Published in final edited form as:

Curr Opin Struct Biol. 2009 December ; 19(6): 666–671. doi:10.1016/j.sbi.2009.09.004.

Adenylate-forming enzymes

Stefan Schmelz and James H. Naismith

Scottish Structural Proteomics Facility and Biomedical Science Research Complex, The University of St Andrews, Scotland KY16 9ST, UK

Abstract

Thioesters, amides and esters are common chemical building blocks in a wide array of natural products. The formation of these bonds can be catalyzed in a variety of ways. For chemists, the use of an activating group is a common strategy and adenylate enzymes are exemplars of this approach. Adenylating enzymes activate the otherwise unreactive carboxylic acid by transforming the normal hydroxyl leaving group into adenosine monophosphate. Recently there have been a number of studies of such enzymes and in this review we suggest a new classification scheme. The review highlights the diversity in enzyme fold, active site architecture and metal coordination that has evolved to catalyze this particular reaction.

Introduction

Adenylation is an elegant biological process used to chemically activate carboxylate substrates by condensing them with ATP to liberate pyrophosphate (the driving force). The resulting carboxylate adenylate is very reactive and would be expected to decompose in water. Necessarily it seems that the enzymes that make the intermediate also catalyze a second step, in which a nucleophile (either amine, alcohol or thiol) reacts with the intermediate to generate the desired product, liberating AMP (Figure 1a). This strategy has a clear conceptual link to the function of common reagents such as acid chlorides or triflates in organic chemistry that are used to create good leaving groups for nucleophilic attack. Adenylate forming enzymes are involved in a variety of metabolic pathways such as ribosomal and non-ribosomal peptide synthesis, fatty acid oxidation or enzyme regulation. Some, as the D-alanyl carrier protein Ligase DltA, are important and potential drug targets [1]. Based on sequence analysis a superfamily of adenylate-forming enzymes was identified and subdivided into three closely related sub-families [2]. The sub families comprise the adenylation domains of non-ribosomal peptide synthetases (NRPS), acyl- or aryl CoA synthetases [3] and luciferase oxidoreductases [4]. Structurally distinct aminoacyl-tRNA synthetases and enzymes involved in NRPS-independent siderophores (NIS) synthesis also catalyze an adenylation reaction [5,6], but have not been integrated into the scheme.

We therefore propose a new classification for adenylate forming enzymes (Figure 1b) that recognizes their common chemistry. The first class constitutes the former superfamily with subclasses Ia, Ib and Ic. The second class comprises aminoacyl-tRNA synthetases with subclasses IIa and IIb, while class III contains the NIS enzymes with subclasses IIIa, IIIb, IIIc and IIId. This review should give the reader an updated overview of adenylate-forming enzymes with focus on their overall fold and active site architecture.

Same reaction but very different domain architectures

All adenylate-forming enzymes catalyze the difficult condensation between the weakly nucleophilic carboxylic acid and weakly electrophilic phosphate. The mechanism proceeds through a negatively pentavalent phosphorus atom which must be stabilized (Figure 1c). Class I adenylate-forming enzymes all have a large N-terminal domain with a small C-terminal domain connected via a flexible hinge (Figure 2a). The relative orientation of the two domains varies presumably as a consequence of the hinge. The N-terminal domain was described as “hammer-and-anvil”, subdivided into three subdomains with an α/β topology, topped by a $\alpha\beta\alpha\beta\alpha$ sandwich and a β -barrel [7]. The C-terminal domain sits on top like a “lid”, is made up of five β -strands and three α -helices. The two domain architecture is consistent for all class I enzymes. The recently deposited structure of the predicted Acyl-CoA synthetase FdrA from *E. coli* (3dmy) shows a very different three domain architecture (Figure 2b) and in the absence of functional studies its identification as Acyl-CoA synthetase may not be secure. Interestingly, the recent structure of fatty acid acyl-AMP ligase from *Mycobacterium tuberculosis* shows only minor changes in sequence and structure from fatty-acid acyl-CoA, but catalyses a different second step [8]. Class II enzymes show no such conserved overall domain arrangement. Subclass IIa aminoacyl-tRNA synthetases (aaRSs) have four domains, a zinc-binding, central catalytic, anticodon-binding and an amino acid specific domain. Some aaRSs (as LeuRS, IleRS or ValRS) have an additional proofreading/editing domain, called CP1 (connective polypeptide 1), which was reported to be important for enzyme fidelity (Figure 2c) [9]. In contrast to class IIa enzymes, subclass IIb consists of an anticodon-binding domain and the central catalytic domain although additional insertions are known (Figure 2d) [10]. Class III enzymes have three domains (Figure 2e, f), an N-terminal domain, a central domain and a C-terminal domain [5].

