "similar structure equals similar biological effect". They are looking for features of the molecule that are necessary for its biological function. Presence of these features enables to establish specific interaction with the target molecule (H-bonds, electrostatic, van der Waals and other interactions). It is necessary to point out that

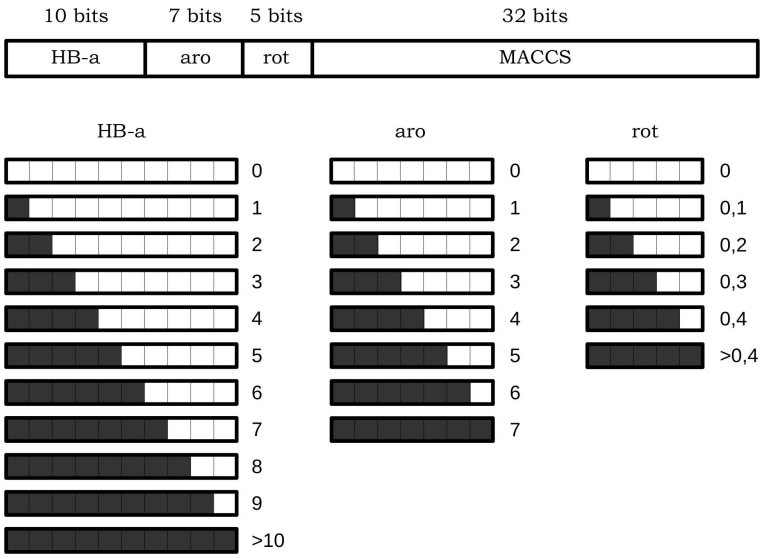



**Figure 2.2:** Example of fingerprint construction. According to Xue [24]. This fingerprint description is combination of four strings. The first (ten bits) represents number of hydrogen bond acceptors. The second one (seven bits) shows number of aromatic bonds in the molecule. The third part (five bits) of the chain is fraction of rotatable bonds. This three strings are combined with the 32-bits MACCS description (described below).

pharmacophore does not represent real functional groups but only abstract electrostatic and steric attributes. [25]

In a real approach, a hypothesis of crucial functional groups and interactions is created. This hypothesis is usually based on X-ray structure of ligand-target complex or structural alignment of known ligands. Features are defined as spheres in space, their position, size and type are set. In the next step, the pharmacophore is tested with using a database of known ligands and non-ligands. During the testing, if the defined pharmacophore is successful (all ligands are included in results and no non-ligand is false positive result in an ideal case), the pharmacophore can be used for screening on a large database. [26]

## 2.3 Structure-based methods

### 2.3.1 Docking

Docking is one of molecular modeling methods which helps to predict interactions of proteins with other molecules. In molecular docking a small molecule is fitted into the protein's active site. These molecules can be proteins, nucleic acids or

small ligands [27, p. 116]. In the theoretical part of our work, we focus on the protein-ligand docking which is very usable as a method of virtual screening.

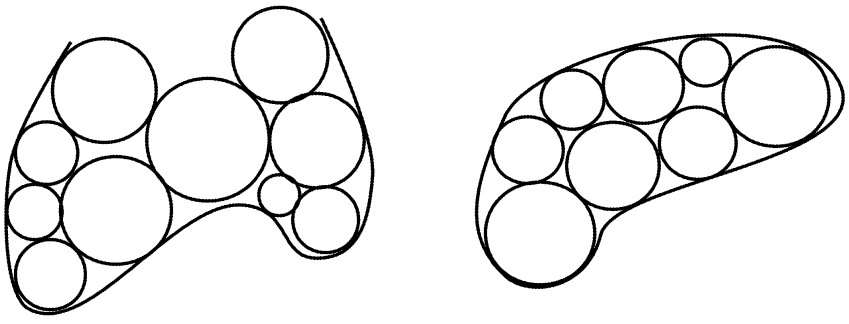
The key-lock theory and induced fit theory are important for understanding this technique. The key-lock theory postulates that a receptor is rigid structure (like a lock) and the ligand has to have accurate shape and size to be able to interact with the receptor (like a key). This theory is important mainly for docking with inflexible receptor.

The second theory rejects the conception of rigid receptor. It supposes that the final protein-ligand structure is a result of their mutual interaction. It is an argument for using docking with flexible receptor. [28]

The docking process has two main parts. The first one is searching for possible poses of ligand and the second one is scoring found poses. Nowadays, there are many searching algorithms. Description of most popular approaches (according to Sousa [29]) is following.

### Geometry-based methods

In the first step, the binding cavity is filled by set of spheres as a representation of atom volumes. Then the ligand is also divided into spheres. (fig. 2.3) In the next step both sets of spheres are combined based on distance compatibility. Newer algorithms use clique-detection algorithm instead of distance-compatible matches. [30]



**(a)** **(b)**

**Figure 2.3:** Geometry-based method. According to Höltje 2003. [28] (a) an example of receptor divided into spheres; (b) an example of ligand divided into spheres

Another possibility is to define characteristic features in the binding site of protein and in the ligand. The distances between features in the receptor are computed

and so are in the ligand. Then the software tries to match features of ligand to compatible features in the receptor. The distances of features are respected in this matching. [31, p. 6]

### Geometric hashing

Hashing is method for processing data. A large data set can be converted to a hash key using hash functions. First, the geometric hash table is created form the ligand. In the second phase, so-called recognition phase, the hash entries are evaluated using matching to the features of the receptor. Best-evaluated hash entries (many times matched hash entries) are selected and analyzed. [32, 31, p. 9]

### Incremental construction methods

Programs using this type of docking algorithm do not dock whole ligand. They cut the ligand along each rotatable bond. Then the first part of the ligand is placed into active site. (There are differences between programs.) Possible poses are scored and the best is chosen for "growing" the ligand.

In the following steps, next parts of ligand are attached and rotated (before that, some programs prepare a list of preferred torsion angles). All possibilities are scored and compared. Other parts are connected until the ligand is complete.

After the growing phase, all results are scored (usually by included scoring function) and compared. Poses with highest score are selected. [28] (fig. 2.4)

### Genetic algorithms

This method tries to find best solution using principles of evolution biology (inheritance, hybridization, mutation and natural selection). Every solution is encoded into linear information (called chromosome or genome). Zero generation of solutions is usually created randomly. Each individuum in the population is evaluated and the evaluation determines its contribution to next generation. Individuums of each next generation are found through hybridization, mutations and copying of chromosomes of the previous generation. This procedure is repeated until the result fulfills set criteria. [28]

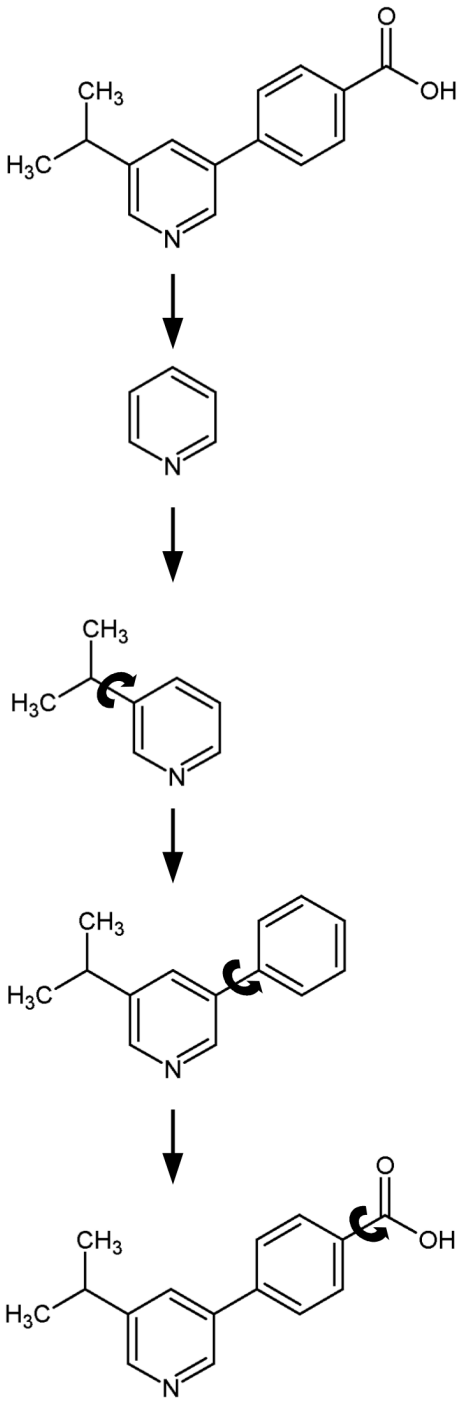



**Figure 2.4:** Incremental Construction Method. According to Höltje 2013. [28] 1. Ligand is divided into fragments along rotatable bonds. 2. First fragment is docked into receptor. 3. Other fragments are step by step attached and rotated.

#### Simulated annealing and Monte Carlo simulations

Simulated annealing is an often used method of molecular simulation. During the simulation run, the temperature of the system is decreasing. Consequence of this process, the system is approaching to the local minimum. It means that the result is influenced mainly by the starting position.

The main disadvantage of the simulated annealing is solved by the Monte Carlo algorithm. The movement of the system is not strictly towards the lower energy but randomly. The energy of the new state is calculated after each step. The new state is accepted if its energy is lower or if the value of probability of acceptance is lower then a randomly calculated number between 0 and 1. The probability of acceptance is calculated throw the Boltzmann equation:

$$P = e^{(-\frac{\Delta E}{kT})} \tag{2.3}$$

where $\Delta$E is energy difference from previous step, k is the Boltzmann constant and T is the thermodynamic temperature.

Advantage of the method is the possibility to go over energetic barriers. Monte Carlo is a stochastic method, there is no guarantee of completely exhausting search. [33, 31, p. 16]

#### Conformational libraries

Computation with pregenerated conformational libraries is a next way how to include ligand flexibility. Possible conformations of ligand are generated from small fragments and internal energy of each conformation is evaluated. Suitable conformations are docked into the receptor by rigid-docking algorithm. [34]

### 2.3.2 Scoring

#### Empirical scoring functions

Binding energy is calculated as a sum of particular increments such as ionic interactions, hydrogen bonds, aromatic interactions, lipophilicity, desolvation, ligand flexibility and others. The effects of parameters on total score are described by coefficients. Values of coefficients are optimized by multilinear regression using experimental data from training set. [35]

An example of this type of scoring is London dG in software MOE (Molecular Operating Environment). It calculates the binding energy with following formula:

$$\Delta G = c + E_{flex} + \sum_{h-bonds} c_{hb} f_{hb} + \sum_{metal-lig} c_m f_m + \sum_{atomsi} \Delta D_i \tag{2.4}$$

where $E_{flex}$ is a topological estimate of ligand entropy, $f_{hb}$ and $f_m$ are measures of geometric imperfections of protein–ligand and metal–ligand interactions, $\Delta D_i$ is the desolvation energy and c, $c_{hb}$ and $c_m$ are empirical determined coefficients. The coefficient are determined using training set with 400 protein-ligand complexes. [36]

**Force field-based scoring functions**

Energy calculation of non-binding interaction is based on classical molecular mechanics principles. There are many formulas for computing binding force-field energy but most of them have some similar base. It is usually a sum of increments like energy of angles, bonds, dihedrals and non-bonded interactions. [37]

Non-boded interaction energy is usually represented as the Lennard-Jones potential:

$$E = \sum_{i=1}^{lig} \sum_{j=1}^{rec} \left(\frac{A_{ij}}{r^{12}} - \frac{B_{ij}}{r^6} + 332 \frac{q_i q j}{D r_{ij}}\right) \tag{2.5}$$

where $A_{ij}$ and $B_{ij}$ are van der Waals repulsion and attraction parameters, $r_{ij}$ is the distance between atoms i and j, $q_i$ and $q_j$ are the point charges on atoms i and j, D is the dielectric function. [28, p. 158]

For example Merck molecular force field:

$$E = \sum E_{bond} + \sum E_{angle} + \sum E_{dihedral} + \sum E_{stretch} + \sum E_{oop} + \sum E_{vdW} + \sum E_{Coulomb} \tag{2.6}$$

where $E_{bond}$ is the energy of bond-stretching energy, $E_{angle}$ is the angle-bending energy, $E_{dihedral}$ is the torsional energy, $E_{stretch}$ is the stretch-bend energy, $E_{oop}$ is the Out-of-Plane-bending energy, $E_{vdW}$ is the van der Waals energy and $E_{Coulomb}$ is the electrostatic interaction energy. [38]

This function is well studied and it has got physical basis. It is relatively fast. Main disadvantage is leaving out of entropy influence on the protein-ligand affinity. [28, p. 158]

An example is GBVI/WSA dG scoring in software MOE. It uses the MMFF94x and AMBER99 force field. [36]

### Knowledge-based scoring functions

This method is based on knowledge of protein-ligand structures mainly from X-ray crystallography. The distances of atom types in protein and ligand are measured and we presume that the more frequent distances are more favorable. The interaction free energy is then described as a function of atom distance based on statistical analysis of known complexes.

$$A(r) = -k_B T \ln g_{ij}(r) \tag{2.7}$$

where $k_B$ is a Boltzmann constant, T is the thermodynamic temperature, $g_{ij}$ is the distribution function for atom-type pair ij and r is the distance between atoms i and j. [39]

### Consensus scoring

This type of scoring is based on combination of at least two methods. It is a way how to get less false positive results with high score value. [40]

# 3 Methods

The database Enamine (4,103,115 molecules) and Chembridge (1,022,400 molecules) were used for the screening. Pharmacophoric search, molecular similarity and docking were selected for the high throughput screening phase. In the second phase, the hits were docked into SIRT6 in more steps with different algorithms and settings. Finally, 44 molecules where recommended for *in vitro* testing (fig. 3.1).

