



Published in final edited form as:

*Cell Cycle*. 2008 May ; 7(10): 1467–1472.

## Analysis of DBC1 and its homologs suggests a potential mechanism for regulation of Sirtuin domain deacetylases by NAD metabolites

Vivek Anantharaman and L. Aravind\*

National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA

### Abstract

Deleted in Breast Cancer-1 (DBC1) and its paralog CARP-1 are large multi-domain proteins, with a nuclear or perinuclear localization, and a role in promoting apoptosis upon processing by caspases. Recent studies on human DBC1 show that it is a specific inhibitor of the sirtuin-type deacetylase, Sirt1, which deacetylates histones and p53. Using sensitive sequence profile searches and HMM-HMM comparisons we show that the central conserved globular domain present in the DBC1 and its homologs from diverse eukaryotes is a catalytically inactive version of the Nudix hydrolase (MutT) domain. Given that Nudix domains are known to bind nucleoside diphosphate sugars and NAD, we predict that this domain in DBC1 and its homologs binds NAD metabolites such as ADP-ribose. Hence, we propose that DBC1 and its homologs are likely to regulate the activity of SIRT1 or related deacetylases by sensing the soluble products or substrates of the NAD-dependent deacetylation reaction. The complex domain architectures of the members of the DBC1 family, which include fusions to the RNA-binding S1-like domain, the DNA-binding SAP domain and EF-hand domains, suggest that they are likely to function as integrators of distinct regulatory signals including chromatin protein modification, soluble compounds in NAD metabolism, apoptotic stimuli and RNA recognition.

### Introduction

DBC1 was originally identified as a gene homozygously deleted in breast cancer and some other tumors<sup>1</sup>. It encodes a large protein, which has been shown to act in a pro-apoptotic fashion upon processing by caspases<sup>2</sup>. Its close paralog in vertebrates CARP-1 (CCAR1) is also a pro-apoptotic gene that up-regulates p21<sup>WAF1/CIP1</sup> and down-regulates cyclin B1<sup>3</sup>. DBC1 and CARP-1 proteins share an N-terminal coiled-coil (zipper) region and multiple conserved globular domains, including a C-terminal inactive EF-hand module that is unlikely to bind Ca<sup>2+</sup> ions<sup>2</sup>. Homologous proteins with conservation spanning the above regions shared by DBC1 and CARP-1 (the DBC1 family) are also found in plants and the slime mold *Dictyostelium*, suggesting that the family emerged prior to the divergence of the crown group of eukaryotes. Both DBC1 and CARP-1 have been shown to localize to the nucleus or the nuclear envelope, but they might also localize to the cytoplasm especially during apoptosis<sup>2,3</sup>. However, their exact biochemical roles remained unclear until two recent concomitant reports on human DBC1. These studies showed that DBC1 is a specific inhibitor of the enzyme SIRT1, which mediates NAD-dependent deacetylation of proteins such as histones and p53<sup>4,5</sup>. Thus, by inhibiting the anti-apoptotic SIRT1, DBC1 is able to promote apoptotic pathways via p53 hyperacetylation. It might also counter the repressive action of SIRT1 mediated lysine deacetylation of histones<sup>4,5</sup>. While the conserved coiled-coil region in DBC1 has been shown

\*for correspondence: aravind@mail.nih.gov, Phone: 301-594-2445.

to be required for its interaction with SIRT1 (Fig. 1), the roles of the conserved globular domains present in the DBC1 family are unclear. Given that these domains are conserved throughout the eukaryotic crown group we were interested in investigating if they might throw light on the conserved regulatory processes mediated by the DBC1 family of proteins.

In this study, using sensitive sequence profile analysis methods and structural comparisons, we show that the largest central conserved globular domain present in the DBC1 family is an inactive version of the Nudix hydrolase (MutT) fold, and is likely to bind NAD derivatives such as ADP-ribose. Based on this we predict that DBC1 and its homologs are likely to regulate the activity of SIRT1 or related deacetylases by sensing the soluble products or substrates of the NAD-dependent deacetylation reaction.

## Results and Discussion

### Identification of an inactive version of the Nudix (MutT) domain in the DBC1 family

To computationally investigate the possible biochemical roles of DBC1 and CARP-1 we sought to systematically identify the conserved domains found in them. Due to the presence of large non-globular and compositionally biased regions in these proteins we first used the SEG program to objectively identify their globular domains (Materials and Methods). To determine the relationships of these predicted globular regions, we then queried them against a library of sensitive position-specific-score matrices (PSSMs) and hidden Markov models (HMM) of previously characterized domains<sup>6</sup>. As a result, it became clear that DBC1, CARP-1 and all other members of the family shared two previously uncharacterized globular regions in addition to the C-terminal inactive EF-hand modules<sup>7</sup>. All members of the DBC1 family from animals, and one of the two paralogs from the chlorophyte *Volvox* also contain an N-terminal S1-like RNA-binding domain (Fig. 2). Further, CARP-1 (but not DBC1) and all invertebrate members of the family possess a DNA-binding SAP domain inserted between the two uncharacterized globular domains shared by all members of the DBC1 family (Fig. 2).

The larger of the two predicted uncharacterized globular domains shared by the DBC1 family is located just C-terminal to the conserved coiled-coil region, which, in DBC1, was shown to be required for the interaction of with SIRT1<sup>4, 5</sup> (Fig. 1). We searched the non-redundant database using a PSSM derived from an alignment that included all representatives of this predicted domain from the DBC1 family with the PSI-BLAST program. This search recovered several members of the Nudix hydrolase superfamily prior to convergence (e.g. diadenosine tetraphosphatase, gi: 4502125; NUDT8, gi: 32469515; *Pyrobaculum* Pcal\_0237, gi: 126458862; expect-value= $10^{-2}$ ). To further test the validity of this possible relationship we performed a HMM-HMM comparison with HHpred program, using the HMM of the above domain of the DBC1 family in a search against a panel of HMMs derived using the globular domains in the PDB database as seeds (Materials and Methods). The top hits in this search were the Nudix hydrolases with highly significant p-values ( $p=10^{-3}$ – $10^{-9}$ ). A reciprocal PSI-BLAST search with a PSSM prepared from an alignment of the Nudix domain including diverse members of the superfamily against a database of predicted proteins from all completely sequenced eukaryotic genomes also recovered members of the DBC1 family (e-value  $10^{-2}$ – $10^{-4}$ ). Prediction of the secondary structure with the Jpred program by using a combination of residue frequencies in a column, HMM and PSSM derived from the alignment of this domain of the DBC1 family showed that the conserved core contained four conserved strands and a helix between strand 2 and strand 3 (Fig. 1). This secondary structure corresponds exactly to the  $\beta$ -grasp fold found in the Nudix hydrolases<sup>8</sup>, which taken together with the results of the sequence profile analysis suggests that this conserved domain in the DBC1 family adopts the same fold as the Nudix hydrolases.

