Minireview

Inferences about the catalytic domain of P-type ATPases from the tertiary structures of enzymes that catalyze the same elementary reaction

Irina N. Smirnova, Vladimir N. Kasho, Larry D. Faller*

CURE: Digestive Diseases Research Center, Department of Medicine, University of California at Los Angeles School of Medicine, and Wadsworth Division, Department of Veterans Affairs Medical Center, Building 115, Room 324, 11301 Wilshire Boulevard, West Los Angeles, CA 90073, USA

Received 1 June 1998

Abstract The machinery to catalyze elementary reactions is conserved, and the number of solved enzyme structures is increasing exponentially. Therefore, structures of enzymes that catalyze phosphate transfer are reviewed, and a supersecondary structure connecting the Walker A sequence to another sequence containing functional amino acids is proposed as an additional signature for the active site. The new signature is used to infer the identity of the P-loop in P-type biological pumps and may be useful in predicting targets for site-directed mutagenesis in other enzymes of unknown structure like the AAA family and ABC transporters.

Key words: Active transport; Enzyme structure; Phosphate transfer; P-type pump; AAA family; ABC transporter

1. Introduction

Relatively little is known about the structures of the molecules responsible for transporting metabolites across cell membranes because they are integral membrane proteins that are difficult to crystallize. However, members of the subclass capable of generating ion gradients (primary or active transporters) are also enzymes, and enzymes are thought to evolve by recruiting the chemistry to catalyze a rudimentary reaction step [1]. Therefore, structures in the Brookhaven Protein Data Bank (PDB) for enzymes that carry out the same elementary reaction can be used to develop testable models for the catalytic domains of biological pumps. In the case of P-type pumps, ATP hydrolysis is catalyzed by initially transferring the γ-phosphate to a carboxyl group on the enzyme [2]. In this review, we infer the identity of the functional analogue of the Walker A sequence [3] in P-type pumps from a discussion of solved enzyme structures and propose a supersecondary structural signature for Mg^{2+}-dependent catalysis of phosphate (P_i) transfer.

2. Published proposal for the MgATP binding site of P-type ATPases

A divalent cation is required for the enzymatic activity of P-pumps [4]. The first clue to the identity of the amino acids that coordinate Mg^{2+} was discovered by measuring the effects of site-directed mutagenesis of aspartic acid in the conserved amino acid sequence DPPR [5] on catalysis of ^{18}O exchange by Na^+/K^+-ATPase [6]. The conclusion that the carboxyl group coordinates Mg^{2+} and forms a ternary complex with P_i was tested by searching the PDB for homologous sequences with the postulated function in enzymes that catalyze P_i transfer. Adenylate kinase (AK) has been proposed as a model for the nucleotide binding domain of P-type pumps [7], and in the structure of the yeast enzyme co-crystallized with Mg^{2+} and the substrate analogue P_1,P_5-bis(adenosine-5′-)pentaphosphate (Ap_5A), the side-chain carboxyl group of aspartic acid in the sequence DGFPR and the γ-phosphate that is transferred from ATP to AMP are in the second coordination sphere of the hydrated Mg^{2+} ion [8]. DGFPR and another conserved peptide (TGDMLR) reach around Ap_5A like two hands with the thumbs (e.g. positively charged arginine side chains) holding AMP in position to react with ATP by hydrogen bonds to the phosphate oxygens and the forefingers (negatively charged aspartic acid side chains) holding Mg^{2+} in a ternary, outer-sphere electrostatic complex with the polyphosphate chain of ATP.

P-type pumps have a conserved sequence $TGD(X)_yK(R)$ homologous to the TGDMLR sequence in the AMP site of yeast AK that has been implicated in function by site-directed mutagenesis [9]. The AMP site is also between two relatively immobile domains of AK with more hydrogen bonds to adenosine than the ATP site, which is completed by large movements of a third domain after ATP binds [10]. Therefore, the AMP site of AK was proposed as a model for the MgATP binding site of P-type pumps with DPPR and $TGD(X)_yK(R)$ functioning like DGFPR and TGDMLR [11]. The model was tested by energy minimization calculations showing that the 'hand' motif of DGFPR and DMLR in AK is predicted for both DPPR and $D(X)_yK(R)$, despite different numbers ($y = 1$, 2, or 5) of intervening amino acids between the charged residues depending upon the ion(s) transported and the source of the pump (Table 1).

