Published in final edited form as:

Curr Opin Struct Biol. 2009 December ; 19(6): 658–665. doi:10.1016/j.sbi.2009.09.008.

Markers of Fitness in a Successful Enzyme Superfamily

Karen N. Allen^{1,*} and Debra Dunaway-Mariano^{2,*}

¹Department of Chemistry, Boston University, 590 Commonwealth Avenue, Boston, MA 02215-2521, USA

²Department of Chemistry and Chemical Biology, University of New Mexico, Albuquerque, NM, 87131, USA

Summary

Haloalkanoic acid dehalogenase (HAD) superfamily members serve as the predominant catalysts of metabolic phosphate ester hydrolysis in all three kingdoms of life. Collectively, the known structural, bioinformatic, and mechanistic data offer a glimpse of the variety of HAD enzymes that have evolved in the service of metabolic expansion. Factors that have contributed to superfamily dominance include a chemically versatile nucleophile, stability of the core superfold, structural modularity of the chemistry and specificity domains, conformational coupling conferred by the topology of the inserted specificity elements, and retention of a conserved mold for stabilization of the trigonal bipyramidal transition state.

Introduction

The haloalkanoic acid dehalogenase superfamily (HADSf) is a large, ubiquitous enzyme superfamily (>19,000 unique sequences) represented in the proteomes of organisms from all three superkingdoms of life and within each cell by numerous members (e.g. 28 in *E. coli*; 30 in *Mycobacterium tuberculosis*, 84 in *Caenorhabditis elegans*, 169 in *Arabidopsis thaliana*, 183 in *Homo sapiens*). Although not restricted to phosphotransferases, the HADSf is dominated by ATPases (~20%) and to an even larger extent, by phosphate monoester hydrolases (“phosphatases”) (~79%). The phosphoryl group is used in numerous contexts in bacterial and animal cells. It provides a “chemical handle” for binding interactions of metabolites with enzymes and as a “recognition element” for differential binding in protein-protein interactions. The cell employs a veritable arsenal of phosphatases in order to remove the phosphate group from phosphorylated proteins (in signal transduction and protein degradation), lipids, polysaccharides, nucleic acids, and metabolites and to do so with high catalytic proficiency (1). In Nature, the HADSf catalytic scaffold has become the major tool for phosphate removal (2-4).

HADSf phosphatases have colonized numerous disparate biological functions, which vary in their degree of essentiality to the cell. The four main categories of function are 1) primary and secondary metabolism, 2) regulation of metabolic pools, 3) cell housekeeping, and 4) nutrient uptake. For example, HADSf phosphatases that function in primary metabolism include

© 2009 Elsevier Ltd. All rights reserved.

*Corresponding author: drkallen@bu.edu

*Corresponding author: dd39@unm.edu

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

phosphoserine phosphatase of the L-serine biosynthetic pathway(5) and histidinol phosphate phosphatase of the L-histidine biosynthetic pathway (6). HADSF phosphatases functioning in secondary metabolic pathways are exemplified by 2-keto-3-deoxy-9-phospho-D-glycero-D-galacto-nononate phosphatase of the bacterial heptose-containing capsular polysaccharide biosynthesis pathway (7-8) and D-glycero-D-manno-heptose-1 β ,7-bisphosphate phosphatase of the bacterial Lipid A pathway(9-10).The human deoxyribonucleotidases(11), which function to balance dNTP pools and to prevent dNTP accumulation, are cited as examples of HADSF phosphatase “regulators”. The HADSF phosphatase housekeepers are used for specific repair of damaged nucleic acids (e.g., removal of the 3' phosphate at DNA strand breaks(12) and proteins (e.g., removal of the phosphate from a phosphogluconated Lys(13)), but are most commonly used to rid the cell of phosphorylated metabolites whose accidental accumulation via stalled pathways, uptake from the environment, or chemical transformation can result in catabolite repression (e.g., 2-keto-3-deoxygalactonate-6-phosphate(14) and 2-deoxyglucose-6-phosphate(15) or misincorporation (e.g., of unnatural nucleotides into DNA (16). The periplasmic HAD phosphatases of Gram negative bacteria assist nutrient uptake through the removal of phosphate groups from environmental metabolites(17).

