AAA+ proteins: diversity in function, similarity in structure

Jamie Snider and Walid A. Houry[1]
1 King's College Circle, Medical Sciences Building, Department of Biochemistry, University of Toronto, Toronto, Ontario, Canada, M5S 1A8

Abstract

The AAA+ (ATPases associated with various cellular activities) superfamily of proteins represents a distinct lineage of the larger class of P-loop NTPases. Members of this superfamily use the power of nucleotide binding and hydrolysis to direct molecular remodelling events. All AAA+ proteins share a common core architecture, which, through various sequence and structural modifications, has been adapted for use in a remarkably diverse range of functions. The following mini-review provides a concise description of the major structural elements common to all AAA+ proteins in the context of their mechanistic roles. In addition, the evolutionary and functional diversity of this superfamily is described on the basis of recent classification studies.

The P-loop NTPases

The energy obtained from the hydrolysis of nucleotides is fundamental to a myriad of biological processes, and indeed to life itself. Cells have evolved various mechanisms for harnessing this energy and directing it towards useful work. One class of proteins using such a mechanism is the P-loop NTPases, an abundant class of nucleotide binding/hydrolysing proteins, which play critical roles in a vast array of cellular functions.

These proteins are found in all three major domains of life, including the prokaryotic (i.e. organisms lacking a true nucleus) Archaea and Bacteria, as well as the Eukarya (i.e. eukaryotes, possessing a true nucleus and membrane-bound organelles derived from bacterial endosymbionts) [1,2]. In fact, roughly 5–10% of proteins encoded by fully sequenced prokaryotic and eukaryotic genomes completed thus far are predicted to contain a P-loop NTPase domain [3]. P-loop NTPases are defined by the presence of the nominal P-loop, a conserved nucleotide phosphate-binding motif, also referred to as the Walker A motif ($GX_4GK[S/T]$), and a second, more variable region, called the Walker B motif ($\Psi\Psi\Psi\Psi[D/E]$, where Ψ is a hydrophobic residue). Both the Walker A and Walker B motifs are important for binding/interaction with nucleotides, which are typically ATP or GTP, and Mg^{2+} [4–6]. The P-loop NTPases share a common $\alpha\beta\alpha$ core domain structure, consisting of a parallel β-sheet sandwiched between two sets of α-helices [7]. Although not universally true, P-loop NTPases most commonly catalyse the hydrolysis of the $\beta-\gamma$ phosphate bond of the bound nucleotide, utilizing the energy released from this reaction to direct conformational changes in other molecules [8].

Key words: ATPase associated with various cellular activities (AAA+), P-loop NTPase, remodelling, Walker A motif, Walker B motif.

Abbreviations used: AAA+, ATPase associated with various cellular activities; ASCE, additional strand catalytic E; KG, kinase-GTPase; PACTT, protease, chelatase, transcriptional activators and transport; RFC1, replication factor C1; STAND, signal transduction ATPases with numerous domains.

[1]**To whom correspondence should be addressed (email walid.houry@utoronto.ca).**

Biochem. Soc. Trans. (2008) **36**, 72–77; doi:10.1042/BST0360072

The P-loop NTPases can be divided into two major structural groups, the KG (kinase-GTPase) group and the ASCE (additional strand catalytic E) group. These groups and representative members are shown in Figure 1. The core structural features for each group are also shown. The KG group is characterized by the adjacency of the Walker B strand and the strand connected to the P-loop, with a characteristic strand order of 54132 in their core β-sheet (Figure 1, lower left-hand panel). The ASCE group contains an additional strand inserted between the strand connected to the P-loop strand and the Walker B strand, giving a core β-sheet strand order of 51432, as well as a catalytically important conserved glutamate residue within the Walker B motif (Figure 1, lower right-hand panel) [9]. Within these two groups, the P-loop NTPases can be divided into numerous distinct lineages (Figure 1) [8–10]. One of the major P-loop lineages, and the main focus of this mini-review, is the AAA+ (ATPase associated with various associated activities) proteins of the ASCE structural group.