3. Tentative alignment of P-type ATPase and AK sequences

A problem with equating the DPPR and TGDHPITAK peptides of Na^+/K^+-ATPase with the DGFPR and TGDMLR peptides of yeast AK is that the sequences do not align properly. The DPPR sequence precedes the TGDHPITAK sequence in Na^+/K^+-ATPase, whereas DGFPR follows TGDMLR in yeast AK. One way of aligning the region in P-type ATPase containing the DPPR and $TGD(X)_yK(R)$ sequences with the primary structure of AK was discovered by comparing the distances between conserved sequences in P-type pumps and AK. The SwissProt database was searched [12] and secondary structural predictions [13,14]

*Corresponding author. Fax: (1) (310) 268-4963.
E-mail: lfaller@ucla.edu

PII: S0014-5793(98)00760-1

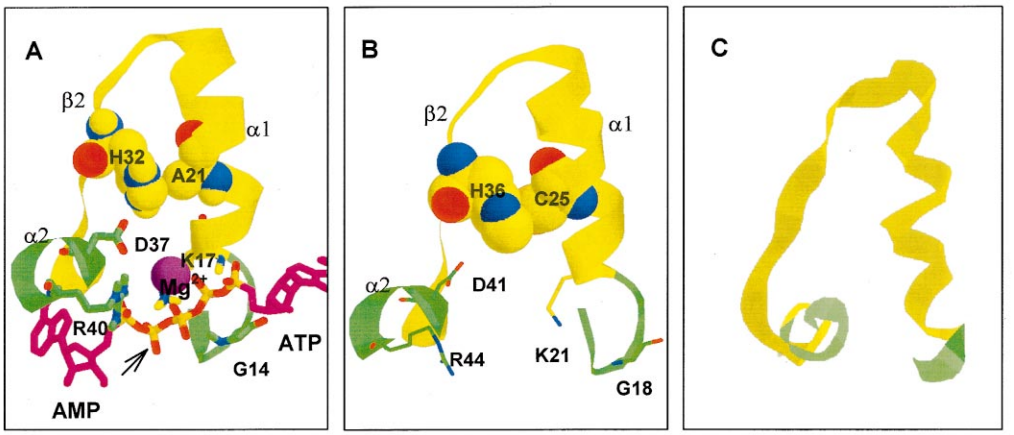

Fig. 1. The observed supersecondary structure of the catalytic hairpin in AK is predicted for P-type pumps. A: RASMOL 2.6 representation of the sequence G11-R40 (Table 1) in yeast AK co-crystallized with Ap_5A and Mg^{2+} [8]. Ribbon representation of polypeptide chain with last amino acid in GPPG sequence (G14) and side chains that interact with substrate (stick representation) indicated by sticks and space filling representation of hydrophobic contacts. B: Structure of the sequence G15–R44 in pig AK crystallized without nucleotide or a divalent cation present [20]. C: Fold predicted by energy minimization with Chem3DPro 3.5 and Chem3DUltra 4.0 software for the sequence D586-K618 in Na^+/K^+-ATPase with the secondary structural predictions in Table 1 imposed upon the model.

and alignments [15] were obtained by submitting the largest cytosolic domain of sheep α1 Na^+/K^+-ATPase (K342–E779) and yeast AK to the European Molecular Biology Laboratory in Heidelberg, Germany, as query sequences.

Twenty amino acids separate both the sequences that align with DPPR and TGDHPITAK in P-type pumps and the sequences that align with GPPG and TGDMLR in adenylate kinases. GPPG is in the Walker A sequence of AK, which differs from the consensus sequence $G(A)(X)_4G(A)KT(S)$ [16] by another G instead of T or S after the strictly conserved K [17]. Table 1 shows examples of a tentative alignment of the sequence homologous to DPPR in pumps with the Walker A sequence of AK. In solved structures, the Walker A sequence forms a phosphate binding loop (P-loop). The P-loop is always preceded by a β-strand, and K begins an α-helix. The predicted secondary structure of P-type pumps puts DPPR in a loop preceded by a strand (not shown), and the predicted pattern of secondary structures helix-loop(or turn)-strand in the conserved spacer (tentatively X_{20}) is the same as the secondary structural motif observed in crystals of AK (Fig. 1).

4. Tertiary structure of spacer

Fig. 1A shows that the architecture of the spacer (yellow) in yeast AK co-crystallized with Mg^{2+} and Ap_5A [8] resembles an α,β-hairpin juxtaposing functional peptides (green), so the supersecondary structure will be called the 'catalytic hairpin'. The α-helical and extended arms of the hairpin are in a plane 'latched' together by contacts between hydrophobic side chains (space filling). The AMP site is on the left [18] with the 'extra' phosphate of the substrate analogue Ap_5A indicated by an arrow. The P-loop is on the right beginning with G11 through G16 in green. The hydrogen bonds between backbone nitrogens and oxygens in the polyphosphate chain of ATP characteristic of P-loops begin with G14 and continue through T19. K17 is at the start of the α-helical arm of the catalytic hairpin with its ε-amino group (below Mg^{2+}) hydrogen-bonded to the β- and γ-phosphates of ATP. The positively charged side chain of K17 assumes different conformations, also interacting with oxygen of phosphate in the transition state (extra phosphate), and is thought to facilitate catalysis by moving with the negative charge as the γ-phosphate is transferred [8]. At the other end of the catalytic hairpin, TG (yellow) makes a turn and the charged side chains of DMLR (green) in α2 are arranged in the hand motif. The guanidino group of R40 is hydrogen-bonded to a phosphate oxygen of AMP, and D37 is close enough to Mg^{2+} in the water-filled (waters not shown) cavity to be in an outer coordination sphere of the hydrated divalent cation. The β- and γ-phosphates of ATP are in the second coordination sphere of