Nudix hydrolases hydrolyze diverse nucleoside diphosphates and contain a characteristic active site defined by a Gx5Ex7Rx4E motif on the conserved helix of their fold<sup>9, 10</sup> (Fig. 1). The acidic residues in this motif primarily participate in chelating Mg<sup>2+</sup> ions and as a general base in catalysis<sup>11</sup> (Fig. 1). The conserved arginine, while not directly participating in substrate contact, interacts with surrounding residues and is thereby critical for maintaining the substrate-binding pocket. Interestingly, the multiple alignment of the above-detected Nudix domain in the DBC1 family with other catalytically active members of the superfamily showed that the former only conserved the arginine, but not the other key acidic residues in the active site motif (Fig 1· 2). This observation indicates the Nudix domain is likely to be catalytically inactive in all members of the DBC1 family. However, the strict conservation of the arginine belonging to the active-site motif suggested it might stabilize the binding pocket typical of the Nudix superfamily even in the DBC1 family. To further investigate this possibility, we used the crystal structures of diadenosine tetraphosphatase (PBD: 1su2 and 1ktg)<sup>12</sup> and an ADP-ribose pyrophosphatase (PBD: 1nqz)<sup>13</sup> as templates to construct a structural model for the inactive Nudix domain of DBC1. It must be stressed that such homology based model for distantly related sequences are only rough approximations of the actual structure. Nevertheless, in the absence of the actual structure it can be used in conjunction with the conservation pattern to explore certain aspects, such as the general features of the binding pocket in the DBC1 family. Importantly, the conservation of a key tryptophan, small residue (usually proline) suggested that the inactive DBC1 Nudix domain preserves the characteristic “outflow” which interrupts the second strand of the β-grasp fold and forms the base of the substrate binding pocket (Fig. 1· 2)<sup>8</sup>. Further the model suggests that all the key residues lining the predicted active-site pocket are of comparable size to those observed in the active Nudix hydrolases, and are typically polar or positive charged. Hence, the Nudix domain of the DBC1 family, while catalytically inactive, would still be able to accommodate soluble ligands with a nucleoside diphosphate moiety.

### Functional implications for the predicted ligand-binding Nudix domain in the DBC1 family

The above discovery of a potential nucleoside diphosphate-binding Nudix domain in the DBC1 family has considerable functional implications in light of the recent experimental evidence for DBC1 being a SIRT1 inhibitor. The sirtuins, like Sirt1, catalyze deacetylation of acetyllysine in histones and other proteins by using NAD, which in the process is converted to O-acetyl ADP-ribose (OAAR), while releasing nicotinamide<sup>14</sup>. Both ADP sugars and NAD are well-known substrates of nudix hydrolases: for example, in animals NUDT9 is a specific ADP ribose hydrolase, whereas NudC from proteobacteria shows a substrate preference for NAD<sup>10, 11</sup>. Given this precedence, we suggest that the inactive Nudix domain found in the DBC1 family is likely to bind either the substrate (NAD) or the product (OAAR) of the SIRT1 enzyme. In particular by binding the ADP-ribose derivatives produced by the SIRT1 reaction DBC1 could sense the presence of Sirtuin deacetylase activity and specifically down-regulate it. This suggestion is supported by close proximity of the SIRT1 -binding coiled-coil region and the inactive Nudix domain (Fig. 1). In addition to the sirtuins, action of other NAD-dependent enzymes, such as poly-ADP-ribose polymerase which are also involved in apoptosis, might also alter concentrations of NAD metabolites sensed by the DBC1 family<sup>15</sup>. In this context, it is of interest to note that the calcium channel TRPM2 (NUDT9L1) also contains an intracellular Nudix domain with a partially disrupted active site<sup>16</sup>. This Nudix domain has low turnover and mediates channel-opening through ADP-ribose-binding and its closing through gradual hydrolysis of the bound ADP-ribose. Presence of the conserved inactive EF-hand module in the DBC1 proteins suggests that, although they might not bind Ca<sup>2+</sup>, they might hetero-dimerize with other functional EF-hand proteins activated by the Ca<sup>2+</sup> influx due to TRPM2-opening in response to high ADP-ribose concentration arising from elevated NAD metabolism.

Further, the presence of S1-like RNA-binding domains in several members of the DBC1 family suggests that they might additionally interact with RNA. Given that different S1-like domain proteins interact with both miRNAs<sup>17</sup> and mRNAs<sup>18</sup>, it would be of interest to investigate if the regulation chromatin modification by the DBC1 family is also regulated by the presence of transcripts or miRNAs. Similarly, the presence of a SAP domain in some members of the DBC1 family, like CARP-1, suggests that these versions are likely to localize to scaffold or matrix attachment sites<sup>19</sup>. Hence, in vertebrates, where both DBC1 and CARP-1 are present, there might be a compartmentalization of their roles -- the former functioning in the nuclear interior and the latter functioning in the nuclear periphery. Sequence profile searches initiated with different starting queries corresponding to the second uncharacterized globular domain in the DBC1 showed that it was an  $\alpha$ -helical domain restricted to the DBC1 family. It might mediate an additional conserved interaction of this family with as yet unknown partners.