The AAA+ proteins

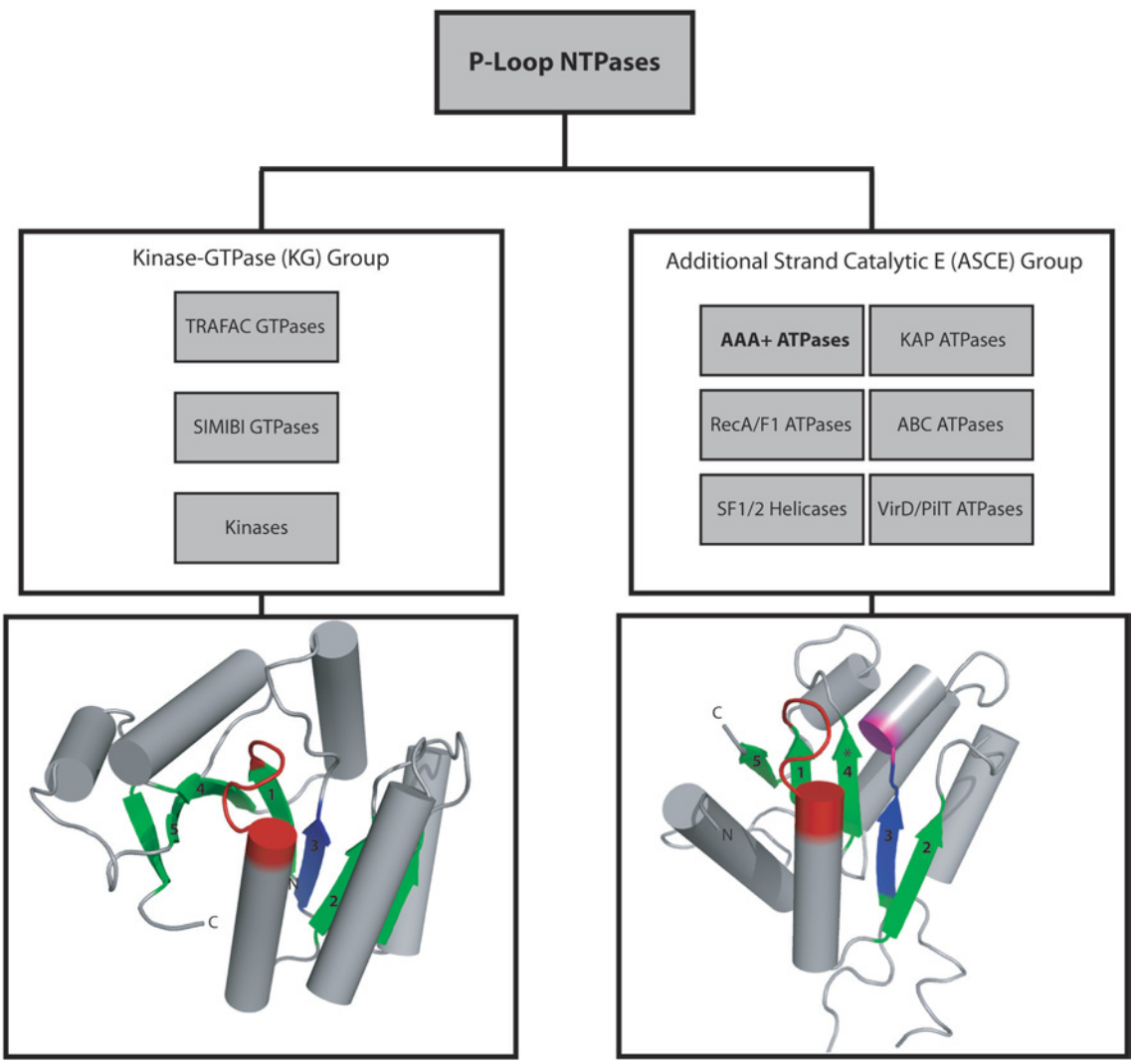
General structure

AAA stands for 'ATPases associated with various cellular activities', and, as the name implies, was first used to describe a class of ATP-hydrolysing enzymes with a range of functional roles [11]. Among other processes, AAA proteins were found to be involved in protein degradation, vesicular fusion, peroxisome biogenesis and the assembly of membrane complexes [2]. Subsequent work showed that AAA proteins are actually a subset of a much larger superfamily of ATPases, now referred to as AAA+ [12].

