Electronic Supplementary Material (ESI) for Medicinal Chemistry Communications

Binding free energy calculations and biological testing of novel thiobarbiturates as inhibitors of the human NAD$^+$ dependent histone deacetylase Sirt2

Electronic supplementary information

UrszulaUciechowska[a], Jörg Schemies[b], Michael Scharfe[a], Michael Lawson[a], Kanin Wichapong[a], Manfred Jung[b], and Wolfgang Sippl[a]*

[a] Department of Pharmaceutical Chemistry, Martin-Luther Universität Halle-Wittenberg, 06120 Halle/Saale, Germany. E-mail: wolfgang.sippl@pharmazie.uni-halle.de; Tel: +49-345-55-25040; Fax: +49-345-55-27355

[b] Institute of Pharmaceutical Sciences, Albert-Ludwigs-Universität Freiburg (Germany)

Contents of supplementary information

1.Computational methods

All calculations were performed on a Pentium IV 2.2 GHz based Linux cluster (20 CPUs). The GOLD software (Cambridge Crystallographic Data Centre)[1] was used for docking, whereas the calculation of all molecular descriptors and the analysis of the docking results were carried out in MOE2008.10 (Chemical Computing Group).[2]

1.1. Virtual Screening

Virtual Screening (VS) was performed using the Chembridge database to identify novel Sirt2 inhibitors. MACCS key fingerprints were used to search the database for compounds similar to the most active thioabarbiturates from our previous work (compounds **3** and **6**, Figure 1). Applying a Tanimoto coefficient of 0.70 we identified 637 compounds which were subsequently filtered by applying the following criteria: MW > 500, logP< 4, topological polar surface area TPSA < 140 $Å^2$. The resulting 510 molecules were docked in the Sirt2 binding pocket using the GOLD program. 129 compounds showed a Goldscore higher than 40 and were further analysed using calculated molecular interaction fields of the Sirt2 binding pocket. 14 thiobarbiturate derivatives were manually selected after visual inspection of their binding mode and their biological data were predicted using the MM-PB/SA and LIE models. The selected compounds were all among the first 33% of the 129 top-ranked docking hits.

1.2. Ligand Docking

The crystal structure of human Sirt2 was taken from the Protein Data Bank (PDB ID: 1J8F, chain B) and was used for docking calculations with the GOLD 4.0 program as already described in previous studies.[3,4] The protein preparation was performed using the MOE2008.10 program. This involves hydrogen addition and restrained minimization of the protein. Similar to our previous study[3] crystal water molecules located in the narrow cavity nearby the active side were used for analysis. We docked all compounds which were firstly rescored with the scoring functions available in GOLD program and only the top-ranked docking poses were taken for further analysis. To further support the docking mode the molecular interaction fields of the binding pocket were calculated and compared with the location of the docking poses. As an example the favourable interaction field calculated with the hydrophobic methyl probe (C3, MOE2008.10) is shown in comparison with the docking pose of compound **16** in Figure S1.

