## Title

Selectivity hot-spots of sirtuin catalytic cores

## Authors

Marco Daniele Parenti[1,†], Santina Bruzzone[2], Alessio Nencioni[3] and Alberto Del Rio[1,4*]

## Affiliations

[1]Department of Experimental, Diagnostic and Specialty Medicine (DIMES), Alma Mater Studiorum, University of Bologna, Via S.Giacomo 14, 40126 Bologna, Italy

[2]Department of Experimental Medicine, Section of Biochemistry, and CEBR, University of Genoa, V.le Benedetto XV 1, 16132 Genoa, Italy

[3]Department of Internal Medicine, University of Genoa, V.le Benedetto XV 6, 16132 Genoa, Italy

[4]Institute of Organic Synthesis and Photoreactivity (ISOF), National Research Council (CNR), Via P. Gobetti 101, 40129 Bologna, Italy

## Contact Information

*correspondence: alberto.delrio@gmail.com

1

## Abstract

Sirtuins are $NAD^+$-dependent deacetylases with several biological roles in DNA regulation, genomic stability, metabolism, longevity and immune cell functions. Numerous disease conditions are linked to sirtuins including metabolic disorders, inflammatory and autoimmune processes and cancer. Few specific small molecule modulators have been reported to date, while the need to identify selective ligands would be crucial for the development of active pharmaceutical ingredients for new targeted therapies but also as a tool for dissecting the biological roles of sirtuin family members. Here we report a comprehensive study aimed to classify and identify the selectivity hot-spots for targeting the catalytic cores of human sirtuins with small molecule modulators. Our selectivity analysis suggests that catalytic cores can be divided in different clusters that can constitute the basis for the development of selective ligands. The ensemble of hot-spot information is expected to be helpful to devise new selective chemicals targeting sirtuin family members.

## Introduction

Sirtuins are a family of highly conserved enzymes from bacteria to mammals[1] responsible for the control of several biological functions such as longevity, metabolism and DNA regulation, and involved in several disease conditions including cancer, inflammation, diabetes and obesity, and neurodegenerative diseases[2–4]. In humans, the sirtuin family comprises seven proteins, named SIRT1-SIRT7, characterized by different cellular and tissue localization, targets and catalytic activities. These enzymes were originally described as class-III family histone deacetylases, since its founding member, Sir2 in yeast, was found to be able to deacetylate histones H3 and H4[5]. Indeed, human sirtuins are able to catalyze the deacetylation of a wide range of proteins in various subcellular compartments by removal of an acetyl moiety of the ε-amino group of the lysine residue in the target protein[6,7]. The reaction yields the deacetylated

protein, nicotinamide, and 2'-O-acetyl-ADP-ribose. Other catalytic activities have been ascribed

to sirtuins such as mono-ADP-ribosyltransferase activity for SIRT4[8] and SIRT6[9] and

demalonylase and desuccinylase for SIRT5[10]. SIRT6 was also found to hydrolytically active to

remove long-chain fatty-acyl groups, including myristoyl and palmitoyl groups, a process known

as lysine deacylation[11]. Most recently, long-chain deacylation was found to be an intrinsic

activity of most sirtuins[12]. The broad involvements of sirtuins in several physiological and

pathological conditions have aroused great interest for the development of small molecule

modulators of this family of enzymes to be used as possible therapeutics, especially for

metabolic and aging-related conditions[13,14]. In the last decade, several three-dimensional

structures of sirtuin homologs have been deposited in the Protein Data Bank, ranging from

archaea to humans, providing a major understanding of structural features, catalytic

mechanisms and substrate specificity of these enzymes[15]. Several bacterial Sir2 structures, and

human SIRT1, SIRT2, SIRT3, SIRT5 and SIRT6 are available whereas, presently, no structures

has been solved for human SIRT4 and SIRT7 yet[16]. Despite this large amount of structural

information, only a relatively small number of active compounds have been identified to date,

often showing limited potency and isoform selectivity[3]. Similarly, few three-dimensional

structures of sirtuin have been solved with activators or inhibitors, reflecting the complexity of

the catalytic machinery and the subtle mechanisms of pharmacological modulation of these

enzymes.  In particular, EX527 (**1**, Figure 1), an inhibitor that combines good potency with

significant isoform selectivity was co-crystallized with SIRT1[17] and SIRT3[18]; SIRT3 structure is

also available in complex with three potent but not selective inhibitors (**2**-**4**, Figure 1) identified

from a large library screening[19], with SRT1720 (**5**, Figure 1), a compound that was described as

SIRT1 activator and SIRT3 inhibitor[20], and also with 4'-bromo-resveratrol[21] (**6**, Figure 1). SIRT5

was solved in complex with suramin (**7**, Figure 1), a potent inhibitor of SIRT1/SIRT2[22], and the

crystal structure of human SIRT2 was solved in complex with a macrocyclic peptide inhibitor[23].

Very recently, the SIRT3/SIRT5 complexes with piceatannol and resveratrol (**8**-**9**, Figure 1), two

of the known sirtuin activating compounds (STAC[24]), revealed the mechanism of direct sirtuin

activation[25]. While these three-dimensional structures highlighted the basis for inhibitor/activator

binding and laid the foundations for the rational design of new and more potent compounds,

additional efforts are needed to shed light on the more complex problem of isoform selectivity.

The development of specific modulators able to interact with a single sirtuin isoform, or at least

with a small subset, could be extremely useful to better understand the biological role of each

sirtuin and, at the same time, could facilitate the identification of new therapeutic agents. In fact,

the small molecules so far reported as sirtuin activators or inhibitors are often tested against a

limited set of sirtuin isoforms, usually SIRT1/2/3, and their selectivity profile is usually explained

by using approximate models. A previous study by Schlicker *et al.*[26] described the structure-

based identification of new classes of isoform specific inhibitors, showing that three-dimensional

sirtuin structures in the non-inhibited state could be successfully used to highlight differences

between isoforms and could help the development of new selective inhibitors. However, the

selectivity determinants for the modulation of the sirtuin catalytic cores are still unclear.

In this study, we aimed to gain insights into the isoform specificity of different sirtuins and to

identify the functional hot-spots that could be used to guide structure-based design efforts for

the identification of new selective sirtuin modulators. To this end, we developed a

comprehensive selectivity model based on sequence and structure-based alignments, and

characterized sirtuins through a detailed analysis of the pharmacophoric properties of the

catalytic site in terms of residue differences of each family member. The model was built by

including all three-dimensional structures of the human sirtuin family available in the PDB,

supplemented by homology models for those isoforms that have not been solved yet.