In addition to the conserved $\alpha\beta\alpha$ core domain structure, and the Walker A and B motifs of the P-loop NTPases, the AAA+ proteins contain a number of other conserved distinguishing features. All these features are found within a 200–250-amino-acid region referred to as the AAA+

Figure 1 | **P-loop NTPases**

Major divisions and representative structures from the P-loop NTPase KG and ASCE structural groups. The top two panels show the two major structural groups (KG and ASCE) and major classes/superfamilies which lie within them. The lower left- and right-hand panels show representative NTPase core domains of the KG and ASCE groups respectively. β-Strands are represented by arrows and are shown in green. The strands of the conserved 'core' β-sheet structure are numbered in order of occurrence in the primary sequence. Helices are represented by cylinders and are shown in grey. The Walker A and B motifs are shown in red and blue respectively. Lower left: P-loop NTPase core domain from *Thermus aquaticus* Ffh protein (KG structural group) [42]. Lower right: P-loop NTPase core domain from *Saccharomyces cerevisiae* RFC1 clamp loader protein (ASCE structural group) [14]. The additional strand between the Walker A- and Walker B-associated β-strands is marked with an asterisk (*). The position of the catalytic glutamic acid residue is shown in pink.

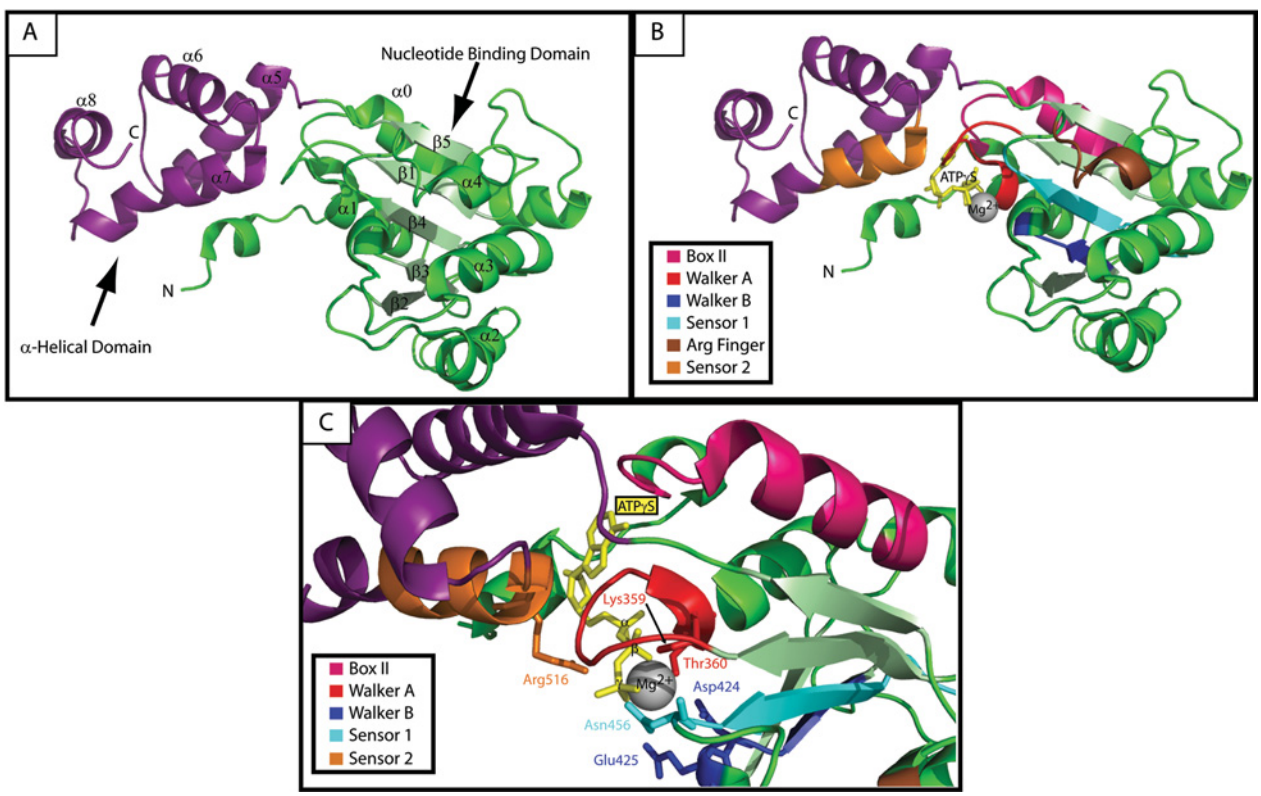


'module' [12,13]. Figure 2 shows a representative of an AAA+ module from the RFC1 (replication factor C1) clamp loader protein of *Saccharomyces cerevisiae* [14]. The structure of the AAA+ module of this protein represents the 'basic' core structure of AAA+ modules without insertions or modifications. Various major features are marked in Figure 2. Like many other members of the ASCE structural group, AAA+ proteins function as oligomeric rings, with a hexameric arrangement being the most common [2].

