PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

STRUCTURE NOTE

Crystal structure of the N-terminal domain of human SIRT7 reveals a three-helical domain architecture

Anu Priyanka,[1] Vipul Solanki,[1] Raman Parkesh,[2] and Krishan Gopal Thakur[1]*

[1] Structural Biology Laboratory, G. N. Ramachandran Protein Centre, CSIR-Institute of Microbial Technology, Chandigarh-160036, India

[2] Computational Medicinal Chemistry and Fluorescent Chemosensors Laboratory, G. N. Ramachandran Protein Centre, CSIR-Institute of Microbial Technology, Chandigarh-160036, India

ABSTRACT

Human SIRT7 is an NAD^+ dependent deacetylase, which belongs to sirtuin family of proteins. SIRT7, like other sirtuins has conserved catalytic domain and is flanked by N- and C-terminal domains reported to play vital functional roles. Here, we report the crystal structure of the N-terminal domain of human SIRT7 ($SIRT7^{NTD}$) at 2.3 Å resolution as MBP-$SIRT7^{NTD}$ fusion protein. $SIRT7^{NTD}$ adopts three-helical domain architecture and comparative structural analyses suggest similarities to some DNA binding motifs and transcription regulators. We also report here the importance of N- and C-terminal domains in soluble expression of SIRT7.

Proteins 2016; 84:1558–1563.

Key words: Human SIRT7; sirtuins; histone deacetylase; maltose binding protein; x-ray crystallography; peptide mass fingerprinting.

INTRODUCTION

Sirtuins are evolutionarily conserved NAD^+-dependent class III histone deacetylases (HDAC) implicated in several age related human diseases such as cancer, diabetes, cardiovascular disease and neurodegenerative disorders.[1] Seven mammalian sirtuins, SIRT1-SIRT7, have been discovered which show diversity and complexity in their cellular localization and functions.[2] SIRT7 is the only sirtuin reported localized in nucleolus where it associates with RNA polymerase I and histones.[3] It positively regulates RNA polymerase I mediated transcription[3] and plays a role in ribosome biogenesis and protein synthesis.[4]

SIRT7 like other sirtuins has conserved catalytic NAD^+ binding domain and unique N- and C-terminal regions containing nuclear and nucleolar localization signals respectively.[5] There have been several reports highlighting the importance of N- and C-terminal regions in sirtuins. It has been shown that extended N- and C-termini of yHst2 helps in oligomerization and autoregulation of acetyl-lysine catalysis and NAD^+ dependent catalytic activity respectively.[6] The N- and C-terminal domains enhance the catalytic activity of SIRT1.[7] It has also been reported that N- and C-terminal regions of SIRT7 mediate interactions with the C-terminal region of

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: Council of Scientific and Industrial Research, India grant to KGT and RP. Grant sponsor: Department of Biotechnology, The Innovative Young Biotechnologist Award (DBT-IYBA), India grant to KGT.
Conflict of Interest: Authors declare that they have no conflict of interest.
*Correspondence to: Krishan Gopal Thakur, CSIR-Institute of Microbial Technology, Sector 39A, Chandigarh-160036, India. E-mail: krishang@imtech.res.in
Received 18 March 2016; Revised 2 June 2016; Accepted 2 June 2016
Published online 10 June 2016 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/prot.25085

Myb binding protein (Mybbp1a) which plays an important role in regulating several transcription factors.[8]

These findings emphasize the importance of unique N- and C-terminal regions in mediating protein-protein interactions, protein localization, catalysis, oligomerization and hence play an important role in biological processes mediated by sirtuins. Crystal structures of the conserved catalytic domains of human sirtuins, excluding SIRT7 and SIRT4, are currently available but there is limited structural information on the N- and C-terminal domains of most of the sirtuins. Thus, it becomes imperative to characterize these domains to gain useful insights in understanding sirtuin biology. Here, we report 2.3 Å resolution crystal structure of the SIRT7^{NTD} solved as a MBP-SIRT7^{NTD} fusion protein. We also show here that N- and C-terminal domains of SIRT7 aid in protein solubility.