Table 1
Tentative alignment of P-type pumps with AK

Protein		Predicted (pumps) or observed (AK) secondary structure				
classification	specific example	loop	α-helix	turn	strand	$TGD(X)_yK(R)$
P-type pump	Na^+/K^+ATPase (sheep)	**DPPR***AA*	*VPDAVGKCRS*	*AGI*	*KVIMV*	TG**D**HPITA**K**
	H^+/K^+-ATPase (pig)	**DPVR***AT*	*VPDAVLKCRT*	*AGI*	*RVIMV*	TG**D**HPITA**K**
	PM Ca^{2+}-ATPase (pig)	**DPVR***PE*	*VPDAIKKCQR*	*AGI*	*TVRMV*	TG**D**NINTA**R**
	SR Ca^{2+}-ATPase (rabiit)	**DPPR***KE*	*VMGSIQLCRD*	*AGI*	*RVIMI*	TG**D**N**K**
AK	yeast	**GPPG***AG*	*KGTQAPNLQERF*	*HA*	*AHLA*	TG**D**ML**R**
	pig	**GGPG***SG*	*KGTQCEKIVQKY*	*GY*	*THLS*	TG**D**LL**R**

Bold type denotes amino acids in the predicted loop of pumps aligned with the first four amino acids in the Walker A sequence $G(X)_4GKG$ of AK and amino acids in the TGD sequence with charged side chains that form the thumb and forefinger of the hand motif. The 20 intervening amino acids are italicized. SR: sarcoplasmic reticulum; PM, plasma membrane.

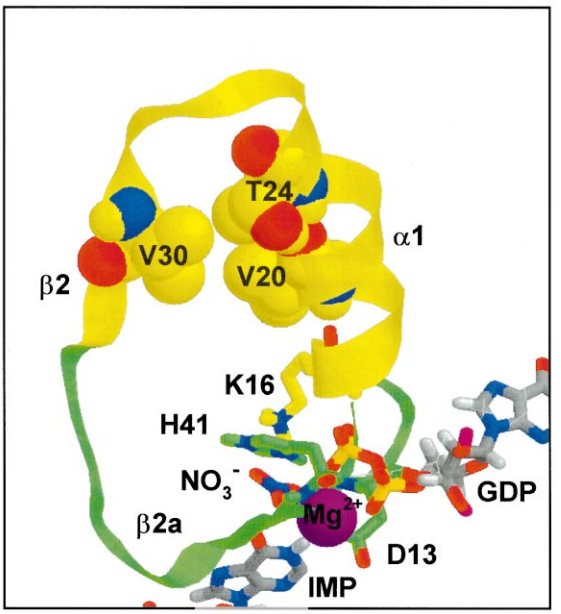

Fig. 2. Example of an enzyme that catalyzes P_i transfer with a hand motif in the P-loop. RASMOL 2.6 representation of the structure of the sequence D13-H41 in AMPSase from *E. coli* with Mg^{2+}, IMP, GDP, NO_3^- (blue and red), and hadacidin (not shown) bound [21]. The hydrogen bonds between polypeptide chain nitrogens and phosphate oxygens characteristic of the P-loop (green right) are formed between D13, G15, K16, G17 and either the β-phopshate of GDP or NO_3^- (D13). Mg^{2+} coordinates α- (2.35 Å) and β- (2.28 Å) phosphate oxygens of GDP.

Mg^{2+}. The positive charge of the divalent cation is thought to promote catalysis by neutralizing the negative charge developed on a pentacoordinate transition state formed during P_i transfer [19].

The structure of pig AK crystallized without Mg^{2+} or a substrate analogue present [20] in Fig. 1B shows that the architecture of the catalytic hairpin is essentially the same even when the contacts with $MgAp_5A$ are lost and the amino acids in more than half of the 20 positions are different (Table 1). The importance of the hydrophobic latch to the structure of the hairpin is suggested by conservation of histidine in the β-strand arm of the hairpin and replacement of the alanine found in yeast by another amino acid capable of hydrophobic bonding (cysteine) in the α-helical arm of the pig enzyme.