Notably, the AAA+ module consists of two discrete domains. The first corresponds to the P-loop NTPase $\alpha\beta\alpha$ nucleotide-binding core domain, and consists of a five-stranded parallel β-sheet, with a 51432 strand order, sandwiched between α-helices (Figure 2A, green) [13]. The second domain, unique to AAA+ proteins, consists of a bundle of four α-helices (Figure 2A, purple). The sequence of this domain is much less conserved across AAA+ proteins than the nucleotide-binding domain, but all appear to share a

Figure 2 | Structure of the AAA+ module of *Saccharomyces cerevisiae* RFC1

(**A**) Overall view of the RFC1 AAA+ module from *Saccharomyces cerevisiae* RFC1 [14]. The nucleotide-binding domain is shown in green. The C-terminal α-helical domain is shown in purple. β-Strands and α-helices are labelled (β1–β5 and α0–α8 respectively). (**B**) Overall view of the AAA+ module showing major motifs (coloured and labelled as described in the inset) and bound adenosine 5′-[γ-thio]triphosphate (ATPγS; yellow sticks) and Mg^{2+} (grey sphere). (**C**) Close-up view of nucleotide-binding/catalytic site. Side chains of key residues are labelled and coloured as described in the inset.



common core fold consisting of two helical hairpins arranged in a left-handed superhelical structure [13,15]. A nucleotide bound by the $\alpha\beta\alpha$ domain is sandwiched between these two domains. In oligomeric structures, the nucleotide also faces the $\alpha\beta\alpha$ domain of the neighbouring subunit [13,16], and, hence, in some AAA+ proteins, the nucleotide is required for proper oligomerization.

The AAA+ module contains several signature motifs. The first motif is Box II (Figures 2B and 2C, pink), which maps to a putative α-helical structure in an extended N-terminal region (α0), and is a defining feature of the AAA+ protein class [2,12]. Residues in this motif, on the basis of their proximity to the bound nucleotide, have been proposed to play a role in adenine recognition although, notably, this motif is not always conserved [12].

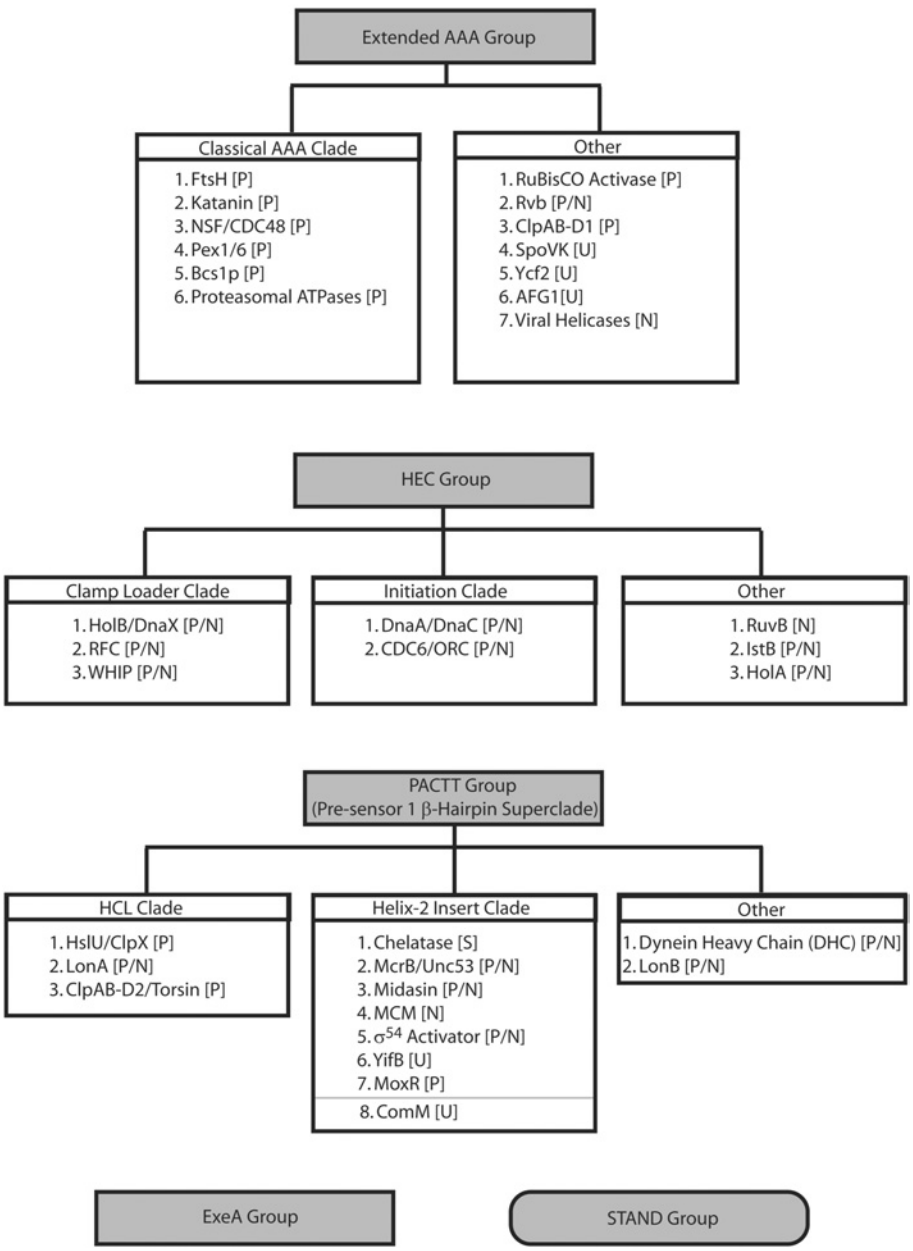
The next conserved feature is the Walker A motif (Figures 2B and 2C, red) which maps to the region between strand β1 and helix α1. The conserved lysine and threonine/serine residues of this motif are proposed to be important in binding the β- and γ-phosphates of bound ATP substrate and Mg^{2+} ion respectively. The lysine residue is also believed to be important in maintaining the proper conformation of the P-loop via hydrogen-bonding with the main-chain oxygens of the loop [4,5]. The Walker A motif is followed by two less conserved motifs, referred to as Box IV and IV' (not shown) [12].