Differences in terms of key residues in the active site and the effects of conformational

variations were analyzed and the results reported in simple tables and graphs that illustrate at a

glance residues that most likely are involved in selectivity and the pharmacophoric features that should be considered in the design of new selective ligands.

## Results and Discussion

### *Sequence and structural comparison of the catalytic cores*

Sirtuins contain a conserved catalytic domain of approximately 250 amino acids, responsible for NAD$^+$ and substrate binding and deacetylase activity. Additionally, these proteins possess N- and C-terminal regions, variable in length and sequence, whose role in sirtuin biology is still unclear, although an effect on substrate specificity was speculated[15].

The structure of the catalytic core consists of a large Rossmann-fold domain, typical of NAD$^+$-binding proteins, a small zinc-binding domain, and a number of flexible loops that keep together the two domains. The binding sites of cofactor and substrate are located in a wide cleft at the interface between the large and the small domains. The sequence alignment of the catalytic core region (deacetylase domain) of the human sirtuin family (Figure 2) reveals a good level of sequence conservation, with the highest identity observed for residues responsible of cofactor binding and catalytic activity. A principal component analysis (PCA) of this alignment (Figure 3) clearly shows the presence of three different clusters: a first one including SIRT1, SIRT2 and SIRT3, a second one composed by SIRT6, SIRT7 and, slightly isolated, SIRT4, while SIRT5 completely differs from the other sirtuins. Consistently, the calculated pairwise sequence identity is the highest among the first cluster and between SIRT6 and SIRT7 (between 40% and 51%), while it is less than 30% for any other sequence pair. These results are in agreement with the molecular phylogenetic analysis results published by R.Frye[27]; this paper classify the seven human sirtuin genes in four different classes: SIRT1, SIRT2, and SIRT3 are class I, SIRT4 is class II, SIRT5 is class III, and SIRT6 and SIRT7 are class IV. The short motifs of conserved

amino acids present within the sirtuin core domain that characterize the different classes of sirtuins are in agreement with the sequence alignment reported in figure 2.

The large domain is characterized by a high level of structural and sequence similarity among various sirtuins; it is formed by a central β-sheet surrounded by six α-helices, except for SIRT2, which exhibits an additional long insertion of approximately 20 residues forming an additional α-helix. This domain possesses many of the specific requirements for NAD$^+$-binding, such as the well known G-X-G motif[28], important for the recognition of the phosphate group, and charged residues to bind the two ribose groups. The NAD$^+$ adenine base binds the C-terminal half of this domain while the nicotinamide enters the N-terminal part, as typical in inverted Rossmann-fold domains[29]. The small domain is the most diverse region in terms of primary sequence and conformation; it is formed by two long insertion in the Rossmann-fold domain, and consist of a three-stranded antiparallel β-sheet, a variable α-helical region, and a Zn$^{2+}$ cation coordinated by the sulphydryl group of four strictly conserved cysteine residues. The presence of the zinc ion is required to ensure enzyme functionality[30], although it does not directly participate in the catalytic mechanism of sirtuins. This domain shows the greatest variability in terms of three-dimensional structure among human sirtuins, and its relative position to the large domain depends also on the presence of substrate and on NAD$^+$ conformation[31]. Several variations are found in the primary sequence of the small domain. First, SIRT6 and SIRT7 contain a deletion which results in the lack of the helix bundle found in other sirtuins. This bundle is replaced by a short loop forming few interactions with the zinc-binding module. The presence of this unique feature may provide a possible explanation for the conformation of the zinc-binding domain in these two sirtuins, which, in turn, may be responsible for the observed lower catalytic activity of SIRT6[32]. SIRT6 also shows a ten-residue insertion between the second set of cysteines, resulting in a further long loop. Secondly, the two mitochondrial sirtuins SIRT4 and SIRT5 exhibit a unique insertion in the small domain comprising a short helix and a 16 residues loop, that may be

important for subcellular localization[22]. Finally, SIRT1 contains a 5 residues insertion adjacent to the last zinc-binding cysteine providing a distinctive small loop in this domain. The remarkable diversity in the small domain may have an important role in regulating key properties, such as substrate specificity and enzyme localization, highlighting this domain as an attractive potential binding site for selective sirtuin modulators. The cofactor binding loop, often referred to as the "flexible loop", is the largest of the four loops linking together the large and the small domains, and it is one of the most flexible regions of the enzyme. This loop appears to be disordered in the unliganded sirtuin structures but it undergoes significant conformational changes upon binding of $NAD^+$ or other reaction intermediates and could adopts multiple conformations depending on the bound ligand[15]. When $NAD^+$ is bound, this loop adopts a fairly "open" conformation, and several residues mediate important interactions between the nicotinamide moiety of the cofactor and the enzyme. When the active site is not occupied by the nicotinamide moiety, as in the case of the reaction product 2'-O-acetyl-ADP-ribose (2'-OAADPr), the loops assume a more "closed" conformation, with a partial occlusion of the nicotinamide binding pocket, precluding $NAD^+$ to bind in a productive conformation. In human sirtuins, the conformational changes in the cofactor binding loop are observed in three-dimensional structures of SIRT3 and SIRT5 whereas in SIRT6,the cofactor binding loop is replaced by a single helix that appears to be ordered in both complexes with ADPR and 2'-N-acetyl-ADP-ribose (2'-NAADPr), indicating that the binding pocket is less flexible and that its conformation is not susceptible to changes upon binding of different ligands[32].

### *Structural comparison of the binding sites*

The $NAD^+$ binding site can be divided into three different pockets: an adenine binding pocket (pocket A), a nicotinamide ribose binding pocket (pocket B) and a nicotinamide moiety binding pocket (pocket C). Several cofactor interactions within these binding sites are generally conserved among different sirtuins. For instance, the adenine base establishes several van der

Waals interactions including the two conserved glycines and various hydrogen bonds formed by the adenine nitrogens and polar residues such as glutamate, serine or threonine, while the hydroxyls of the adenine ribose mainly interact with a conserved asparagine residue. The phosphate group and the nicotinamide ribose moiety show a complex interaction network affected in both cases by the conformational variability of cofactor binding loop[15]. In particular, the nicotinamide ribose ring can adopt two slightly different conformations that may have implications in the supposed catalytic mechanism[33]. In the same way, conformational changes also occur in the pocket C, where the nicotinamide group can assume either a "non-productive" conformation, where it binds outside of the C pocket in a conformation that is not compatible with acetyl-lysine binding and the deacetylation reaction, or a "productive" conformation, in which the presence of an acetyl-lysine drives the nicotinamide moiety in the C pocket to establish interactions with invariant key residues. The substrate binding site is placed in a cleft between the large and the small domains. The backbone of the substrate peptide forms a $\beta$-strand-like interaction, known as $\beta$ staple with two loops in the enzyme, one within the Rossmann-fold domain and one that links together the two domains, while the acetyl-lysine side chain makes several interactions within an hydrophobic tunnel and an hydrogen bond between the N-atom and a backbone carbonyl of a conserved valine residue (leucine for SIRT6). The correct formation of $\beta$ staple interactions and acetyl-lysine binding tunnel are induced by the binding of the substrate peptide to its pocket by means of a rigid-body rotation of the small domain relative to the large domain[34].