MATERIALS AND METHODS

Cloning, overexpression, and purification of MBP-SIRT7^{NTD}

The full-length codon optimized human *sirt7* was synthesized commercially (GenScript). Among various constructs of SIRT7, only MBP-SIRT7^{NTD} encompassing the region 5-73, produced diffraction quality crystals suitable for structural analyses. The detailed cloning, protein purification protocols and engineered linker sequence in the pMAL-c2X vector are available in the Supporting Information (Supporting Information Figs. S1, S2, and refer to the section "cloning, expression and purification of SIRT7 constructs" in the Supporting Information).

Peptide mass fingerprinting

The in-gel trypsin digestion of the desired protein bands resolved on 15% SDS PAGE was performed as per the protocol mentioned in the kit (InGel^{TM} Blue kit, G-Biosciences). The peptide mass data obtained using MALDI-TOF (AB SCIEX 5800) was further analysed using Protein Analysis Worksheet (PAWS) software (Freeware edition for windows 95/98/NT/2000, ProteoMetrics, LLC, NY).

Crystallization, data collection, and structure determination

The crystallization trials of MBP-SIRT7^{NTD} were set up using commercial screens from Hampton Research and Molecular Dimension using 96 well vapour diffusion sitting drop trays at 291 K. Initial hits were observed within 2–3 days of setting up of crystallization trials. Conditions were further optimized to produce diffraction quality crystals. The crystals harvested for data collection were obtained by mixing 2 µL of protein with 2 µL of reservoir solution containing 0.1*M* Sodium cacodylate, pH 5.5, and 25% w/v PEG 4000.

Table I
Data Collection and Refinement Statistics

Data collection	
Space group	P 1 21 1
Cell dimensions	
a, *b*, *c* (Å)	60.49, 49.33, 65.68
α, β, γ (°)	90.00, 101.20, 90.00
Resolution (Å)	48.60–2.33
R_{merge}	0.112 (0.391)
$I/\sigma I$	8.1 (3.0)
Completeness (%)	98.8 (96.4)
Redundancy	3.7 (3.5)
Wavelength (Å)	0.97947
Refinement	
Resolution (Å)	48.60–2.33
No. of unique reflections	16,201
R_{free} test set (%)	5.07
R_{work}/R_{free}	0.183/0.237
Atom count	
Protein	3318
Water	217
Ligand	23
B-factor	
Protein	33.2
Water	32.9
Ligand	24.6
r.m.s.d.[a]	
Bond lengths (Å)	0.002
Bond angles (°)	0.644
Structure Validation	
Ramachandran plot statistics	
Most favored (%)	97.87
Allowed regions (%)	2.13
Disallowed region (%)	0

Values in the parentheses correspond to the highest resolution shells.
[a]r.m.s.d., root mean square deviation.

X-ray diffraction data were collected on synchrotron beam line at PX-BL21, RRCAT, Indore, India. Data were processed using iMOSFLM[9] and scaled using SCALA.[10] PHASER[11] was used to solve structure by molecular replacement using crystal structure of MBP (PDB ID: 1ANF) as a search template. Several iterative rounds of manual model building using COOT[12] and refinement using PHENIX.REFINE[13] were used to generate model with R/R_{free} of 0.183/0.237 respectively. Molprobity was used for structure validation.[14] Data collection and refinement statistics are given in Table I. Molecular graphics figures were created using Pymol and Chimera.[15,16]

RESULTS

Region connecting SIRT7^{NTD} and catalytic domain is prone to proteolytic degradation

Full-length SIRT7 with N-terminal 6xHis-tag could be purified using Ni-NTA based affinity chromatography. We repeatedly observed two intense bands in the elution fractions [Supporting Information Fig. S3(A,B)]. To confirm the identity of the proteins both upper and lower bands were excised and subjected to in-gel trypsin

digestion followed by peptide mass fingerprinting. PAWS analysis suggested that upper band corresponds to the full length SIRT7 [Supporting Information Fig. S4(A,B)] and the lower band corresponds to the N-terminal region of SIRT7 [Supporting Information Fig. S4(C,D)]. This data suggested that may be a long flexible linker susceptible to proteolytic cleavage connects N-terminal region and the catalytic core. Since, the construct was having N-terminal 6xHis-tag so we could copurify N-term domain along with the full length protein. Moreover, in the gel filtration profile we could see SIRT7 eluting as a monomer [Supporting Information Fig. S3(C)]. Based on peptide mass fingerprinting followed by PAWS analysis and secondary structure analysis using Psipred,[17] N-terminal domain of SIRT7, encompassing region 5–73 of SIRT7 was cloned in an engineered pMAL-c2X vector to yield MBP-SIRT7^{NTD} fusion protein. The resulting fusion protein could be purified to homogeneity and was further used for crystallization trials.