The Walker B motif is associated with the β3 strand (Figures 2B and 2C, blue) and has the consensus sequence $\Psi\Psi\Psi\Psi$DE, including the conserved catalytic glutamate residue characteristic of the ASCE group of P-loop NTPases. The carboxy group of this residue is believed to act as a catalytic base, abstracting a proton from a molecule of water, thereby priming it for a nucleophilic attack on the γ-phosphate of bound ATP [9,13]. The conserved aspartate residue is involved in the co-ordination of the Mg^{2+} ion [4,6].

Immediately after the Walker B motif is the Sensor 1 motif present on strand β4 (Figures 2B and 2C, cyan). The Sensor 1 motif contains a conserved polar residue, which is asparagine in the RFC1 structure, although threonine, serine or histidine is sometimes observed. This residue has been shown to be functionally important, and may interact with the γ-phosphate of ATP either directly, potentially acting as a 'sensor' of nucleotide binding/hydrolysis, or indirectly via a water molecule, possibly helping to properly orient the water molecule for nucleophilic attack on the substrate [17,18]. It is

Figure 3 | **Classification of AAA+ proteins**

Division of the AAA+ proteins into groups (grey boxes), clades (unfilled boxes) and families (numbered). The STAND group is shown in a grey rounded rectangle to reflect uncertainty as to its classification within the AAA+ lineage. The Figure is based upon data from two published AAA+ classification efforts performed using both sequence and structural information [2,15]. For each family, the major known or putative target(s) are listed in square brackets beside the family name: P, protein; N, nucleic acid; S, small molecule; U, unknown.



important to note that the Sensor 1 motif is not strictly unique to the AAA+ proteins, but is also found in certain other divisions of the ASCE structural class of P-loop NTPases [2].

After Sensor 1 is Box VII which contains a conserved arginine residue that is near the N-terminus of strand $\beta 5$ (Figure 2B, brown) [12]. In most AAA+ proteins, this residue is oriented towards the ATP-containing active site of a neighbouring subunit and is proposed to act as an 'arginine finger', interacting with the γ-phosphate of the nucleotide in the neighbouring subunit. The arginine finger has been shown to be necessary for ATP hydrolysis by playing an important role in intersubunit communication/catalysis [17,19–21]. This arginine finger is also found in some other ASCE group members, particularly those which form rings [22].