***Structural superposition of available three-dimensional structures***

All available three-dimensional structures of human sirtuins were downloaded from Protein Data Bank and were supplemented by models obtained from the Swiss-Model database of SIRT4 and SIRT7, whose crystallographic structures are not available yet[35–38]. With the exception of these two sirtuins, all other isoforms have between four and ten distinct three-dimensional

structures solved. A total of 42 structures, in unbound form or in complex with different

cofactors, substrates and/or small molecule modulators (Table S1) were retrieved from the PDB

database. Structures were prepared and structurally aligned (see experimental section) in order

to gain more insight into the differences and similarities among human sirtuins. After the

structural alignment, a structure-based sequence alignment was obtained for the residues falling

into a range of 4Å from a reference ligand. This allowed a close analysis of variations in the

active site residues likely to be important for binding affinity and selectivity of new ligands. This

kind of structural alignment could bring at the same time the information concerning

conformational variations of the same residue in a different position or mutations of the residue

among different sirtuin isoforms, i.e. different kind of residues in the same position. Figure S1

shows the alignment of active site residues for all 42 available sirtuin structures in which all

residues are aligned only considering structural superposition. Within the same sirtuin isoform,

conformational variations of the catalytic core are mainly detectable in the cofactor-binding loop,

while other residues remain almost unchanged. In the same way, conformational changes

across different sirtuin isoforms were identified mainly in the same loop; however, assuming that

every sirtuin isoform can adopt the same active site conformation depending on the bound

ligand, the conformational variations appears not to be relevant in terms of ligand selectivity.

Conversely, the comparison across various isoforms is mainly characterized by specific point

aminoacidic differences which are relevant in terms of binding selectivity and are located on the

whole structural alignment. In order to highlight the variations by specifically focusing on the

catalytic core residues, a representative three-dimensional structure was selected for each

sirtuin isoform and a simplified structural alignment was generated in Figure 4. The

representative structures were selected so as to obtain a set of three-dimensional structures

with similar closed-form conformations, first considering the conformation of homology models

generated for SIRT4 and SIRT7. The whole alignment was then divided in pockets A, B, C and

substrate pocket, as defined above, to better depict the selectivity hot-spots.

Each residue was classified according to the estimated importance to generate ligand selectivity: conserved residues, low significance and high significance residues (Figure 4). Conserved residues are understandably not relevant for selectivity while some aminoacidic differences were classified as low significant for selectivity when the variation was limited to similar residues (e.g. valine/leucine) or when the side chain of the residues was directed outside the active site and therefore not prone to affect the shape and the properties of the catalytic core and the putative binding of a ligand. The nicotinamide-ribose binding pocket (Figure 4B) shows the higher degree of conservation and only one residue is classified as important for the selectivity, although the aminoacidic difference is limited, i.e. F to Y, and involving only two sirtuins, namely SIRT4 and SIRT7. Conversely, the structural alignments of adenine binding pocket (Figure 4A) and nicotinamide binding pocket (Figure 4C) reveal a higher degree of difference, with 5 and 8 significant residues, respectively. The substrate binding pocket is characterized by a significant conservation (Figure 4D) from SIRT1 to SIRT5, while SIRT6 and especially SIRT7 show differences that are the results of gaps sequence; nonetheless, four residues are predicted to have some effects on selectivity in this region.

Overall, we identified 18 residues in the active site (red plus in Figure 4) that could have a significant impact on ligand selectivity and that should be taken into account while designing new sirtuin modulators. The pharmacophoric features and the position in the active site relative to the cofactor and a reference inhibitor (SRT1720) was depicted in Figure 5 while the complete information about aminoacidic differences for all possible couples of sirtuin isoforms was reported in a double-entry table in Figure S2 that encodes similarities and differences among the whole sirtuin family.

Complementary to the results obtained with the sequence alignment of the deacetylase domain, the structural superposition of sirtuin active sites reveal that catalytic cores can be divided into different clusters. SIRT1, SIRT2 and SIRT3 (group 1) are characterized by high similarity in the

active sites, with few different residues located mainly in the adenine pocket and in the

substrate pocket. Likewise, SIRT6 and SIRT7 (group 2) show significant similarity, with a total of

6 variable residues, while both SIRT4 and SIRT5 (group 3) show several aminoacidic variations

compared to all other sirtuins. Based on these observations, it is possible to hypothesize the

design of selective ligands for sirtuins belonging to different identified groups, while we predict

that the selectivity within a single group is harder to achieve. For example, inhibitors of SIRT6 or

SIRT7 could be likely selective versus SIRT1 or SIRT2 due to the important aminacidic

differences that characterize the active site of these isoforms. On the contrary, ligands resulting

in a high selectivity for SIRT1 over SIRT2, or SIRT6 over SIRT7, are expected to be difficult to

obtain on the basis of the intrinsic similarities in their aminoacidic composition. Certainly this

similarity does not preclude the possibility to obtain ligand selective that exploit other structural

features such as preferential conformational behavior and specific shape of the catalytic core

(see below). Figure 5 depict these concepts showing the difference in the active site among

sirtuins belonging to the same group (Figure 6A and Figure 6B) and belonging to different

groups (Figure 6C). From the pharmacophoric point of view, the sirtuin active site (Figure 5) is

characterized by a high number of hydrophobic interactions (green boxes), especially in

proximity of the reference ligand SRT1720 that lie in the nicotinamide and the substrate binding

pocket. This is especially evident for SIRT1, SIRT2 and SIRT3 (group 1, as defined above).

Taking into consideration the differences among each sirtuin isoform, these interactions are

primarily replaced with polar interactions (blue boxes) or with sequence gaps (gray boxes), the

latter implying the presence of small but significant variations in the shape and dimension of the

catalytic core. On the contrary, the adenine binding pocket and the nicotinamide-ribose binding

pocket show mainly polar or charged features (blue and purple/red respectively), often involving

hydrogen bonding with the cofactor molecule. In these cases, the differences among sirtuin

isoforms involve charge shifting (positive to negative or *vice versa*) or substitutions with

11

hydrophobic features; in both cases, the result is a variation of the surface properties of the active site.

