Subsystem: Coenzyme A Biosynthesis

Andrei Osterman
The Burnham Institute and FIG,

I. Subsystem Introduction

Coenzyme A is a universal and essential cofactor in all forms of cellular life acting as a principal acyl carrier in numerous biosynthetic, energy-yielding, and degradative pathways {Begley et al. 2001}. Earlier bioinformatics analysis of Coenzyme A biosynthesis revealed a number of interesting variations between species {Gerdes et al,, 2002; Osterman and Overbeek, 2003; Genschel, 2004}. A current version of this subsystem in SEED covers 263 diverse genomes (239 bacteria, 13 archaea and 7 eukaryotes).

II. Subsystem notes, functional variants, open problems and conjectures

(The following slides provide abbreviations of functional roles and diagrams.)

A five-step pathway from pantothenate (vitamin B_5) to Coenzyme A is the universal component of the subsystem conserved in the majority of species. The most variable aspect of this pathway is pantothenate kinase (PANK). All of the three known nonorthologous forms of PANK can be detected in bacterial genomes, and, in some cases, two alternative forms are present in the same organism. One of these forms (PANK3) was recently identified in *B.subtilis* (gene coaX) and later characterized in details and confirmed in other bacterial species (such as *H.pylori)*, by Dr. E. Strauss (J.Biol. Chem. in press). PANK2, characteristic of all eukaryotes, was predicted {Daugherty et al., 2002} and subsequently verified {Choudhry et al, 2003} as the only PANK in all Staphylococci. No archaeal gene for PANK has been identified so far, although reasonable candidates may be proposed, such as PAE3407 from *Pyrobaculum aerophilum*. Members of this uncharacterized family of putative GHMP-like kinases are conserved in all archaea and have a tendency to cluster on the chromosome with other genes involved in CoA biosynthesis (see an illustration). For example, in case of *P. aerophilum*, this gene is next to PAE3409 and PAE3410 coding for KPRED and KPHMT, respectively. This conjecture was independently made by {Genschel, 2004}. Based on a long-range sequence similarity analysis, another protein family conserved in archaea (eg PAE1629 of *P. aerophilum*) was proposed for the role of DPCK (see COG0237 at NCBI). We have not included these candidates in this version of subsystem to emphasize a necessity of experimental verification

II. Subsystem notes, functional variants, open problems and conjectures

(continued)

Several examples illustrating major functional variants of the subsystem are outlined below and illustrated in the following slides. A semi-automated identification and analysis of the variations in this subsystem were described by us in {Y.Ye et al, 2005, published in ISMB 05) .

A de novo pantothenate biosynthesis, which is present in many bacteria (variant 1 and 2), fungi (variant 2) and archaea (variant 3) is functionally replaced by salvage in higher multicellular eukaryotes and a number OF bacterial pathogens (variant 4).

The most radically truncated version of subsystem (variant 7) is observed in all Chlamydiaceae and Rickettsiaceae. These intracellular pathogens, are most likely dependent on the salvage of the last CoA precursor (dephospho-CoA) from the eukaryotic host.

A sub-set of genes present in a small group of bacterial pathogens (such as *Mycoplasma penetrans* and *Treponema palidum*, variant 6) may be rationalized via a "pantetheine shunt". This hypothetical route assumes salvage of the host's pantetheine, a possible product of a poorly understood CoA catabolism. Although this variant appears to be rarely used for CoA biogenesis, it may be a rather common route of pantetheine recycling.

A disrupted pattern (missing PANK, PPCS and PPCDC) observed in *Buchnera aphidicola* suggests another interesting possibility, a metabolic exchange between this obligate intracellular endosymbiont and aphid host cells. According to this hypothesis, pantothenate produced but not utilized by *B. aphidicola,* may be fed directly into the universal pathway of the host. The latter may "pay back" by providing phosphopantetheine intermediate required for the last two steps of CoA synthesis in *B. aphidicola.*

Among open (missing gene) problems within this subsystem, one may notice the absence of gene candidate for aspartate decarboxylase (ASPDC) in ~ 45 bacterial and fungal genomes with an otherwise complete set of genes for the de novo synthesis (e.g. variants 2 and 5). Possible interpretations of this observation are: (i) the presence of an alternative non-orthologous and presently unknown form of ASPDC (missing gene) or (ii) the existence of an alternative source of beta-alanine ("missing" pathway). Many aspects of CoA biosynthesis are still unclear in archaea.