N- and C-terminal domains of SIRT7 help in soluble expression

Based on the above observation several variants of SIRT7 encompassing different regions were cloned in expression vectors to yield 6xHis-tag and MBP fusion variants. Most of the constructs expressed well but proteins could not be purified from the soluble fraction. We even could not get the conserved catalytic domain expressed in soluble form. Our data suggests that both N- and C-terminal domains of SIRT7 assist in soluble expression of SIRT7. The detailed information regarding the expression profile of constructs created and tested can be found in the Supporting Information (Supporting Information Fig. S1).

The SIRT7^{NTD} adopts a three-helical domain architecture

There is no sequence homolog of SIRT7^{NTD} available in PDB but structure could be solved at 2.3 Å resolution by molecular replacement using MBP as a search model. After initial rounds of refinement a clear electron density for the helices could be seen for the region corresponding to SIRT7^{NTD}. The final model includes 5–61 residues of SIRT7^{NTD}, −368 to −10 residues of MBP (for ease of visualization and structural analysis residues corresponding to fusion tag, including linker region, are numbered from −368 to 0 and residues corresponding to SIRT7^{NTD} are numbered from 5–61), and one molecule of maltose bound to the active site of MBP. SIRT7^{NTD} consist of three helices, α1 to α3 and a short 3_{10} helix from Ala40 to Glu42 connects α1 and α2. α1 is the longest helix consisting of 35 amino acids covering residue range from Gly5 to Arg37. The other two short helices α2 and α3, covering Ala45 to Leu51 and Ser54 to Glu60 respectively, lie almost perpendicular to the α1 [Fig. 1(A,B)]. The α1 is predominately rich in positively charged residues and a major part of the helix is solvent exposed [Fig. 1(C)]. Residues Leu29, Val32, Leu36, and Ile35 from α1, Leu50 and Leu51 from α2 and Leu57, Val58, and Leu61 from α3 assemble to form a small leucine rich hydrophobic core [Fig. 1(D)]. The side chains of several residue pairs, Arg11/Glu15, Glu22/Arg26, Glu23/Arg26, Glu23/Arg30, Glu27/Arg30, Glu27/Gln31, Ser33-HOH-Arg37, Gln31/Arg34, Arg34/Glu47, Lys38/Glu42, Lys38/Glu47, Ser44/Glu47, Glu46/Arg49, and Arg28/Ser54 participate in forming potential hydrogen bonds/salt bridges. An interesting network of these potential hydrogen bonds/salt bridges observed in the molecule is shown in [Fig. 1(B)]. The residues in the linker region "AAAQTNAAHM" adopt "HHHHHHHHCT" secondary structure conformation. N-terminal region of α1 in SIRT7^{NTD} interacts with MBP through several potential hydrogen bonds, water mediated interactions and other nonbonded interactions [Fig. 2(B)]. Notably, Arg16 interacts with –Thr318 and –Asp316 through side-chain-side-chain interactions and salt-bridge respectively. Glu10 interact with −Ser298 through side-chain-side-chain interactions. The water bridges are observed between Arg8-HOH140-(–Ala33), Ser9-HOH97-(–Ser34), Glu10-HOH145-(−Ser198), Glu10-HOH145-HOH152-(−Pro323), in MBP-SIRT7^{NTD} [Fig. 2(B)]. The Arg rich region from Leu61 to Gln73 (LQGRSRRREGLKR), coding for nuclear localization signal[5] could not be modelled due to poor electron density suggesting high disorder in the region.