There are approximately 900 enzyme superfamilies which differ in size as well as in biological and functional ranges. Through evolution, enzyme superfamilies increase in size as new members are added to take on novel biochemical functions, typically through the process of gene duplication and natural selection. Because such a large fraction of biomolecules possess a phosphate group it might be argued that the HADSF is large because of the high cellular demand for catalyzed phosphoryl transfer. Indeed, roughly 35-40% of the bacterial metabolome is comprised of phosphorylated metabolites (15). However, this does not explain why the HADSF metabolic phosphatases out-number those from superfamilies that also catalyze phosphoryl transfer. Over the past nine years the examination of the HADSF from a structural and mechanistic viewpoint reveals the features which may have allowed adaptation of the superfamily into a multitude of metabolic niches.

The HADSF Catalytic Scaffold-A Self-contained Unit for Phosphoryl Transfer

The chemistry shared by all phosphatase members of the HADSF occurs in two partial reactions (Figure 1). The first is phosphoryl transfer from the substrate to the Asp nucleophile and the second is phosphoryl transfer from the aspartylphosphate intermediate to a water molecule (18). A second Asp residue (denoted Asp+2 because it is positioned two residues from the Asp nucleophile) functions as a general acid/base residue to protonate the leaving group in the first partial reaction and deprotonate the water nucleophile in the second partial reaction(19). The transition states that are formed in these two partial reactions are essentially the same, with the exception of the “OR” leaving group in the first versus the “OH” of the second. As compared to other superfamilies such as the alkaline phosphatase, phosphoglycerate mutase, and tyrosine-specific phosphatase superfamilies that catalyze phosphoryl transfer using Ser, His, and Cys nucleophiles, respectively, the HADSF use of a catalytic Asp provides both a good nucleophile and a good leaving group that functions at low and high pH. The use of the divalent metal cofactor Mg²⁺, common to all phosphoryltransferases of the HADSF, allows close approximation of the anionic nucleophile to the substrate phosphomonoester dianion via electrostatic stabilization, aids in the relative positioning of the nucleophile and substrate phosphoryl group, and provides charge neutralization of the transition state. The mechanistic machinery for phosphoryl transfer is formed by four-loops that position “core residues” that are conserved throughout the HADSF family, and that provide the means of identification of HADSF members via amino-acid sequence alignments. The core residues are the loop 1 Asp, loop 2 Ser/Thr, loop 3 Lys/Arg and loop 4 Asp/Glu. The loop 1 Asp nucleophile together with the loop 4 Asp/Glu bind the Mg²⁺ cofactor, which along with the loop 3 Arg/Lys, bridges the Asp nucleophile and its targeted phosphate electrophile. The loop 2 Ser/Thr also forms a ligand

to the transferring phosphoryl group. These four active site loops are located at the concave surface of one face of the core domain, which consists of a Rossmann fold. Substrates bind with the phosphoryl group oriented toward the Rossmann core and leaving group projected toward solvent or accessorization domains (called cap domains) that provide binding residues for the leaving group moiety (Figure 2).

The structural placement of the HADSF core catalytic residues in a single Rossmann domain provides several of the features that can be considered advantageous to the adaptation of the superfamily. First, the core domain itself is sufficient to provide phosphoryl-transfer activity. This is supported by the fact that phosphoryl transfer is observed in magnesium-dependent phosphatase-1 (MDP-1) ($k_{\text{cat}}/K_m = 5 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$ toward arabinose 5-phosphate (13)) which has the smallest cap insert (a single β -strand) pinned back from the core catalytic residues (20). Thus, the core domain of the HADSF can be considered to be a “modular” phosphoryl transfer domain. Second, the Rossmann fold is the most populated fold among α/β -folds, and contains a “super-site” (defined as a common location of the active site among proteins sharing a common fold but not function) near the folding nucleus (21); this feature ensures stability of the active site against thermal fluctuations (22). Global structural stability of superfolds such as the Rossmann fold is hypothesized to promote protein evolvability by stabilizing mutations that afford new function (23). Finally, the location of the core residues catalyzing phosphoryl transfer at the C-termini of the β -strands of the repeating β - α units, produces an open active site which can be accessed by macromolecules (producing protein phosphatases) but can also be easily adapted to small molecules using a cap-domain inserted between secondary structural elements of the repeated units.