It is worth mentioning that compound selectivity for SIRT1 vs. SIRT2 was studied by previous works and several examples of fairly selective inhibitors were reported in the last few years[39–45]. The rationale for this selectivity was investigated using classical structural-activity relationships (SAR) and, occasionally, with the help of molecular modeling techniques that were used to predict the active site residues involved. The same techniques were recently applied by our group for the discovery of selective SIRT6 inhibitors[46]. Nevertheless, despite the amount of structural information appeared in the literature in the last years, the mechanisms underlining ligand selectivity towards cognate sirtuin isoform is still not fully understood. In fact, the selectivity could be the result of a combination of different effects not related to the single residue modifications in the active site but, for instance, different kinetic profiles, conformational changes of the proteins, presence of allosteric sites as well as subcellular location of protein targets and their level of expression in specific tissues. As an example, the isoform selectivity of EX-527 (**1**, Figure 1) appears to be based on kinetic differences of catalysis, suggesting that selectivity could also vary under different physiological conditions, such as substrate availability [18]. However, such effects are difficult to be estimated *a priori* and experimental determinations are needed elucidate selective behaviors, whereas aminoacidic composition that alter active site properties could be considered in structural analyses, like the one herein presented, that could be used for the rational design of new small-molecule modulators or for the development of analog compounds using standard SARs.

The structural model of human sirtuin family that we obtained clearly suggests several functional hot-spots that could be exploited to improve selectivity of newly identified sirtuin ligands. Nonetheless, it is important to note that approximations to generate this model have been taken into account and further experimental structural insights could contribute to refine our

conclusions. In particular, it is important to note that conformational variations of the proteins, e.g the open and closed forms, could play a role in term of ligand selectivity while our model is based on the principle that all sirtuin family members can adopt both forms depending on the bound substrate, ligand and cofactor. Similarly, an experimental validation is still awaited for SIRT4 and SIRT7 that still have no three-dimensional structure representative in the Protein Data Bank and, for the time being, they needed to be modeled by homology. In the same way, the presence of allosteric sites, or ligands able to induce isoform-specific conformational variations have not been described yet. Finally, the model was built by including residues within a distance range compatible with reference ligands and cofactors (i.e. SRT1720 and $NAD^+$) taken from the few crystallographic structures available; clearly, the discovery of additional ligands and their three-dimensional structure in complex with sirtuins might provide additional information on new interacting residues extending outside the range encompassed by our model.

During the revision process of this paper, a study by Rumpf et al.[47], reporting the crystal structures of SIRT2 solved in complex with selective inhibitors SirReal1 (PDB code 4RMI) and SirReal2 (PDB codes 4RMG, 4RMH) was published. In this study, the authors discovered an "open-locked conformation" for sirtuins upon ligand rearrangement resulting in a highly selective and potent SIRT2 inhibitor. A selectivity pocket is exploited by these inhibitors, and none of other inhibitors whose binding modes have been elucidated by means of X-ray crystallography show similar binding features. This results confirm, as noted above, that conformational flexibility could play a role on ligand selectivity; however, as reported by the same study, SIRT1 and SIRT3 are likely able to adopt a similar conformation as observed in the SIRT2–SirReal2 complexes that would allow binding of SirReal2, as demonstrated by high-quality homology models, supporting our assumption that all sirtuin family members can adopt a specific conformation depending on the bound substrate, ligand and cofactor. The higher binding affinity

13

of SirReal2 for SIRT2 could be hence ascribed to few aminoacidic differences in hypothetical selectivity pocket that changes its shape and surface characteristics. To complete and integrate our starting model, the residues involved in binding of these selective SIRT2 inhibitors, excluded those already described in Figure 4, were reported in an additional structural alignment (Figure S3).

## Conclusions

The broad involvement of sirtuins in several pathological conditions has raised a strong interest in the development of specific inhibitors or activators to explore disease models that are dependent on sirtuin functions. Selective modulators, able to modify the enzymatic activity of single sirtuin isoforms could be used as chemical probes to elucidate the role of each sirtuin isoform in biological activities but also as a starting point to develop more efficient targeted therapies. Our study is the first effort to summarize in a comprehensive structural model the selectivity hot-spots among sirtuin family members, providing a simple tool to estimate the selectivity of a small molecule ligand between two or more isoforms. Starting from a putative binding mode of a ligand into the active site of one sirtuin isoform, experimental (i.e. X-ray or NMR) or molecular modeling (docking, molecular dynamics) techniques, could be used in combination to our model to allow understanding at a glance the active site residues that most likely are involved in ligand selectivity and how to design suitable pharmacophoric variations to improve the activity profiles. Although based on approximations and simplifications, this model constitute valuable tool to better understand the complexity of the sirtuin machinery and to advance the search for selective small-molecule modulators.

## Experimental Section

**Sequence alignment and principal component analysis.** The sequences of the seven human sirtuins were retrieved from Uniprot database as reported in Table 1. The multiple

14

sequence alignment was generated using T-Coffee method[48] implemented in Jalview[49], followed by a manual revision to ensure the maximum accuracy. The alignment was then exported in FASTA format and converted to CSV format, using a bash shell script, to make it suitable for further steps. The PCA analysis was carried out using Simca-P software version 11.2 (Umetrics, Sweden). All variables with zero variance (conserved residues) were automatically excluded from the model and the first two principal components were calculated and used to generate the score plot reported in Figure 3.

**Protein preparation.** The crystal structures of human sirtuins (Table S1) were retrieved from PDB Database and submitted to a standard preparation procedures (protein preparation wizard) as included in the software package Maestro (Version 9.3, Schrödinger, LLC). Any water molecules, ions and crystallization mediums were removed (except for zinc ion) while substrates, ligands and cofactors were kept, if present. Each structure was optimized using PROPKA and minimized until an RMSD of 0.2 Å.

**Structure-based alignment.** The prepared structures were then aligned using the Protein Structure Alignment tool included in Maestro, using backbone as reference atoms. The Multiple Sequence Viewer tool was then used to align sequences according to structure superposition, and to generate an alignment based only on structure superposition and not on sequence conservation. The structural alignment was restricted by selecting residues falling into a range of 4 Å from reference ligands, namely the SRT1720 and Carba-NAD structures taken from PDB code 4BN5; these reference ligands are able to fill up completely all active site pockets, and therefore were chosen among the other available structures. Due to the large number of three-dimensional structure included in the model, and to the high level of complexity of the structural alignment, the results obtained of automated tools were excluded to avoid inaccuracies and processed thorough a manual revision of the alignment. The alignment reported in Figure 4 was generated starting from the global alignment by selecting 7 representative structures, one for

each sirtuin isoform. In particular, the following PDB codes were selected: 4I5I (SIRT1), 3ZGV

(SIRT2), 3GLT (SIRT3), 3RIY (SIRT5), 3K35 (SIRT6), supplemented by the homology models

for SIRT4 and SIRT5.