Fig. 1C shows that a fold predicted (see legend) for the sodium pump sequence aligned with the catalytic hairpin of AK in Table 1 resembles the structures in Fig. 1A,B. Amino acids with apolar side chains capable of stabilizing the structure by hydrophobic bonding (valine or isoleucine) are conserved near the middle of the predicted α-helical and extended secondary structures (Table 1). Therefore, orienting ATP in the active site of Na^+/K^+-ATPase so that DHPITAK is the functional analogue of DMLR positions the putative P-loop (DPPR) containing oppositely charged side chains arranged in the hand motif [11] near Mg^{2+} and the γ-phosphate (cf. positions of G11 at start of ribbon and G14 in Fig. 1A).

5. Hand motifs in the P-loop

To see if charged amino acids in the P-loop can interact with both Mg^{2+} and P_i, the structures of enzymes in the PDB that catalyze P_i transfer were searched for precedents. Fig. 2 shows the structure of adenylosuccinate synthetase (AMPSase) crystallized with Mg^{2+} and substrates or substrate analogues bound [21]. AMPSase catalyzes adenylosuccinate synthesis from IMP by initially transferring the γ-phosphate of GTP to the 6-keto group of inosine. The supersecondary structure of the catalytic hairpin (Fig. 2), the role of polypeptide chain nitrogens in the P-loop, and coordination of Mg^{2+} by the polyphosphate chain (figure legend) are essentially the same as in AK.

One new feature of AMPSase is an extended P-loop [22] containing the conserved amino acid sequence DEGK with oppositely charged side chains (D13 and K16) separated by two amino acids (Table 2) folded into the hand motif (Fig. 2). The same role played by K17 in yeast AK is suggested by the position of K16 at the start of the α-helical arm of the cata-

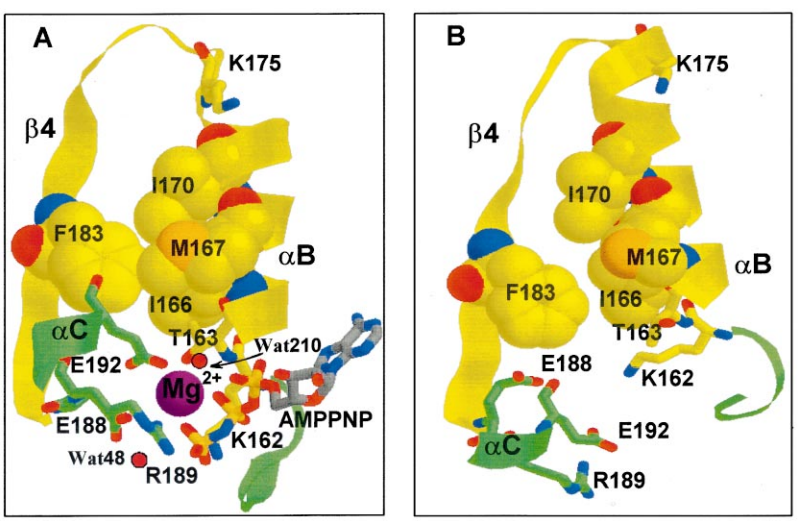

Fig. 3. Example of a conformational change in the catalytic hairpin. A: RASMOL 2.6 representation of catalytic hairpin (yellow) in a closed conformation of the β-subunit of bovine MF_1 with Mg^{2+} and the chemically inert substrate analogue AMP-PNP bound. The P-loop is on the right with K162 (side chain behind polyphosphate chain) beginning helix αB. B: Catalytic hairpin in open conformation of MF_1 β-subunit with nothing bound.

Table 2
Alignment of positively charged amino acid in P-loops and signature helix-turn-strand secondary structures