Box VII is followed by two subtle motifs, Box VII' and VII'' (not shown), and then the Sensor 2 motif (Figures 2B

and 2C, orange). All of these motifs are in the second helical bundle domain of the AAA+ module. The Sensor 2 motif, which is on the third helix of the second domain ($\alpha 7$), contains a conserved arginine residue that interacts with the γ-phosphate of the bound ATP substrate. The function of this arginine residue appears to be somewhat divergent, and it has been implicated in a variety of roles including ATP binding, hydrolysis, sensing and intersubunit interaction. It is also believed to be involved in mediating movement of the C-domain relative to the N-domain of the AAA+ module during ATP hydrolysis [12,13,16]. Motion between these two domains is proposed to be important in the generation of mechanical force which affects substrate molecules and functional partners [13].

Evolution, classification, and diverse functions of AAA+ proteins

Sequence and structural analysis indicate that the AAA+ superfamily is very ancient and underwent considerable divergence before the appearance of the last common ancestor of the Eukarya, Bacteria and Archaea domains of life [2,12]. Various phylogenetic studies using sequence and structural information have shown that the AAA+ superfamily can be divided into numerous smaller families [2,12,15,23,24].

Figure 3 shows some of the major groups, clades and families of AAA+ proteins, and represents an attempt to integrate the results of two classification studies [2,15] with the inclusion of new families and reassignment of certain families in accordance with the more recent work [15]. The major groups are the extended AAA group, the HEC (helicases and clamp loaders) group, the PACTT (protease, chelatase, transcriptional activators and transport) group, the ExeA group and the STAND (signal transduction ATPases with numerous domains) group. The STAND group is shown in a grey rounded rectangle due to uncertainty as to its classification within the AAA+ lineage [8,15].

It is clear from these studies, and the work of others, that the AAA+ superfamily has undergone a remarkable level of diversification. A large number of families has evolved, each recognizing their own unique substrates and interaction partners, and playing a role in distinct cellular processes. For instance, within the extended AAA group alone, AAA+ families are involved in diverse processes including protein unfolding and degradation (FtsH, proteasomal ATPases, ClpA) [25–27], protein disaggregation (ClpB) [28], microtubule disassembly (katanin) [29,30], membrane fusion [NSF (*N*-ethylmaleimide-sensitive factor)] [31,32], the ubiquitin system [CDC48 (cell division cycle 48)] [33], peroxisome biogenesis (Pex1/6) [34,35], cytochrome assembly (Bsc1p) [36,37], regulation of enzymatic activity (Rubisco activase) [38], helicase activity (Rvb, viral helicases) [39,40] and bacterial sporulation (SpoVK) [41].

The functional diversity of the AAA+ superfamily has been made possible through modification and adaptation of the core AAA+ architecture, via the introduction of new domains and motifs, creating an array of molecular machines capable of harnessing the power of nucleotide binding and hydrolysis for use in the remodelling of a wide range of substrates including proteins, nucleic acids and even, in the case of the chelatase family of the PACTT group, small molecules. Thus the AAA+ module appears to represent a remarkably adaptable piece of 'molecular engineering', which Nature has adopted for a wide-ranging use.

Final thoughts

Although an extensive amount of research has been performed on members of the AAA+ superfamily, there are still some families about which we know very little. In addition, our understanding of the underlying mechanisms employed by various AAA+ families is limited at best. Obtaining a thorough understanding of the role of these different families and the fundamental differences and commonalities between the mechanisms they employ is essential for helping us develop an integrated view of AAA+ systems and a greater comprehension of the fundamental workings of these molecular machines which are so crucial to life. The future of AAA+ research promises to be exciting and highly rewarding, and will provide us with invaluable information about a host of biological systems.

J.S. is the recipient of a graduate fellowship from the Natural Sciences and Engineering Research Council of Canada. W.A.H. is a Canadian Institutes of Health Research New Investigator. This research was supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada to W.A.H.