Structure based bioinformatics analysis using Profunc server[18] provided potential hits for two DNA binding templates. One of the hits, Mad-Max, proto-oncogenic transcription factor (PDB ID: 1NLW) had E-value of 0.954 [Fig. 1(F)].[19] There are reports highlighting the association of Mad-Max complexes with mSin3 which recruits HDAC1 and HDAC2 onto the target genes for transcriptional repression.[20,21] This observation is interesting as physical interaction of SIRT7 and Myc, a structurally similar transcription factor, has been reported. Myc facilitates recruitment of SIRT7 onto the promoters of ribosomal proteins for repressing the expression of ribosomal protein and suppression of ER stress and prevents fatty liver disease.[22] So far there is no report suggesting DNA binding property of SIRT7. But structural similarity of SIRT7^{NTD} with the DNA binding region of Mad-Max transcription factor and need of chromatin DNA for deacetylase activity reported recently[23] opens up a new possibility of investigating such molecular interactions. Structural similarity based search performed using PDBeFOLD[24] with SIRT7^{NTD} from amino acid residues 21–61, gave ICAT (Inhibitor of β-catenin and TCF-4) helical domain (PDB IDs 1LUJ, 1T08, 1MLE) as the top hit. ICAT shares 18% sequence identity, 0.66 Q score and RMSD of 1.21 over 38 Cα atoms (PDB ID: 1LUJ) [Fig. 1(E)]. Helical domain of ICAT reportedly binds 10–12 armadillo repeats of β-catenin and negatively regulates Wnt signalling by disrupting interactions