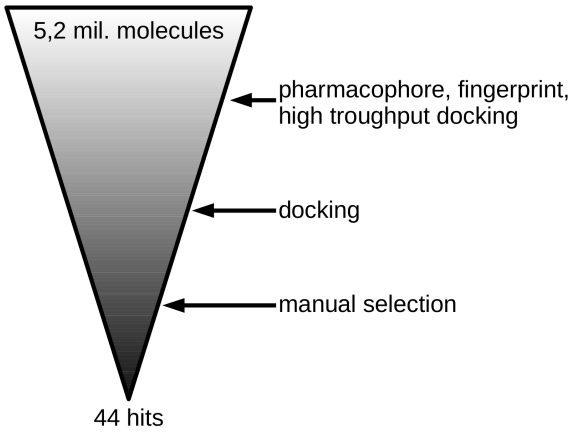



**Figure 3.1:** Scheme of the screening process

## 3.1 Protein preparation

For the work, the structure of SIRT6 from Protein Data Bank (PDB code 3K35 [41]) was used. All components excluding chain A were removed. The structure was processed using function Protonate 3D of software MOE. Value of pH was set to 7.0, temperature 300 K and salt concentration 0.1 mol/l. Then, the structure was checked (Ramachandran plot, bonds length, angles and dihedrals) and no problems were found.

## 3.2 Small molecules preparation

**EX-527, quercetin, sirtinol, resveratrol.** 3D structures of these compounds were found in public database Pubchem [42, 43, 44, 45], downloaded in sdf-format and imported do MOE-database file.

**Inhibitors published by Parenti, 2014 [46].** 2D structures were build in software ChemSketch [47]. Preparation file for computing (adding hydrogens, preliminary structure optimization) was done in software Avogadro [48]. 3D coordinates were computed in software Gaussian09 [49] (density functional method B3LYP/6-31G(d) Opt). Final structures were transferred to mol-file by OpenBabel 2.3.1 [50] and imported to MOE-database-file.

**Database Enamine.** The database Enamine was downloaded as 2D structures. The conversion to 3D was done in LigPrep [51] (pH 7.0 $\pm$ 2.0, generate tautomers, Force field OPLS_2005). Then, the output was filtered using Canvas [52] (criteria: Mw $< 700$, LogP $< 7$, HBa $< 15$, HBd $< 15$). Output included 4,094,462 structures.

**Database Chembridge.** The 2D structures of compounds were downloaded from http://zinc.docking.org/catalogs/chbr in format prepared for docking. [53].

## 3.3 Docking in software MOE

The target for docking was pocket for binding of natural inhibitor nicotimamide, so called C pocket.

### 3.3.1 Software verification

The molecule EX-527 was re-docked to SIRT1 for verification of the docking function of software MOE. Their co-crystal structure was gained from public database (PDB identification 4I5I [54]). Then the protein was prepared for docking through general procedure described in section 3.1 on page 19. Docking was done under these parameters: placement Triangle Matcher, first re-scoring London dG, refinement Forcefield and second re-scoring GBVI/WSA dG; MMFF94x force field was used.

Six out of ten best scored poses were similar to those published by Zhao *et al.* [54]. The software MOE was found suitable for this project.

### 3.3.2 Docking protocol development

Different settings of docking task in software MOE were tested. Main focus was to the best definition of docking site. Other point of interest was spatial position of Phe62 side chain. All protocols were tested through docking of EX-527, S-sirtinol, resveratrol and quercetin and following observing similarities with pose of EX-527 in SIRT1 (PDB code 4I5I [54]).

As most suitable protocol was selected definition of docking side through selecting residues Ser54, Asn112 and His131. The residue Phe62 was left in the original position (unlike to protocol published by Parenti *et al.* [46]).

### 3.3.3 Screening using docking

Docking protocol described above was applied for screening in the database . Docking parameters were set for fast screening (placement Triangle Matcher, re-scoring London dG, without refinement, without ligand flexibility). Because of cpu-time demands, only first of eight parts of the database was screened. From 476,814 molecules in 4,768,014 poses was selected 301 best scored. This molecules was re-docked under similar parameters with allowed bonds rotation. As result, 3010 poses were gotten. Twenty best molecules were re-scored with more accurate setting (placement Triangle Matcher, re-scoring 1 London dG, refinement Forcefield, re-scoring 2 GBVI/WSA dG and London dG, allowed bonds rotation).

Set of 92 compounds offered by collaborating research group was screened through the docking in software MOE. The database was prepared for docking – adding hydrogen atoms (pH 7.4), converting in 3D structures (software MOE, function Rebuild3D), removing inorganic ions and other small structures. Molecules were docked into SIRT6 (placement Triangle Matcher, re-scoring 1 London dG, without refinement, allowed bonds rotation). 25 best scored molecules were re-docked with more accurate parameters (placement Triangle Matcher, re-scoring 1 London dG, refinement Forcefield, re-scoring 2 GBVI/WSA dG and London dG, allowed bonds rotation). 10 compounds were selected and recommended for *in vitro* testing. Selection was based on score values and observed possible interactions.

## 3.4 Pharmacophore

The ligands of SIRT6 published by Kokkonen *et al.* (EX-527, quercetin, resveratrol and S-sirtinol) [3] were used to define pharmacophore. First, flexible alignment of these compounds was done. The fourth pose was chosen because of good correlation of oxygen atoms. The pharmacophore shown on figure 3.2 was defined with following parameters:

| | | |
|---|---|---|
| F1 | Aro | 1.5 |
| F2 | Don/Acc | 1.3 |
| F3 | Don/Acc | 1.3 |
| F4 | Don/Acc | 1.2 |
| F5 | Don | 1.2 |
| C1 | At least one F[2,3] | |
| C2 | At least one F[4,5] | |

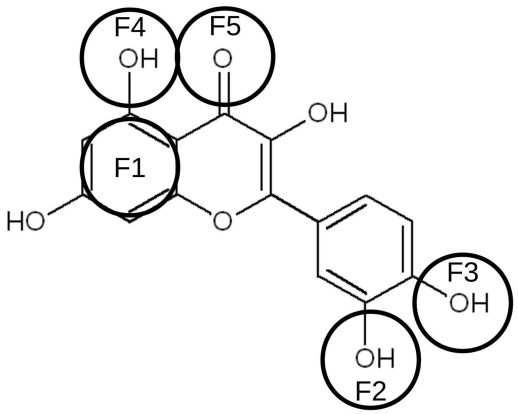



**Figure 3.2:** 2D representation of pharmacophore

The pharmacophore was tested by using COX-1 ligands (from DuD database [55]) and SIRT6 ligands (EX-527, sirtinol, resveratrol, suramin and quercetin). Result included 4 molecules – quercetin, resveratrol, suramin and ZINCC00012342.

This pharmacophore was used for searching through the database Enamine and 3,362 conformations were considered to be suitable. These compounds were docked into SIRT6 (placement Triangle Matcher, re-scoring London dG, without refinement,

without ligand flexibility). The outcome was 33,600 poses. Unique molecules with better score than quercetin were selected and erroneous compounds were eliminated (mainly structures with incorrectly interpreted sulfonic group). These 151 compounds were re-docked with allowed bound rotation in ligand molecule. Twenty best scored molecules were re-docked with more accurate parameters (placement Triangle Matcher, re-scoring 1 London dG, refinement Forcefield, re-scoring 2 GBVI/WSA dG and London dG, allowed bonds rotation).

The same pharmacophore was used for screening in the database Chembridge and 2,387 conformations were obtained. In the next step, the function Wash of the software MOE was used and for each structure, tautomers were enumerated up to limit of 10 and the results included 6,150 structures. This database was docked into SIRT6 (placement Triangle Matcher, re-scoring London dG, without refinement, allowed bonds rotation) and 61,500 poses were obtained. Thirty best scored unique molecules were selected and re-docked with more accurate parameters (placement Triangle Matcher, re-scoring 1 London dG, refinement Forcefield, re-scoring 2 GBVI/WSA dG and London dG, allowed bonds rotation).

## 3.5 Molecular similarity

The screening using fingerprint was done in the software MOE [36]. For computing descriptors, the method BIT_MACC was used. Known ligands published by Kokkonen *et al.* [3] (EX-527, quercetin, resveratrol and S-sirtinol) and Parenti *et al.* [46] (ID 5, 9, 16, 17) were selected as pattern molecules.

During searching, the overlap level was set individual for each pattern molecule to get circa 800 hits. Together, it were 7,659 molecules. These potential ligands were docked into SIRT6 (placement Triangle Matcher, re-scoring London dG, without refinement, without ligand flexibility) and the outcome was 76,330 poses. Unique molecules that scored better than quercetin (95 entries) were re-docked with allowed bonds rotation and 950 poses were obtained. Molecules that scored better than quercetin were selected (22 molecules) and used for final re-docking (placement Triangle Matcher, re-scoring 1 London dG, refinement Forcefield, re-scoring 2 GBVI/WSA dG and London dG, allowed bonds rotation).

Similar process was done with the database Chembridge. The overlap level was set differently to get circa 500 similar molecules for each known inhibitor (together 4,124 molecules). The three step docking run with the same parameters as described in previous paragraph. Results included 41,280 poses after first docking. For second computation, 46 molecules that scored better than quercetin were selected and this computation gave 1,380 poses. Best 20 molecules were re-docked in the third step.

## 3.6 Molecular similarity of fragments

In contrast to previous procedure, fragments of known ligand were used instead of whole molecules. Together, eleven fragments were created from inhibitors published by Parenti *et al.* [46]. The same method (BIT_MACC) was used for comparing chemical structure. During docking, the same three-steps process was used.

In the database Enamine, overlap was set on level to get circa 800 hits for each fragment. It was 8,079 hits in sum. After first docking, there were 29,200 poses. 264 molecules were re-docked in the second phase with result of 2,512 poses. For the third step, 50 compounds were selected.

During search in the database Chembridge, the overlap level was set to find circa 500 hits and 4,110 molecules entered to the docking phase. The first docking computation gave 41,030 poses. For the next step, 59 molecules that scored better than quercetin were selected and 1,770 poses were gained. In the third step, 20 best scored molecules were re-docked.