## Conclusions

The recent finding that DBC1 interacts with the histone and p53 deacetylase Sirt1 has provided the first hints regarding the actual biochemical functions of these conserved multi-domain proteins. We present evidence that the DBC1 family contains a catalytically inactive, ligand-binding version of the Nudix domain located immediately C-terminal to the coiled-coil region which interacts with Sirt1. We hence speculate that the DBC1 family is likely to sense soluble products of NAD metabolism, such as ADP-sugars produced by the action of sirtuins, thereby regulating these deacetylases. This would imply that the Nudix domains of the DBC1 family are functionally analogous to the Macro domains found in histone macroH2A1, certain sirtuins and Swi2/Snf2 ATPases<sup>6</sup>. Certain Macro domains are enzymes which convert ADPR-1'' monophosphate to ADPR<sup>20</sup>, whereas others are catalytically inactive high-affinity ADPR-binding domains<sup>15</sup>. The latter version of the Macro domain in macroH2A1 binds OAADPR formed by Sirt1 action and might thereby regulate local heterochromatin structure<sup>21</sup>. The complex domain architectures of the DBC1 proteins suggest that they are likely to act in linking a variety of distinct regulatory process including chromatin protein deacetylation, NAD metabolism, apoptotic signaling and RNA metabolism. Given the involvement of members of the DBC1 family in tumorigenesis and apoptosis, the prediction of a soluble ligand-binding domain in the makes them attractive as potential targets for therapeutic intervention with small molecules.

## Material and Methods

The non-redundant (NR) database of protein sequences (National Center for Biotechnology Information, NIH, Bethesda) was searched using the PSI-BLAST programs<sup>22</sup>. Profile searches using the PSI-BLAST program were conducted either with a single sequence or a sequence with a PSSM used as the query, with a profile inclusion expectation (E) value threshold of 0.01, and were iterated until convergence<sup>22</sup>. For all compositionally biased queries the correction using composition-based statistics was used in the PSI-BLAST searches<sup>23</sup>. Multiple alignments were constructed using the Kalign program<sup>24</sup>, followed by manual correction based on the PSI-BLAST results. The multiple alignment was used to create a HMM using the Hmmbuild program of the HMMER package<sup>25</sup>. It was then optimized with Hmmlcaliberate and the resulting profile was used to search a database of completely sequenced genomes using the Hmmsearch program of the HMMER package<sup>25</sup>. HMM-HMM comparison searches were performed using the HHpred program<sup>26, 27</sup>. The JPRED program<sup>28</sup> and the COILS program were used to predict secondary structure. Globular domains were predicted using the SEG program with the following parameters: window size 40, trigger complexity=3.4; extension complexity=3.75<sup>29</sup>.

The Swiss-PDB viewer<sup>30</sup> and Pymol programs<sup>31</sup> were used to carry out manipulations of PDB files. The model was generated using SWISS-MODEL<sup>32</sup> (<http://swissmodel.expasy.org/>). Briefly, this process consisted of threading the human DBC1 Nudix domain on to the known structures of Nudix domains (templates) using the multiple sequence alignment and the alignment window of the SWISS-PDB viewer program. This alignment of the DBC1 Nudix domain to the template structures was then submitted for modeling to the SwissModel server. Energy minimization of the modeled domain was carried out using the GROMOS 43B1 force field incorporated in SwissModel.

#### Acknowledgements

The work by the authors was supported by the intramural funds of the National Library of Medicine, NIH.

#### References

1. Hamaguchi M, Meth JL, von Klitzing C, Wei W, Esposito D, Rodgers L, Walsh T, Welch P, King MC, Wigler MH. DBC2, a candidate for a tumor suppressor gene involved in breast cancer. *Proc Natl Acad Sci U S A* 2002;99:13647–52. [PubMed: 12370419]
2. Sundararajan R, Chen G, Mukherjee C, White E. Caspase-dependent processing activates the proapoptotic activity of deleted in breast cancer-1 during tumor necrosis factor-alpha-mediated death signaling. *Oncogene* 2005;24:4908–20. [PubMed: 15824730]
3. Rishi AK, Zhang L, Yu Y, Jiang Y, Nautiyal J, Wali A, Fontana JA, Levi E, Majumdar AP. Cell cycle- and apoptosis-regulatory protein-1 is involved in apoptosis signaling by epidermal growth factor receptor. *J Biol Chem* 2006;281:13188–98. [PubMed: 16543231]
4. Kim JE, Chen J, Lou Z. DBC1 is a negative regulator of SIRT1. *Nature* 2008;451:583–6. [PubMed: 18235501]
5. Zhao W, Kruse JP, Tang Y, Jung SY, Qin J, Gu W. Negative regulation of the deacetylase SIRT1 by DBC1. *Nature* 2008;451:587–90. [PubMed: 18235502]
6. Iyer LM, Anantharaman V, Wolf MY, Aravind L. Comparative genomics of transcription factors and chromatin proteins in parasitic protists and other eukaryotes. *Int J Parasitol* 2008;38:1–31. [PubMed: 17949725]
7. Lewit-Bentley A, Rety S. EF-hand calcium-binding proteins. *Curr Opin Struct Biol* 2000;10:637–43. [PubMed: 11114499]
8. Burroughs AM, Balaji S, Iyer LM, Aravind L. Small but versatile: the extraordinary functional and structural diversity of the beta-grasp fold. *Biol Direct* 2007;2:18. [PubMed: 17605815]
9. Koonin EV. A highly conserved sequence motif defining the family of MutT-related proteins from eubacteria, eukaryotes and viruses. *Nucleic Acids Res* 1993;21:4847. [PubMed: 8233837]
10. McLennan AG. The Nudix hydrolase superfamily. *Cell Mol Life Sci* 2006;63:123–43. [PubMed: 16378245]
11. Mildvan AS, Xia Z, Azurmendi HF, Saraswat V, Legler PM, Massiah MA, Gabelli SB, Bianchet MA, Kang LW, Amzel LM. Structures and mechanisms of Nudix hydrolases. *Arch Biochem Biophys* 2005;433:129–43. [PubMed: 15581572]
12. Ranatunga W, Hill EE, Mooster JL, Holbrook EL, Schulze-Gahmen U, Xu W, Bessman MJ, Brenner SE, Holbrook SR. Structural studies of the Nudix hydrolase DR1025 from *Deinococcus radiodurans* and its ligand complexes. *J Mol Biol* 2004;339:103–16. [PubMed: 15123424]
13. Kang LW, Gabelli SB, Bianchet MA, Xu WL, Bessman MJ, Amzel LM. Structure of a coenzyme A pyrophosphatase from *Deinococcus radiodurans*: a member of the Nudix family. *J Bacteriol* 2003;185:4110–8. [PubMed: 12837785]
14. Denu JM. The Sir 2 family of protein deacetylases. *Curr Opin Chem Biol* 2005;9:431–40. [PubMed: 16122969]
15. Karras GI, Kustatscher G, Buhecha HR, Allen MD, Pugieux C, Sait F, Bycroft M, Ladurner AG. The macro domain is an ADP-ribose binding module. *Embo J* 2005;24:1911–20. [PubMed: 15902274]
16. Naziroglu M. New molecular mechanisms on the activation of TRPM2 channels by oxidative stress and ADP-ribose. *Neurochem Res* 2007;32:1990–2001. [PubMed: 17562166]