ATP binding

In class I enzymes the active site is formed by the interface between the domains, which are connected by a flexible hinge [11]. The domains significantly alter their relative orientations to adjust the catalytic site depending whether the protein is catalyzing adenylate formation or the second reaction. For example in human acyl-CoA synthetase [12] a conserved catalytic lysine important for adenylation is moved out after adenylate formation to facilitate thioester formation [12, 13]. Class II enzymes locate the active site with the central catalytic domain which undergoes conformational changes upon substrate binding. Recently Weimer, Shane et al. identified a series of conserved motions in class IIa enzymes [14]. Class III enzymes employ all three domains to make a deep cavity in the middle of the protein [5], which, apart from limited ordering of surrounding loops upon substrate binding, do not alter their structure upon substrate binding. Class I and IIa enzymes bind ATP in a Rossmann-fold, but class IIb and class III do not. Rather, in subclass IIb aaRSs ATP is bound by three conserved motifs within a seven folded anti-parallel β -sheet [15,16], while in class III enzymes it is bound by a novel fold, which shares parts of the cAPK nucleotide triphosphates binding fold [5].

Role of metals

All adenylate-forming enzymes are reported to be Mg^{2+} dependent [17], but in at least one case Mn^{2+} can substitute for Mg^{2+} [18]. In each structure with ATP bound, a Mg^{2+} is reported to coordinate the phosphate. The divalent metal ion will neutralize the charge of ATP, stabilize the transition state and neutralize the leaving pyrophosphate. With such an acknowledged central role, it is surprising that the coordination geometry and even more so the number of Mg^{2+} ions varies across the superfamily and even within classes. The NRPS

synthetase DltA (class Ia) (3fcc) shows a $\beta - \gamma$ coordination [19]. Whilst Mg^{2+} in tryptophanyl-tRNA synthetase from *G. stearothermophilus* (1m83) (class IIa) is coordinated by $\alpha - \beta - \gamma$ phosphates unlike the human homologue (2qui) [20], where an $\alpha - \beta$ coordination is found. In AcsD (2w02), a class IIIa enzyme, a single Mg^{2+} coordinates the $\alpha - \gamma$ phosphate, a geometry found only in seven other structures consisting of kinases and the ribonucleotide transformylase PurT (1kj9). In contrast lysyl-tRNA synthetase, a subclass IIb enzyme, binds 3 metal ions, coordinating the α and β phosphate and the β and γ phosphates [21].

Catalytic residues

At the reaction “schwerpunkt” (α phosphate of ATP) it is possible to identify chemical similarities of surrounding amino acids. In every structure an Arg, Lys or His (Figure 4) residue contacts the α phosphate but not always in a conserved position. In many but not all enzymes an additional positively charged residue surrounds the phosphate. Class I enzymes use either Arg (Acetyl-CoA synthetase [22]), Lys (GrsA and firefly luciferases) or His (firefly luciferase) near the α phosphate, a highly conserved threonine is also found here in all these enzymes. In class III two conserved residues, His and Arg, are close to the α phosphate but neither of these residues do superpose with those in class I. A strictly conserved arginine found in all class IIa or IIb type tRNA synthetases (R262 in LysU) does superimpose with R305 in AcsD. These residues (along with metal ion(s)) polarize the α phosphate of ATP enhancing the electrophilicity of the phosphorus. The positive charge will also polarize the negative charge of the incoming carboxylate increasing its nucleophilicity. The surrounding positive charges crucially will stabilize the negatively pentavalent transition state.