Protein		Observed (in tertiary) or predicted (from primary) secondary structure					Ref.
family	example	loop	α-helix	turn	strand	functional peptide	
Nucleotide kinase	AK (yeast)	**GPP***GAG*	**K**G**T**QAPNLQERF	HA	AHLA	TG**D**ML**R**	[8]
	GK (yeast)	GPS*GTG*	**K**STLLKKLFAEY	PDS	FGFSV	SSTT**R**TP**R**AG**E**	[32]
G protein	p21 (human)	GAG*GVG*	**K**SALTIQLIQ	N	HFVD	E**Y**DP**T**	[42]
Nitrogenase-like	AMPSase (*E. coli*)	GTQWG**D***EG*	**K**GKIVDLLTER	A	KYVVR	YQGGHNAG**H**	[21]
	DTBS (*E. coli*)	GTD**T***EVG*	**KT**VASCALLQAAKAA	GYR	TAGY	**K**PVA**S**GSEKTPEGLRNS**D**	[43]
RecA-like	MF_1 (bovine, beta)	GGA*GVG*	**KT**VLIMELINN	VAKAHGG	YSVFAGV	GE**R**TR**E**	[28]
	MF_1 (bovine, alpha)	GDR*QTG*	**KT**SIAIDTIINQKRFN	DGTDEKKK	LYCIYVAI	GQKEST	[28]
PCK	PCK (*E. coli*)	GLS*GTG*	*K*TTLS	TDPK	RRLI	G**DD**E	[25]
P-type	Na,K-ATPase (sheep)	**D**PP	**R**AAVPDAVGKCRS	AGI	KVIMV	TG**D**HPITA**K**	[5]
MAGUK	EM55 (human)	GASGVG	**R**SHIKNALLSQ	NPEK	FVYPV	PYTT**R**PP**R**KS**E**	[30]
	SP93 (human)	GPM	**K**DRINDDLISE	FPDKFG	SCV	PHTT**R**PK**R**DYE	[31]
AAA	PEX-N (rat)	GPPGSG	**K**TTAVTAACSRL	GLH	LLKV	PCSSLCADSSRT	[34]
	PEX6-C (rat)	GPPGTG	**K**TLLAKAVATE	CSLT	FLSV	**K**GP**E**LINMYVGQ	[34]
	PEX6-N (yeast)	TTNNVG	**K**ATMVRFASKY	LGI	HLLE	IDCLSLTSNSRQ	[35]
	PEX6-C (yeast)	GPPGTG	**K**TLMAKAIATN	FSLN	FFSV	**K**GP**E**LLNMYIGE	[35]
ABC	HisP (*S.t.*)	GSSGSG	**K**STFLRCINF	LEKPSEG	AIIV	NG**Q**NINLV**R**	[36]
	RibsA-N (*E. coli*)	GENGAG	**K**STMMKVLTG	IYTRDAG	TLLW	LGKETTFTGPKS	[38]
	RibsA-C (*E. coli*)	GLMGAG	**R**TELMKVLYG	ALPRTSG	YVTLD	GHEVVTRSPQDG	[38]

Bold type indicates alignment of K or R in P-loop, amino acids with side chain functional groups known to interact with one or more participants in P_i transfer, and conserved targets for mutagenesis in functional peptides of proteins with unknown structures. Amino acids in P-loops of known structures that hydrogen-bond via backbone nitrogens to phosphate oxygens of substrate are italicized. The closed conformation of MF_1 β-subunit is aligned. DTBS, dethiobiotin synthetase; p21, H-*ras* oncogene protein p21.

lytic hairpin (yellow ribbon) and proximity of the ε-amino group to the β-phosphate of GDP (2.65 Å) and NO_3^- (3.26 Å). The NO_3^- ion is located near both the β-phosphate of GDP and the 6-keto group of IMP in the position presumably occupied by P_i in the transition state. Mg^{2+} is close enough to NO_3^- (2.57 Å) and the carboxyl group of D13 (3.3 Å) to form a ternary complex with D13 functioning like D89 in AK [8]. Site-directed mutagenesis experiments [23] indicate that the carboxylic acid side chain of D13 is essential for catalytic activity, and the mutation K16Q reduces the rate significantly (65%). Therefore, Table 2 shows an alternative alignment of the DPPR sequence in Na^+/K^+-ATPase with the Walker A sequence in which R589 is assumed to play the same role as K in other enzymes that catalyze P_i transfer, D586 can function like D13 in AMPSase, and the number of amino acids in the catalytic hairpin varies from 13 to 25.

Other significant differences between AMPSase and AK are that the extended arm (β2) of the catalytic hairpin is followed after a loop by an additional extended strand (β2a) instead of an α-helix (α2 in Fig. 1A) and there is no analogue of the DMLR sequence. However, the enzymes are alike in that the peptides connected to the P-loop contain functional amino acids. Site-directed mutagenesis experiments indicate that H41 in β2a, which is hydrogen-bonded to the β-phosphate of GDP and near NO_3^- (3.9 Å), is essential for catalytic activity [23].

6. Functional peptide connected to P-loop by catalytic hairpin

The variety of peptides containing amino acids that bind the acceptor and/or participate in catalysis connected to the P-loop by the catalytic hairpin was examined by searching for additional structures of enzymes with reactants or their analogues bound. Some examples, grouped according to the structural classification of proteins (SCOP) [24], are listed in the top part of Table 2. All belong to the P-loop containing nucleotide triphosphate hydrolase fold except phosphoenolpyruvate carboxykinase (PCK), which has the consensus Walker A sequence but a new protein fold [25]. Recently, it was proposed that P-type pumps belong to the haloacid dehalogenase (HAD) superfamily [26]. However, the assignment is based on alignment and superposition of only three of the five conserved sequences that characterize P-type pumps [5] with sequences containing amino acids in the active site of L-2-haloacid dehalogenase II without considering function. Members of the HAD family do not bind nucleotide, and the elementary reaction catalyzed is esterification of a group on the enzyme instead of P_i transfer.