## Acknowledgments

## References

1       C. B. Brachmann, J. M. Sherman, S. E. Devine, E. E. Cameron, L. Pillus and J. D. Boeke,
        *Genes Dev.*, 1995, **9**, 2888–902.

2       S. Libert, K. Pointer, E. L. Bell, A. Das, D. E. Cohen, J. M. Asara, K. Kapur, S. Bergmann,
        M. Preisig, T. Otowa, K. S. Kendler, X. Chen, J. M. Hettema, E. J. van den Oord, J. P.
        Rubio and L. Guarente, *Cell*, 2011, **147**, 1459–72.

3       S. Bruzzone, M. D. Parenti, A. Grozio, A. Ballestrero, I. Bauer, A. Del Rio and A.
        Nencioni, *Curr. Pharm. Des.*, 2013, **19**, 614–23.

4       T. Finkel, C.-X. Deng and R. Mostoslavsky, *Nature*, 2009, **460**, 587–91.

5       M. Braunstein, R. E. Sobel, C. D. Allis, B. M. Turner and J. R. Broach, *Mol. Cell. Biol.*,
        1996, **16**, 4349–56.

6       R. H. Houtkooper, E. Pirinen and J. Auwerx, *Nat. Rev. Mol. Cell Biol.*, 2012, **13**, 225–38.

7       F. Andreoli and A. Del Rio, *Drug Discov. Today*, 2014.

8       M. C. Haigis, R. Mostoslavsky, K. M. Haigis, K. Fahie, D. C. Christodoulou, A. J. Murphy,
        D. M. Valenzuela, G. D. Yancopoulos, M. Karow, G. Blander, C. Wolberger, T. A. Prolla,
        R. Weindruch, F. W. Alt and L. Guarente, *Cell*, 2006, **126**, 941–54.

9       G. Liszt, E. Ford, M. Kurtev and L. Guarente, *J. Biol. Chem.*, 2005, **280**, 21313–20.

10      J. Du, Y. Zhou, X. Su, J. J. Yu, S. Khan, H. Jiang, J. Kim, J. Woo, J. H. Kim, B. H. Choi,
        B. He, W. Chen, S. Zhang, R. A. Cerione, J. Auwerx, Q. Hao and H. Lin, *Science*, 2011,
        **334**, 806–9.

11      H. Jiang, S. Khan, Y. Wang, G. Charron, B. He, C. Sebastian, J. Du, R. Kim, E. Ge, R.
        Mostoslavsky, H. C. Hang, Q. Hao and H. Lin, *Nature*, 2013, **496**, 110–3.

12      J. L. Feldman, J. Baeza and J. M. Denu, *J. Biol. Chem.*, 2013, **288**, 31350–31356.

13      L. Guarente, *Nature*, 2006, **444**, 868–74.

14      J. C. Milne and J. M. Denu, *Curr. Opin. Chem. Biol.*, 2008, **12**, 11–7.

15      B. D. Sanders, B. Jackson and R. Marmorstein, *Biochim. Biophys. Acta*, 2010, **1804**,
        1604–16.

16      F. Andreoli, A. J. M. Barbosa, M. D. Parenti and A. Del Rio, *Curr. Pharm. Des.*, 2013, **19**,
        578–613.

17      X. Zhao, D. Allison, B. Condon, F. Zhang, T. Gheyi, A. Zhang, S. Ashok, M. Russell, I.
        Macewan, Y. Qian, J. a Jamison and J. G. Luz, *J. Med. Chem.*, 2013.

18      M. Gertz, F. Fischer, G. T. T. Nguyen, M. Lakshminarasimhan, M. Schutkowski, M.
        Weyand and C. Steegborn, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, E2772–81.

19      J. S. Disch, G. Evindar, C. H. Chiu, C. A. Blum, H. Dai, L. Jin, E. Schuman, K. E. Lind, S.
        L. Belyanskaya, J. Deng, F. Coppo, L. Aquilani, T. L. Graybill, J. W. Cuozzo, S. Lavu, C.
        Mao, G. P. Vlasuk and R. B. Perni, *J. Med. Chem.*, 2013, **56**, 3666–79.

20      G. T. T. Nguyen, S. Schaefer, M. Gertz, M. Weyand and C. Steegborn, *Acta Crystallogr.
        D. Biol. Crystallogr.*, 2013, **69**, 1423–32.

21      G. T. T. Nguyen, M. Gertz and C. Steegborn, *Chem. Biol.*, 2013, **20**, 1375–85.

22      A. Schuetz, J. Min, T. Antoshenko, C.-L. Wang, A. Allali-Hassani, A. Dong, P. Loppnau,
        M. Vedadi, A. Bochkarev, R. Sternglanz and A. N. Plotnikov, *Structure*, 2007, **15**, 377–89.

23      K. Yamagata, Y. Goto, H. Nishimasu, J. Morimoto, R. Ishitani, N. Dohmae, N. Takeda, R.
        Nagai, I. Komuro, H. Suga and O. Nureki, *Structure*, 2014, **22**, 345–52.

24      J. C. Milne, P. D. Lambert, S. Schenk, D. P. Carney, J. J. Smith, D. J. Gagne, L. Jin, O.
        Boss, R. B. Perni, C. B. Vu, J. E. Bemis, R. Xie, J. S. Disch, P. Y. Ng, J. J. Nunes, A. V
        Lynch, H. Yang, H. Galonek, K. Israelian, W. Choy, A. Iffland, S. Lavu, O. Medvedik, D.
        A. Sinclair, J. M. Olefsky, M. R. Jirousek, P. J. Elliott and C. H. Westphal, *Nature*, 2007,
        **450**, 712–6.

25      M. Gertz, G. T. T. Nguyen, F. Fischer, B. Suenkel, C. Schlicker, B. Fränzel, J.
        Tomaschewski, F. Aladini, C. Becker, D. Wolters and C. Steegborn, *PLoS One*, 2012, **7**,
        e49761.