## 3.7 Docking in software Schrödinger Glide

Hits from pharmacophore and fingerprint search were docked using the software Glide [56] in standard precision mode. The grid box was defined through selecting residues Arg63, Asn114 and His131 and its size was 25 Å; the ligand flexibility was allowed. The results were compared with results from previous experiments (docking in MOE software) and best poses were used in the next computations.

## 3.8 Final selection

Results from all screening computations in the databases Enamine and Chembridge were visually inspected with focus on possible interactions with the receptor, chemical stability of compound, reality of the protonation state and size of the compound. Together, 53 compounds were pre-selected. These were docked in software Glide [56] (grid box defined trough selecting residues Arg63, Asn114 and His131, box size 25, standard precision, flexible ligand) and Induced Fit [57] (box center set trough residues Arg63, Asn114 and His131, box size set automatically). All data were organized and final selection of molecules for *in vitro* testing was done. This selection was based on score values made by different scoring functions (London dG, GBVI/WSA dG, Glide score and Glige score after included fit docking), predicted interactions and structure variability of compounds. 44 molecules were recommended for *in vitro* testing.

## 3.9 *In vitro* testing

SIRT6 has a weak deacetylase function. Commercially available kit for sirtuins uses non-natural substrate of SIRT6. Because of this, the method was modified by Kokkonen *et al.* [3].

Expression vector hSIRT6-pGEX-6P3 containing the coding sequence of the human SIRT6 was kindly provided by prof. Katrin Chua (Stanford, USA). GST-tagged SIRT6 was produced by fermentation in Escherichia coli BL21(DE3)-pRARE. After induction with 0.1 mM IPTG for 20 h at +16 °C, the soluble overexpressed protein was affinity purified on glutathione agarose (Sigma). The purity of the GST-SIRT6 was $> 80$ %, as detected by SDS-PAGE.

The substrate used in testing for SIRT6 is (Ac-RYQK (Ac)-AMC) which mimic the biological deacetylation site H3K56. In each well, the reaction mixture contained 320 μM substrate, 3 mM NAD, in-house recombinant SIRT6 (4.5 μg/well) and SIRT assay buffer. All compounds were tested at 200 μM concentration. During compound screening the final DMSO concentration was 5 %. To initiate the reaction, SIRT6 was added to the reaction wells. The reactions and the remaining SIRT6 were incubated at 37 °C for 90 minutes. After that, 50 μl of developer solution

(6 μg/μl trypsin, 40 mM NA) was added to all wells and SIRT6 to the control wells. The plates were incubated for 30 minutes at room temperature and the fluorescence was measured with excitation and emission wavelengths of 380 and 440 nm, respectively, using EnVision 2104 Multilabel Reader (PerkinElmer, Finland). In the initial substrate profiling assays, the substrates were used at 400 μM concentration and SIRT6 was from a commercial source (Cayman Chemical Company). The results were calculated after the subtraction of the background fluorescence in the control wells.

# 4 Results

## 4.1 Virtual screening

Together 34 of compounds from Enamine and Chembridge databases were recommended for in vitro testing. The selection was based on docking score values and visual inspection of calculated poses – predicted interactions (especially H-bonds and ionic interactions), estimated stability of the compound and reality of protonation state were observed.

The tables 4.1 and 4.2 show compounds recommended for *in vitro* testing from the database Enamine and Chembridge. Several scoring values were calculated for the compounds in order to compare them (see tables 4.1 and 4.2).

Figures represent predicted interactions of SIRT6 with selected compounds from the database Enamine (figures 4.1 to 4.17) and Chembridge (figures 4.18 to 4.34). Figure legends include description of this interactions. Very frequent is ionic interaction of Arg63 with negative charged functions (usually sulfonamide, carboxylic or phenolic group). H-bonds are created most frequent with Arg63, H-interactions with Ser214 and backbone of Ala51, Thr213 and Trp69 are also usual.

| # | Enamine ID | London dG | GBVI/WSA dG | Induced Fit | Glide score |
|---|---|---|---|---|---|
| 1 | Z18857029 | -17,06 | -8,40 | -12,65 | -6,09 |
| 2 | Z24458677 | -14,71 | -9,19 | -11,33 | -6,65 |
| 3 | Z24531360 | -14,62 | -7,64 | -12,88 | -5,77 |
| 4 | Z255749124 | -14,24 | -7,88 | -12,79 | -6,16 |
| 5 | Z296717150 | -18,22 | -6,97 | -12,83 | -6,44 |
| 6 | Z324857542 | -16,57 | -5,37 | -10,73 | -6,50 |
| 7 | Z333042742 | -16,06 | -6,78 | -12,36 | -6,27 |
| 8 | Z426369012 | -16,08 | -8,10 | -9,76 | -6,42 |
| 9 | Z48488437 | -16,99 | -8,15 | -13,00 | -6,86 |
| 10 | Z49533201 | -16,75 | -7,06 | -11,22 | -6,08 |
| 11 | Z51124092 | -18,11 | -8,36 | -12,08 | -6,97 |
| 12 | Z54120220 | -16,94 | -7,76 | -9,86 | -7,01 |
| 13 | Z55358136 | -16,18 | -7,90 | -11,50 | -6,97 |
| 14 | Z55359425 | -16,19 | -8,02 | -11,73 | -8,20 |
| 15 | Z56773856 | -13,64 | -8,87 | -12,65 | -8,31 |
| 16 | Z56836580 | -17,47 | -7,14 | -11,18 | -6,45 |
| 17 | Z56846171 | -17,36 | -6,88 | -11,65 | -5,00 |

**Table 4.1:** Score values of compounds from database Enamine

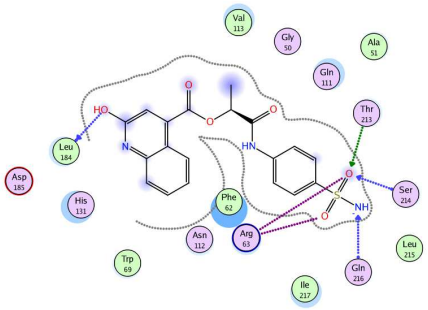



**Figure 4.1:** Compound #1 in SIRT6 – ionic interaction of sulfonamide group with Arg63, H-bonds with Leu184, Thr213, Ser214 and Gln216

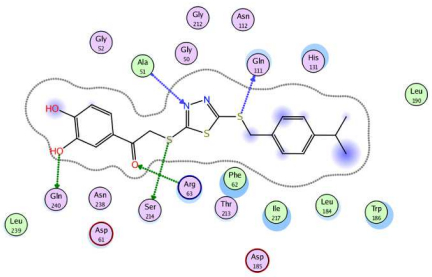



**Figure 4.2:** Compound #2 in SIRT6 – H-bonds Ala51, Arg63, Gln111, Ser214 and Gln240

| # | ZINC code | London dG | GBVI/WSA dG | Induced Fit | Glide score |
|---|---|---|---|---|---|
| 18 | ZINC67758065 | -14,603 | -8,018 | -11,578 | -7,810 |
| 19 | ZINC02779379 | -14,373 | -7,371 | -10,983 | -6,364 |
| 20 | ZINC13955124 | -12,745 | -7,285 | -11,015 | -6,535 |
| 21 | ZINC03146430 | -14,986 | -7,885 | -12,205 | -6,331 |
| 22 | ZINC05019574 | -12,305 | -7,104 | -9,544 | -6,192 |
| 23 | ZINC02873537 | -14,878 | -6,791 | -11,560 | -7,276 |
| 24 | ZINC04833719 | -11,937 | -6,704 | -10,664 | -6,242 |
| 25 | ZINC05101886 | -11,526 | -6,685 | -11,497 | -7,179 |
| 26 | ZINC22015281 | -17,792 | -7,608 | -11,485 | -4,966 |
| 27 | ZINC71639627 | -17,583 | -6,348 | -8,483 | -5,386 |
| 28 | ZINC04697420 | -16,701 | -6,910 | -9,544 | -6,192 |
| 29 | ZINC01233560 | -14,610 | -6,886 | -11,205 | -6,220 |
| 30 | ZINC02900986 | -15,699 | -7,363 | -10,795 | -5,903 |
| 31 | ZINC67758065 | -15,010 | -7,144 | -11,920 | -7,081 |
| 32 | ZINC97301462 | -14,686 | -6,845 | -10,513 | -7,001 |
| 33 | ZINC67758065 | -14,488 | -7,735 | -11,565 | -8,260 |
| 34 | ZINC01225795 | -17,258 | -7,603 | -11,400 | -6,338 |

**Table 4.2:** Score values of compounds from database Chembridge

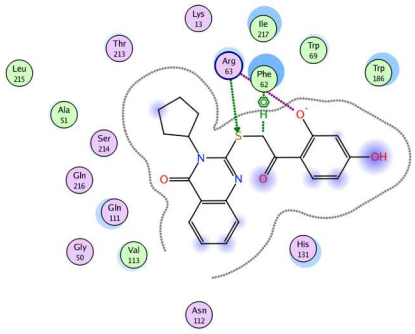


**Figure 4.3:** Compound #3 in SIRT6 – ionic interaction of phenolic oxygen with Arg63, H-bonds with Arg63, pi-interaction with Phe62

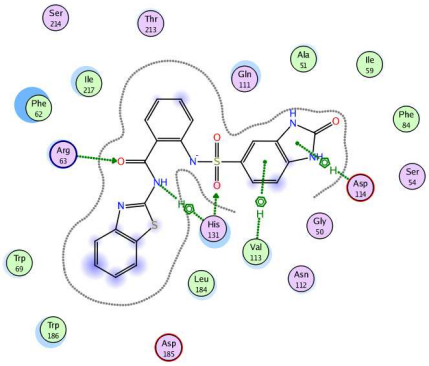


**Figure 4.4:** Compound #4 in SIRT6 – H-bonds with Arg63 and His131, pi-interactions with Val113, His131 and Asp134

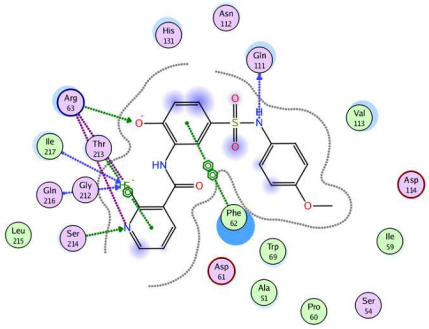


**Figure 4.5:** Compound #5 in SIRT6 – ionic interaction of thiophenolic group with Arg63, H-bonds with Arg63, Gln111, Gly212, Ser214, Gln216 and Ile217, pi-interaction with Thr213

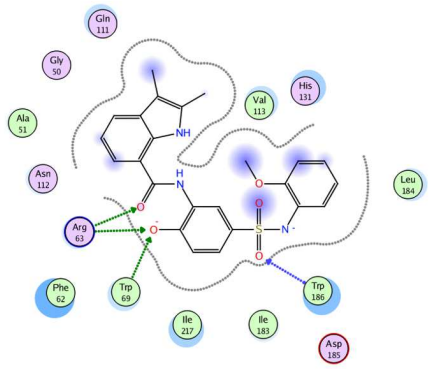


**Figure 4.6:** Compound #6 in SIRT6 – H-bonds with Arg63, Trp69 and Trp186

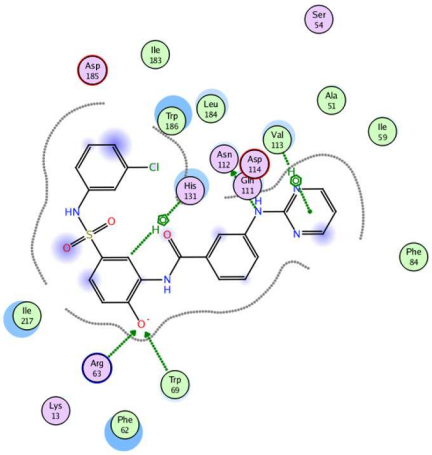



**Figure 4.7:** Compound #7 in SIRT6 – H-bonds with Arg63, Trp69 and Asn112, pi-interactions with Val113 and His131

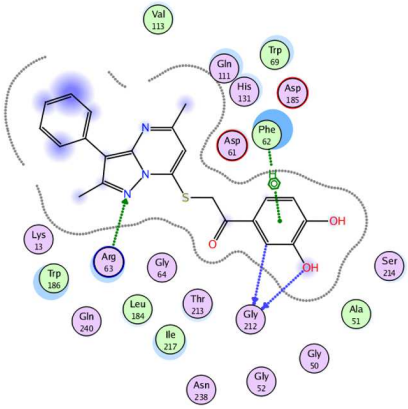



**Figure 4.8:** Compound #8 in SIRT6 – H-bonds with Phe62, Arg63 and Gly212

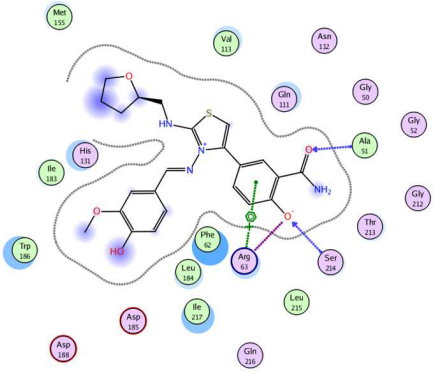



**Figure 4.9:** Compound #9 in SIRT6 – ionic interaction with Arg63, H-bonds with Ala51 and Ser214, cation-pi-interaction with Arg63