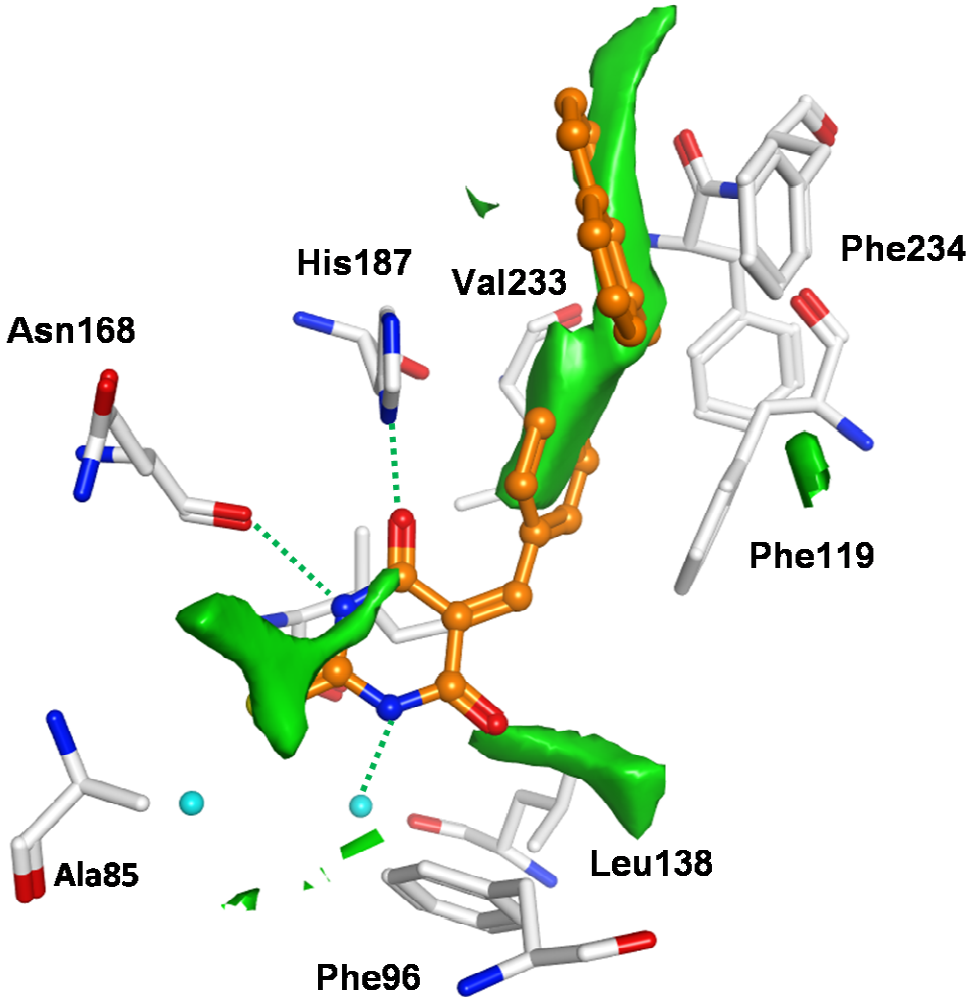

Figure S1. Favourable interaction region for a hydrophobic methyl probe (coloured green, contour level -2.6 kcal/mol) at the Sirt2 binding pocket. The most favourable interaction field is observed in the acetyl lysine binding channel of Sirt2. In comparison the docking pose of inhibitor **16** (coloured orange) is shown.

1.3 Molecular Dynamics Simulation

Molecular dynamics (MD) were carried out using AMBER 9.0 and the AMBER 1999 force field.[5,6] The initial structure of the Sirt2-inhibitor complex was taken for each compound from the GOLD docking study. The ligand force fields parameters were taken from the general Amber force field (GAFF), whereas AM1 ESP atomic partial charges were assigned to the inhibitors. The complexes were soaked in a box of TIP3P water molecules with a margin of 10 Å. Prior to the free MD simulations, two steps of relaxation were carried out; in the first step, we kept the protein fixed with a constraint of 500 kcal $mol^{-1}Å^{-1}$. In the second step, the inhibitor structures were relaxed for 0.5 ps, during which the protein atoms were

restrained to the X-ray coordinates with a force constant of 500 kcal $mol^{-1}Å^{-1}$. In the final step, all restraints were removed and the complexes were relaxed for 1 ps. The temperature of the relaxed system was then equilibrated at 300K through 20 ps of MD using 2 fs time steps. A constant volume periodic boundary was set to equilibrate the temperature of the system by the Langevin dynamics[7] using a collision frequency of 10 ps^{-1} and a velocity limit of 5 temperature units. During the temperature equilibration routine, the complex in the solvent box was restrained to the initial coordinates with a weak force constant of 10 kcal $mol^{-1}Å^{-1}$. The final coordinates of the temperature equilibration routine (after 20 ps) was then used to complete a 1 ns molecular dynamics routine using 2 fs time steps, during which the temperature was kept at 300 K. For the Langevin dynamics a collision frequency of 1 ps^{-1} and a velocity limit of 20 temperature units were used. The pressure of the solvated system was equilibrated at 1 bar at a certain density in a constant pressure periodic boundary by an isotropic pressure scaling method employing a pressure relaxation time of 2 ps. The time step of the free MD simulations was 2 fs with a cut-off of 9 Å for the non-bonded interaction, and SHAKE[8] was employed to keep all bonds involving hydrogen atoms rigid. Electrostatic interactions were computed using the Particle Mesh Ewald method.[9] The MD simulations of the Sirt2-inhibitor complexes were performed in total for 5 ns and gave low root mean square deviation (RMSD) values for the protein and bound inhibitor. Exemplarily the RMSD plot of the simulation of compound **3** is shown in Figure S2.