1. Functional Roles, Abbreviations, Subsets and Alternative Forms of Enzymes

Subsystem: Coenzyme A Biosynthesis

Alternative forms | **Subsets of roles**

	Subsystem: Coenzyme A Biosynthesis			
	1	ASPDC	Aspartate 1-decarboxylase (EC 4.1.1.11)	De novo biosynthesis
	2	KPHMT	3-methyl-2-oxobutanoate hydroxymethyltransferase (EC 2.1.2.11)	
*KPRED	3	KPRED	2-dehydropantoate 2-reductase (EC 1.1.1.169)	
*KPRED	4	KARED	Ketol-acid reductoisomerase (EC 1.1.1.86)	
	5	PBAL	Pantoate--beta-alanine ligase (EC 6.3.2.1)	
	6	PANF	Pantothenate:Na+ symporter (TC 2.A.21.1.1)	Universal pathway
*PANK	7	PANK	Pantothenate kinase (EC 2.7.1.33)	
*PANK	8	PANK2	Pantothenate kinase type II, eukaryotic (EC 2.7.1.33)	
*PANK	9	PANK3	Pantothenate kinase type III, CoaX-like (EC 2.7.1.33)	
	10	PPCS	Phosphopantothenoylcysteine synthetase (EC 6.3.2.5)	
	11	PPCDC	Phosphopantothenoylcysteine decarboxylase (EC 4.1.1.36)	
*PPAT	12	PPAT	Phosphopantetheine adenylyltransferase (EC 2.7.7.3)	
*PPAT	13	PPAT2	Phosphopantetheine adenylyltransferase, type II eukaryotic (EC 2.7.7.3)	
	14	DPCK	Dephospho-CoA kinase (EC 2.7.1.24)	

2. Subsystem spreadsheet summary

Subsystem: Coenzyme A Biosynthesis

modifies from the SEED display to illustrate identified functional variants

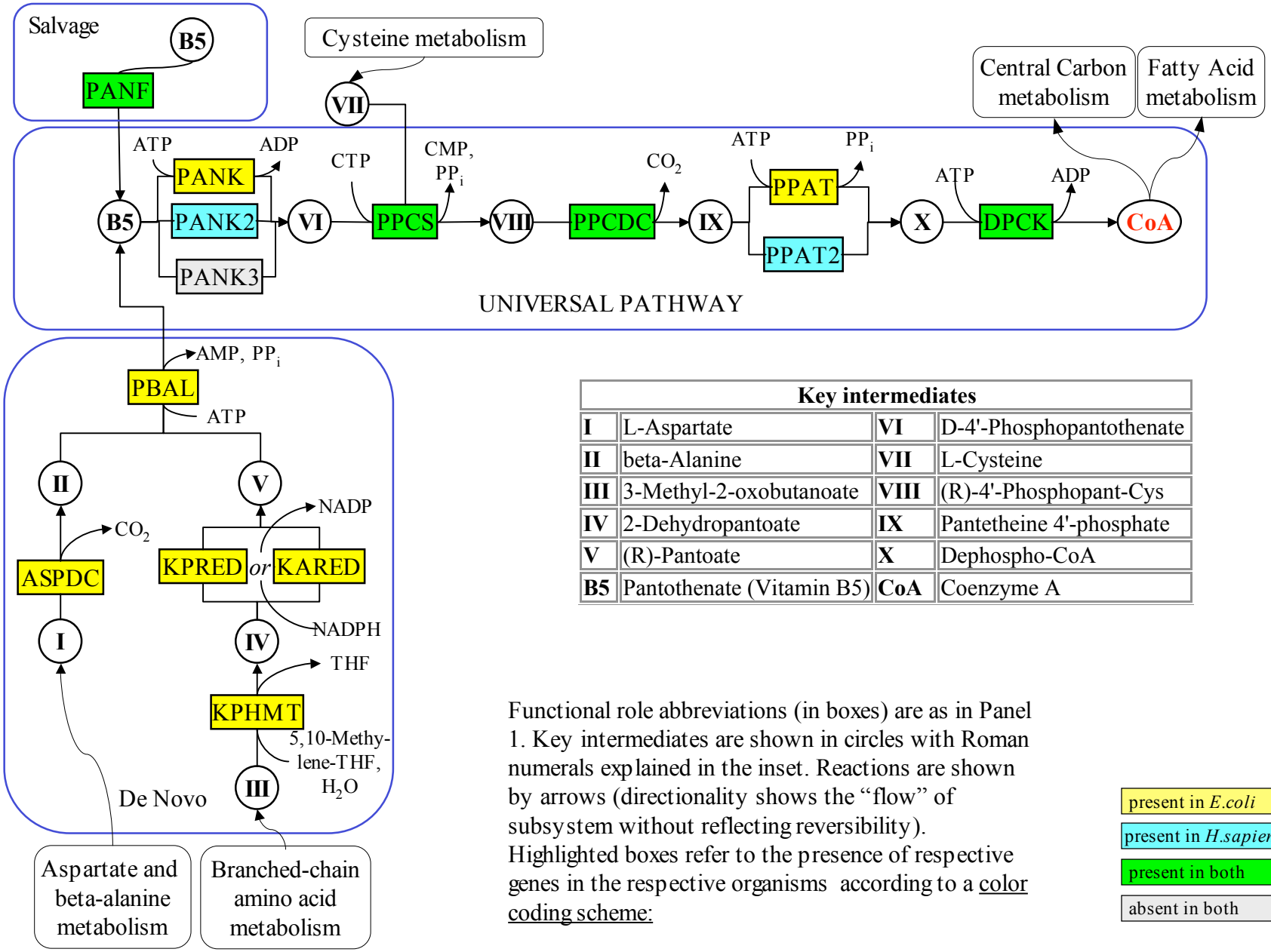
Examples/ number of genomes	variant	ASPDC	KPHMT	alternatives		PBAL	PANF	alternatives			PPCS	PPCDC	alternatives		DPCK
				KPRED	KARED			PANK	PANK3	PANK2			PPAT	PPAT2	
De novo, complete /100		+	+	+ or/and +		+	±	+or/and +	or	+	+	+	+		+
Esc.coli K12	1a	*panD*	*panB*	*panE*	*ilvC*	*panC*	*panF*	*coaA*			*coaBC*	*coaBC*	*coaD*		*coaE*
Dein.radiod..	1b	900	2793		1701	1347	659		643		761	761	824		2074
Staph.aur.	1c	2477	2479	2480	1926	2478	1771+			1999	1089	1089	1004		1561
De novo (ASPDC=?)/45		?	+	+ or/and +		+	±	+or/and +	or	+	+	+	+ or +		+
Shew. onei.	2a	?	810	3466	3948	809		199			3851	3851	4255		389
Geob.metalli.	2b	?	179	2031	2898	178	1470		1816		2004	2004	67		1465
Sach. cerev.	2c	?	377	755	3888	2637				1397	2697	3297+		2314	1076
De novo, archaea/8		?	+	+ or/and +		?	±	?			+	+		+	?
Pyrob.aeroph.	3	?	2402	2401		?		?			831	831		608	?
Salvage of B5/40							+/?	+or/and +	or	+	+	+	+ or +		+
Strep.pneum	4a				403			741			1109	1110	1781		873
Therm.teng.	4b				15				2203		1410	1410	1388		814
Hom.sap	4c						11370+			10185+	12518	11601		12914	12914+
Truncated pathways/20		±	±	±		±		±					±		+
Buch.aphid	5	?	196		566	195							554		202
Trep.pallid.	6								430				283		296
Chlam.trach.	7														504