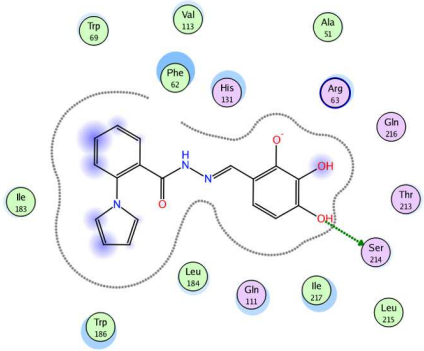



**Figure 4.10:** Compound #10 in SIRT6 – H-bond with Ser214

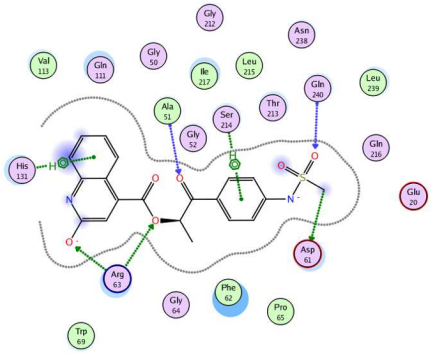



**Figure 4.11:** Compound #11 in SIRT6 – H-bonds with Ala51, Asp61, Arg63 and Gln240, pi-interactions with His131 and Ser214

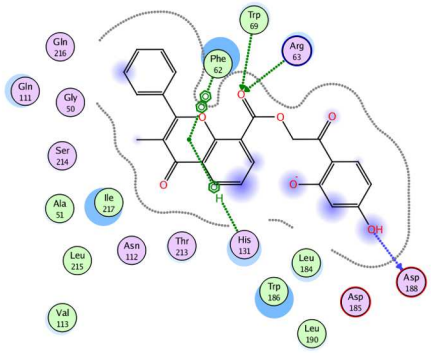



**Figure 4.12:** Compound #12 in SIRT6 – H-bonds with Arg63, Trp69 and Asp188, pi-interactions with Phe62 and His131

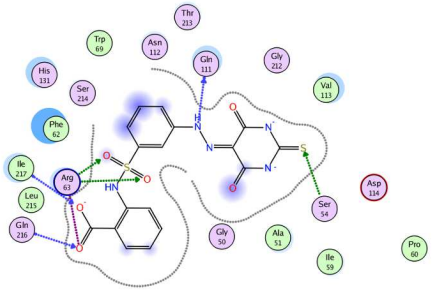



**Figure 4.13:** Compound #13 in SIRT6 – ionic interaction of carboxylic group with Arg63, H-bonds with Ser54, Arg63, Gln111, Gln216 and Ile217

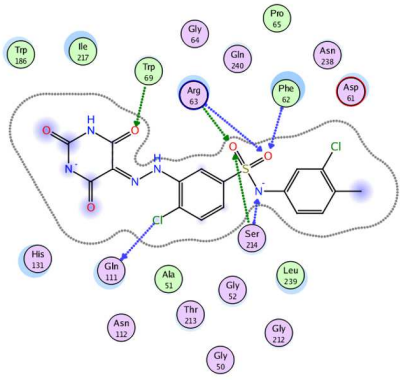



**Figure 4.14:** Compound #14 in SIRT6 – H-bonds with Phe62, Arg63, Trp69, Gln111 and Ser234

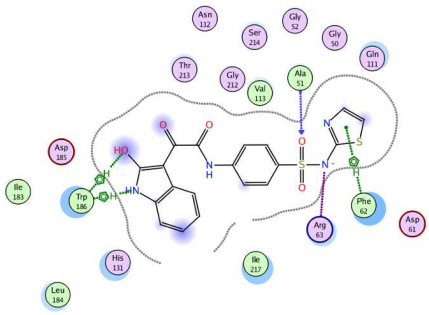



**Figure 4.15:** Compound #15 in SIRT6 – ionic interaction of sulphonamide group with Arg63, H-bond with Ala51, pi-interactions with Phe62 and Trp186

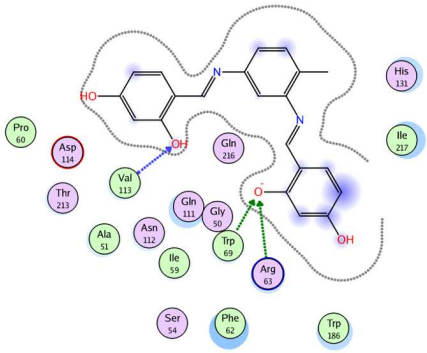



**Figure 4.16:** Compound #16 in SIRT6 – H-bonds with Arg63, Trp69 and Val113

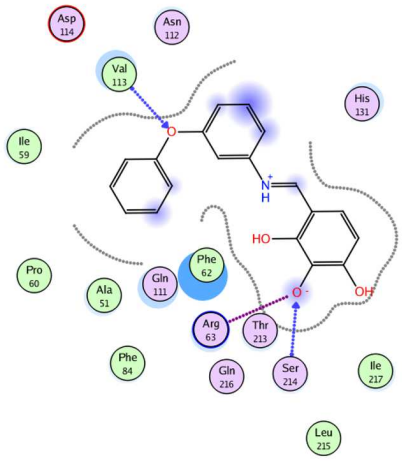



**Figure 4.17:** Compound #17 in SIRT6 – ionic intraction of phenolic group with Arg63, H-bonds with Val113 and Ser214

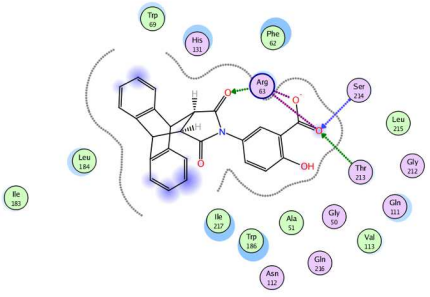



**Figure 4.18:** Compound #18 in SIRT6 – ionic interaction of carboxylic group with Arg63, H-bonds with Arg63, Thr213 and Ser234

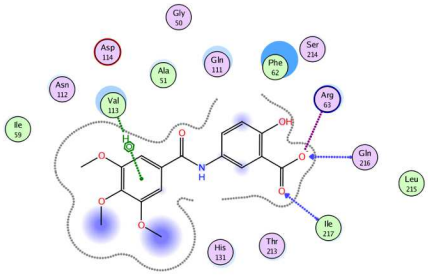



**Figure 4.19:** Compound #19 in SIRT6 – ionic interaction of carboxylic group with Arg63, H-bonds with Arg63, Gln216 and Ile217, pi-interaction with Val113

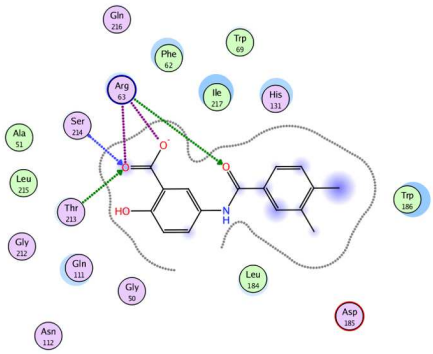



**Figure 4.20:** Compound #20 in SIRT6 – ionic interaction of carboxylic group with Arg63, H-bonds with Arg63, Thr213 and Ser214

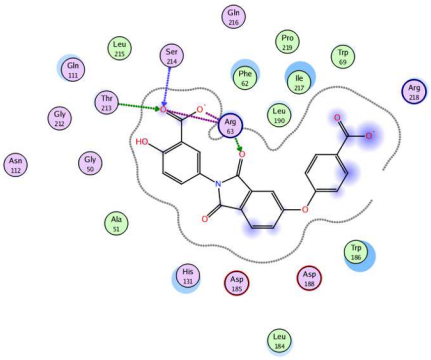



**Figure 4.21:** Compound #21 in SIRT6 – ionic interaction of carboxylic group with Arg63, H-bonds with Arg63, Thr213 and Ser214

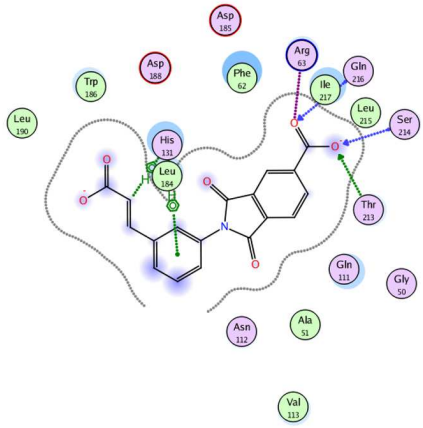



**Figure 4.22:** Compound #22 in SIRT6 – ionic interaction of carboxylic group with Arg63, H-bonds with Thr213, Ser214 and Gln216, pi-interactions with His131 and Leu184

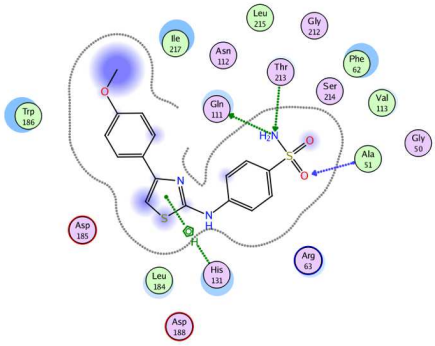



**Figure 4.23:** Compound #23 in SIRT6 – H-bonds with Ala51, Gln111 and Thr213, pi-interaction with His131

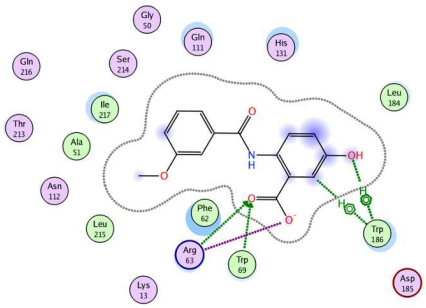



**Figure 4.24:** Compound #24 in SIRT6 – ionic interaction of carboxylic group with Arg63, H-bonds with Arg63 and Trp69, pi-interactions with Trp186