The HADSF Cap Domains- Adaptation to a Diverse Array of Substrates

The accessorization of the core domain has allowed diversification toward substrates and provides a natural division of the HADSF into three major structural classes (C0, C1 and C2 (with C2 subdivided into C2a and C2b) based on the location and topology of the insert (24) (Figure 3). There is a continuum in the length of the inserts; the C0 HADSF members have the shortest inserts and are thus considered the most basic of the HAD members; yet each is unique due to the specialization of the insert (24-25), which can provide specificity determinants or an oligomerization motif (*vide infra*). The HADSF fold is distinguished from other Rossmann folds by two key structural motifs: 1) immediately following strand β_1 is a unique, six-residue structural motif which assumes a nearly complete single-helical turn and 2) downstream of this helical turn there is a β -hairpin formed by two strands projecting from the core of the domain. The helical turn, by allowing alternatively tightly or loosely wound helical conformations, induces a movement in the β -hairpin immediately juxtaposed to the active site. Because the type C1 structures add the cap domain to the β -hairpin itself, this automatically generates alternating conformational states (for e.g. (26-27)). Although the C2 caps have a different insertion site, together, the movements of the helical turn- β -hairpin and C2 cap allow conformational dynamics of the core and cap domains that comprise the active site (see for example (28)).

Compared to the C0 members then, the C1 and C2 HADSF members are distinguished in that their sequence inserts fold into sizeable domains that move to encapsulate the phosphoryl transfer site and the substrate leaving group. This advance brings a new level of sophistication to HAD function as the cap domain adds additional contact surface for substrate selection and an effective means to desolvate the catalytic site. In addition, this modularity of the phosphoryl-transfer core domain and the specificity-conferring cap domain is possibly the structural basis for the adaptation of the scaffold to a wide range of phosphate esters with retention of catalytic efficiency. It has been argued that evolutionary success of proteins is linked to the ability to accumulate neutral mutations which afford promiscuous activities that are later favored by

selection pressure(29). The cap domain provides a fertile ground for the mutation of residues to bind the “OR” moiety of the substrate without effecting the phosphoryl transfer machinery provided by the core. The vestiges of such evolution under low selective pressure may be observed in *E. coli* NagD and other subfamily C2 members which show the common trait that substrate specificity and catalytic efficiencies are low ($k_{\text{cat}}/K_m \sim 1 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$) and the boundaries defining physiological substrates are somewhat overlapping (they share the ability to catabolize other related secondary metabolites)(30). The modularity of the cap domain also makes it tempting to speculate that Nature may have employed the insertion of entire cap domains in order to confer new activities in the HADSF. However, bioinformatic sequence analysis argues against this model because homologs of cap domains do not occur in extant organisms and phylogenetic tree construction shows that the cap domains evolved from one another by insertion of short segments of secondary structure (Figure 4).

Capless not Hapless

Two of the first capless members of the HADSF to be structurally characterized are phage T4 polynucleotide kinase (31) and MDP-1(20). Both of these capless members work on large, macromolecular substrates and show conservation of residues that form a surface surrounding the active site entrance. The substrate itself, in these cases acts as a cap by excluding bulk solvent. This mechanism differs from that used in the capless subfamily typified by KDOP phosphatase. Although these C0 members do not sport an accessorizing cap domain, the short insert region at the β -hairpin confers the ability to oligomerize, providing core residues from the adjacent protomer to bind the substrate “OR” moieties (8). Collectively, the most simplistic prediction from this structural finding in the assignment of function would be that monomeric capless members act on macromolecules, while capped members and oligomeric capless members, which have circumscribed, “pseudo-capped” binding sites, act on small molecules (Figure 5). However, recent examples belie this assumption. Histidinol phosphate phosphatase, a member which is most closely related to the C2 subfamily but is capless in that it bears only a minimal insert in the core domain, utilizes the insert region to provide binding determinants to the substrate leaving group. Together the histidinol moiety and surrounding residues block solvent entry to the active site(6). On the other hand, two capped C1 members have been shown to act on macromolecular substrates: Eya, which serves to dephosphorylate the C-terminal tyrosyl residue of the DNA damage-related histone H2A.X(32) and chronofin which dephosphorylates the Ser3 of the actin-binding factor cofilin(33). Obviously, the capped proteins are accessible to at least the termini of large proteins and assemblies. These examples serve to highlight the substrate range allowed by the use of inserts in addition to domains to produce oligomers, provide binding determinants, and permit dynamics.