26   C. Schlicker, G. Boanca, M. Lakshminarasimhan and C. Steegborn, *Aging (Albany. NY).*, 2011, **3**, 852–72.

27   R. A. Frye, *Biochem. Biophys. Res. Commun.*, 2000, **273**, 793–8.

28   C. R. Bellamacina, *FASEB J.*, 1996, **10**, 1257–69.

29   G. S. Prasad, V. Sridhar, M. Yamaguchi, Y. Hatefi and C. D. Stout, *Nat. Struct. Biol.*, 1999, **6**, 1126–31.

30   J. Min, J. Landry, R. Sternglanz and R. M. Xu, *Cell*, 2001, **105**, 269–79.

31   J. L. Avalos, J. D. Boeke and C. Wolberger, *Mol. Cell*, 2004, **13**, 639–648.

32   P. W. Pan, J. L. Feldman, M. K. Devries, A. Dong, A. M. Edwards and J. M. Denu, *J. Biol. Chem.*, 2011, **286**, 14575–87.

33   H. Yuan and R. Marmorstein, *J. Biol. Chem.*, 2012, **287**, 42428–35.

34   M. S. Cosgrove, K. Bever, J. L. Avalos, S. Muhammad, X. Zhang and C. Wolberger, *Biochemistry*, 2006, **45**, 7511–21.

35   K. Arnold, L. Bordoli, J. Kopp and T. Schwede, *Bioinformatics*, 2006, **22**, 195–201.

36   M. Biasini, S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, F. Kiefer, T. G. Cassarino, M. Bertoni, L. Bordoli and T. Schwede, *Nucleic Acids Res.*, 2014, **42**, W252–8.

37   N. Guex, M. C. Peitsch and T. Schwede, *Electrophoresis*, 2009, **30 Suppl 1**, S162–73.

38   F. Kiefer, K. Arnold, M. Künzli, L. Bordoli and T. Schwede, *Nucleic Acids Res.*, 2009, **37**, D387–92.

39   P. Di Fruscia, K.-K. Ho, S. Laohasinnarong, M. Khongkow, S. H. B. Kroll, S. A. Islam, M. J. E. Sternberg, K. Schmidtkunz, M. Jung, E. W.-F. Lam and M. J. Fuchter, *Medchemcomm*, 2012.

40   M. Fridén-Saxin, T. Seifert, M. R. Landergren, T. Suuronen, M. Lahtela-Kakkonen, E. M. Jarho and K. Luthman, *J. Med. Chem.*, 2012, **55**, 7104–13.

41   M. A. Khanfar, L. Quinti, H. Wang, S. H. Choi, A. G. Kazantsev and R. B. Silverman, *Eur. J. Med. Chem.*, 2014, **76**, 414–26.

42   S. S. Mahajan, M. Scian, S. Sripathy, J. Posakony, U. Lao, T. K. Loe, V. Leko, A. Thalhofer, A. D. Schuler, A. Bedalov and J. A. Simon, *J. Med. Chem.*, 2014, **57**, 3283–94.

43   F. Medda, R. J. M. Russell, M. Higgins, A. R. McCarthy, J. Campbell, A. M. Z. Slawin, D. P. Lane, S. Lain and N. J. Westwood, *J. Med. Chem.*, 2009, **52**, 2673–82.

44      D. Rotili, V. Carafa, D. Tarantino, G. Botta, A. Nebbioso, L. Altucci and A. Mai, *Bioorg. Med. Chem.*, 2011, **19**, 3659–68.

45      T. Suzuki, M. N. A. Khan, H. Sawada, E. Imai, Y. Itoh, K. Yamatsuta, N. Tokuda, J. Takeuchi, T. Seko, H. Nakagawa and N. Miyata, *J. Med. Chem.*, 2012, **55**, 5760–73.

46      M. D. Parenti, A. Grozio, I. Bauer, L. Galeno, P. Damonte, E. Millo, G. Sociali, C. Franceschi, A. Ballestrero, S. Bruzzone, A. Del Rio and A. Nencioni, *J. Med. Chem.*, 2014, **57**, 4796–804.

47      T. Rumpf, M. Schiedel, B. Karaman, C. Roessler, B. J. North, A. Lehotzky, J. Oláh, K. I. Ladwein, K. Schmidtkunz, M. Gajer, M. Pannek, C. Steegborn, D. a. Sinclair, S. Gerhardt, J. Ovádi, M. Schutkowski, W. Sippl, O. Einsle and M. Jung, *Nat. Commun.*, 2015, **6**, 6263.

48      C. Notredame, D. G. Higgins and J. Heringa, *J. Mol. Biol.*, 2000, **302**, 205–17.

49      A. M. Waterhouse, J. B. Procter, D. M. A. Martin, M. Clamp and G. J. Barton, *Bioinformatics*, 2009, **25**, 1189–91.