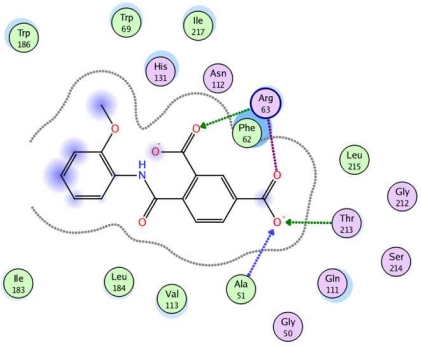



**Figure 4.25:** Compound #25 in SIRT6 – ionic interaction with Arg63, H-bonds with Ala51, Arg63 and Thr213

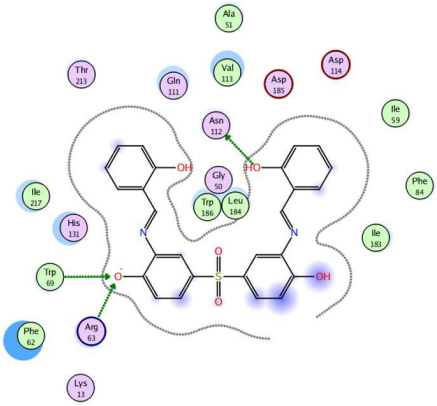



**Figure 4.26:** Compound #26 in SIRT6 – H-bonds with Arg63, Trp69 and Asn112

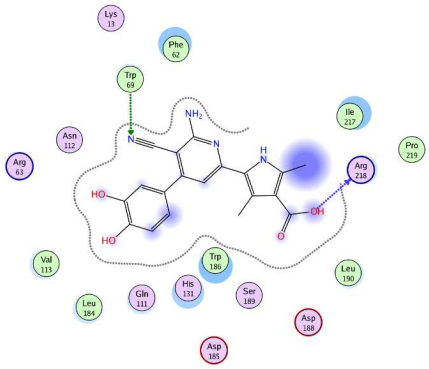


**Figure 4.27:** Compound #27 in SIRT6 – H-bonds with Arg63 and Trp69

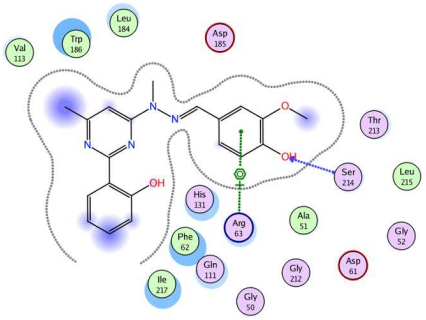


**Figure 4.28:** Compound #28 in SIRT6 – H-bond with Ser214, cation-pi-interaction with Arg63

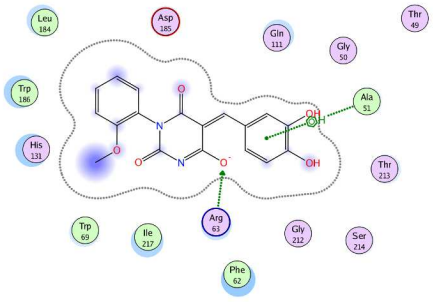


**Figure 4.29:** Compound #29 in SIRT6 – H-bond with Arg63, pi-interaction with Ala51

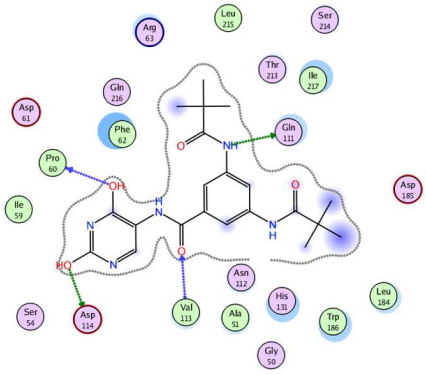


**Figure 4.30:** Compound #30 in SIRT6 – H-bonds with Pro60, Gln111 and Asp114

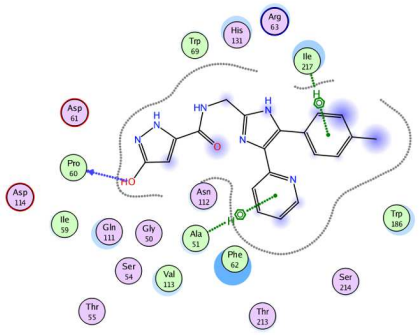



**Figure 4.31:** Compound #31 in SIRT6 – H-bond with Pro60, pi-intractions with Ala51 and Ile217

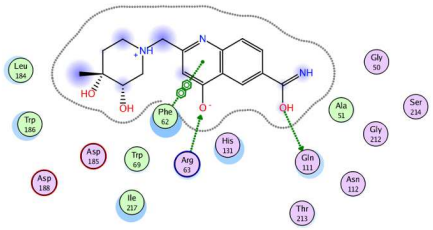



**Figure 4.32:** Compound #32 in SIRT6 – H-bonds with Arg63 and Gln111, pi-intraction with Phe62

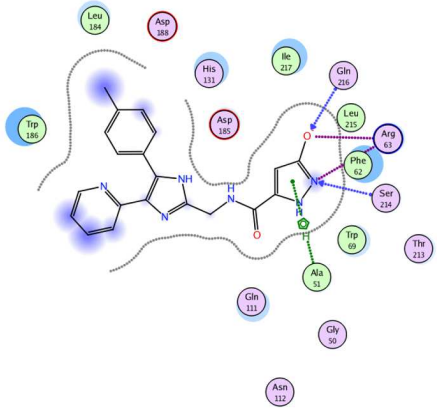



**Figure 4.33:** Compound #33 in SIRT6 – ionic interaction of phenolic group with Arg63, H-bonds with Ser214 and Gln216, pi-interaction with Ala51

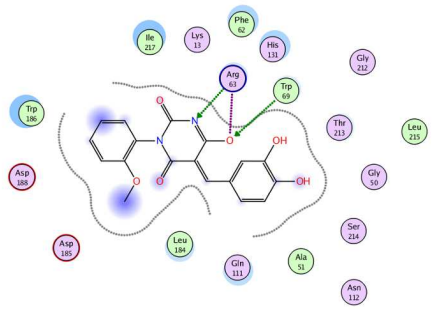



**Figure 4.34:** Compound #34 in SIRT6 – ionic interaction of imide group with Arg63, H-bonds with Arg63 and Trp69

Figures 4.35 to 4.44 shows structures of 10 selected compounds from the set offered by collaborating research group. Figures and figure legends includes representation and description of predicted interactions with SIRT6. The table 4.3 shows score values of this molecules.

| # | London dG |
|---|---|
| 35 | -16,0930 |
| 36 | -16,2868 |
| 37 | -16,1741 |
| 38 | -15,3426 |
| 39 | -15,1900 |
| 40 | -15,0790 |
| 41 | -14,5962 |
| 42 | -14,5928 |
| 43 | -14,4555 |
| 44 | -14,3617 |

**Table 4.3:** Score values of compounds #35 to #44

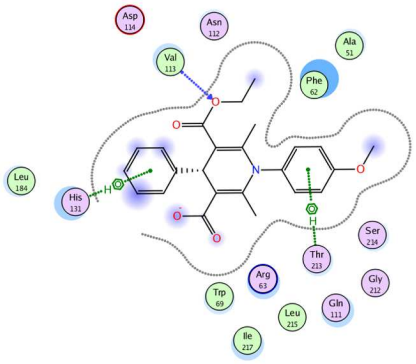



**Figure 4.35:** Compound #35 in SIRT6 – H-interaction with Val113, pi-interactions with His131 and Thr213

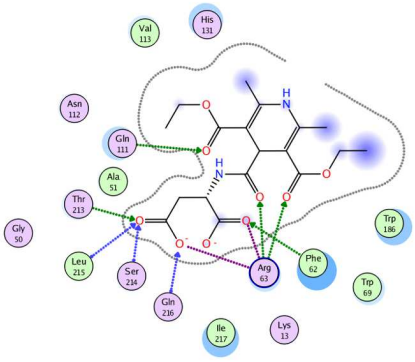



**Figure 4.36:** Compound #36 in SIRT6 – ionic interactions of carboxylic groups with Arg63, H-bonds with Phe62, Arg63, Gln111, Thr213, Ser214, Leu215 and Gln216

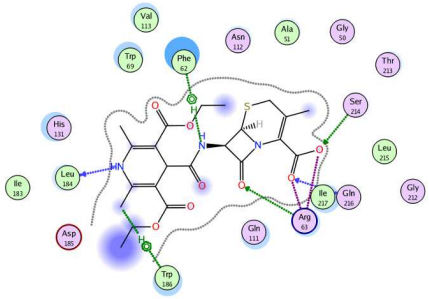


**Figure 4.37:** Compound #37 in SIRT6 – ionic interaction of carboxylic group with Arg63, H-bonds with Arg63, Leu184, Ser214 and Gln216, pi-interactions with Phe62 and Trp186

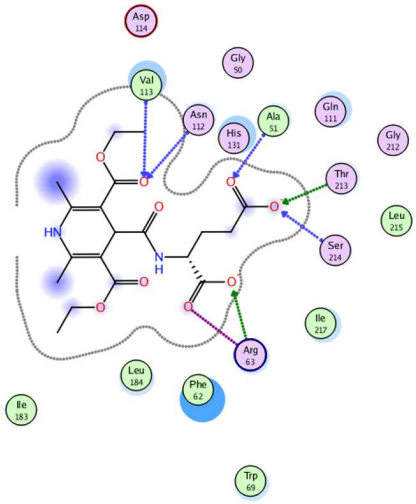


**Figure 4.38:** Compound #38 in SIRT6 – ionic interaction of carboxylic group with Arg63. H-bonds with Ala51, Arg63, Asn112, Val113, Thr213 and Ser214

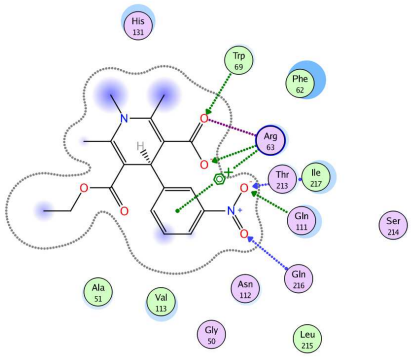


**Figure 4.39:** Compound #39 in SIRT6 – ionic interaction of carboxylic group with Arg63, H-bonds with Arg63, Trp69, Gln111, Thr213, Gln216 and Ile217, cation-pi-interaction with Arg63

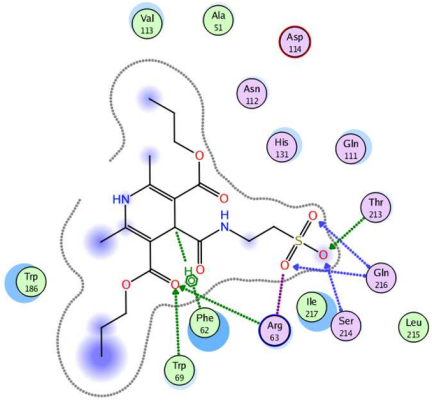


**Figure 4.40:** Compound #40 in SIRT6 – ionic intraction of sulphonic group with Arg63, H-bonds with Arg63, Trp69, Thr213, Ser214 and Gln216, pi-interaction with Phe62

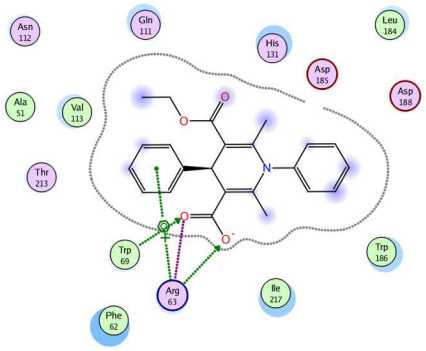



**Figure 4.41:** Compound #41 in SIRT6 – ionic interaction of carboxylic group with Arg63, H-bond with Trp69, cation-pi-interaction with Arg63

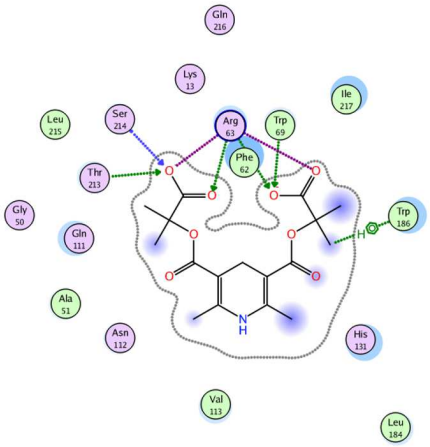



**Figure 4.42:** Compound #42 in SIRT6 – ionic interactions of carboxylic groups with Arg63, H-bonds with Arg63, Trp69, Thr213 and Ser214, pi-interaction with Trp186