A conserved trigonal bipyramidal mold for catalysis

While the acquisition of substrate specificity requires restructuring of the active site to be tailored to a single substrate, this necessitates the operation of the same catalytic motif in different structural contexts. For this to occur, the catalytic scaffold must not rely on interaction with the “OR” group for transition-state stabilization. Indeed, the Mg^{2+} cofactor and the conserved catalytic residues appear to form a steric and electrostatic mold that stabilizes the trigonal-bipyramidal transition state common to both partial reactions. Such a catalytic stratagem is supported by the finding that in high-resolution structures of a wild-type hexose phosphate phosphatase complexed with tungstate, the tungstate is observed to assume a rare pentavalent trigonal coordination geometry (Figure 6) in which the Asp nucleophile and an oxygen atom occupy the apical positions and that elements that stabilize the trigonal bipyramidal transition states are stringently conserved among liganded HADSF structures currently available(34). The essential nature of the Rossmann fold for catalytic function is

evidenced by the use of backbone amide groups in the stabilization of the trigonal-bipyramidal transition.

Conclusions

The variation in the HADSF in terms of substrate range contrasts with a relatively simplistic unifying catalytic stratagem for the transfer of the phosphoryl group. The use of an open catalytic site supporting all necessary components for stabilizing a trigonal bipyramidal “mold” for the transition state of catalysis provides a portable unit for phosphoryl transfer. It is the accessorization of this module that has produced the variety inherent in the superfamily. As new structural and mechanistic data are uncovered, the unifying principles for the functional assignment of members of unknown function will be refined. Additional information is still needed about the structural “fit” necessary between cap and core and the variations in thermodynamic stability of the different cap assemblages to fully exploit this knowledge in the understanding of enzyme evolution and in efforts in the field of protein engineering.

Acknowledgments

We gratefully acknowledge Dr. Nicholas Silvaggi for the preparation of figures. This work is supported by NIH GM061099 to K.N.A. and D.D.-M.

References and Recommended Reading

Papers of particular interest, published within the period of review, have been highlighted as:

• of special interest

•• of outstanding interest

1. Wolfenden R, Snider MJ. The depth of chemical time and the power of enzymes as catalysts. *Acc Chem Res* 2001;34(12):938–45. [PubMed: 11747411]
2. Collet JF, van Schaftingen E, Stroobant V. A new family of phosphotransferases related to P-type ATPases. *Trends Biochem Sci* 1998;23(8):284. [PubMed: 9757826]
3. Koonin EV, Tatusov RL. Computer Analysis of Bacterial Haloacid Dehalogenases Defines a Large Superfamily of Hydrolases with Diverse Specificity: Application of an Iterative Approach to Database Search. *Journal of Molecular Biology* 1994;244(1):125–32. [PubMed: 7966317]
4. Aravind L, Galperin MY, Koonin EV. The catalytic domain of the P-type ATPase has the haloacid dehalogenase fold. *Trends Biochem Sci* 1998;23(4):127–9. [PubMed: 9584613]
5. Collet J-F, Gerin I, Rider MH, Veiga-da-Cunha M, Van Schaftingen E. Human - 3-phosphoserine phosphatase: sequence, expression and evidence for a phosphoenzyme intermediate. *FEBS Letters* 1997;408(3):281–4. [PubMed: 9188776]
6. Rangarajan ES, Proteau A, Wagner J, Hung MN, Matte A, Cygler M. Structural snapshots of *Escherichia coli* histidinol phosphate phosphatase along the reaction pathway. *J Biol Chem* 2006;281(49):37930–41. [PubMed: 16966333]
- 7. Wang L, Lu Z, Allen KN, Mariano PS, Dunaway-Mariano D. Human symbiont *Bacteroides thetaiotaomicron* synthesizes 2-keto-3-deoxy-D-glycero-D- galacto-nononic acid (KDN). *Chem Biol* 2008;15(9):893–7. [PubMed: 18804026] This paper defines a biosynthesis pathway for an unusual sialic acid derivative
8. Lu Z, Wang L, Dunaway-Mariano D, Allen KN. Structure-function analysis of 2-keto-3-deoxy-D-glycero-D-galactononate-9-phosphate phosphatase defines specificity elements in type C0 haloalkanoate dehalogenase family members. *J Biol Chem* 2009;284(2):1224–33. [PubMed: 18986982] [PMCID: 2613609]
9. Wu J, Woodard RW. *Escherichia coli* YrbI is 3-deoxy-D-manno-octulosonate 8-phosphate phosphatase. *J Biol Chem* 2003;278(20):18117–23. [PubMed: 12639950]