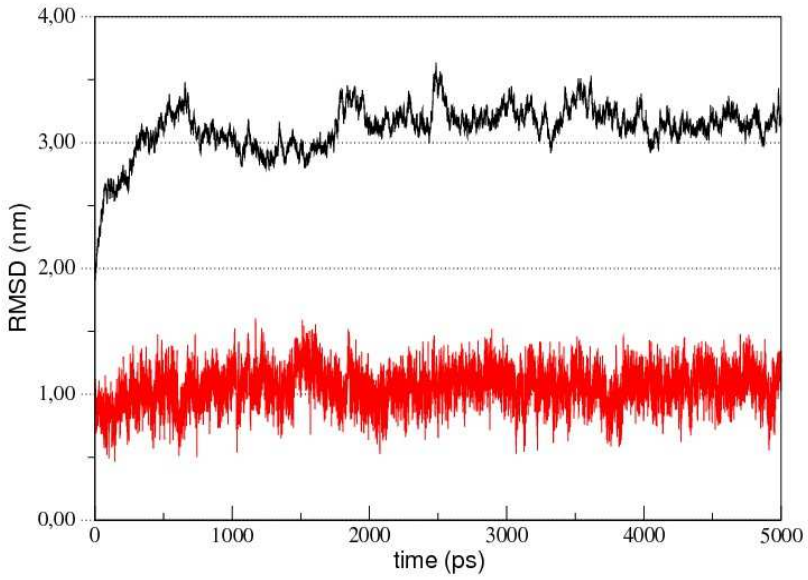

Figure S2. RMSD plot representing human Sirt2 (black) and bound compound **3** (red).

Simulations of free ligands in water were performed to estimate energies necessary for calculating binding affinities using LIE approach. Standard MD simulations were run on a ligand in water. Box of water molecules (TIP3P) were used with a minimum distance between the ligand and the boundaries of 20 Å. MD simulations for (1 ns) were carried out similar as described above for receptor-ligand complex, but no restraints were applied. Based on generated snapshots we ran single step minimization just to get the VDW and electrostatic contribution and an average value was taken.

1.4. MM-PB/SA Calculations

The MM-PB/SA and MM-GB/SA methods are characterized by the use of Poisson-Boltzmann (PB) and Generalized Born[10] models to compute the electrostatic component of the solvation free energy. The binding free energy of the protein-ligand complex is approximated by the following equation:

$$\Delta G = \Delta H - T\Delta S \quad \textbf{(3)}$$

T is the temperature of the system at 300 Kelvin. The binding free energy (ΔG) of the protein-ligand complex is computed as:

$$\Delta G = G_{\mathbf{complex}} - [\, G_{\mathbf{protein}} + G_{\mathbf{ligand}} \,] \quad \textbf{(4)}$$

In equation 4, $G_{\mathbf{complex}}$ is the absolute free energy of the complex, $G_{\mathbf{protein}}$ is the absolute free energy of the protein, and $G_{\mathbf{ligand}}$ is the absolute free energy of the ligand. We extracted 100 snapshots (at time intervals of 2 ps) for each species (complex, protein and ligand). The enthalpy term in equation 1 is dissected into sub-energy terms:

$$H_{\mathbf{tot}} = H_{\mathbf{gas}} + G_{\mathbf{solv}} \quad \textbf{(5)}$$

$$H_{\mathbf{gas}} = E_{\mathbf{el}} + E_{\mathbf{vdW}} + E_{\mathbf{int}} \quad \textbf{(6)}$$