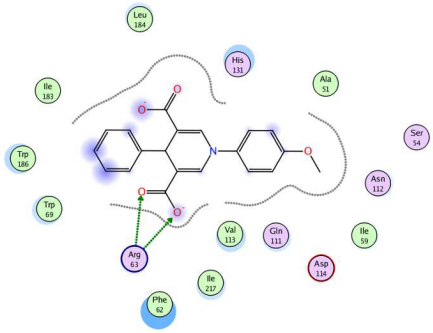



**Figure 4.43:** Compound #43 in SIRT6 – H-bonds with Arg63

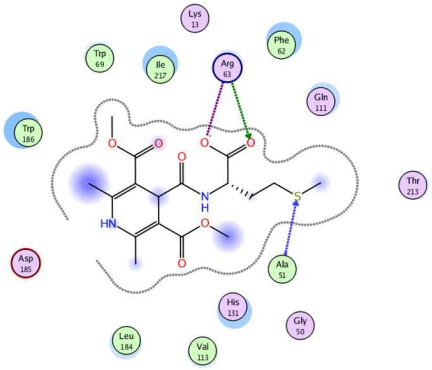



**Figure 4.44:** Compound #44 in SIRT6 – ionic interaction of carboxylic group with Arg63, H-bonds with Arg63 and Ala51

## 4.2 *In vitro* testing

Together 11 compounds from the set of 44 recommended (or similar in case of unavailability) have been tested so far. The inhibition values and standard deviations of the measurement are in the table 4.4:

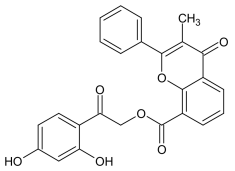
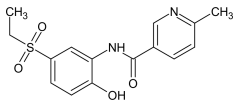
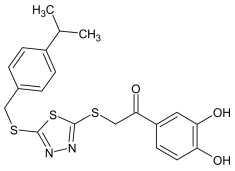
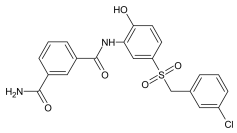
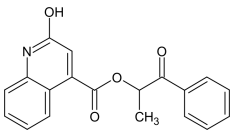
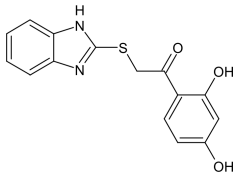
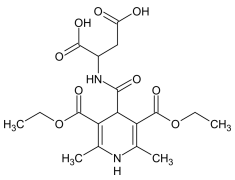
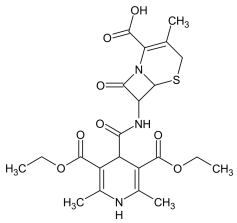
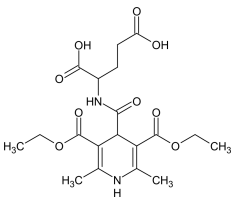
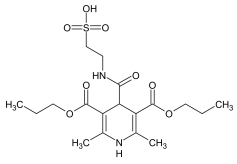
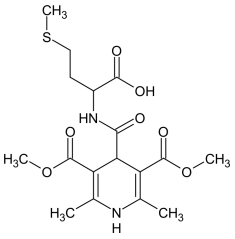
| | inhib. (%) | Sd | | inhib. (%) | Sd |
|---|---|---|---|---|---|
|  | 5.94 | 6.48 |  | -1.75 | 13.31 |
|  | 3.35 | 2.46 |  | 50.47 | 5.31 |
|  | -10.25 | 16.47 |  | -11.70 | 26.47 |
|  | 34.06 | 3.46 |  | 50.76 | 3.74 |
|  | 37.08 | 26.16 |  | 50.48 | 17.34 |
|  | 93.06 | 7.30 | | | |

**Table 4.4:** Inhibition values of tested compounds; inhib. = inhibition of SIRT6, Sd = standard deviation

# 5 Discussion

## 5.1 Methods

The screening process can be divided into three phases. The first one can be called high-throughput phase. The second one was docking in MOE and the third one docking computations for final selection.

Together four methods were selected for the high-throughput phase – docking, pharmacophoric search, fingerprint and fingerprint of fragments. All of them have some results and some hits in the set of 44 selected molecules.

### 5.1.1 Docking

The crystal structure of SIRT6 with adenosine-5-diphosphoribose (ADP) – PDB code 3K35 – were used for docking. We tested the X-ray structure of SIRT6 using the C-pocket as active site for docking and found out it suitable. In earlier published study of SIRT6 inhibitors Parenti *et al.* [46] used X-ray structure of SIRT6 for docking. In the work published by Kokkonen *et al.* [58] the homology model of SIRT6 as target for docking. This was due to the fact that they docked substrate-mimicking inhibitors, peptides and pseudopeptides into the substrate-binding site.

In SIRT6 the C-pocket is a putative binding site for inhibitors. To identify bioactive conformation Parenti *et al.* [46] modified the side chain conformation of Phe62 in the C-pocket of SIRT6. The new orientation of Phe62 was based on the X-ray structure of SIRT2 with nicotinamide. We tried make a similar change and we used the rotamer-library to select probable position of Phe62 side chain. The modified orientation of Phe62 were evaluated by docking EX-527 and comparing the interactions with X-ray of SIRT1. The original orientation of Phe62 were found to be best. Thus we kept the original orientation of the putative site as it is presented in X-ray structure.

MOE was selected for most docking runs but some dockings were performed with Schrödinger's Maestro. This was due to the fact that MOE were able to redocked the co-crystallized EX-527 into SIRT1 successfully, and the catalytic core of SIRT1 has high homology with SIRT6.

Docking of large databases is time-consuming. Other faster methods can be used for high-throughput screening like fingerprint and pharmacophore screening. However, hits from high-throughput docking was most frequent in the final set of molecules which were recommended for *in vitro* testing.

### 5.1.2 Pharmacophore

The pharmacophoric search was less computing-time demanding method. Unfortunately, this method was less successful. It gave lot of compounds which were rejected during the docking phase. The reason of this failure can be in imperfect definition of the pharmacophore. There are a few known inhibitors for SIRT6 and none of them are really potent. In addition no X-ray structure with bound ligand for SIRT6 is published. Thus there is no experimental evidence about the interactions needed for a potent inhibitor. This method has high demands on knowledge and experience of the researcher.

### 5.1.3 Fingerprint and fingerprint of fragments

The method BIT_MACC was selected for the descriptor computation. It was a bit more computing-time demanding than pharmacophoric search. The possibility of researcher's intervention was much lower in comparison to pharmacophore definition. The results of fingerprint approach is dependent on the selected template and it results similar compounds that templates. This might cause problems to identify a novel scaffolds.

In general, this technique gave lot of hits to the final set of 44 compounds.

## 5.2 *In vitro* results

After the careful visual inspection of the virtual screening results, 11 compounds were selected for in vitro testing. The sirtuin inhibition was determined at 200 $\mu$M as previously described with Fluor de Lys SIRT6 assay [3]. Two of tested compounds showed moderate inhibition ($> 30\%$), whereas four compounds showed $> 50\%$ inhibition of SIRT6 at 200 $\mu$M. While evaluating the data from the inhibitory activities

measurements, we observed that many of the tested compounds displayed autofluorescence and/or were colorful. If compounds are autofluorescent the observed inhibition can be underestimated. Generally, this is not a problem in sirtuin assays, since it is observed in the compound background wells and therefore it can be taken into account in the calculation of results. If compounds can be colorful, especially causing them to absorb the blue region wavelengths used in the assay, leading to exaggerated observed inhibition. In some cases a compound can also have both of these properties, and then the combined effect is difficult to take into consideration. As both problems were observed with tested compounds all compounds were sent to our collaborator Dr. Ruin Moaddel in National Institute of Health (NIH) to identify possible false positives. Dr. Ruin Moaddel uses assay with SIRT6 coated magnetic beads (SIRT6-MB) to identify the inhibition or activation of SIRT6. [59]

# 6 Conclusions

All aims of the work, which were set in the introduction part, were achieved.

In the theoretical part (chapter 2), most frequent used methods of ligand- and structure based virtual screening were described. Writing of the review helped to select and optimize used *in silico* methods during the experimental work.

Method of virtual screening suitable for this work were selected, optimized and used to identify small molecules to inhibit deacetylase activity of SIRT6.

Small molecules with inhibition effect on the deacetylase activity of SIRT6 were found using *in silico* methods. Together 44 compounds were recommended for *in vitro* testing. 11 compounds have been tested so far and four of them show significant inhibition activity on SIRT6. The hit rate is 36 %. The activity of true positive hits should be verified and then selected compounds can be used as lead structures for further design of SIRT6 inhibitors.

# Bibliography

[1] ALBERTS, Bruce. *Molecular biology of the cell.* 5th ed. New York: Garland Science, 2008, 1268 p. ISBN 08-153-4106-7.

[2] KAWAHARA, Tiara L. A.; MICHISHITA, Eriko; ADLER, Adam S.; DAMIAN, Mara; BERBER, Elisabeth; LIN, Meihong; MCCORD, Ron A.; ONGAIGUI, Kristine C. L.; BOXER, Lisa D.; CHANG Howard Y.; CHUA, Katrin F.. SIRT6 Links Histone H3 Lysine 9 Deacetylation to NF-$\kappa$B-Dependent Gene Expression and Organismal Life Span. *Cell* [online]. 2009, vol. 136, issue 1, pp. 62–74 [viewed 2015-04-13]. DOI: 10.1016/j.cell.2008.10.052. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0092867408014463. ISSN 0092-8674.

[3] KOKKONEN, Piia; RAHNASTO-RILLA, Minna; MELLINI, Paolo; JARHO, Elina; LAHTELA-KAKKONEN Maija; KOKKOLA, Tarja. Studying SIRT6 regulation using H3K56 based substrate and small molecules. *European Journal of Pharmaceutical Sciences* [online]. 2014, vol. 63, pp. 71–76 [viewed 2014-12-08]. DOI: 10.1016/j.ejps.2014.06.015. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0928098714002759. eISSN 1879-0720.

[4] MARKS, Paul A.; XU, Wen-Shan. Histone deacetylase inhibitors: Potential in cancer therapy. *Journal of Cellular Biochemistry* [online]. 2009, vol. 107, issue 4, pp. 600–608 [viewed 2015-04-13]. DOI: 10.1002/jcb.22185. Available from: http://doi.wiley.com/10.1002/jcb.22185. ISSN 0730-2312.

[5] SAUVE, Anthony A. Sirtuin chemical mechanisms. *Biochimica et Biophysica Acta (BBA) – Proteins and Proteomics* [online]. 2010, vol. 1804, issue 8, pp. 1591–1603 [viewed 2015-05-02]. DOI: 10.1016/j.bbapap.2010.01.021. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1570963910000361. ISSN 1570-9639.

[6] PAN, Patricia W.; FELDMAN, Jessica; DEVRIES, Mark K.; DONG, Aiping; EDWARDS Aled; DENU, John M. Structure and Biochemical

Functions of SIRT6. *Journal of Biological Chemistry*. 2011, vol. 286, issue 16, pp. 14575–14587. DOI: 10.1074/jbc.M111.218990. Available from: http://www.jbc.org/cgi/doi/10.1074/jbc.M111.218990. ISSN 0021-9258.

[7] YANG, Bo; ZWAANS, Bernadette M. M.; ECKERSDORFF, Mark; LOMBARD, David B. The sirtuin SIRT6 deacetylates H3 K56Ac in vivo to promote genomic stability. *Cell Cycle* [online]. 2014, vol. 8, issue 16, pp. 2662–2663 [viewed 2015-04-13]. DOI: 10.4161/cc.8.16.9329. Available from: http://www.tandfonline.com/doi/abs/10.4161/cc.8.16.9329. ISSN 1538-4101.