10. Biswas T, Yi L, Aggarwal P, Wu J, Rubin JR, Stuckey JA, et al. The tail of KDSC: Conformational changes control the activity of a haloacid dehalogenase superfamily phosphatase. *J Biol Chem*. 2009
11. Rinaldo-Matthis A, Rampazzo C, Reichard P, Bianchi V, Nordlund P. Crystal structure of a human mitochondrial deoxyribonucleotidase. *Nat Struct Biol* 2002;10:779–87. [PubMed: 12352955]
12. Kim Y, Yakunin AF, Kuznetsova E, Xu X, Pennycooke M, Gu J, et al. Structure- and function-based characterization of a new phosphoglycolate phosphatase from *Thermoplasma acidophilum*. *J Biol Chem* 2004;279(1):517–26. [PubMed: 14555659]
13. Fortpied J, Maliekal P, Vertommen D, Van Schaftingen E. Magnesium-dependent phosphatase-1 is a protein-fructosamine-6-phosphatase potentially involved in glycation repair. *J Biol Chem* 2006;281(27):18378–85. [PubMed: 16670083]
14. Wang, L.; Allen, KN.; Dunaway-Mariano, D. unpublished data.
- 15. Kuznetsova E, Proudfoot M, Gonzalez CF, Brown G, Omelchenko MV, Boroza I, et al. Genome-wide analysis of substrate specificities of the *Escherichia coli* haloacid dehalogenase-like phosphatase family. *J Biol Chem* 2006;281(47):36149–61. [PubMed: 16990279] High throughput metabolic screens are used to categorize the functions of all HADSF members in a single bacterial strain.
16. Titz B, Hauser R, Engelbrecher A, Uetz P. The *Escherichia coli* protein YjjG is a house-cleaning nucleotidase in vivo. *FEMS Microbiol Lett* 2007;270(1):49–57. [PubMed: 17286574]
17. Passariello C, Forleo C, Micheli V, Schippa S, Leone R, Mangani S, et al. Biochemical characterization of the class B acid phosphatase (AphA) of *Escherichia coli* MG1655. *Biochim Biophys Acta* 2006;1764(1):13–9. [PubMed: 16297670]
18. Collet JF, Stroobant V, Van Schaftingen E. Evidence for phosphotransferases phosphorylated on aspartate residue in N-terminal DXDX(T/V) motif. *Methods Enzymol* 2002;354:177–88. [PubMed: 12418225]
19. Lahiri SD, Zhang G, Dunaway-Mariano D, Allen KN. Caught in the act: the structure of phosphorylated beta-phosphoglucomutase from *Lactococcus lactis*. *Biochemistry* 2002;41(26):8351–9. [PubMed: 12081483]
20. Peisach E, Selengut JD, Dunaway-Mariano D, Allen KN. X-ray crystal structure of the hypothetical phosphotyrosine phosphatase MDP-1 of the haloacid dehalogenase superfamily. *Biochemistry* 2004;43(40):12770–9. [PubMed: 15461449]
21. Russell RB, Sasieni PD, Sternberg MJ. Supersites within superfolds. Binding site similarity in the absence of homology. *J Mol Biol* 1998;282(4):903–18. [PubMed: 9743635]
22. Mirny LA, Shakhnovich EI. Universally conserved positions in protein folds: reading evolutionary signals about stability, folding kinetics and function. *J Mol Biol* 1999;291(1):177–96. [PubMed: 10438614]
- 23. DePristo MA, Weinreich DM, Hartl DL. Missense meanderings in sequence space: a biophysical view of protein evolution. *Nat Rev Genet* 2005;6(9):678–87. [PubMed: 16074985] Comprehensive overview of the biophysical effects of protein mutation and selection.
24. Burroughs AM, Allen KN, Dunaway-Mariano D, Aravind L. Evolutionary genomics of the HAD superfamily: understanding the structural adaptations and catalytic diversity in a superfamily of phosphoesterases and allied enzymes. *J Mol Biol* 2006;361(5):1003–34. [PubMed: 16889794]
25. Allen KN, Dunaway-Mariano D. Phosphoryl group transfer: evolution of a catalytic scaffold. *Trends Biochem Sci* 2004;29(9):495–503. [PubMed: 15337123]
26. Dai J, Wang L, Allen KN, Radstrom P, Dunaway-Mariano D. Conformational cycling in beta-phosphoglucomutase catalysis: reorientation of the beta-D-glucose 1,6-(Bis)phosphate intermediate. *Biochemistry* 2006;45(25):7818–24. [PubMed: 16784233]
27. Dai J, Finci L, Zhang C, Lahiri S, Zhang G, Peisach E, et al. Analysis of the structural determinants underlying discrimination between substrate and solvent in beta-phosphoglucomutase catalysis. *Biochemistry* 2009;48(9):1984–95. [PubMed: 19154134]
28. Fieulaine S, Lunn JE, Borel F, Ferrer JL. The structure of a cyanobacterial sucrose-phosphatase reveals the sugar tongs that release free sucrose in the cell. *Plant Cell* 2005;17(7):2049–58. [PubMed: 15937230] PMID: 1167551
- 29. Bloom JD, Romero PA, Lu Z, Arnold FH. Neutral genetic drift can alter promiscuous protein functions, potentially aiding functional evolution. *Biol Direct* 2007;2:17. [PubMed: 17598905]