$H_{\mathbf{gas}}$ is the potential energy of the solute which is determined as the sum of van der Waals (E_{vdw}), electrostatic (E_{el}) and internal energies (E_{int}) in gas phase by using the SANDER module of AMBER 9.0.[11] $G_{\mathbf{solv}}$ is the solvation free energy for transferring the solute from vacuum into solvent and is a sum of electrostatic ($G_{\mathbf{el}}$) and non-electrostatic (hydrophobic) contributions ($G_{\mathbf{nonel}}$) as shown in equation 7:

$$G_{\mathbf{solv}} = G_{\mathbf{el}} + G_{\mathbf{nonel}} \quad \textbf{(7)}$$

G_{el} in equation 7 was computed at 0.15 M salt concentration by the PBSA module of Amber 9.0 by dividing implicitly solvated solute species into 0.4Å cubic grid points and summing up the electrostatic potentials computed at each grid point. Electrostatic potential $Ø_{(r)}$ at a grid point *r* that is not at the solvent-solute boundary was computed by a linear Poisson Boltzmann (PB) equation, which is a three dimensional vector differential equation as shown in equation 8:

$$\nabla \varepsilon_{(r)} \nabla Ø_{(r)} = -4\pi . \rho_{(r)} \quad \textbf{(8)}$$

In equation 8 $\varepsilon_{(r)}$ is the dielectric constant ($\varepsilon = 1$ for the solute interior and $\varepsilon = 80$ for implicit PB water) and $\rho_{(r)}$is the charge density. The grid point potentials were then summed up for each atom ***i*** to yield atomic potentials $Ø_i$**.** The PB implicit solvent molecules at the solute-solvent boundary were allowed to energetically converge over 1000 iterations before the single-point Poisson computations (PBSA) were applied for each snapshot. The total entropy (S_{tot}), as formulated in equation 9 arose from changes in the degree of freedom:

$$S_{tot} = S_{trans} + S_{rot} + S_{vib} \quad \textbf{(9)}$$

In equation 9 (S_{trans}) is the translational, (S_{rot}) the rotational, (S_{rot}), and the vibrational (S_{vib})] entropy of each species.

Considering all absolute energy terms as given in equation 2, the binding free energy ΔG takes the following form:

$$\Delta G_{binding} = [\Delta H_{gas} + \Delta G_{solv}] - T\Delta S_{tot} \quad \textbf{(10)}$$

Parameter/topology files used in MM-PB/SA computations were prepared for the complex, the protein and the inhibitors using the LEAP module. Snapshots extracted from trajectories were pre-minimized in the gas phase by the SANDER module using a conjugate gradient method until the root-mean-square-deviation of the elements of the gradient vector was less than 10^{-4}kcal/mol^{-1}Å^{-1}. Frequencies of the vibrational modes were computed at 300 K for these minimized structures including all snapshot atoms and using a harmonic approximation of the energies. The energy contributions to the free energy of ligand binding of the 14 novel thiobarbiturates are listed in Table S1.

Table S1: Energy contributions to the free energy of binding of the 14 novel thiobarbiturate compounds (test set) obtained using the MM-PB/SA models. ΔE_{el} and ΔE_{vdw} are the electrostatic and van der Waals energies of binding, respectively, ΔE_{GBSA} are contributions to the solvation free energy, ΔH_{tot} is the enthalpy of binding, $T\Delta S_{tot}$ is the entropy of binding, and ΔG_{calc} is the calculated binding free energy. ΔG_{exp} values were calculated by the following equation: ΔG_{exp}=-RT ln(pIC_{50}).