17. Moss EG, Lee RC, Ambros V. The cold shock domain protein LIN-28 controls developmental timing in *C. elegans* and is regulated by the *lin-4* RNA. *Cell* 1997;88:637–46. [PubMed: 9054503]
18. Abaza I, Gebauer F. Functional domains of *Drosophila* UNR in translational control. *Rna*. 2008
19. Aravind L, Koonin EV. SAP - a putative DNA-binding motif involved in chromosomal organization. *Trends Biochem Sci* 2000;25:112–4. [PubMed: 10694879]
20. Shull NP, Spinelli SL, Phizicky EM. A highly specific phosphatase that acts on ADP-ribose 1"-phosphate, a metabolite of tRNA splicing in *Saccharomyces cerevisiae*. *Nucleic Acids Res* 2005;33:650–60. [PubMed: 15684411]
21. Kustatscher G, Hothorn M, Pugieux C, Scheffzek K, Ladurner AG. Splicing regulates NAD metabolite binding to histone macroH2A. *Nat Struct Mol Biol* 2005;12:624–5. [PubMed: 15965484]
22. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997;25:3389–402. [PubMed: 9254694]
23. Schaffer AA, Aravind L, Madden TL, Shavirin S, Spouge JL, Wolf YI, Koonin EV, Altschul SF. Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements. *Nucleic Acids Res* 2001;29:2994–3005. [PubMed: 11452024]
24. Lassmann T, Sonnhammer EL. Kalign, Kalignvu and Mumsa: web servers for multiple sequence alignment. *Nucleic Acids Res* 2006;34:W596–9. [PubMed: 16845078]
25. Eddy SR. Profile hidden Markov models. *Bioinformatics* 1998;14:755–63. [PubMed: 9918945]
26. Soding J. Protein homology detection by HMM-HMM comparison. *Bioinformatics* 2005;21:951–60. [PubMed: 15531603]
27. Soding J, Biegert A, Lupas AN. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* 2005;33:W244–8. [PubMed: 15980461]
28. Cuff JA, Barton GJ. Application of multiple sequence alignment profiles to improve protein secondary structure prediction. *Proteins* 2000;40:502–11. [PubMed: 10861942]
29. Wootton JC. Non-globular domains in protein sequences: automated segmentation using complexity measures. *Comput Chem* 1994;18:269–85. [PubMed: 7952898]
30. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 1997;18:2714–23. [PubMed: 9504803]
31. DeLano, WL. The PyMOL Molecular Graphics System. San Carlos, CA, USA: DeLano Scientific; 2002.
32. Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res* 2003;31:3381–5. [PubMed: 12824332]