In all three classes point mutations of these positively charged residues show decreased enzyme activity. In class I enzymes mutations of the conserved lysine (K529R in firefly luciferase) facing the α phosphate of ATP caused ~625 fold reduction of activity. Beside lowered activity, mutants of H245 in firefly luciferase showed shifts in light emission spectra [23]. In class II aaRSs K233 mutants (second lysine in the conserved ATP binding KMSKS motif) of tyrosyl-tRNA synthetase, have been reported to deteriorate the ability to form tyrosine-adenylate [24, 25]. Similar results were reported for AcsD, a class IIIa enzyme, losing activity after mutation of positively charged R305 and H444 [5].

The pyrophosphate leaving group in class III appears to be retained after adenylation formation in a pocket which extends beyond the γ phosphate. In class I enzymes pyrophosphate is released after forming the intermediate reported for the adenylation domain of gramicidin S synthetase a class I enzyme [7]. In class II leucyl- or tyrosyl-tRNA synthetase pyrophosphate is reported to be released after amino acid adenylation is formed to mediate binding of tRNA for the second step of reaction [26, 27].

Adenylation enzymes have important and diverse roles in metabolic pathways of prokaryotes and eukaryotes. We suggest that an inclusive definition of the superfamily based on the commonality of the adenylation chemistry may be more useful than one limited to three dimensional structure. Recent reports of adenylation-forming enzymes show that the same chemistry can be catalyzed by very different structures. Despite the profound differences in structure and active site architecture, there are clear chemical similarities in the environment that surrounds the α phosphate. This chemical conservation is of course driven by the requirement to stabilize the pentavalent negatively charged transition state. From a chemical standpoint, the nucleophilic attack of the α or γ phosphate by a hydroxyl are closely related. To this point, it was noted that the class III adenylation enzymes are related to kinases. Chemical conservation has been observed in other enzyme families, proteases being perhaps the best known example. Akin to the proteases, the adenylation

enzymes have utilized different spatial arrangements of residues to stabilize the transition state. It is likely that the requirement to activate carboxylic acids for condensation is particularly ancient (thioester, amide and ester bonds are extremely common chemical building blocks) and this may have driven the convergent evolution seen in the structures of the adenylate forming enzymes. Furthermore they are not only interesting enzymes, but have utilities in commercial significant biotransformation or biosynthetic chemistry [28-31].