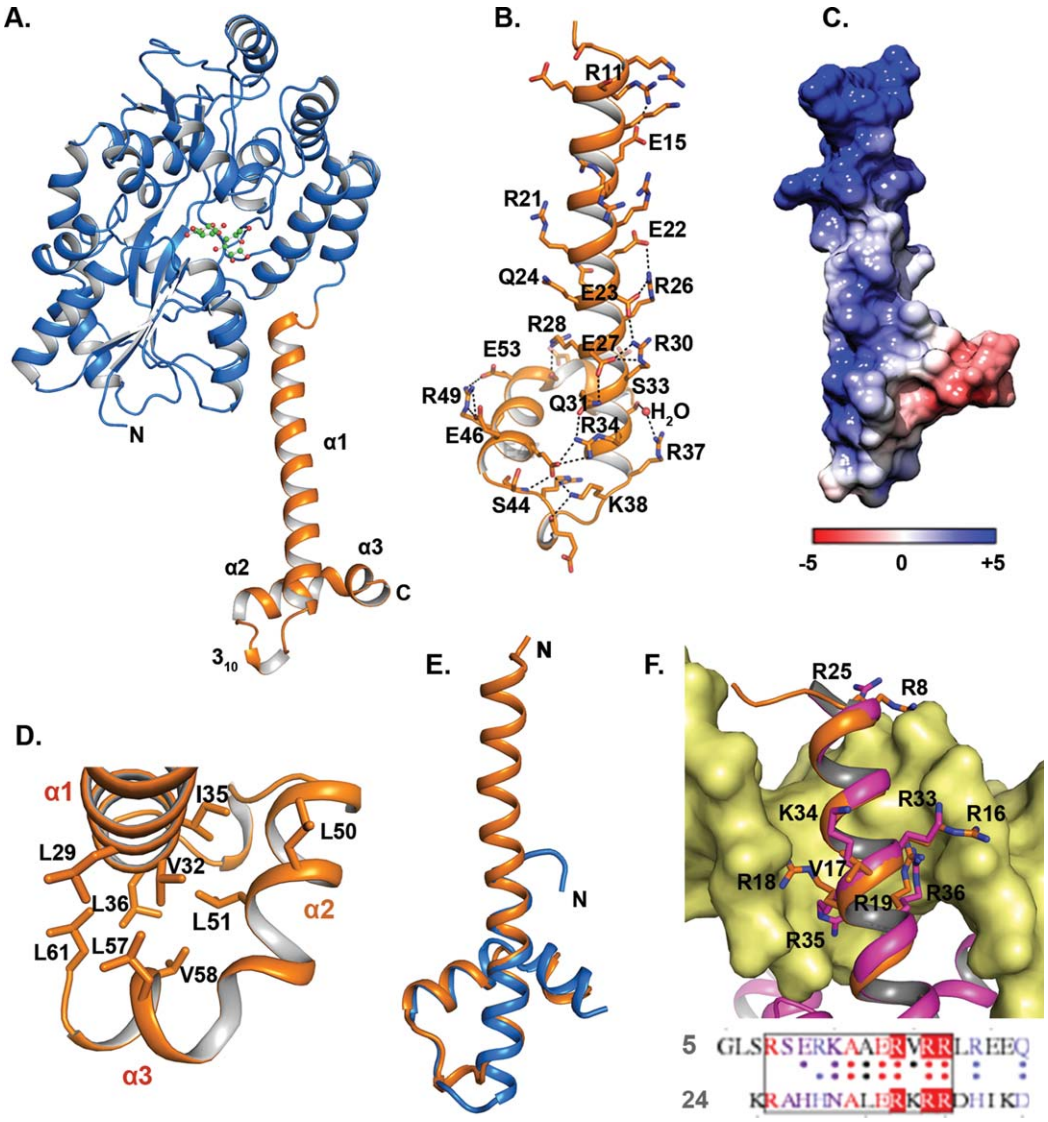

Figure 1

Three dimensional structure of the N-terminal domain of SIRT7. **A**: Cartoon representation of MBP-SIRT7^{NTD} fusion protein. MBP and SIRT7^{NTD} are highlighted in blue and orange colors, respectively. SIRT7^{NTD} adopts a three-helical domain architecture. Maltose bound to the active site of MBP is shown in ball and stick representation. **B**: Potential hydrogen bonds/salt bridges, shown in black broken line, are populated on one phase of the helix. **C**: Electrostatic potential surface representation of SIRT7^{NTD}. **D**: Close view of the Leu rich hydrophobic core. **E**: Structural superposition of SIRT7^{NTD} (orange) with the ICAT helical domain of β-cateninn/ICAT complex (blue) (PDB ID 1LUJ). **F**: Structural superposition of SIRT7^{NTD} (orange) with the DNA binding region of Mad-Max proto-oncogenic transcription factor (magenta) bound to DNA (pale yellow). The positively charged residues in the vicinity of DNA are shown in stick representation. Lower panel shows the sequence alignment of the DNA binding motif.

between β-catenin and TCF family of transcription factors.[25,26] The canonical Wnt pathway is involved in various kinds of cancers.[27] SIRT7 also reportedly plays role in several kinds of cancers.[28] Structural similarity of SIRT7^{NTD} with ICAT also raises possibility of involvement of the former in facilitating interactions of SIRT7 with different histone/nonhistone targets which might play role in normal physiological processes and diseased conditions like cancer. These structural insights need further investigations to validate relevance in physiological conditions.

Crystal packing analysis

MBP-SIRT7^{NTD} crystallized in space group P2$_1$ with one molecule in the asymmetric unit. We performed crystal packing analysis to investigate role of SIRT7^{NTD} in forming crystal contacts and to explore any potential self association. Our analysis suggests that each SIRT7^{NTD} molecule interacts with three symmetry related molecules [Fig. 2(A)]. All the three crystal contacts were observed between SIRT7^{NTD} and MBP molecules. No