[8] YUAN, Jian; PU, Mintie; ZHANG, Zhiguo; LOU, Zhenkun. Histone H3-K56 acetylation is important for genomic stability in mammals. *Cell Cycle* [online]. 2014, vol. 8, issue 11, pp. 1747–1753 [viewed 2015-04-13]. DOI: 10.4161/cc.8.11.8620. Available from: http://www.tandfonline.com/doi/abs/10.4161/cc.8.11.8620. ISSN 1538-4101.

[9] KUGEL, Sita; MOSTOSLAVSKY, Raul. Chromatin and beyond: the multitasking roles for SIRT6. *Trends in Biochemical Sciences* [online]. 2014, vol. 39, issue 2, pp. 72–81 [viewed 2015-04-14]. DOI: 10.1016/j.tibs.2013.12.002. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0968000413002016. ISSN 0968-0004.

[10] ZWAANS, Bernadette M. M; LOMBARD,David B. Interplay between sirtuins, MYC and hypoxia-inducible factor in cancer-associated metabolic reprogramming. *Disease Models*. 2014, vol. 7, issue 9, pp. 1023–1032. DOI: 10.1242/dmm.016287. Available from: http://dmm.biologists.org/cgi/doi/10.1242/dmm.016287. eISSN 1754-8411.

[11] STAHURA, Florence; BAJORATH, Jürgen. New Methodologies for Ligand-Based Virtual Screening. *Current Pharmaceutical Design* [online]. 2005, vol. 11, issue 9, pp. 1189–1202 [viewed 2014-09-17]. DOI: 10.2174/1381612053507549. Available from: http://www.eurekaselect.com/openurl/content.php?genre=article&issn=1381-6128&volume=11&issue=9&spage=1189. ISSN 1381-6128.

[12] SPIWOK, Vojtěch; KRÁLOVÁ, Blanka. Virtuální screening léčiv. *Chemické listy* [online]. 2009, vol. 108, no. 1, pp. 52–55 [viewed 2014-09-17]. WOS: 000262772600008. Available from: http://www.chemicke-listy.cz/docs/full/2009_01_52-55.pdf. ISSN 1213-7103.

[13] VERMA, Jitender; KHEDKAR, Vijay; COUTINHO, Evans. 3D-QSAR in Drug Design – A Review. *Current Topics in Medicinal Chemistry* [online]. 2010, vol. 10, issue 1, pp. 95–115 [viewed 2015-04-07]. DOI: 10.2174/156802610790232260. Available from: http://www.eurekaselect.com/85600/article/3d-qsar-drug-design-review. ISSN 1568-0266.

[14] KUBINYI, Hugo. QSAR and 3D QSAR in drug design Part 1. *Drug Discovery Today* [online]. 1997, vol. 2, issue 11, pp. 457–467 [viewed 2014-11-02]. DOI: 10.1016/S1359-6446(97)01079-9. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1359644697010799. ISSN 1359-6446.

[15] DEMALE, Manoj G.; HARKE, Sanjay N.; KHAN, Firoz A. Kalan; SHINDE, Devanand B.; SANGSHETTI, Jaiprakash N. Recent Advances in Multidimensional QSAR (4D–6D): A Critical Review. *Mini-Reviews in Medicinal Chemistry* [online]. 2014, vol. 14, issue 1, pp. 35–55 [viewed 2014-11-02]. DOI: 10.2174/13895575113136660104. Available from: http://www.eurekaselect.com/117463/article. ISSN 1389-5575.

[16] CRAMER, Richard D., PATTERSON, David E.; BUNCE, Jeffrey D. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *Journal of the American Chemical Society* [online]. 1988, vol. 110, issue 18, pp. 5959–5967 [viewed 2014-11-02]. DOI: 10.1021/ja00226a005. Available from: http://pubs.acs.org/doi/abs/10.1021/ja00226a005. ISSN 0002-7863.

[17] TROPSHA, Alexander. Best Practices for QSAR Model Development, Validation, and Exploitation. *Molecular Informatics* [online]. 2010, vol. 29, 6–7, pp.

476–488 [viewed 2015-05-02]. DOI: 10.1002/minf.201000061. Available from: http://doi.wiley.com/10.1002/minf.201000061. ISSN 1868-1743.

[18] KUBINYI, Hugo. QSAR and 3D QSAR in drug design Part 2: applications and problems. *Drug Discovery Today* [online]. 1997, vol. 2, issue 12, pp. 538–546 [viewed 2014-11-15]. DOI: 10.1016/S1359-6446(97)01084-2. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1359644697010842. ISSN 1359-6446.

[19] VOGT, Martin; BAJORATH, Jürgen. Chemoinformatics: A view of the field and current trends in method development. *Bioorganic & Medicinal Chemistry* [online]. 2012, vol. 20, issue 18, pp. 5317–5323 [viewed 2015-02-24]. DOI: 10.1016/j.bmc.2012.03.030. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0968089612002209. ISSN 0968-0896.

[20] XUE, Ling; GODDEN, Jeffrey W.; BAJORATH, Jürgen. Database Searching for Compounds with Similar Biological Activity Using Short Binary Bit String Representations of Molecules. *Journal of Chemical Information and Modeling* [online]. 1999, vol. 39, issue 5, pp. 881–886 [viewed 2015-02-26]. DOI: 10.1021/ci990308d. Available from: http://pubs.acs.org/cgi-bin/doilookup/?10.1021/ci990308d. ISSN 0095-2338.

[21] SHERIDAN, Robert P.; KEARSLEY, Simon K. Why do we need so many chemical similarity search methods?. *Drug Discovery Today* [online]. 2002, vol. 7, issue 17, pp. 903–911 [viewed 2015-03-03]. DOI: 10.1016/S1359-6446(02)02411-X. Available from: http://linkinghub.elsevier.com/retrieve/pii/S135964460202411X. ISSN 1359-6446.

[22] PATTERSON, David E.; CRAMER, Richard D.; FERGUSON, Allan M.; CLARK, Robert D.; WEINBERGER, Laurence E. Neighborhood Behavior: A Useful Concept for Validation of "Molecular Diversity" Descriptors. *Journal of Medicinal Chemistry* [online]. 1996, vol. 39, issue 16, pp.

3049–3059 [viewed 2015-03-03]. DOI: 10.1021/jm960290n. Available from: http://pubs.acs.org/doi/abs/10.1021/jm960290n. ISSN 0022-2623.

[23] WILLETT, Peter. Similarity-based virtual screening using 2D fingerprints. *Drug Discovery Today* [online]. 2006, vol. 11, 23–24, pp. 1046–1053 [viewed 2015-04-07]. DOI: 10.1016/j.drudis.2006.10.005. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1359644606004193. ISSN 1359-6446.

[24] XUE, Ling; BAJORATH, Juergen. Molecular descriptors in chemoinformatics, computational combinatorial chemistry, and virtual screening. *Combinatorial Chemistry* [online]. 2000, vol. 3, issue 5, pp. 363–372 [viewed 2015-05-07]. DOI: 10.2174/1386207003331454. Available from: http://www.eurekaselect.com/openurl/content.php?genre=article. ISSN 1386-2073.

[25] WERMUTH, Camille G.; GANELLIN, Robin C.; LINDBERG, Per; MITSCHER, Lester. Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998). *Pure and applied chemistry. Chimie pure et appliqueé* [online]. 1998, vol. 70, issue 5, pp. 1129–1143 [viewed 2015-03-17]. DOI: 10.1351/pac199870051129. Available from: http://www.degruyter.com/view/j/pac.1998.70.issue-5/pac199870051129/pac199870051129.xml. ISSN 1365-3075.

[26] HUNG, Che-Lun a Chi-Chun CHEN. Computational Approaches for Drug Discovery. *Drug Development Research* [online]. 2014, vol. 75, issue 6, pp. 412–418 [viewed 2015-03-17]. DOI: 10.1002/ddr.21222. Available from: http://doi.wiley.com/10.1002/ddr.21222. eISSN 1098-2299.

[27] KUCHAŘ, Miroslav. *Výzkum a vývoj léčiv: studijní program: syntéza a výroba léčiv*. Vyd. 1. Praha: Vydavatelství VŠCHT, 2008. 167 p. ISBN 978-80-7080-677-7.

[28] HÖLTJE, Hans-Dieter. *Molecular modeling: basic principles and applications*. $2^{nd}$ ed. Weinheim: Wiley-VCH, 2003, 228 p. ISBN 35-273-0589-0.

[29] SOUSA, Sergio F.; RIBEIRO, Antonio J. M.; COIMBRA, J. T. S.; NEVES, R. P. P.; MARTINS, S. A.; MOORTHY, Nshn; FERNANDES P. A.; RAMOS, M. J. Protein-Ligand Docking in the New Millennium – A Retrospective of 10 Years in the Field. *Current Medicinal Chemistry*. 2013, vol. 20, no. 18, pp. 2296–2314. DOI: 10.2174/0929867311320180002. ISSN 0929-8673.

[30] MOUSTAKAS, Demetri T.; LANG, P. Therese; PEGG, Scott; PETTERSEN, Eric; KUNTZ, Irwin D.; BROOIJMANS, Natasja; RIZZO, Robert C. Development and validation of a modular, extensible docking program: DOCK 5. *Journal of Computer-Aided Molecular Design* [online]. 2006, vol. 20, 10–11, pp. 601–619 [viewed 2014-11-01]. DOI: 10.1007/s10822-006-9060-4. Available from: http://link.springer.com/10.1007/s10822-006-9060-4. ISSN 0920-654X.

[31] BOYD, Donald B.; LIPKOWITZ, Kenny B. *Reviews in computational chemistry, Volume 17*. New York: Wiley, 2001, 432 p. ISBN 04-713-9845-4.

[32] FISCHER, Daniel; LIN, Shuo Liang; WOLFSON, Haim L.; NUSSINOV, Ruth. A Geometry-based Suite of Molecular Docking Processes. *Journal of Molecular Biology* [online]. 1995, vol. 248, issue 2, pp. 459–477 [viewed 2014-11-01]. DOI: 10.1006/jmbi.1995.0234. Available from: http://www.idealibrary.com/links/doi/10.1006/jmbi.1995.0234. ISSN 0022-2836.

[33] GOODSELL, David S.; OLSON, Arthur J. Automated docking of substrates to proteins by simulated annealing. *Proteins: Structure, Function, and Genetics* [online]. 1990, vol. 8, issue 3, pp. 195–202 [viewed 2014-11-01]. DOI: 10.1002/prot.340080302. Available from: http://doi.wiley.com/10.1002/prot.340080302. ISSN 0887-3585.

[34] BROOIJMANS, Natasja; KUNTZ, Irwin D. Molecular Recognition and Docking Algorithms. *Annual Review of Biophysics and Biomolecular Structure* [online]. 2003, vol. 32, issue 1, pp. 335–373 [viewed 2015-02-16]. DOI: 10.1146/annurev.biophys.32.110601.142532. Available from: http://www.annualreviews.org/doi/abs/10.1146/annurev.biophys.32.110601.142532. ISSN 1056-8700.

[35] ELDRIDGE, Matthew D.; MURRAY, Christopher W.; AUTON, Timothy R.; PAOLINI, Gaia V.; MEE, Roger P. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *Journal of Computer-Aided Molecular Design* [online]. 1997, vol. 11, issue 5, pp. 425–445 [viewed 2014-10-29]. DOI: 10.1023/A:1007996124545. Available from: http://link.springer.com/10.1023/A:1007996124545. ISSN 0920-654X.

[36] Molecular Operating Environment (MOE); Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2013.

[37] GRINTER, Sam; ZOU, Xiaoqin. Challenges, Applications, and Recent Advances of Protein-Ligand Docking in Structure-Based Drug Design. *Molecules* [online]. 2014, vol. 19, issue 7, pp. 10150–10176 [viewed 2014-11-01]. DOI: 10.3390/molecules190710150. Available from: http://www.mdpi.com/1420-3049/19/7/10150/. ISSN 1420-3049.

[38] HALGREN, Thomas A. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry* [online]. 1996, vol. 17, 5–6, pp. 490–519 [viewed 2014-11-01]. DOI: 10.1002/(SICI)1096-987X(199604)17:5/6490::AID-JCC13.0.CO;2-P. Available from: http://doi.wiley.com/10.1002/%28SICI%291096-987X%28199604%2917%3A5/6%3C490%3A%3AAID-JCC1%3E3.0.CO%3B2-P. ISSN 0192-8651.