References

1. Du L, He Y, Luo Y. Crystal structure and enantiomer selection by D-alanyl carrier protein ligase DltA from *Bacillus cereus*. *Biochemistry*. 2008; 47:11473–11480. [PubMed: 18847223]
2. Fulda M, Heinz E, Wolter FP. The fadD gene of *Escherichia coli* K12 is located close to rnd at 39.6 min of the chromosomal map and is a new member of the AMP-binding protein family. *Mol Gen Genet*. 1994; 242:241–249. [PubMed: 8107670]
3. Jogl G, Tong L. Crystal structure of yeast acetyl-coenzyme A synthetase in complex with AMP. *Biochemistry*. 2004; 43:1425–1431. [PubMed: 14769018]
4. Conti E, Stachelhaus T, Marahiel MA, Brick P. Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *Embo J*. 1997; 16:4174–4183. [PubMed: 9250661]
- * 5. Schmelz S, Kadi N, McMahon SA, Song LJ, Oves-Costales D, Oke M, Liu HT, Johnson KA, Carter LG, Botting CH, et al. AcsD catalyzes enantioselective citrate desymmetrization in siderophore biosynthesis. *Nature Chemical Biology*. 2009; 5:174–182. This paper reports the chemical and structural characterization of the first class III member of adenylate-forming enzyme. The enzyme makes AMP-citrate which is then condensed with L-serine to make a citrate ester. The paper highlights the similarity in both chemistry and structure with kinases and adenylate-forming enzymes.
6. Chang KH, Xiang H, Dunaway-Mariano D. Acyl-adenylate motif of the acyl-adenylate/thioester-forming enzyme superfamily: a site-directed mutagenesis study with the *Pseudomonas* sp. strain CBS3 4-chlorobenzoate:coenzyme A ligase. *Biochemistry*. 1997; 36:15650–15659. [PubMed: 9398293]
7. Conti E, Franks NP, Brick P. Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes. *Structure*. 1996; 4:287–298. [PubMed: 8805533]
- * 8. Arora P, Goyal A, Natarajan VT, Rajakumara E, Verma P, Gupta R, Yousuf M, Trivedi OA, Mohanty D, Tyagi A, et al. Mechanistic and functional insights into fatty acid activation in *Mycobacterium tuberculosis*. *Nat Chem Biol*. 2009; 5:166–173. [PubMed: 19182784] The first structure of a new class of fatty acid acyl-AMP ligase. This protein activates fatty acid forming an adenylate intermediate, which is then coupled to an acyl carrier protein in an NRPS like biosynthetic pathway. This novel protein differs in function (but only slightly in sequence) from the conventional fatty acid acyl-CoA ligase which couples to CoA. The study identifies the key determinants that control the final outcome of the reaction.
9. Mascarenhas AP, Martinis SA. Functional segregation of a predicted “hinge” site within the beta-strand linkers of *Escherichia coli* leucyl-tRNA synthetase. *Biochemistry*. 2008; 47:4808–4816. [PubMed: 18363380]
10. Xie W, Nangle LA, Zhang W, Schimmel P, Yang XL. Long-range structural effects of a Charcot-Marie-Tooth disease-causing mutation in human glycyl-tRNA synthetase. *Proc Natl Acad Sci U S A*. 2007; 104:9976–9981. [PubMed: 17545306]
- * 11. Wu R, Reger AS, Lu X, Gulick AM, Dunaway-Mariano D. The mechanism of domain alternation in the acyl-adenylate forming ligase superfamily member 4-chlorobenzoate:coenzyme A ligase. *Biochemistry*. 2009; 48:4115–4125. [PubMed: 19320426] In this study the main focus is on the conformational domain changes in the two different functional states of class I adenylate forming enzymes. Mutational and functional studies on the hinge region show that one particular residue is maintaining the 140 degree domain shift moving two important catalytic residues in and out.
- * 12. Kochan G, Pilka ES, von Delft F, Oppermann U, Yue WW. Structural snapshots for the conformation-dependent catalysis by human medium-chain acyl-coenzyme A synthetase