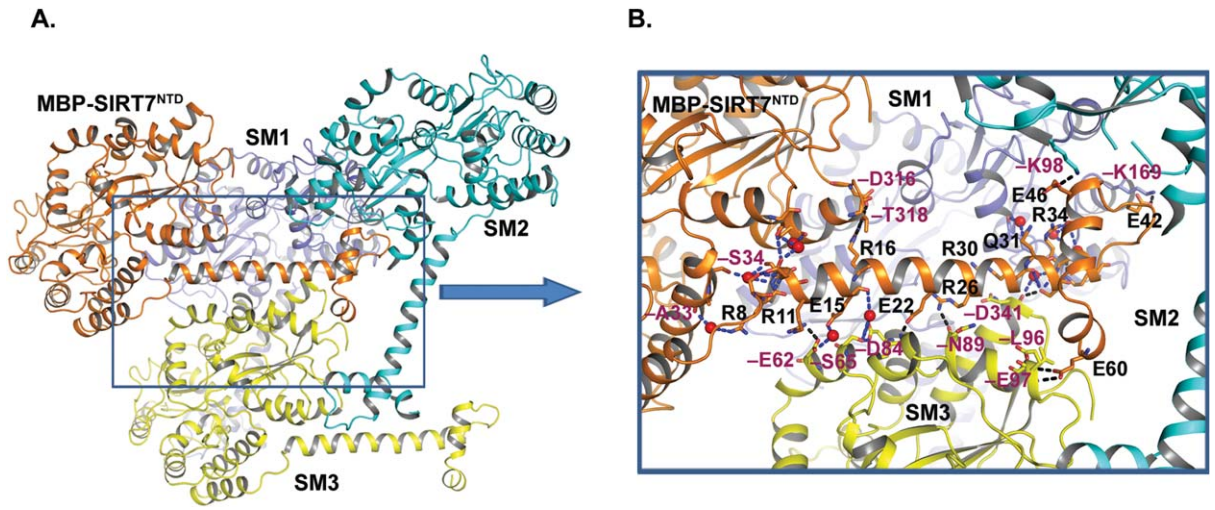

Figure 2

Crystal packing analysis. **A**: Symmetry mates, SM1-SM3, lying within 4 Å distance of SIRT7^{NTD} and involved in crystal contact are shown in different colours. **B**: Closer view of the interactions at the crystal packing interfaces. Residues participating in the interactions from asymmetric unit and the symmetry mates are shown in stick representation and are labeled in black and magenta colors, respectively. Water molecules are represented as red spheres.

intermolecular interactions were observed between the SIRT7^{NTD} molecules hence suggesting SIRT7^{NTD} do not self associate. The potential interactions at the crystal contact interface are shown in [Fig. 2(B)]. The side chains of Arg11, Glu22 and Arg30 of SIRT7^{NTD}, interact with side chains of −Glu62, −Lys283 and −Asp341, respectively in molecule 3. The side-chain of Arg26, Ser33 of SIRT7^{NTD} interact with main chain of −Asn89, −Asp341 respectively, and the side-chain of Glu60 interacts with main chain of −Leu96 and −Glu97 in molecule 3. Crystal contacts in molecule 1 are formed by bifurcated hydrogen bond between –Asp164 and Arg37. There is salt bridge between Glu42 and –Lys169 of molecule 1, and between Glu46 and –Lys98 of molecule 2. Several water mediated interactions also stabilize the crystal contacts [Fig. 2(B)].

DISCUSSION

Research in recent years has implicated SIRT7 in various types of disease cancers thus making SIRT7 as one of the potential drug targets.[28] Although, sirtuins are attractive drug targets for several diseases but they all share a conserved catalytic domain hence making it challenging to design specific modulators. Therefore, it becomes imperative to exploit allosteric site or structurally distinct N- and C-terminal regions of sirtuins which play role in protein-protein interactions and modulating catalytic activity. The N-terminal regions of sirtuins appear to adopt structurally distinct architecture as suggested by distinct combinations of predicted secondary structural elements and poor sequence conservation (Supporting Information Fig. S3). Structural analysis of SIRT7^{NTD} suggests structural similarity to regulators of transcription factors and possible role in DNA binding as well. While our manuscript was under preparation Tong *et al.* reported that SIRT7 requires both N- and C-terminal regions for its DNA activated deacetylase activity.[23] We hope this study would help provide useful insights to understand role of SIRT7^{NTD} mediated physiological functions and will also aid in rational drug design.

ACCESSION NUMBER

The coordinates and structure factors have been deposited at the Protein Data Bank, with accession code 5IQZ.

ACKNOWLEDGMENTS

The authors thank Gundeep Kaur for help in data collection at RRCAT, Indore. They thank Dr. Ravindra D. Makde, Mr. Ashwani Kumar and Dr. Biplab Ghosh, staff of PX-BL21 beamline, RRCAT, Indore, India for their assistance and support in data collection. AP and VS are graduate students recipient of UGC fellowship.

REFERENCES

1. Haigis MC, Sinclair DA. Mammalian sirtuins: biological insights and disease relevance. Annu Rev Pathol 2010;5:253–295.
2. Frye RA. Phylogenetic classification of prokaryotic and eukaryotic Sir2-like proteins. Biochem Biophys Res Commun 2000;273:793–798.