**A**

Secondary Structure  
 DBC1\_Hsap\_40548408  
 CCAR1\_Hsap\_119574686  
 00016490:G:001\_Tnig\_47223307  
 LOC568087\_Drer\_68370963  
 LOC564155\_Drer\_68366944  
 LOC575098\_Spur\_115974585  
 LOC655829\_Tcas\_91076266  
 ENSANG00000018266\_Agam\_58380568  
 LOC409891\_Amel\_110764054  
 Dpu11000010289\_Dpul  
 Caps1000013887\_Caps  
 lst-3\_Cele\_71966159  
 Bml\_48755\_Bmal\_158591281  
 LOC100121576\_Nvit\_156550661  
 EMB1579\_Acha\_42568895  
 PHYPADRAFT\_170003\_Fpat  
 Vcar1000014345\_Vcar  
 Vcar1000004204\_Vcar  
 DGL124\_ddis\_66820685  
 ndx-4\_Cele\_51316979  
 nudE\_Ecol\_49176450  
 nudF\_Ecol\_16131274  
 nudF\_Ecol\_15803581  
 nudK\_Ecol\_16130392  
 nudD\_Ecol\_89108871  
 NUDT7\_Hsap\_4502125  
 NUD9\_Hsap\_20455184  
 TRPM2\_Hsap\_109730277  
 NUDT10\_Hsap\_41393549  
 DCP2\_Hsap\_31542498  
 NUDT8\_Hsap\_68565920  
 NUDT7\_Hsap\_68565858  
 idi\_Ecol\_16130791  
 Rv1700\_Mtub\_15608838  
 DR\_1184\_Drad\_15806203  
 DTPHydrolase\_Lang\_159162286  
 TTHA1800\_Tthe\_55981849  
 ysdF\_Ecol\_15803349  
 ysdF\_Pae\_15595533  
 Rv3908\_Mtub\_15611044  
 Dr1025\_Drad\_49259148  
 Consensus/85%  
 Key residues:

```

-----EEEEEEEE-----EEEEEE-----EEE-EEE-EE-----H---HHHHHHHHHHHH-----EEEEEE
339 PEHPLKIQVFLGRK EE-EAVLV GG-EWS--PSLDGLD-PQADPQVLRVTAICCAQQTGID---LSGCT
461 FQHPARILVFLVGMK GKDEAMAI GG-HWS--PSLDGPD-PEKDPVSLIKTAICCKALITGID---LSVCT
489 FQHPARILVFLVGMK GKDEAMAI GG-HWS--PSLDGAD-PERDPSVLIKTAICCKALITGID---LSLCT
767 ESLEPEVEFVFLVGMK GKDEAMAI GG-HWS--PSLDGAD-PHDASVLIKTAICCKALITGID---LSLCT
305 AVHPSSTIKFLIVDS GG-EQRLI GG-HWS--PEADGAM-PAKDSLIVMTAVICLQKSGID---LSACT
460 FQHPTRLINFLVGMK GKDEAMAI GG-HWS--PSLDGAD-PVNDTSLIKTAITTKALAGID---LTACT
424 FVHPARILVFLVGLR GKNETMAI GG-PWS--PSLDGDN-PDKDPSVLIKTAITCKALITGID---LSSCT
164 YLHPTRMIFLIVGTR GKNETMAV GG-PWS--PSLDGAD-PDTPKVMRTAITCKALITGID---LSACS
430 YVHPTRLINFLVGLR GKNETMAI GG-PWS--PSLDGPN-PEKDPVSLIRTAITCKALITGID---LSSCT
397 FVHPARILVFLVGLR GKNETMAI GG-PWS--PSLDGAD-PKSDPVLIRTAITCKALITGID---LSSCT
461 FQHPTRLIFLVLGK GK-EPMAI GG-PWS--PSLDGAE-PLKDPRLVINTAITTRGLITGID---LSSCT
438 PQSLKNLIFLIVGAR GK-ETMI GG-SWS--PQDGDG-PNS-ATTMIRTAITTKSLITGID---LSSVS
439 SVSLKRCNLVFGVTR NK-EMMAI GG-AWS--PSLDGDN-PETDPQVVRTAIVTRVALITGID---LSRCP
428 YIHPTRLINFLVGLR GKNETMAI GG-PWS--PSLDGPN-PDKDPSVLIKTAITCKALITGID---LNSCT
595 IPHICNLIFLVLK DK-SPMAI GG-SWD--PT-DGMD-PSVQSSLIQTMLHSHKQLHLD---LSNCR
999 PAHMHNLVFLVLRK DRSAIMAI GG-FWD--PELDGPD-PDQDGLIKTAICTREYLITGID---LQCK
929 GAHVLHALVFAARE 23 QRVLGAL GG-TWD--PELDGCD-PESEDGLIKTAICTREYLITGID---LSACA
559 YQHPRLKMLLVQKR 22 GGNEVAV GG-CWE--PV-DGCH-PARDAALVATAVTTRRATVITGID---LSPCT 13
603 DPSITKIKFLVSRK DK-DISCI GG-TWL--PDLGEG-VND-PQTLINTAITTKYETDITGID---LSGCK
1 MVVKAAGLVYRRLA GKIEFLLL 4 PPHHTP-PK-GHVD-PGEDW--QAARTEKEANITKEQLTIE 5 \Active
125 PQIAPCLIVAIR 6 NG-VHTV-LA-GFVE-VGETL--QAVAEVMESEGIK---VKNLR 5 \NUDIX
45 TNREAVNIVPIV--- -DDHLILI 7 ESYELGF-SK-GLID-PGESVY---EAANREIKVEEAGI---ANDLT 1
55 ERGHAVALVFPDF--- -VRDEVLVI 11 ET-PHLEEMVAGMIE-EGESVE---DVARREIAEAGII---VWRKT 1
43 DRNGAIIILVYNT--- -KPKTIVLI 13 SG-QLIE-SC-AGLL-DNDEPE---VCRREIAEAGIPE---VQVSR 1
15 TPLVSLDFIVEN--- -SRGFLIG 7 GG-YWFV-PG-GRVQ-KDETL---AAPERLIMAEIGLVR-ITAGQ 3
1 MALRACCLIFERRCL 6 NAIEFLLL 4 GIHHTP-PK-GHVE-PGEDDL---ETALITEQEALAEAGLITIE 6
169 HAADPIITPKRDS 11 HILQVVI 4 CG-EMAI-PG-GHVD-PKFKIS---ATKAEVEEALNSL---QRPSA 16
396 HTLYPMVTMRRERD 9 KMLEVIV 3 LSEHVAL-PG-GSRP-PGEMLP---RYKKTILIRBHWSE---ENLKI 1
95 FKRAACLFERSE--- -REDEVLLV 5 PD-RWIV-PG-GMPE-PEPEPG---GAAVEVVEEAGK-LQRL-
TLENVLLV 5 KS-GWGF-PK-GKVN-KEEAPH---DCAAVEVEEETGG---IKDYI 3
29 PASAAILVFLSVR--- -GVPAALYT 10 KG-DVSP-PG-GKCD-PADQDVV---HTALRETRLEGLIA---VPEEH 1
37 YNKYSVLFLVAKV 39 KGLHLFT 9 PG-EVCF-PG-GKRD-PIDMDA---ATALREAEQVSRVPHQVEV 1
30 RLHLAFSSVLEN--- -AKGLLVT 9 PG-VWNSVC-GHPQ-LGESNE---DAVIRCRVLEIQT---ITPTE 1
40 VEHFGAVLIVAMD--- -DNGNIPV 8 RR-LWEL-PA-GLDVAEGPPH---LTAARELEVEEAGV---ASTWQ 1
32 HYRRAAVLVATREA 12 DPRVLLV 8 KG-QIAF-PG-GSLD-AGETPT---QAALREAEVADTP---AAVTL 2
12 GYRRNVGICLNM--- -NDKKIFAA 4 IPDAWQ-PQ-GSID-EGEDPR---NAATRELEVEEAGVVAIEV 1
1 -MLGAGGVVNA--- -KREVLIL 3 MG-FWVF-PK-GHPE-PGESLE---EAAVEVMEEIVRAE-VLLPL 3
6 GYRNVGICLNM--- -RQCQWMA 3 GHSWQF-PQ-GGLN-PESAB---QAVLEPEGLSRKRVLLIA 3
6 GFRNVGICLNM--- -EAQILFA 3 NQDQFP-PQ-GGLN-DRETFE---SALYRELEVEEAGDVAIIA 3
63 HETSAGLVTDGID 3 AQVAALIG 5 GRLLNSL-PK-GHIE-LGTAEE---QTALEVAETELRSLVAALG 3
11 VEIRAACVVLIN--- -ERCDLILV 11 AG-LMHI-PS-GAVE-DGENTP---DAAVEACEETGVR---VRVPK 4 /
.....h.hh.....p.hhh.....s.w..P..s..p.s.p.....shhRp..b.uhp.....h.....
# # * * *

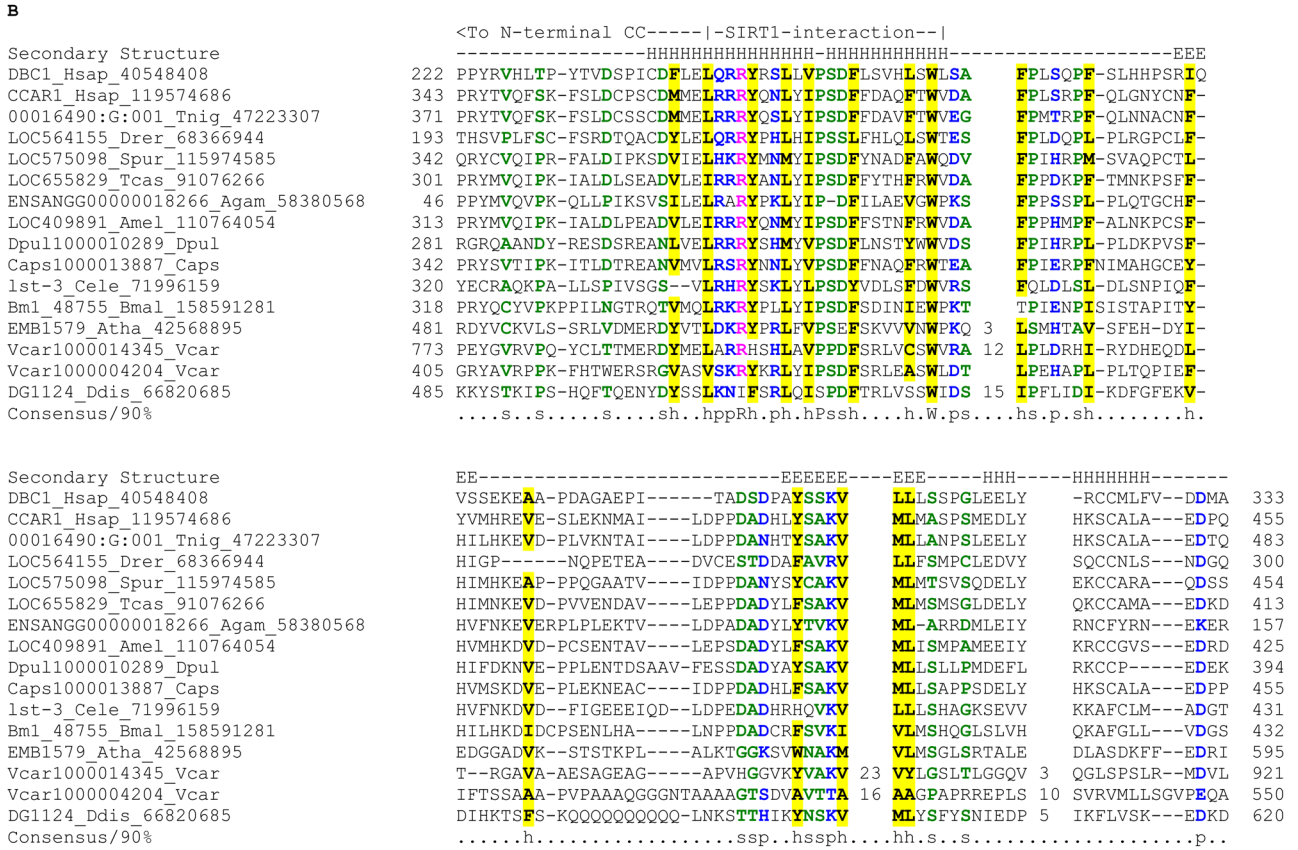
```