- ACSM2A. *J Mol Biol.* 2009; 388:997–1008. [PubMed: 19345228] An elegant study which traps substrates of the enzyme along the reaction mechanism. The structures reveal the conformational changes in domain structure that accompany the switch from adenylate formation to thioester condensation.
- * 13. Osman KT, Du L, He Y, Luo Y. Crystal structure of *Bacillus cereus* D-alanyl carrier protein ligase (DltA) in complex with ATP. *J Mol Biol.* 2009; 388:345–355. [PubMed: 19324056] A second paper from this lab (see also Du et al., *Biochemistry*, 2008) on the D-alanyl carrier protein ligase. This paper focuses on the mechanism, in particular the conserved lysine residue.
 - ** 14. Weimer KM, Shane BL, Brunetto M, Bhattacharyya S, Hati S. Evolutionary basis for the coupled-domain motions in *Thermus thermophilus* leucyl-tRNA synthetase. *J Biol Chem.* 2009; 284:10088–10099. [PubMed: 19188368] This paper discusses a normal mode analysis of aminoacyl-tRNA synthetases which identifies the large domain movements required for catalysis. The analysis was then extended to look for residues which control this motion. The analysis identifies residues which are predicted to have co-evolved at the domain interfaces.
 - 15. Cusack S. Eleven down and nine to go. *Nat Struct Biol.* 1995; 2:824–831. [PubMed: 7552701]
 - 16. Crepin T, Yaremchuk A, Tukalo M, Cusack S. Structures of two bacterial prolyl-tRNA synthetases with and without a cis-editing domain. *Structure.* 2006; 14:1511–1525. [PubMed: 17027500]
 - 17. Airas RK. Magnesium dependence of the measured equilibrium constants of aminoacyl-tRNA synthetases. *Biophys Chem.* 2007; 131:29–35. [PubMed: 17889423]
 - 18. Belrhali H, Yaremchuk A, Tukalo M, Berthet-Colominas C, Rasmussen B, Bosecke P, Diat O, Cusack S. The structural basis for seryl-adenylate and Ap4A synthesis by seryl-tRNA synthetase. *Structure.* 1995; 3:341–352. [PubMed: 7613865]
 - * 19. Yonus H, Neumann P, Zimmermann S, May JJ, Marahiel MA, Stubbs MT. Crystal structure of DltA. Implications for the reaction mechanism of non-ribosomal peptide synthetase adenylation domains. *J Biol Chem.* 2008; 283:32484–32491. [PubMed: 18784082] The modification of lipoteichoic acid by D-alanine plays a crucial role in the cell walls of Gram⁺ bacteria. This paper reports the structure of the enzyme and addresses the question of multiple conformational states observed during turnover. More work was carried in two other studies on *Bacillus cereus* homologues by Du et al., *Biochemistry*, 2008 and Osman, KT et al, *J. Mol. Biol.*, 2009.
 - 20. Shen N, Zhou M, Yang B, Yu Y, Dong X, Ding J. Catalytic mechanism of the tryptophan activation reaction revealed by crystal structures of human tryptophanyl-tRNA synthetase in different enzymatic states. *Nucleic Acids Res.* 2008; 36:1288–1299. [PubMed: 18180246]
 - 21. Desogus G, Todone F, Brick P, Onesti S. Active site of lysyl-tRNA synthetase: structural studies of the adenylation reaction. *Biochemistry.* 2000; 39:8418–8425. [PubMed: 10913247]
 - 22. Gulick AM, Starai VJ, Horswill AR, Homick KM, Escalante-Semerena JC. The 1.75 Å crystal structure of acetyl-CoA synthetase bound to adenosine-5'-propylphosphate and coenzyme A. *Biochemistry.* 2003; 42:2866–2873. [PubMed: 12627952]
 - 23. Branchini BR, Magyar RA, Murtiashaw MH, Anderson SM, Helgerson LC, Zimmer M. Site-directed mutagenesis of firefly luciferase active site amino acids: a proposed model for bioluminescence color. *Biochemistry.* 1999; 38:13223–13230. [PubMed: 10529195]
 - 24. Fersht AR, Knill-Jones JW, Bedouelle H, Winter G. Reconstruction by site-directed mutagenesis of the transition state for the activation of tyrosine by the tyrosyl-tRNA synthetase: a mobile loop envelopes the transition state in an induced-fit mechanism. *Biochemistry.* 1988; 27:1581–1587. [PubMed: 3284584]
 - 25. First EA, Fersht AR. Mutation of lysine 233 to alanine introduces positive cooperativity into tyrosyl-tRNA synthetase. *Biochemistry.* 1993; 32:13651–13657. [PubMed: 8257698]
 - 26. Kuratani M, Sakai H, Takahashi M, Yanagisawa T, Kobayashi T, Murayama K, Chen L, Liu ZJ, Wang BC, Kuroishi C, et al. Crystal structures of tyrosyl-tRNA synthetases from Archaea. *J Mol Biol.* 2006; 355:395–408. [PubMed: 16325203]
 - 27. Lin CS, Irwin R, Chirikjian JG. Kinetic studies of leucyl transfer RNA synthetase from baker's yeast. Order of addition of substrates and release of products. *J Biol Chem.* 1975; 250:9299–9303. [PubMed: 172506]
 - 28. Meier JL, Burkart MD. The chemical biology of modular biosynthetic enzymes. *Chem Soc Rev.* 2009; 38:2012–2045. [PubMed: 19551180]