3. Ford E, Voit R, Liszt G, Magin C, Grummt I, Guarente L. Mammalian Sir2 homolog SIRT7 is an activator of RNA polymerase I transcription. Genes Dev 2006;20:1075–1080.
4. Tsai YC, Greco TM, Cristea IM. Sirtuin 7 plays a role in ribosome biogenesis and protein synthesis. Mol Cell Proteomics 2014;13:73–83.
5. Kiran S, Chatterjee N, Singh S, Kaul SC, Wadhwa R, Ramakrishna G. Intracellular distribution of human SIRT7 and mapping of the nuclear/nucleolar localization signal. FEBS J 2013;280:3451–3466.
6. Zhao K, Chai X, Clements A, Marmorstein R. Structure and autoregulation of the yeast Hst2 homolog of Sir2. Nat Struct Biol 2003;10: 864–871.
7. Pan M, Yuan H, Brent M, Ding EC, Marmorstein R. SIRT1 contains N- and C-terminal regions that potentiate deacetylase activity. J Biol Chem 2012;287:2468–2476.
8. Karim MF, Yoshizawa T, Sato Y, Sawa T, Tomizawa K, Akaike T, Yamagata K. Inhibition of H3K18 deacetylation of Sirt7 by Myb-binding protein 1a (Mybbp1a). Biochem Biophys Res Commun 2013;441:157–163.
9. Battye TG, Kontogiannis L, Johnson O, Powell HR, Leslie AG. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr D 2011;67:271–281.
10. Collaborative Computational Project N. The CCP4 suite: programs for protein crystallography. Acta Crystallogr D 1994;50:760–763.
11. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr 2007;40:658–674.
12. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D 2004;60:2126–2132.
13. Echols N, Grosse-Kunstleve RW, Afonine PV, Bunkoczi G, Chen VB, Headd JJ, McCoy AJ, Moriarty NW, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Adams PD. Graphical tools for macromolecular crystallography in PHENIX. J Appl Crystallogr 2012;45: 581–586.
14. Chen VB, Arendall WB, III, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D 2010;66:12–21.
15. Delano WL. The PyMOL Molecular Graphics System. 2002.
16. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera–a visualization system for exploratory research and analysis. J Comput Chem 2004;25:1605–1612.
17. Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol 1999;292:195–202.
18. Laskowski RA, Watson JD, Thornton JM. ProFunc: a server for predicting protein function from 3D structure. Nucleic Acids Res 2005; 33:W89–93. Web Server issue):-.
19. Nair SK, Burley SK. X-ray structures of Myc-Max and Mad-Max recognizing DNA. Molecular bases of regulation by proto-oncogenic transcription factors. Cell 2003;112:193–205.
20. Laherty CD, Yang WM, Sun JM, Davie JR, Seto E, Eisenman RN. Histone deacetylases associated with the mSin3 corepressor mediate mad transcriptional repression. Cell 1997;89:349–356.
21. Hassig CA, Fleischer TC, Billin AN, Schreiber SL, Ayer DE. Histone deacetylase activity is required for full transcriptional repression by mSin3A. Cell 1997;89:341–347.
22. Shin J, He M, Liu Y, Paredes S, Villanova L, Brown K, Qiu X, Nabavi N, Mohrin M, Wojnoonski K, Li P, Cheng HL, Murphy AJ, Valenzuela DM, Luo H, Kapahi P, Krauss R, Mostoslavsky R, Yancopoulos GD, Alt FW, Chua KF, Chen D. SIRT7 represses Myc activity to suppress ER stress and prevent fatty liver disease. Cell Rep 2013;5:654–665.
23. Tong Z, Wang Y, Zhang X, Kim DD, Sadhukhan S, Hao Q, Lin H. SIRT7 Is Activated by DNA and Deacetylates Histone H3 in the Chromatin Context. ACS Chem Biol, in press.
24. Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr D 2004;60:2256–2268.
25. Graham TA, Clements WK, Kimelman D, Xu W. The crystal structure of the beta-catenin/ICAT complex reveals the inhibitory mechanism of ICAT. Mol Cell 2002;10:563–571.
26. Tago K, Nakamura T, Nishita M, Hyodo J, Nagai S, Murata Y, Adachi S, Ohwada S, Morishita Y, Shibuya H, Akiyama T. Inhibition of Wnt signaling by ICAT, a novel beta-catenin-interacting protein. Genes Dev 2000;14:1741–1749.
27. Polakis P. Wnt signaling and cancer. Genes Dev 2000;14:1837–1851.
28. Kiran S, Anwar T, Kiran M, Ramakrishna G. Sirtuin 7 in cell proliferation, stress and disease: rise of the Seventh Sirtuin!. Cell Signal 2015;27:673–682.