Cpd.	ΔE_{el}	ΔE_{vdw}	ΔE_{sol}	ΔH_{tot}	$T\Delta S_{tot}$	ΔG_{calc}	ΔG_{exp}	pIC_{50}
11	-18.54	-34.70	30.06	-23.17	-16.78	-6.39	-7.11	5.23
12	-11.63	-37.42	24.66	-24.39	-22.91	-1.48	-7.11	5.23
13	-6.53	-37.38	20.27	-23.64	-20.89	-2.75	-7.21	5.30
14	-16.21	-33.30	24.42	-25.09	-19.89	-5.20	-7.80	5.74
15	-17.41	-42.30	32.26	-27.91	-22.31	-5.60	-7.34	5.40
16	-16.03	-43.68	31.05	-28.66	-18.25	-10.41	-7.59	5.58
17	-4.25	-32.16	11.07	-25.34	-19.49	-5.85	-7.91	5.82
18	-19.19	-39.51	33.10	-25.60	-17.06	-8.54	-7.15	5.26
19	-16.06	-34.72	27.67	-23.11	-16.59	-6.52	-7.69	5.19
20	-6.81	-43.13	24.03	-25.90	-18.28	-7.62	-7.05	5.66
21	-15.31	-41.75	30.00	-25.06	-15.61	-9.45	-7.36	5.42
22	-14.26	-40.19	28.63	-25.82	-19.33	-6.49	-7.54	5.55
23	-14.62	-38.58	28.26	-24.95	-18.25	-6.70	-7.41	5.46
24	-13.92	-43.33	32.31	-24.91	-17.42	-7.49	-7.12	5.24

1.5. LIE Calculations

A number of different computational approaches have been developed over the years for predicting binding free energies. Recently Aqvist *et al.* proposed a method for estimation binding affinities, known as linear interaction energy approximation LIE.[12] This is a semi-empirical method, based on linear response theory, which is less computationally expensive than free energy perturbation methods (FEP). This method relies on the simulation of several different states, mostly unphysical, whereas the LIE approach uses the initial and final states of the binding process, which is the free and bound state of the ligand. LIE is also faster than FEP and MM-PB/SA uses an explicit solvent model, which means that desolvation can be reasonably handled. These advantages make LIE a very useful tool in structure-based lead

optimization, which helps to understand detailed interactions between the lead compounds and their receptor, and to estimate binding affinities.

The concept of the LIE approach is to separately calculate the van der Waals and electrostatic interaction energies for the ligand in water and the ligand in complex with solvated protein.[12] Then, averages of interaction energies between the ligand and its surrounding are analyzed. The LIE equation is known as:

$$\Delta G_{calc} = \alpha \Delta E_{vdw} + \beta \Delta E_{elec} + \gamma = \alpha (E_{B\_vdw} - E_{F\_vdw}) + \beta (E_{B\_ele} - E_{F\_ele}) + \gamma \quad \textbf{(11)}$$

In equation 11 the Δ term indicates the change in energy from the ligand free and bound state (E_{bound}-E_{Free}). The α, β and γ are LIE empirical parameters, determined by comparing calculated and experimentally estimated binding affinities. Obtaining suitable values α and β has been subject of several investigations described recently. Aqvist *et al.* have found that α≈0.5 and β≈1.043 gave the best results for correlating the calculated binding free energies with the experimental values.[12,13] On the other hand Jones-Herzog and Jorgensen observed that α≈0.5 and β≈0.161 are not optimal for sulfonamide inhibitors with human thrombin.[14] Proper fitting parameters must be determined by comparing calculated and experimentally estimated binding affinities, they also depend on the investigated system and the force field. The γ parameter is a so-called additional constant, which sometimes needs to be added in order to obtain reasonable binding free energy predictions.[15] In this study the γ constant was set to zero and also used as γ≠0 in order to compare the predicted affinities.

For the LIE approach the MD simulations over 1 ns of receptor-bound ligands as well as ligands free in water were carried out to obtain the van der Waals and electrostatic interaction energies between the ligand and its surroundings. We run standard MD simulations on the protein-ligand complex in a solvated system and the ligand in a system where it is just solvated. After MD simulations we got snapshots of each ligand in the different environments which were used for calculating electrostatic and van der Waals energies. Energies were averaged over frames taken from the last 100 ps of the trajectories. Fitting parameters α and β, were determined by comparing calculated energies and experimentally estimated binding affinities using linear regression analysis.