29. Fisch KM, Gillaspay AF, Gipson M, Henrikson JC, Hoover AR, Jackson L, Najar FZ, Wagele H, Cichewicz RH. Chemical induction of silent biosynthetic pathway transcription in *Aspergillus niger*. *J Ind Microbiol Biotechnol*. 2009
30. Watanabe K, Hotta K, Praseuth AP, Koketsu K, Migita A, Boddy CN, Wang CC, Oguri H, Oikawa H. Total biosynthesis of antitumor nonribosomal peptides in *Escherichia coli*. *Nat Chem Biol*. 2006; 2:423–428. [PubMed: 16799553]
31. Challis GL. Mining microbial genomes for new natural products and biosynthetic pathways. *Microbiology*. 2008; 154:1555–1569. [PubMed: 18524911]

Figure 2.

Members of the different classes of adenylate-forming enzymes, N- and C-terminus are marked with blue and red spheres, respectively. The N-terminal and C-terminal domain are colored in green and yellow, respectively **a class I:** non-ribosomal independent Gramicidin synthetase N-terminal adenylation domain in complex with AMP and phenylalanine (PDB code: 1amu) AMP white carbons, phenylalanine grey carbons; **b class I:** crystal structure of a predicated acyl-CoA synthetase from E.coli (3dmy); **c class II:** LeuRS form T. thermophilus (1obc) a subclass IIa aminoacyl tRNA synthetase; **d class II:** monomer of human GlyRS (2pme) a subclass IIb aminoacyl tRNA synthetase; **e class III** Achromobactin synthetase protein D (AcsD) in complex with ATP, monomer (PDB code: 2w02)

Figure 3.

Active site alignment of different classes of adenylate-forming enzymes. Alignment was carried out on α -phosphate of ATP or AMP. **a** Alignment of class I enzymes of Acetyl-CoA synthetase (2p2f), gramicidin S synthetase 1 (1amu), Japanese firefly luciferase (2d1q) and 4-Chlorobenzoyl-CoA ligase (3cw8). **b** Alignment of class II enzymes tryptophanyl-tRNA synthetase (1m83), human tryptophanyl-tRNA synthetase (2qui) and lysyl-tRNA synthetase (1e24), alignment of α -phosphate of ATP is shown, ATP in grey line representation is ATP of 2qui. **c** Alignment of class III enzyme: AcsD colored in blue (2w02) and AlcC colored in green, Mg ion in AcsD and AlcC show an α - γ phosphate coordination, while AlcC coordinates a second Mg ion in the α - β position (green). In contrast to class I and II water in class III enzymes is close to the α phosphate.

Figure 4.

Coordination of α -phosphate by positively charged residues in different classes of adenylate-forming enzymes. In each member at least one positively charged residue, as lysine, arginine or histidine is present. Acetyl-CoA synthetase (2p2f) and Japanese firefly luciferase (2d1q) are class I, human tryptophanyl-tRNA synthetase (2qui) and lysyl-tRNA synthetase (1e24) are class II, AcsD (2wo2) and AlcC (alcc) are class III adenylate-forming enzymes. ATP colored in grey and shown in line representation is ATP of AcsD.