Machinery for catalyzing P_i transfer to water instead of the phosphate group of AMP (AK) or the 6-keto group of IMP (AMPSase) is shown in Fig. 3. The structure of the catalytic hairpin of the β-subunit of mitochondrial F_1-ATPase (MF_1) was chosen to illustrate P_i transfer to a different acceptor because a conformational change in the hairpin affecting function is also illustrated. Three conformations, corresponding to the open, loose, and tight states of the binding-change mechanism [27], are observed when the enzyme is crystallized from a solution containing Mg^{2+}, ADP, and β,γ-imidoATP (AMP-PNP) [28]. Fig. 3A shows an additional contact between the hydroxyl group of T163 in the consensus Walker A sequence and Mg^{2+}, but otherwise the P-loop/substrate contacts, the coordination of Mg^{2+} by substrate, and the supersecondary structure of the catalytic hairpin in the closed conformations of the β-subunit (ADP or AMP-PNP bound) are not significantly different than in Figs. 1A and 2.

The functional peptide connected to the P-loop by the catalytic hairpin is completely different with the guanidino group of R189 in a turn after the extended arm (β4) of the catalytic hairpin close enough to the γ-phosphate to help K162 (side

chain behind polyphosphate chain) facilitate transfer and E192 in the following helix (αC) close enough to a water molecule (Wat210, only oxygen shown) and Mg^{2+} to be an outer-sphere ligand of the divalent cation. E188 is believed to activate Wat48 for nucleophilic attack on the γ-phosphorus of ATP.

Fig. 3B shows that the catalytic hairpin is folded differently in the open conformation of the β-subunit (nothing bound). An additional helical turn including K175 forms, and the hydrophobic latch opens by F183 in the extended arm (β4) moving from interdigitation with I166, M167, and I170 in αB (Fig. 3A) to contact only with I166. The catalytic hairpin widens and twists slightly changing the relative orientation of the charged residues, so that they are no longer positioned to bind MgADP.

7. Conserved features of catalytic hairpin signature

Glycine-rich loop [29] is another name for the Walker A sequence that explains why P-type ATPases are cited as examples of ATP-binding proteins lacking the signature P-loop [17]. However, the conserved supersecondary structure connecting the P-loop to another functional peptide provides an additional criterion for locating the active site in enzymes that catalyze P_i transfer. Figs. 1–3 illustrate the general features of the catalytic hairpin deduced from more than 25 structures in the PDB. (1) The secondary structural motif is helix-turn-strand; (2) the length is approximately 20 amino acids; (3) the arms of the hairpin lie in a plane; (4) hydrophobic bonds latch the arms together; and (5) the arms are approximately the same length, so that amino acids in the P-loop and in the peptide following the hairpin can interact with Mg^{2+}-nucleotide triphosphate and/or the phosphoryl group acceptor. The catalytic hairpin signature is not conserved in all enzymes with Walker A sequences that catalyze P_i transfer, e.g. myosin. However, the helix-turn-strand motif is predicted for many enzymes that catalyze ATP hydrolysis or contain the Walker A sequence, and the bottom half of Table 2 lists examples selected to illustrate predictions for enzymes of unknown tertiary structure.

8. Refined model for MgATP binding site of P-type ATPase

The catalytic hairpin signature can be used to locate nonconsensus P-loops as illustrated by the proposed alignment of the positively charged amino acid in the Na^+/K^+-ATPase sequence DPPR with the conserved lysine of known structures in Table 2. The refined model for P-type pumps in which pump and AK sequences are aligned (Fig. 1A) preserves the positions and functional roles of amino acids in the Walker A sequence portrayed in Figs. 1–3. (1) Backbone nitrogens of residues from 3–4 before to 1–2 after the positive charge (italicized in Table 2) hydrogen-bond to the phosphate oxygens of the substrate; (2) a positively charged amino acid begins the α-helical arm of the catalytic hairpin; and (3) the positive charge polarizes the polyphosphate chain facilitating P_i transfer. Aligning R589 in Na^+/K^+-ATPase with K17 in yeast AK aligns D586 with the aspartate in the P-loop of AMPSase (D13) that is close enough to coordinate the Mg^{2+} required for P_i transfer (Fig. 2). The catalytic hairpin connects DPPR to HPITAK for which the hand motif that binds AMP in AK (Fig. 1A) is predicted [11].