References

1 Martin, J., Gruber, M. and Lupas, A.N. (2004) Coiled coils meet the chaperone world. Trends Biochem. Sci. **29**, 455–458
2 Iyer, L.M., Leipe, D.D., Koonin, E.V. and Aravind, L. (2004) Evolutionary history and higher order classification of AAA+ ATPases. J. Struct. Biol. **146**, 11–31
3 Koonin, E.V., Wolf, Y.I. and Aravind, L. (2000) Protein fold recognition using sequence profiles and its application in structural genomics. Adv. Protein Chem. **54**, 245–275
4 Walker, J.E., Saraste, M., Runswick, M.J. and Gay, N.J. (1982) Distantly related sequences in the α- and β-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. **1**, 945–951
5 Saraste, M., Sibbald, P.R. and Wittinghofer, A. (1990) The P-loop: a common motif in ATP- and GTP-binding proteins. Trends Biochem. Sci. **15**, 430–434
6 Gorbalenya, A.E. and Koonin, E.V. (1989) Viral proteins containing the purine NTP-binding sequence pattern. Nucleic Acids Res. **17**, 8413–8440
7 Milner-White, E.J., Coggins, J.R. and Anton, I.A. (1991) Evidence for an ancestral core structure in nucleotide-binding proteins with the type A motif. J. Mol. Biol. **221**, 751–754
8 Leipe, D.D., Koonin, E.V. and Aravind, L. (2004) STAND, a class of P-loop NTPases including animal and plant regulators of programmed cell death: multiple, complex domain architectures, unusual phyletic patterns, and evolution by horizontal gene transfer. J. Mol. Biol. **343**, 1–28
9 Leipe, D.D., Koonin, E.V. and Aravind, L. (2003) Evolution and classification of P-loop kinases and related proteins. J. Mol. Biol. **333**, 781–815
10 Leipe, D.D., Wolf, Y.I., Koonin, E.V. and Aravind, L. (2002) Classification and evolution of P-loop GTPases and related ATPases. J. Mol. Biol. **317**, 41–72