Secondary Structure  
 DBC1\_Hsap\_40548408  
 CCAR1\_Hsap\_119574686  
 00016490:G:001\_Tnig\_47223307  
 LOC568087\_Drer\_68370963  
 LOC564155\_Drer\_68366944  
 LOC575098\_Spur\_115974585  
 LOC655829\_Tcas\_91076266  
 ENSANG00000018266\_Agam\_58380568  
 LOC409891\_Amel\_110764054  
 Dpu11000010289\_Dpul  
 Caps1000013887\_Caps  
 lst-3\_Cele\_71966159  
 Bml\_48755\_Bmal\_158591281  
 LOC100121576\_Nvit\_156550661  
 EMB1579\_Acha\_42568895  
 PHYPADRAFT\_170003\_Fpat\_162670778  
 Vcar1000014345\_Vcar  
 Vcar1000004204\_Vcar  
 DGL124\_ddis\_66820685  
 ndx-4\_Cele\_51316979  
 nudE\_Ecol\_49176450  
 nudF\_Ecol\_16131274  
 nudF\_Ecol\_15803581  
 nudK\_Ecol\_16130392  
 nudD\_Ecol\_89108871  
 NUDT7\_Hsap\_4502125  
 NUD9\_Hsap\_20455184  
 TRPM2\_Hsap\_109730277  
 NUDT10\_Hsap\_41393549  
 DCP2\_Hsap\_31542498  
 NUDT8\_Hsap\_68565920  
 NUDT7\_Hsap\_68565858  
 idi\_Ecol\_16130791  
 Rv1700\_Mtub\_15608838  
 DR\_1184\_Drad\_15806203  
 DTPHydrolase\_Lang\_159162286  
 TTHA1800\_Tthe\_55981849  
 ysdF\_Ecol\_15803349  
 ysdF\_Pae\_15595533  
 Rv3908\_Mtub\_15611044  
 Dr1025\_Drad\_49259148  
 Consensus/85%  
 Key residues:

```

EEEE-----EEEEEEEE-----EEEEEEEE.HHHHHHHHHHHHHHHHHHHHH-----
KWRFAEQYLQPG----- PPRRLQTVVVLVDWVTIM ---PTLEEALCAQQAEEAAPPTQEAQGETE 462
QWRFAEIRYHRPEETHKGRV VPAHVETVVLFDVWVHCL ---PTRSEWELSRGYKQLDDEEKDDGAEKE 592
QWRFAEIRYHRPEETHKGRV VPAHVETVVLFDVWVHCL ---PTRSEWELSRRLREQLAEKLSAERKAEAD 620
QWRFAEIRYHRPEETHKGRV VPAHVETVVLFDVWVHCL ---PTRSEWELSRGLKQLAEKLLAERKAEAD 898
QWRFAEIRYHRPEETHKGRV VPAHVETVVLFDVWVHCL ---PSEWELSLDQLESLDSEVDFHPS 425
QWRLELSYRPEETHKGRV IPARVETVVLFDVWVHCL ---PTKVESELADLYRKTLSNKLAAVDSRDK 591
QWRVELLYRRESETHKGRA VPARVETVVLFDVWVSCV ---PTRLEWDLGHQSYKQKLEKRLKADLEIGG 555
LWRYVDIYRRTESYHGRLL IPFRVETVVLFDVWVSCV ---PTEGNEKTOQYQQLERLSSSSNSGD 295
QWRLELYRRAETHKSGR 1 VPSRVEVILFDVWVSCV ---PIKLEWDLGLQSYKQKLERLLRAESSPD 562
QWRMAEIRYHRVSS----- MKRSLESVLFDVWVSCV ---PTASQWETIVHYSYQPSSEPEEETKTEET 522
QWRFAEVYHRPEEMHKGRL 18 HFSLEDPKMTVADRHEL ---ECRLLSSKGLKSLIAKLTAKIKSEEE 607
QWFSVQIYRYRA----- DKQRIDHNVYLLDPTQSLA ---LDDACMLAETKIAEQAKLANVAIKI 559
QWRMAEIRYRYRA----- EKDRMDTCCLLDPTDTSGLM ---PTEDCYQLLATLRYKQGNKLAVAQAOKL 561
QWRLELYRRAETHKSGR 1 VPSRVEVILFDVWVSCV ---PTRLEWDLGHLSYKQKLERLLRAESAE 560
HWNFPLEIHDRVGTG----- VFSYKEITVYVLDLSECL ---PSFDWRTQWLAAHKAALTRDRLLSQEVK 720
LWRYVEVYHRLEISGNG---- -LPSHKEITVYVLDLPTIS-CI ---SSGEVQWQWAKARQAKLESDAKKTIV 1125
HWRYCDITMRAASSPTSGS 14 DPQLREVVVPAVPAADACM ---AGFDAPFEAEARQQAFQKQRIAESEARK 1097
QWRFAEVVYLRRAASDGGSS 7 SGLVEERVVLPVDAVAA ---AQPNAAVEAIRARQAVFGEARARTEAE 731
KWKYMEVHYRPEETEE---- -QSYQEITVYVLDITDQ ---PMDVSLKLVPTVSETTSPAPVGTATT 728
LWRYVEVYHRLEISGNG---- -KGGPKSVYVYLAELNPNP 5 HEH-QNWKCELEDAIKIADYAEKMSLLRKF 132 1KTG
FW-PE----- -PQSLMTAFNAEYDSDGI 4 KEL-LEANRYRDDLLPLFPFPGVARRLIEDT 248 1VK6
KLSMA----- -PSYFSKMIIVVADLYPE 6 PEPLQVWR-PLAHMDLLEDPPNEARNSVA 172 1VH2
VLSFAS----- -PGGTSERSSINVGEVDATT 10 NED-IRVHVSRQEQWVEEKGIDNAASVIA 193 1KHZ
LEFLMS----- -PGGVTELIHFPIAEYSDNQ 9 DED-IEVLELFPQSALEMKITGEIRDGKTVLL 180 1VIU
VWQHYDDNF----- -SGTDFTHYVVLGFRFRVS 7 DEQHDDYRWLTSDALLANVHANSRAYFLAE 151 1RYA
NY-VARRK----- ---- -PKTVIYVLAEVKDYD 6 HEH-QAYRWLQWLEGEACQAGKEMKAALQEGH 139 1XSA
HWRYVEVYHRLEISGNG---- -LPSHKEITVYVLDLPTIS-CI ---SSGEVQWQWAKARQAKLESDAKKTIV 1125
PRTNDNAWETVAVSVHFO 16 CDSGASIRWQVDRDRIPLYANHTLQKAAE 1548
NDQPKHRTVYVLTVELL 6 VSGRKRREKVEDAIKVLCHQKPVHAYELEK 142
---NDQLARIYIIPGPKDT 7 REI-RNIEVFSIEKIPCHRNNDMPKSLGLAP 222
DPQKATVIVLAVGGLDIP 7 EEV-DEVFALPLALL-QTQNMQYTHFCRGGH 163
LIDTDLITFVGLDHNF 6 AEVKDVFVPLAYFLHPQVHQHYVTRLGHFR 170
GIVNEVCVFEAARTSAL 4 DEV-MDYWCCLADLVLHGIDTAPWSPWVMV 164 1NFS
PGFSESVVYVLAATGREV 7 EADMTMGWYPIAEAARRVLRGEIVNSAIAG 172 1MQW
PVG-FH-VTEVLGRIPAEP 7 PEV-AQIITPTLAEARVLEFRRTLPDGETE 164 1M0Z
YWLTFDFPFVREKLNQWGS 1 WKQQAQWVLEKFTQDQE 10 PEV-GEWVTPPEQLIDLTVE-FKKPVYKVEL 158 1F3Y
RYVNPK----- -GVREHVHVLIMRGEGAPRL EEGMTGAGVSPPEARALLAFPEDLGLLEVAL 121 1VC9
NWRYKLRVLRVWDTKFV--- -CIGKQKWLQJLVSGDAE 8 PEF-DGWRWYSWYVQVQVFSFRVYRVRVK 149
QWRFAEIRYHRVTHSQSLP - -CIGKQKWLQJLVSGDAE 8 PEF-DGWRWYSWYVQVQVFSFRVYRVRVK 149
WYVFDGR----- -RVHKTVHHYLMRPLGELS 3 LEV-AEVAWVPIRELPRILYADERLAEVAD 198
YLRG----- -PDGVLILRVLALPEPGQ 7 DEI-AEASVSRDFALVAAGQIRMYQTKLF 144 1S2U
.h.h.....p.....h.hhh.....p.....h.....
# # * * *

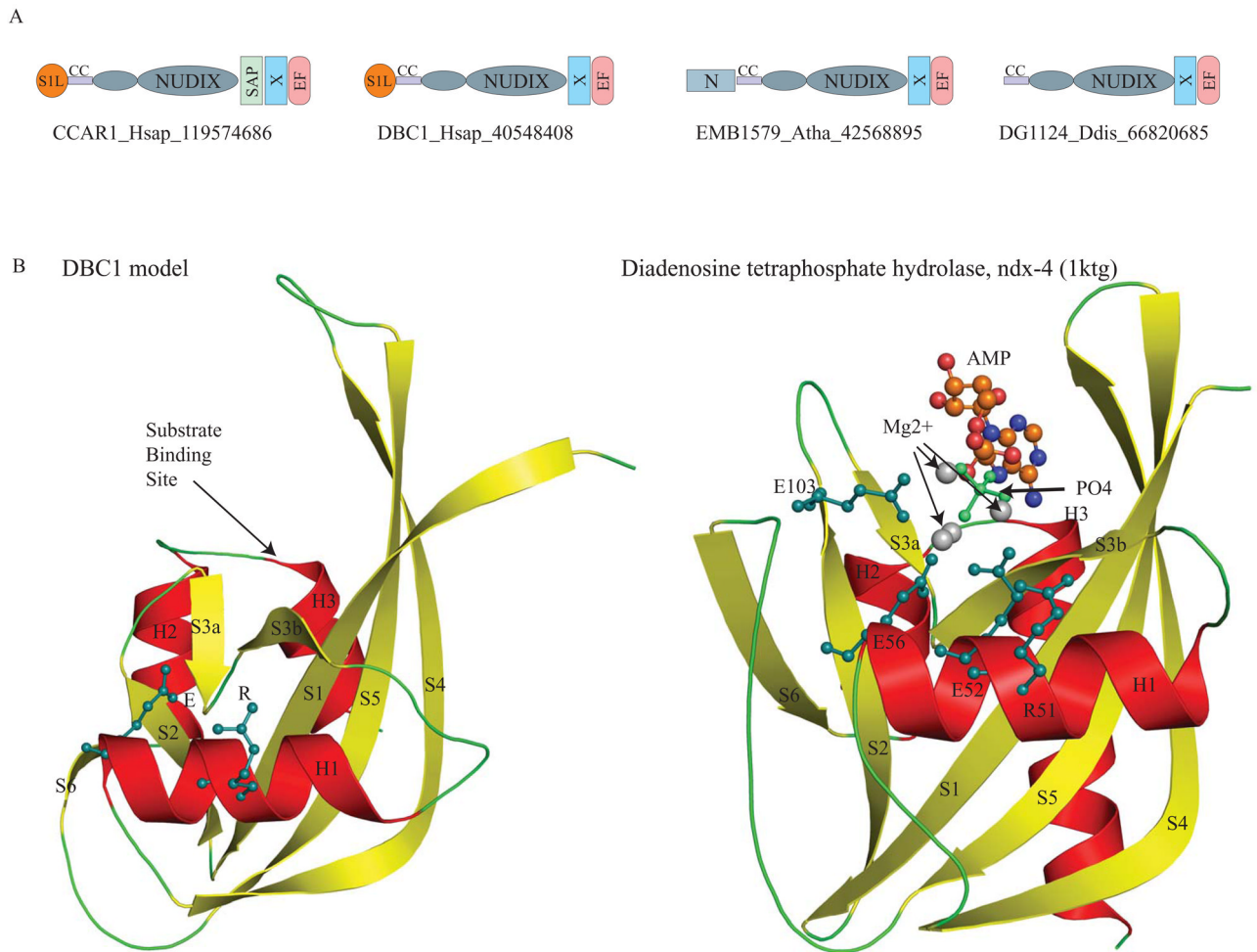
```