PMCID: 1914045The link between the ability of a protein to support mutations that do not affect primary function and to evolve promiscuous functions is demonstrating for cytochrome P450.

30. Tremblay L, Zhang G, Dai J, Dunaway-Mariano D, Allen KN. Structure and Activity Analyses of E. coli K-12 NagD Provide Insight Into the Evolution of Biochemical Function in the HAD Enzyme Superfamily. *Biochemistry* 2005;45:1183–93. [PubMed: 16430214]
31. Galburt EA, Pelletier J, Wilson G, Stoddard BL. Structure of a tRNA repair enzyme and molecular biology workhorse: T4 polynucleotide kinase. *Structure (Camb)* 2002;10(9):1249–60. [PubMed: 12220496]
- 32. Krishnan N, Jeong DG, Jung SK, Ryu SE, Xiao A, Allis CD, et al. Dephosphorylation of the C-terminal tyrosyl residue of the DNA damage-related histone H2A.X is mediated by the protein phosphatase eyes absent. *J Biol Chem* 2009;284(24):16066–70. [PubMed: 19351884]PMCID: 2713548This paper identifies a protein target for the dual-function transcription factor/phosphatase eyes absent.
33. Gohla A, Birkenfeld J, Bokoch GM. Chronophin, a novel HAD-type serine protein phosphatase, regulates cofilin-dependent actin dynamics. *Nat Cell Biol* 2005;7(1):21–9. [PubMed: 15580268]
- 34. Lu Z, Dunaway-Mariano D, Allen KN. The catalytic scaffold of the haloalkanoic acid dehalogenase enzyme superfamily acts as a mold for the trigonal bipyramidal transition state. *Proc Natl Acad Sci U S A* 2008;105(15):5687–92. [PubMed: 18398008]PMCID: 2311376High resolution X-ray crystallographic structures show that vanadate and tungstate are bound to a HAD phosphatase with trigonal bipyramidal geometry. Mutation of the catalytic acid/base shows this residue is important for stabilization as well as catalysis.

Figure 1.

The general catalytic mechanism for phosphohydrolase members of the HADSF. Catalysis proceeds through an aspartylphosphate intermediate.

Figure 2.

The catalytic motifs of the HADSF. A) Composite scheme of the residues involved in forming the phosphoryl-transfer catalytic site in the HADSF. Hydrogen bonds are depicted as dashed lines labeled with average bond lengths in Å (range of bond lengths in parenthesis) for the twelve liganded structures from which this is constructed. B) Position of the catalytic loops in the active site scaffolds (core in black, cap in grey) of C1, C2a and C2b members (coloring the same in A and B, substrate specificity loop in pink).

Figure 3. Topology diagram of the classic HADSF Rossmann core domain. The HAD C0/C1 and C2 cap insertion points are labeled. The signature single helical turn is in pink and β -hairpin is in green.

Figure 4. Schematic of the possible evolutionary path of C1 members from insertion and duplication of secondary structural elements into the β -hairpin of the Rossmann fold (PDB codes in bold).

Figure 5.

Venn diagram of the substrate range of HADSF members. Capped proteins usually have small substrates (1) but may also work on termini of macromolecules (4); oligomerization can allow C0 members to utilize small substrates (3); capless members may utilize large molecule substrates (2) but can also utilize small substrates using the insert segment.

Figure 6. Stereo diagram of the catalytic motifs (colored as in Figure 2) of the HADSF member hexose phosphate phosphatase complexed with tungstate highlights the trigonal bipyramidal geometry supported by the phosphoryl transfer site.