9. Predictions for enzymes of unknown tertiary structure

The classification of the other proteins with unknown tertiary structures listed in Table 2 into families is based on conserved signature sequences including Walker A, implying the capability to bind nucleotide and possibly catalyze P_i transfer. For example, it is not known whether aggregation of membrane-associated guanylate kinases (MAGUK) in order to bind to channels or receptors and serve as signal transducers involves guanylate kinase (GK) activity. Two arginines and a glutamic acid (bold type) in the peptide following the catalytic hairpins of 55 kDa erythrocyte membrane protein (EM55) [30] and the channel associated protein of synapse-110 (CHAPSYN-110) [31] are conserved and align with residues that bind GMP in GK [32]. However, the inference that the GK domain may be inactive because the Walker A motif is defective with R instead of K (EM55) or missing amino acids (CHAPSYN-110) [33] may be incorrect, if experiments confirm the proposed alignment of DPPR in Na^+/K^+-ATPase with the P-loop found in other enzymes that catalyze P_i transfer.

Predicting targets for site-directed mutagenesis in the peptide connected to the P-loop by the catalytic hairpin is more difficult than identifying unusual P-loops because of the different architectures and functional groups illustrated in Figs. 1–3. The α-subunit of MF_1 is included in Table 2 to show that a protein domain containing the α,β-hairpin supersecondary structure may not even function as a catalyst. In the α-subunit, uncharged glutamine is substituted for the reactive glutamate (E188) of the β-subunit and kept from effectively activating water by hydrogen bonding to tyrosine, which replaces phenylalanine, in the extended arm [28].

To illustrate prediction of the active ATP-binding motif in proteins where there are two copies, the example of the AAA (ATPase associated with diverse cellular activities) family included in Table 2 is peroxisome assembly factor-2 (PEX6). Two species (rat [34] and yeast [35]) are compared to show that the positions of the acidic amino acids following the catalytic hairpin in the N-terminal copy (PPEX6-N) are not conserved, whereas the glutamic acid near the end of the extended arm (bold type) in the C-terminal copy (PEX6-C) is conserved in all species examined. The prediction that PEX6-C is the active copy because of the crucial role E188 plays in MF_1 (Fig. 3A) is supported indirectly by mutation of the conserved lysine in the Walker A sequence to alanine without loss of function in rat PEX6-N but with loss of function in rat PEX6-C [34].

Another motif that might be important for catalysis is illustrated by including the histidine permease (HisP) from *Salmonella typhimurium* (*S.t.*) [36] in Table 2 as an example of ATP-binding cassette (ABC) transporters. The sequence QNINLVR in the peptide connected to the P-loop resembles the DNINTAR sequence in plasma membrane P-type Ca^{2+} pumps (Table 1), suggesting Q, which also coordinates Mg^{2+} in known structures, and R as targets for site-directed mutagenesis. A naturally occurring mutation in which proline in the turn of HisP (Table 2) is replaced by serine dramatically reduces histidine transport [37], consistent with disruption of the predicted catalytic hairpin fold. The component of the ribose transporter containing a duplicated ATP-binding motif (RbsA) [38] is also included in Table 2 because the structure of RbsA-N will soon be available [39] to check the prediction

that the signature catalytic hairpin supersecondary structure is conserved in ABC transporters.

The signature helix-turn-strand motif after the Walker A sequence is predicted for all members of the AAA and ABC families examined. AAA proteins [40] and ABC transporters [41] are involved in a great number and variety of ATP-driven processes, and many, like the multidrug resistance (MDR) proteins and cystic fibrosis transmembrane conductance regulator (CFTR), are medically important. Therefore, the new signature motif discovered by reviewing the structures of enzymes that catalyze P_i transfer may be useful in designing site-directed mutagenesis experiments to identify functional amino acids in a number of important proteins with unknown tertiary structures.

Acknowledgements: This work was supported by a Veterans Administration Merit Review Award to L.D.F.