The experimental binding energies (ΔG_{exp}) were calculated from the measured inhibition constants (IC_{50}) by equation 12, using the gas constant (R) and the temperature (T).

$$\Delta G_{exp} = -RT \ln IC_{50} \quad \textbf{(12)}$$

The predictive ability of the MM-PB/SA as well as the LIE models were tested by calculating the predictive correlation coefficient r^2_{pred}, which is calculated by the following formula (13):

$$r^2_{PRED} = (SD - PRESS)/SD \quad \textbf{(13)}$$

In equation 13 SD is the sum of squared deviations between the biological activities of the test set and mean activity of the training set compounds, PRESS is the sum of squared deviations between experimental and predicted activities of the test set compounds.

2. *In vitro* Screening

Fluorescent deacetylase assay.

All compounds purchased from Chembridge Corporation (San Diego, USA) were evaluated for their ability to inhibit recombinant Sirt2 using a homogeneous fluorescent deacetylase assay. The inhibitors were solved in DMSO and 3 µL or less of the inhibitor in DMSO was added to an incubation mixture. A probe with only DMSO was used as a negative control. The assay was carried out in 96-well plates with a reaction volume of 60 µL containing the fluorescent histone deacetylase substrate ZMAL (10.5 µM), NAD^+ (500 µM) and Sirt2. Enzyme volume was dependent on the activity of the preparation. Substrate conversion was between 10-30 % without inhibitor. After incubation time of 4 h at 37 ^oC, the deacetylation reaction was stopped with a solution of trypsin buffer (60 µL) containing trypsin (6 mg mL^{-1}) from bovine pancreas (10 000 BAEE units mg^{-1}) and the sirtuin inhibitor nicotinamide (8 mM). The microplate was incubated with this solution for 20 min at 37 ^oC, the fluorescence intensity was then measured in a plate reader (BMG Polarstar) with a coumarin filter (λ_{exc}=390 nm, λ_{em}= 460 nm). Inhibition rates were determined in reference to the DMSO control. For all determinations at least duplicates were carried out. IC_{50} values were determined using GraphPad Prism software.

References

1. Jones, G., Willett, P., Glen, R. C., Leach, A. R., Taylor, R. *J. Mol. Biol*. 1997, **267**, 727-748.
2. C. C. G. Inc., Molecular Operating Enviroment (MOE2008.10), Montreal, Canada, 2010.
3. Uciechowska, U., Schemies, J., Neugebauer, R., Huda, E., Schmitt, M, Meier, R., Verdin, E., Jung, M., Sippl, W. *Chem.Med.Chem*., 2008, **3**, 1965-1976
4. Neugebauer, R.C., Uchiechowska, U., Meier, R., Hruby, H., Valkov, V., Verdin, E., Sippl, W., Jung, M.*J. Med. Chem.***51**, 2008, 1203-1213.
5. Hornak, H., Abel, R., Okur, A., Strockbine, B., Roitberg, A. and C. Simmerling. *Proteins*, 2006, **65**, 712-25.
6. Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A. & Case, D. *Journal of Computational Chemistry*, 2004, **25**, 1157-1174.
7. Pastor, R.W., Brooks, B.R. &Szabo, A. *Mol. Phys*1988,**65**, 1409-1419.
8. Ryckaert, J.P., Ciccotti, G. &Berendsen, H.J.C. N. *J. Comput. Phys.*, 1977, **23**, 327-341.
9. Darden, T., York, D. & Pedersen, L. *J. Chem. Phys.*, 1993,**103**, 8577-8593.
10. Gohlke, H. & Case, D.A. *J. Comput. Chem.* 2004,**25**, 238-250.
11. Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, Jr., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., Kollman, P.A. *J. Am. Chem. Soc.*, 1995, **117**, 5179-5197.
12. Aqvist, J., Medina, C., Samuelsson, J. E. *Protein Eng*. 1994, **7**, 385-391.
13. Hansson, T., Marelius, J., Aqvist, J. *J. Comput.-Aided Mol. Des.* 1998, **12**, 27-35.
14. Jones-Hertzog, D.K., Jorgensen, W.L. *J. Med. Chem*, 1997, **40**, 1539-1549.
15. Ljungberg, K. B., Marelius, J., Musil, D., Svensson, P., Norden, B., Aqvist. *Eur. J. Pharm. Sci.* 2001, **12**, 441-446.