11 Kunau, W.H., Beyer, A., Franken, T., Gotte, K., Marzioch, M., Saidowsky, J., Skaletz-Rorowski, A. and Wiebel, F.F. (1993) Two complementary approaches to study peroxisome biogenesis in *Saccharomyces cerevisiae*: forward and reversed genetics. Biochimie **75**, 209–224
12 Neuwald, A.F., Aravind, L., Spouge, J.L. and Koonin, E.V. (1999) AAA+: a class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. Genome Res. **9**, 27–43
13 Ogura, T. and Wilkinson, A.J. (2001) AAA+ superfamily ATPases: common structure – diverse function. Genes Cells **6**, 575–597
14 Bowman, G.D., O'Donnell, M. and Kuriyan, J. (2004) Structural analysis of a eukaryotic sliding DNA clamp–clamp loader complex. Nature **429**, 724–730
15 Ammelburg, M., Frickey, T. and Lupas, A.N. (2006) Classification of AAA+ proteins. J. Struct. Biol. **156**, 2–11
16 Ogura, T., Whiteheart, S.W. and Wilkinson, A.J. (2004) Conserved arginine residues implicated in ATP hydrolysis, nucleotide-sensing, and inter-subunit interactions in AAA and AAA+ ATPases. J. Struct. Biol. **146**, 106–112
17 Karata, K., Inagawa, T., Wilkinson, A.J., Tatsuta, T. and Ogura, T. (1999) Dissecting the role of a conserved motif (the second region of homology) in the AAA family of ATPases: site-directed mutagenesis of the ATP-dependent protease FtsH. J. Biol. Chem. **274**, 26225–26232
18 Guenther, B., Onrust, R., Sali, A., O'Donnell, M. and Kuriyan, J. (1997) Crystal structure of the δ' subunit of the clamp-loader complex of *E. coli* DNA polymerase III. Cell **91**, 335–345
19 Rombel, I., Peters-Wendisch, P., Mesecar, A., Thorgeirsson, T., Shin, Y.K. and Kustu, S. (1999) MgATP binding and hydrolysis determinants of NtrC, a bacterial enhancer-binding protein. J. Bacteriol. **181**, 4628–4638
20 Davey, M.J., Indiani, C. and O'Donnell, M. (2003) Reconstitution of the Mcm2-7p heterohexamer, subunit arrangement, and ATP site architecture. J. Biol. Chem. **278**, 4491–4499
21 Johnson, A. and O'Donnell, M. (2003) Ordered ATP hydrolysis in the γ complex clamp loader AAA+ machine. J. Biol. Chem. **278**, 14406–14413
22 Lupas, A.N. and Martin, J. (2002) AAA proteins. Curr. Opin. Struct. Biol. **12**, 746–753
23 Beyer, A. (1997) Sequence analysis of the AAA protein family. Protein Sci. **6**, 2043–2058
24 Frickey, T. and Lupas, A.N. (2004) Phylogenetic analysis of AAA proteins. J. Struct. Biol. **146**, 2–10
25 Gottesman, S. (2003) Proteolysis in bacterial regulatory circuits. Annu. Rev. Cell Dev. Biol. **19**, 565–587
26 Tomoyasu, T., Gamer, J., Bukau, B., Kanemori, M., Mori, H., Rutman, A.J., Oppenheim, A.B., Yura, T., Yamanaka, K., Niki, H. et al. (1995) *Escherichia coli* FtsH is a membrane-bound, ATP-dependent protease which degrades the heat-shock transcription factor σ^{32}. EMBO J. **14**, 2551–2560
27 Smith, D.M., Benaroudj, N. and Goldberg, A. (2006) Proteasomes and their associated ATPases: a destructive combination. J. Struct. Biol. **156**, 72–83
28 Zolkiewski, M. (2006) A camel passes through the eye of a needle: protein unfolding of the Clp ATPases. Mol. Microbiol. **61**, 1094–1100
29 Baas, P.W., Karabay, A. and Qiang, L. (2005) Microtubules cut and run. Trends Cell Biol. **15**, 518–524
30 McNally, F.J. and Vale, R.D. (1993) Identification of katanin, an ATPase that severs and disassembles stable microtubules. Cell **75**, 419–429
31 Sollner, T., Bennett, M.K., Whiteheart, S.W., Scheller, R.H. and Rothman, J.E. (1993) A protein assembly–disassembly pathway *in vitro* that may correspond to sequential steps of synaptic vesicle docking, activation, and fusion. Cell **75**, 409–418
32 Whiteheart, S.W. and Matveeva, E.A. (2004) Multiple binding proteins suggest diverse functions for the *N*-ethylmaleimide sensitive factor. J. Struct. Biol. **146**, 32–43
33 Wang, Q., Song, C. and Li, C.C. (2004) Molecular perspectives on p97-VCP: progress in understanding its structure and diverse biological functions. J. Struct. Biol. **146**, 44–57
34 Erdmann, R., Wiebel, F.F., Flessau, A., Rytka, J., Beyer, A., Frohlich, K.U. and Kunau, W.H. (1991) *PAS1*, a yeast gene required for peroxisome biogenesis, encodes a member of a novel family of putative ATPases. Cell **64**, 499–510
35 Voorn-Brouwer, T., van der Leij, I., Hemrika, W., Distel, B. and Tabak, H.F. (1993) Sequence of the *PAS8* gene, the product of which is essential for biogenesis of peroxisomes in *Saccharomyces cerevisiae*. Biochim. Biophys. Acta **1216**, 325–328
36 Cruciat, C.M., Hell, K., Folsch, H., Neupert, W. and Stuart, R.A. (1999) Bcs1p, an AAA-family member, is a chaperone for the assembly of the cytochrome bc_1 complex. EMBO J. **18**, 5226–5233
37 Nobrega, F.G., Nobrega, M.P. and Tzagoloff, A. (1992) BCS1, a novel gene required for the expression of functional Rieske iron–sulfur protein in *Saccharomyces cerevisiae*. EMBO J. **11**, 3821–3829
38 Portis, Jr, A.R. (2003) Rubisco activase: Rubisco's catalytic chaperone. Photosynth. Res. **75**, 11–27
39 Kanemaki, M., Kurokawa, Y., Matsu-ura, T., Makino, Y., Masani, A., Okazaki, K., Morishita, T. and Tamura, T.A. (1999) TIP49b, a new RuvB-like DNA helicase, is included in a complex together with another RuvB-like DNA helicase, TIP49a. J. Biol. Chem. **274**, 22437–22444
40 Makino, Y., Kanemaki, M., Kurokawa, Y., Koji, T. and Tamura, T. (1999) A rat RuvB-like protein, TIP49a, is a germ cell-enriched novel DNA helicase. J. Biol. Chem. **274**, 15329–15335
41 Fan, N., Cutting, S. and Losick, R. (1992) Characterization of the *Bacillus subtilis* sporulation gene *spoVK*. J. Bacteriol. **174**, 1053–1054
42 Freymann, D.M., Keenan, R.J., Stroud, R.M. and Walter, P. (1999) Functional changes in the structure of the SRP GTPase on binding GDP and Mg^{2+}GDP. Nat. Struct. Biol. **6**, 793–801

Received 7 August 2007
doi:10.1042/BST0360072