**Fig. 1. A) Multiple alignment of Nudix domains of the DBC1 family with known representatives**  
Multiple sequence alignment of the Nudix domain was constructed using Kalign after parsing high-scoring pairs from PSI-BLAST search results. The alignment was refined based on the pairwise alignments produced by the profile-profile searches with the HHpred program against the Nudix profile. The secondary structure from the crystal structures is shown above the alignment with E representing a strand and H a helix. The 85% consensus shown below the alignment was derived using the following amino acid classes: hydrophobic (h: ALICVMYFW, yellow shading); small (s: ACDGNPSTV, green); polar (p: CDEHKNQRST, blue) and its charged subset (c: DEHKR, pink), and big (b: FILMQRWYEK; grey shading). The limits of the domains are indicated by the residue positions, on each end of the sequence. The numbers within the alignment are non-conserved inserts that have not been shown. The key positions involved in stabilizing the substrate-binding cleft and constituting the active site motif of the active version are respectively marked with “#” and “\*” symbols. The sequences are denoted by their gene name followed by the species abbreviation and GenBank Identifier (gi). The PDB ids, when, available are shown to the right end of the alignment. The species abbreviations are: Agam : *Anopheles gambiae*; Amel : *Apis mellifera*; Atha : *Arabidopsis thaliana*; Bmal : *Brugia malayi*; Caps : *Capitella* sp; Cbri : *Caenorhabditis briggsae*; Cele : *Caenorhabditis elegans*; Ddis : *Dictyostelium discoideum*; Dpul : *Daphnia pulex*; Drad : *Deinococcus radiodurans*; Drer : *Danio rerio*; Ecol : *Escherichia coli*; Hsap : *Homo sapiens*; Lang : *Lupinus angustifolius*; Mtub : *Mycobacterium tuberculosis*; Nvit : *Nasonia vitripennis*; Paer : *Pseudomonas aeruginosa*; Ppat : *Physcomitrella patens*; Spur : *Strongylocentrotus purpuratus*; Tcas : *Tribolium castaneum*; Tnig : *Tetraodon nigroviridis*; Tthe : *Thermus thermophilus*; Vcar : *Volvox carteri*  
**B) Multiple alignment of the N-terminal extension in DBC1**