[39] MUEGGE, Ingo; MARTIN, Yvonne C. A General and Fast Scoring Function for Protein–Ligand Interactions: A Simplified Potential Approach. *Journal of Medicinal Chemistry* [online]. 1999, vol. 42, issue 5, pp. 791–804 [viewed 2014]. DOI: 10.1021/jm980536j. Available from: http://pubs.acs.org/doi/abs/10.1021/jm980536j. ISSN 0022-2623.

[40] CHARIFSON, Paul S.; CORKERY, Joseph J.; MURCKO, Mark A.; WALTERS, W. Patrick. Consensus Scoring: A Method for Obtaining Improved Hit Rates from Docking Databases of Three-Dimensional Structures into Proteins. *Journal of Medicinal Chemistry* [online]. 1999, vol. 42, issue 25,

pp. 5100–5109 [viewed 2014-10-29]. DOI: 10.1021/jm990352k. Available from: http://pubs.acs.org/doi/abs/10.1021/jm990352k. ISSN 0022-2623.

[41] PAN, Patricia W., Aiping DONG, Jinrong MIN, Cheryl ARROWSMITH a Aled EDWARDS. Crystal structure of human SIRT6. *To be Published* [online]. 2011 [viewed 2015-09-15]. DOI: 10.2210/pdb3k35/pdb. Available from: http://www.rcsb.org/pdb/explore/explore.do?structureId=3k35.

[42] PubChem Compound Database [online database]. CID 445154. 2005 [viewed 2014-09-12] Available from: http://pubchem.ncbi.nlm.nih.gov/compound/445154.

[43] PubChem Compound Database [online database]. CID 5512209. 2005 [viewed 2014-09-12] Available from: http://pubchem.ncbi.nlm.nih.gov/compound/5512209.

[44] PubChem Compound Database [online database]. CID 5113032. 2005 [viewed 2014-09-12] Available from: http://pubchem.ncbi.nlm.nih.gov/compound/5113032.

[45] PubChem Compound Database [online database]. CID 5280343. 2004 [viewed 2014-09-12] Available from: http://pubchem.ncbi.nlm.nih.gov/compound/5280343.

[46] PARENTI, Marco Daniele; GROZIO, Alessia; BAUER, Inga; GALENO, Lauretta; DAMONTE, Patrizia; MILLO, Enrico; SOCIALI, Giovanna; FRANCESCHI, Claudio; BALLESTRERO, Alberto; BRUZZONE, Santina; RIO, Alberto Del; NENCIONI, Alessio. Discovery of Novel and Selective SIRT6 Inhibitors. *Journal of Medicinal Chemistry* [online]. 2014, vol. 57, issue 11, pp. 4796–4804 [viewed 2015-05-07]. DOI: 10.1021/jm500487d. Available from: http://pubs.acs.org/doi/abs/10.1021/jm500487d. eISSN 1520-4804.

[47] ACD/ChemSketch, version 12.01. Advanced Chemistry Development, Inc., Toronto, On, Canada, www.acdlabs.com, 2010.

[48] Avogadro: an open-source molecular builder and visualization tool. Version 1.1.0. http://avogadro.openmolecules.net/

[49] 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 D.01. 2009.

[50] O'BOYLE, Noel M.; BANCK, Michael; JAMES, Craig A.; MORLEY, Chris; VANDERMEERSCH, Tim; HUTCHISON, Geoffrey R. Open Babel: An open chemical toolbox. *Journal of Cheminformatics* [online]. 2011, vol. 3, issue 1, pp. 33 [viewed 2014-09-20]. DOI: 10.1186/1758-2946-3-33. Available from: http://www.jcheminf.com/content/3/1/33. ISSN 1758-2946.

[51] Schrödinger Release 2014-1: LigPrep, version 2.9, Schrödinger, LLC, New York, NY, 2014.

[52] Schrödinger Release 2014-1: Canvas, version 1.9, Schrödinger, LLC, New York, NY, 2014.

[53] ZINC Catalog ChemBridge [online]. 2014 [viewed 2014-10-12]. Available from: http://zinc.docking.org/catalogs/chbr

[54] ZHAO, Xun; ALLISON, Dagart; CONDON, Bradley; ZHANG, Feiyu; GHEYI, Tarun; ZHANG, Aiping; ASHOK, Sheela; RUSSELL, Marijane; MACEWAN, Iain; QIAN, Yuewei; JAMISON, James A.; LUZ, John Gately. The 2.5 Å Crystal Structure of the SIRT1 Catalytic Domain Bound to Nicotinamide Adenine Dinucleotide (NAD+) and an Indole

(EX527 Analogue) Reveals a Novel Mechanism of Histone Deacetylase Inhibition. *Journal of Medicinal Chemistry* [online]. 2013, vol. 56, issue 3, pp. 963–969 [viewed 2014-09-11]. DOI: 10.1021/jm301431y. Available from: http://pubs.acs.org/doi/abs/10.1021/jm301431y. ISSN 0022-2623.

[55] HUANG, Niu; SHOICHET, Brian K.; IRWIN, John J. Benchmarking Sets for Molecular Docking. *Journal of Medicinal Chemistry* [online]. 2006, vol. 49, issue 23, pp. 6789–6801 [viewed 2014-09-17]. DOI: 10.1021/jm0608356. Available from: http://pubs.acs.org/doi/abs/10.1021/jm0608356. ISSN 0022-2623.

[56] Small-Molecule Drug Discovery Suite 2014-1: Glide, version 6.2, Schrödinger, LLC, New York, NY, 2014.

[57] Small-Molecule Drug Discovery Suite 2014-1: Schrödinger Suite 2014-1 Induced Fit Docking protocol; Glide version 6.2, Schrödinger, LLC, New York, NY, 2014; Prime version 3.5, Schrödinger, LLC, New York, NY, 2014.

[58] KOKKONEN, Piia; RAHNASTO-RILLA, Minna; KIVIRANTA, Päivi H.; HUHTINIEMI, Tero; LAITINEN, Tuomo; POSO, Antti; JARHO, Elina; LAHTELA-KAKKONEN, Maija. Peptides and Pseudopeptides as SIRT6 Deacetylation Inhibitors. *ACS Medicinal Chemistry Letters* [online]. 2012, vol. 3, issue 12, pp. 969–974 [viewed 2015-04-19]. DOI: 10.1021/ml300139n. Available from: http://pubs.acs.org/doi/abs/10.1021/ml300139n. ISSN 1948-5875.

[59] SINGH, Nidhi; RAVICHANDRAN, Sarangan; SPELMAN, Kevin; FUGMANN, Sebastian D.; MOADDEL, Ruin. The identification of a novel SIRT6 modulator from Trigonella foenum-graecum using ligand fishing with protein coated magnetic beads. *Journal of Chromatography B* [online]. 2014, vol. 968, pp. 105–111 [viewed 2015-05-08]. DOI: 10.1016/j.jchromb.2014.03.016. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1570023214001871. ISSN 1570-0232.

# List of abbreviations

| | |
|---|---|
| ADP | adenosine-5-diphosphoribose |
| AMBER99 | Assisted Model Building with Energy Refinement 99 |
| B3LYP | Becke 3-Parameter (Exchange), Lee, Yang and Parr |
| BIT_MACC | bit-packed MACCS structural keys |
| CoMFA | Comparative Molecular Field Analysis |
| CoSA | Comparative Spectral Analysis |
| DMSO | dimethylsulfoxid |
| EC50 | half maximal effective concentration |
| GBVI/WSA | Generalized-Born Volume Integral/Weighted Surface Area |
| GRID | Grid Independent Descriptors |
| GST | glutathione S-transferase |
| H3K56 | histone H3 lysine 56 |
| H3K9 | histone H3 lysine 9 |
| IC50 | half maximal inhibitory concentration |
| IPTG | isopropyl $\beta$-D-1-thiogalactopyranoside |
| $K_i$ | inhibitor binding affinity |
| logP | logarithm of partition-coefficient |
| MACCS | Molecular Access System |
| MMFF94x | Merck Molecular Force Field 94x |
| MOE | Molecular Operating Environment |
| NA | nicotinamide |
| NAD | nicotinamide adenine dinucleotide |
| PDB | Protein Data Bank |
| PLS | Partial Least Squares |
| QSAR | Quantitative Structure-Activity Relationship |
| SDS-PAGE | sodium dodecyl sulfate-polyacrylamide gel electrophoresis |
| SIRT1 | NAD-dependent deacetylase sirtuin-6 |
| SIRT6 | NAD-dependent deacetylase sirtuin-6 |
| SIRT6-MB | SIRT6 coated magnetic beads |

# List of tables

4.1 Score values of compounds from database Enamine . . . . . 28
4.2 Score values of compounds from database Chembridge . . . . . 29
4.3 Score values of compounds #35 to #44 . . . . . 38
4.4 Inhibition values of tested compounds . . . . . 41

# List of figures


1.1 Nucleosome . . . 3
1.2 Mechanism of acetyllysine deacetylation catalyzed by sirtuin . . . 4

2.1 General scheme of current drug design . . . 7
2.2 Example of fingerprint construction . . . 12
2.3 Geometry-based method . . . 13
2.4 Incremental Construction Method . . . 15

3.1 Scheme of the screening process . . . 19
3.2 2D representation of pharmacophore . . . 22

4.1 Compound #1 in SIRT6 . . . 28
4.2 Compound #2 in SITR6 . . . 28
4.3 Compound #3 in SIRT6 . . . 30
4.4 Compound #4 in SIRT6 . . . 30
4.5 Compound #5 in SIRT6 . . . 30
4.6 Compound #6 in SIRT6 . . . 30
4.7 Compound #7 in SIRT6 . . . 31
4.8 Compound #8 in SIRT6 . . . 31
4.9 Compound #9 in SIRT6 . . . 31
4.10 Compound #10 in SIRT6 . . . 31
4.11 Compound #11 in SIRT6 . . . 32
4.12 Compound #12 in SIRT6 . . . 32
4.13 Compound #13 in SIRT6 . . . 32
4.14 Compound #14 in SIRT6 . . . 32
4.15 Compound #15 in SIRT6 . . . 33
4.16 Compound #6 in SIRT6 . . . 33
4.17 Compound #17 in SIRT6 . . . 33
4.18 Compound #18 in SIRT6 . . . 33
4.19 Compound #19 in SIRT6 . . . 34
4.20 Compound #20 in SIRT6 . . . 34
4.21 Compound #21 in SIRT6 . . . 34

4.22 Compound #22 in SIRT6 . . . . . 34
4.23 Compound #23 in SIRT6 . . . . . 35
4.24 Compound #24 in SIRT6 . . . . . 35
4.25 Compound #25 in SIRT6 . . . . . 35
4.26 Compound #26 in SIRT6 . . . . . 35
4.27 Compound #27 in SIRT6 . . . . . 36
4.28 Compound #28 in SIRT6 . . . . . 36
4.29 Compound #29 in SIRT6 . . . . . 36
4.30 Compound #30 in SIRT6 . . . . . 36
4.31 Compound #31 in SIRT6 . . . . . 37
4.32 Compound #32 in SIRT6 . . . . . 37
4.33 Compound #33 in SIRT6 . . . . . 37
4.34 Compound #34 in SIRT6 . . . . . 37
4.35 Compound #35 in SIRT6 . . . . . 38
4.36 Compound #36 in SITR6 . . . . . 38
4.37 Compound #37 in SIRT6 . . . . . 39
4.38 Compound #38 in SIRT6 . . . . . 39
4.39 Compound #39 in SIRT6 . . . . . 39
4.40 Compound #40 in SITR6 . . . . . 39
4.41 Compound #41 in SIRT6 . . . . . 40
4.42 Compound #42 in SITR6 . . . . . 40
4.43 Compound #43 in SIRT6 . . . . . 40
4.44 Compound #44 in SITR6 . . . . . 40