References

[1] Babbitt, P.C. and Gerlt, J.A. (1997) J. Biol. Chem. 272, 30591–30594.
[2] Post, R.L. and Kume, S. (1973) J. Biol. Chem. 248, 6993–7000.
[3] Walker, J.E., Saraste, M., Runswick, M.J. and Gay, N.J. (1982) EMBO J. 1, 945–951.
[4] Skou, J.C. (1957) Biochim. Biophys. Acta 23, 394–401.
[5] Serrano, R. (1989) Annu. Rev. Plant Physiol. Plant Mol. Biol. 40, 61–94.
[6] Farley, R.A., Heart, E., Kabalin, M., Putnam, D., Wang, K., Kasho, V.N. and Faller, L.D. (1997) Biochemistry 36, 941–951.
[7] Taylor, W.R. and Green, N.M. (1989) Eur. J. Biochem. 179, 241–248.
[8] Abele, U. and Schultz, G.E. (1995) Protein Sci. 4, 1262–1271.
[9] Clarke, D.M., Loo, T.W. and MacLennan, D.H. (1990) J. Biol. Chem. 265, 22223–22227.
[10] Gerstein, M., Schulz, G. and Chothia, C. (1993) J. Mol. Biol. 229, 494–501.
[11] Kasho, V.N., Stengelin, M., Smirnova, I.N. and Faller, L.D. (1997) Biochemistry 36, 8045–8052.
[12] Altshul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Nucleic Acids Res. 25, 3389–3402.
[13] Rost, B. and Sander, C. (1993) J. Mol. Biol. 232, 584–599.
[14] Rost, R. and Sander, C. (1994) Proteins 19, 55–72.
[15] Sander, C. and Schneider, R. (1991) Proteins 9, 56–68.
[16] Traut, T.W. (1994) Eur. J. Biochem. 222, 9–19.
[17] Saraste, M., Sibbald, P.R. and Wittinghofer, A. (1990) Trends Biochem. Sci. 15, 430–434.
[18] Müller, C.W. and Schulz, G.E. (1992) J. Mol. Biol. 224, 159–177.
[19] Knowles, J.R. (1980) Annu. Rev. Biochem. 49, 877–919.
[20] Dreusicke, D., Karplus, P.A. and Schulz, G.E. (1988) J. Mol. Biol. 199, 359–371.
[21] Poland, B.W., Fromm, H.J. and Honzatko, R.B. (1996) J. Mol. Biol. 264, 1013–1027.
[22] Alexeev, D., Baxter, R.L. and Sawyer, L. (1994) Structure 2, 1061–1072.
[23] Kang, C., Sun, N., Poland, B.W., Gorrell, A., Honzatko, R.B. and Fromm, H.J. (1997) J. Biol. Chem. 272, 11881–11885.
[24] Murzin, A.G., Brenner, S.E., Hubbard, T. and Chothia, C. (1995) J. Mol. Biol. 247, 536–540.
[25] Matte, A., Tari, L.W., Goldie, H. and Delbaere, L.T.J. (1997) J. Biol. Chem. 272, 8105–8108.
[26] Aravind, L., Galperin, M.Y. and Koonin, E.V. (1998) Trends Biochem. Sci. 23, 127–129.
[27] Boyer, P.D. (1993) Biochim. Biophys. Acta 1140, 215–250.
[28] Abrahams, J.P., Leslie, A.G.W., Lutter, R. and Walker, J.E. (1994) Nature 370, 621–628.
[29] Dreusicke, D. and Schulz, G.E. (1986) FEBS Lett. 208, 301–304.
[30] Ruff, P., Speicher, D.W. and Husain-Chishit, A. (1991) Proc. Natl. Acad. Sci. USA 88, 6595–6599.
[31] Kim, E., Cho, K.-O., Rothschild, A. and Sheng, M. (1996) Neuron 17, 103–113.
[32] Stehle, T. and Schulz, G.E. (1992) J. Mol. Biol. 224, 1127–1141.
[33] Woods, D.F. and Bryant, P.J. (1993) Mechanisms Dev. 44, 85–89.
[34] Tusukamoto, T., Miura, S., Nakai, T., Yokota, S., Shimozawa, N. and Osumi, T. Nature Genet. 11, 395-401.
[35] Erdmann, R., Wiebel, F.F., Flessau, A., Rytka, J., Beyer, A., Frohlich, K.U. and Kunau, W.H. (1991) Cell 64, 499–510.
[36] Higgins, C.F., Haag, P.D., Nikaido, K., Ardeshir, F., Garcia, G. and Ames, G.F. (1982) Nature 298, 723–727.
[37] Shyamala, V., Baichwal, V., Beall, E. and Ames, G.F.-L. (1991) J. Biol. Chem. 266, 18714–18719.
[38] Bell, A.W., Buckel, S.D., Groarke, J.M., Hope, J.N., Kingsley, D.H. and Hermodson, M.A. (1986) J. Biol. Chem. 261, 7652–7658.
[39] Armstrong, S.R., Tabernero, L., Zhang, H., Hermodson, M. and Stauffacher, C. (1998) Biophys. J. 74, A338.
[40] Confalonieri, F. and Duguet, M. (1995) BioEssays 17, 639–650.
[41] Tusnády, G.E., Bakos, É., Váraki, A. and Sarkadi, B. (1997) FEBS Lett. 402, 1–3.
[42] Pai, E.F., Krengel, U., Petsko, G.A., Goody, R.S., Kabsch, W. and Wittinghofer, A. (1990) EMBO J. 9, 2351–2359.
[43] Huang, W., Jia, J., Gibson, K.J., Taylor, W.S., Rendina, A.R., Schneider, G. and Lindqvist, Y. (1995) Biochemistry 34, 10985–10995.