Multiple sequence alignment of the conserved N-terminal extension of the inactive Nudix domains in the DBC1 family was constructed as described above. The helical region that interacts with the SIRT1 in DBC1 is marked. The extended coiled coil region continues N-terminal to this region.



**Fig. 2. A. Domain architectures of the DBC1 family**

The domain architectures found in the DBC1 family are shown. Domain architectures are labeled with a representative gene name, the species abbreviation, and the Genebank identifier (GI) number separated by underscores. The N-terminal extension of the Nudix domain is shown as a small ellipse. Domain abbreviations are: S1L – S1-like OB fold domain; EF – EF hand; CC – coiled coil region; and N – N terminal domain specific to plant members of the DBC1 family.

B. Model of the DBC1 Nudix domain compared with the crystal structure of the catalytically active version ndx-4 (pdb: 1ktgA).

The DBC1 model was constructed using DR1025, a nucleotide pyrophosphatase (PDB: 1SU2), DR1184, a CoA pyrophosphatase (PDB: 1NQZ), both from *Deinococcus* and ndx-4, a diadenosine tetraphosphate hydrolase (PDB: 1KTG) from *Caenorhabditis*, as templates. The crystal structure of ndx-4, a diadenosine tetraphosphate hydrolase (1KTG), is shown for comparison. The conserved active-site residues, the substrate binding site, Magnesium ions, phosphate ion and the substrate ADP are shown.