## Figures and Tables



**1** (EX-527)

**2**

**3**

**4**

**5** (SRT1720)

**6** (bromo-resveratrol)

**7** (suramin)

**8** (piceatannol)

**9** (resveratrol)

**Figure 1.** Ligands that have been solved with sirtuin structures

```
SIRT1  244 EDAVKLLQ--ECKKIIVLTGAGVSVSCGIPDFRSR-DGIYARLAVDFPDLPDPQAMFDIEYFRK 304
SIRT2   65 EGVARYMQSERCRRVICLVGAGISTSAGIPDFRSPSTGLYDNLEK--YHLPYPEAIFEISYFKK 126
SIRT3  126 QDVAELIRARACQRVVVMVGAGISTPSGIPDFRSPGSGLYSNLQQ--YDLPYPEAIFELPFFFH 187
SIRT4   45 KELQRFIT--LSKRLLVMTGAGISTESGIPDYRSEKVGLYARTDR--RPIQHGDFVRSAPIRQR 104
SIRT5   41 ADFRKFFA--KAKHIVIISGAGVSAESGVPTFRGA-GGYWRKWQA--QDLATPLAFAHNP--SR  97
SIRT6   35 WELARLVW--QSSSVVFHTGAGISTASGIPDFRGP-HGVWTMEER--GLAPKFDTTF-------  86
SIRT7   90 RELASAVR--NAKYLVVYTGAGISTAASIPDYRGP-NGVWTLLQK--GRSVSA-ADL-------  140

SIRT1  305 DPRPFFKFAKEIYPGQFQPSLCHKFIAL----SDKEGKLLRNYTQNIDTLEQVAGIQ--RIIQC 362
SIRT2  127 HPEPFFALAKELYPGQFKPTICHYFMRL----LKDKGLLLRCYTQNIDTLERIAGLEQEDLVEA 186
SIRT3  188 NPKPFFTLAKELYPGNYKPNVTHYFLRL----LHDKGLLLRLYTQNIDGLERVSGIPASKLVEA 247
SIRT4  105 Y-WARNFVGWPQ-FSSHQPNPAHWALST----WEKLGKLYWLVTQNVDALHTKAGSR--RLTEL 160
SIRT5   98 V-WEFYHYRREV-MGSKEPNAGHRAIAECETRLGKQGRRVVVITQNIDELHRKAGTK--NLLEI 157
SIRT6   87 -------------ESARPTQTHMALVQ----LERVGLLRFLVSQNVDGLHVRSGFPRDKLAEL 132
SIRT7  141 -------------SEAEPTLTHMSITR----LHEQKLVQHVVSQNCDGLHLRSGLPRTAISEL 186

SIRT1  363 HGSFATASCLI--CKY--KVDCEA-V--------------------------RGDIFNQVV-P 393
SIRT2  187 HGTFYTSHCVSASCRH--EYPLSW-M--------------------------KEKIFSEVT-P 219
SIRT3  248 HGTFASATCTV--CQR--PFPGED-I--------------------------RADVMADRV-P 278
SIRT4  161 HGCMDRVLCLD--CGE--QTPRGV-LQERFQVLNPTWSAEAHGLAPDGDVFLSEEQVRSFQV-P 218
SIRT5  158 HGSLFKTRCTS--CGV--VAE--N-YKS---PICPA--LSGKGA-PEP--GTQDASIPVEKL-P 205
SIRT6  133 HGNMFVEECAK--CKT--QYVRDT-V--------------------------VGTMGLKATGR 164
SIRT7  187 HGNMYIEVCTS--CVPNREYVRVFDV--------------------------TERTALHRHQTGR 223

SIRT1  394 RCP--------RCPADEPLAIMKPEIVFFGEN--L-PEQFHRAMKYDKDEVDLLIVIGSSLKVR 446
SIRT2  220 KCE--------DC----QSLVKPDIVFFGES--L-PARFFSCMQSDFLKVDLLLVMGTSLQVQ 267
SIRT3  279 RCP--------VC-----TGVVKPDIVFFGEP--L-PQRFLLHV-VDFPMADLLLILGTSLEVE 325
SIRT4  219 TCV--------QC-----GGHLKPDVVFFGDT--V-NPDKVDFVHKRVKEADSLLVVGSSLQVY 266
SIRT5  206 RCEEA------GC-----GGLLRPHVVWFGEN--L-DPAILEEVDRELAHCDLCLVVGTSSVVY 255
SIRT6  165 LCTVAKARGLRAC-----RGELRDTILDWEDS--L-PDRDLALADEASRNADLSITLGTSLQIR 220
SIRT7  224 TCH--------KC-----GTQLRDTIVHFGERGTLGQPLNWEAATEAASRADTILCLGSSLKVL 274

SIRT1  447 ---PVALIPSSI-PHEVPQI-LINREPLPHL--------------------HFDVELLGDCDV 484
SIRT2  268 ---PFASLISKA-PLSTPRL-LINKEKAGQSDPFLGMIMGLGGGMDFDSKKAYRDVAWLGECDQ 326
SIRT3  326 ---PFASLTEAV-RSSVPRL-LINRDLVGPL--------------AW--HPRSRDVAQLGDVVH 368
SIRT4  267 ---SGYRFILTAWEKKLPIA-ILNIGPTRSD--------------------DLACLKLNSRCGE 306
SIRT5  256 ---PAAMFAPQVAARGVPVA-EFNTETTPAT--------------------NRFRFHFQGPCGT 295
SIRT6  221 ---PSGNLPLAT-KRRGGRLVIVNLQPTKHD--------------------RHADLRIHGYVDE 260
SIRT7  275 KKYPRLWCMTKP-PSRRPKLYIVNLQWTPKD--------------------DWAALKLHGKCDD 317

SIRT1  485 IINELCHRLGGEYA                              498
SIRT2  327 GCLALAELLGWKKE                              340
SIRT3  369 GVESLVELLGWTEE                              382
SIRT4  307 LLPL------IDPC                              314
SIRT5  296 TLPEALACHENETV                              309
SIRT6  261 VMTRLMKHLGLEIP                              274
SIRT7  318 VMRLLMAELGLEIP                              331
```

**Figure 2.** Sequence alignment of the deacetylase domain of human sirtuins. Residues are colored by sequence identity and boxes indicate special residues: pink for NAD$^+$-binding region, cyan for zinc-binding cysteines, and grey for catalytic histidines.

**Figure 3.** Principal component analysis of sirtuins sequence alignment. The picture reports the score plot obtained from the first two principal components (t[1]/t[2]).

**Figure 4.** Structural alignment of human sirtuins. The alignments were generated using

residues in a range of 4Å from the ligand and considering a three-dimensional superposition.

The residues were color-coded using Taylor classification and numbered following gene

sequences taken from the Uniprot database. The last line of each alignment suggests the