Patterns of genes grouped in functional variants (and sub-variants) are generalized by: "+" - presence of a gene (for a given role) is required; "±" - optional, "?" - function is inferred by pathway analysis but a gene is unknown (can not be projected by homology). In several examples illustrating variants assigned genes are shown by unique FIG IDs. Missing genes inferred by the functional context analysis are shown by "?" (see a discussion). Matching colors highlight genes that occur close to each other on the chromosome. Two examples (red box) are further illustrated by projection on a subsystem diagram.

3. Subsystem diagram

Example: ***E.coli K12 vs Homo sapiens***

Subsystem: Coenzyme A Biosynthesis



Key intermediates			
I	L-Aspartate	**VI**	D-4'-Phosphopantothenate
II	beta-Alanine	**VII**	L-Cysteine
III	3-Methyl-2-oxobutanoate	**VIII**	(R)-4'-Phosphopant-Cys
IV	2-Dehydropantoate	**IX**	Pantetheine 4'-phosphate
V	(R)-Pantoate	**X**	Dephospho-CoA
B5	Pantothenate (Vitamin B5)	**CoA**	Coenzyme A

Functional role abbreviations (in boxes) are as in Panel 1. Key intermediates are shown in circles with Roman numerals explained in the inset. Reactions are shown by arrows (directionality shows the "flow" of subsystem without reflecting reversibility). Highlighted boxes refer to the presence of respective genes in the respective organisms according to a color coding scheme:

- present in *E.coli*
- present in *H.sapiens*
- present in both
- absent in both

III. References

1. Begley, T. P., C. Kinsland, and E. Strauss. 2001. The biosynthesis of coenzyme A in bacteria. Vitam Horm 61:157-71.
2. Choudhry, A. E., T. L. Mandichak, J. P. Broskey, R. W. Egolf, C. Kinsland, T. P. Begley, M. A. Seefeld, T. W. Ku, J. R. Brown, M. Zalacain, and K. Ratnam. 2003. Inhibitors of pantothenate kinase: novel antibiotics for staphylococcal infections. Antimicrob Agents Chemother 47:2051-5.
3. Daugherty, M., B. Polanuyer, M. Farrell, M. Scholle, A. Lykidis, V. de Crecy-Lagard, and A. Osterman. 2002. Complete reconstitution of the human coenzyme A biosynthetic pathway via comparative genomics. J Biol Chem 277:21431-9.
4. Genschel, U. 2004. Coenzyme A Biosynthesis: Reconstruction of the Pathway in Archaea and an Evolutionary Scenario Based on Comparative Genomics. Mol Biol Evol.
5. Gerdes, S. Y., M. D. Scholle, M. D'Souza, A. Bernal, M. V. Baev, M. Farrell, O. V. Kurnasov, M. D. Daugherty, F. Mseeh, B. M. Polanuyer, J. W. Campbell, S. Anantha, K. Y. Shatalin, S. A. Chowdhury, M. Y. Fonstein, and A. L. Osterman. 2002. From genetic footprinting to antimicrobial drug targets: examples in cofactor biosynthetic pathways. J Bacteriol 184:4555-72.
6. Osterman, A., and R. Overbeek. 2003. Missing genes in metabolic pathways: a comparative genomics approach. Curr Opin Chem Biol 7:238-51.