estimated importance of each residue in selectivity: black dot (•) indicates residues conserved in all sirtuins, while black minus (-) and red plus (+) indicates residues respectively with poor or higher probability to be involved in selectivity. A) Structural alignment of residues comprising adenine binding pocket (pocket A). B) Structural alignment of residues comprising nicotinamide-ribose binding pocket (pocket B). C) Structural alignment of residues comprising nicotinamide binding pocket (pocket C). D) Structural alignment of residues comprising substrate binding pocket.
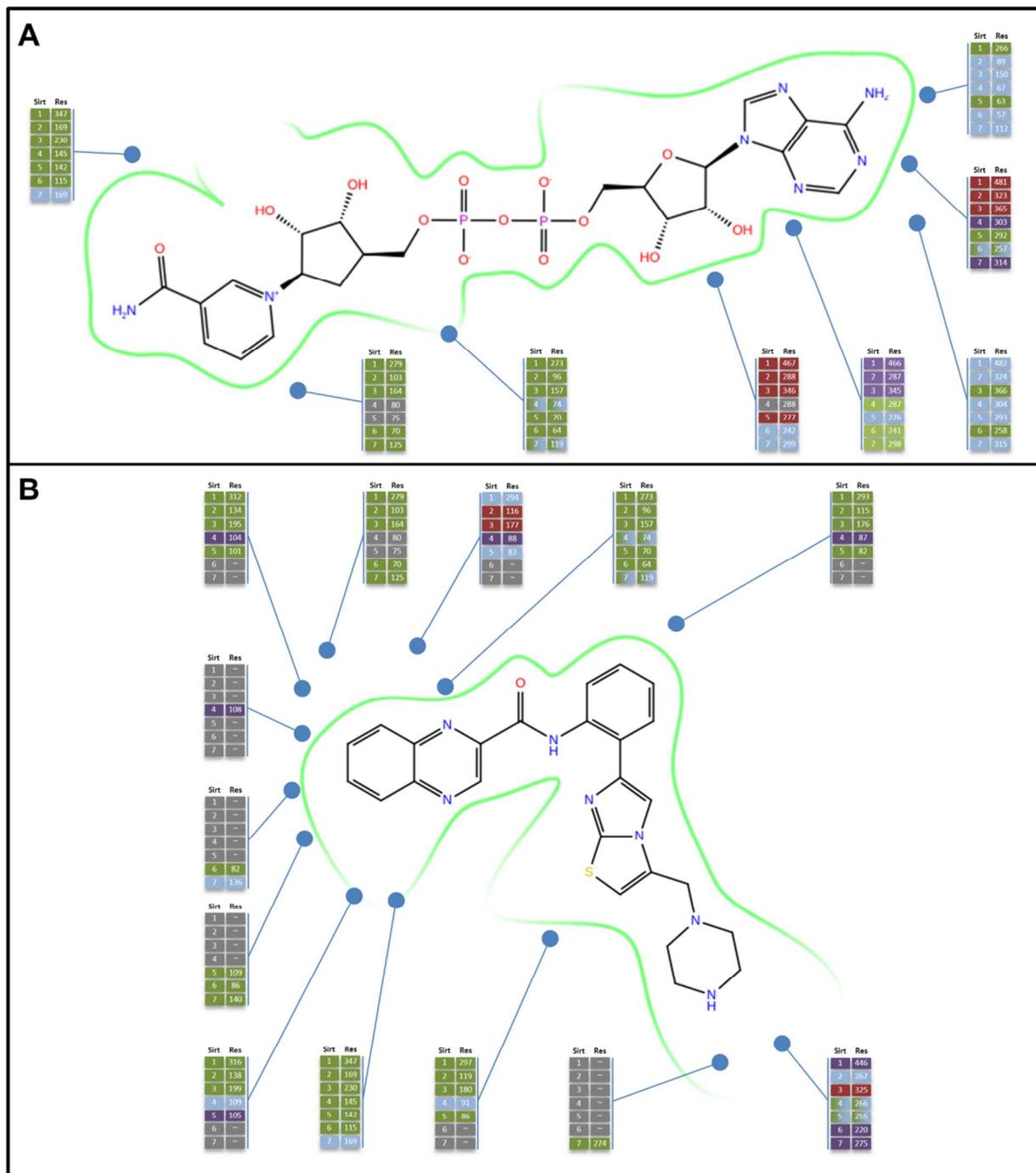
**Figure 5.** Pharmacophoric features of catalytic core residues of sirtuin family. The green line represents the active site surface and the blue dots the position of important residues in the catalytic core. For each dot a small table reports the residue number for each sirtuin, color-coded by pharmacophoric features: green for hydrophobic, blue for polar, purple for positive

charge, red for negative charge and gray for absent residues or glycines. The green/blue cells represent tyrosine residues. Some residues can interact with both ligand and cofactor and are reported in both parts of the figure. A) Ligand interaction diagram of NAD$^+$. B) Ligand interaction diagram of SRT1720.
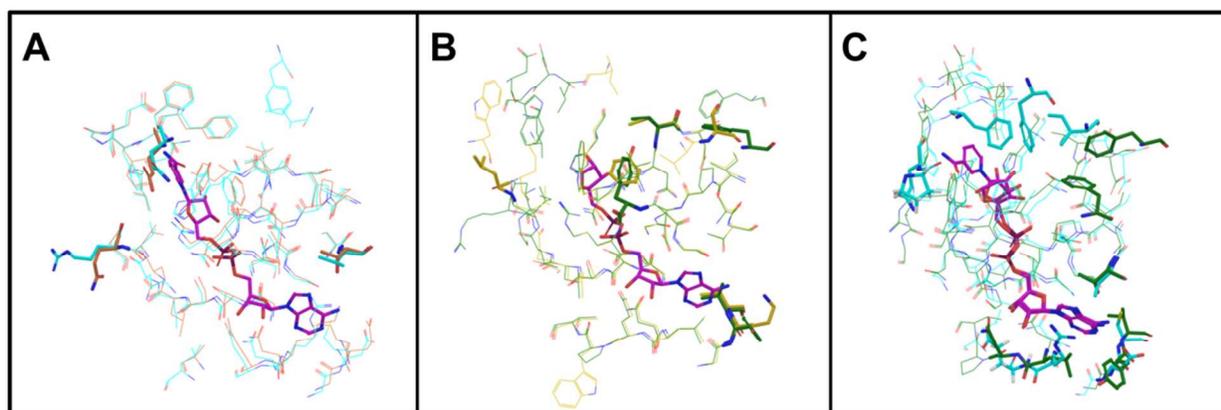


**Figure 6.** Structural comparison of SIRT1, SIRT2, SIRT6 and SIRT7. Residues classified as important for selectivity are drawn in thick tube. Reference ligands are reported to help the identification of active site pockets. A) Structural comparison of SIRT1 (cyan) and SIRT2 (orange). Reference ligand (purple) is NAD$^+$ taken from PDB code 4I5I. B) Structural comparison of SIRT6 (dark green) and SIRT7 (lime green). Reference ligand (purple) is ADPR taken from PDB code 3K35. C) Structural comparison of SIRT1 (cyan) and SIRT6 (dark green), with the same reference ligands of A and B (purple).

**Table 1.** Human sirtuins sequences available in Uniprot database.

| Seq | Uniprot Code | Tot Residues | Catalytic HIS | Nucleotide Binding | Metal Binding | Deacetylase Domain |
|-----|-------------|-------------|--------------|-------------------|--------------|-------------------|
| Sirt1 | Q96EB6 | 747 | 363 | 261-280, 345-348, 440-442, 465-467 | 371,374,395,398 | 244 - 498 |
| Sirt2 | Q8IXJ6 | 389 | 187 | 84-104, 167-170, 261-263, 286-288 | 195,200,221,224 | 65-340 |
| Sirt3 | Q9NTG7 | 399 | 248 | 145-165, 228-231, 319-321, 344-346 | 256,259,280,283 | 126-382 |
| Sirt4 | Q9Y6E7 | 314 | 161 | 62-82, 143-146, 260-262, 286-288 | 169,172,220,223 | 45-314 |
| Sirt5 | Q9NXA8 | 310 | 158 | 58-77, 140-143, 249-251, 275-277 | 166,169,207,212 | 41-309 |
| Sirt6 | Q8N6T7 | 355 | 133 | 52-71, 113-116, 214-216, 240-242 | 141,144,166,177 | 35-274 |

| Sirt7 | Q9NRC8 | 400 | 187 | 107-126, 167-170, 268-270, 297-299 | 195,198,225,228 | 90-331 |

28