Journal of Molecular Catalysis B: Enzymatic 59 (2009) 140–144

Contents lists available at ScienceDirect

Journal of Molecular Catalysis B: Enzymatic

journal homepage: www.elsevier.com/locate/molcatb

Substrate flexibility of the adenylation reaction in the Tyrocidine non-ribosomal peptide synthetase

Michelle L. Schaffer*, Linda G. Otten**

Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, UK

ARTICLE INFO

Article history:
Received 21 August 2008
Received in revised form 20 January 2009
Accepted 3 February 2009
Available online 13 February 2009

Keywords:
Adenylation
Enzymatic promiscuity
Non-ribosomal peptide synthetase (NRPS)
Combinatorial biosynthesis
ATP-PP_i exchange assay
Substrate specificity

ABSTRACT

The initiation module of the biosynthetic pathway of Tyrocidine A, a non-ribosomal peptide antibiotic, selectively binds L-phenylalanine along with ATP and Mg^{2+} catalyzing the formation of an aminoacyl adenylate. The adenylation (activating) reactions of non-ribosomal peptide synthetases are considered the first selectivity filter for correct processing of natural products. Although each adenylation domain is selective for a specific substrate, many of these domains have the ability to adenylate a number of substrates. The overall tolerance of the active site of these domains has not been extensively characterized in terms of substrate kinetics. Exploiting a recently developed ATP-PP_i exchange assay protocol, we were able to screen and fully characterize a panel of amino acids and non-proteinogenic substrates to show the broad specificity of the first L-phenylalanine activating domain of Tyrocidine TycA A1. The results revealed unexpected flexibility in the substrate interactions thought necessary for activation. These results could be universal for adenylation domains and can be exploited for the biosynthesis of new and interesting "natural" products.

1. Introduction

Non-ribosomal peptide synthetases (NRPSs) are multi-domain enzymes that utilize a modular organization to coordinate stepwise catalytic reactions which generate structurally complex and biologically interesting natural products [1–3]. The NRPS machinery is analogous to other mega-enzymes such as polyketide (PKS) [4] and fatty acid synthases (FAS) [3,5]. They follow a similar assembly-line rationale, except that the substrates of the NRPS are in principle mostly amino acids or non-proteinogenic amino acid derivatives, whereas the PKS and FAS systems use acyl-coenzyme A (CoA) building blocks for assembly [4,5]. In each case, after activation of the substrate the modular construction utilizes a thioester to covalently link the substrate to the assembly line for the iterative generation of fatty acids, polyketides or peptides [4–6]. The first catalytic step in non-ribosomal peptide biosynthesis is the recognition and activation of a monomer by the adenylation (A) domain (Fig. 1a) where an aminoacyl adenylate is formed and then transferred to the thiol of a 4′-phosphopantetheine (Ppant) arm of the corresponding peptidyl carrier protein (PCP) domain [1,7] (Fig. 1a). This thioester loading reaction directs the peptide bond formation between amino acids tethered on consecutive PCP domains by the subsequent condensation domain [2,4,8]. The adenylation reaction is considered the first selectivity filter for the activation of the correct monomer building block. Each adenylation domain contains a "specificity-conferring code" for its selected substrate [9–11]. The next selectivity hurdle is recognition of adjoining covalently attached substrates by the condensation domain for the biosynthesis of the correct linear products [4,6,8]. A prototypical NRPS initiation module will contain at least an adenylation domain, but usually also a PCP domain (Fig. 1a).

Peptide-based natural products are quite complex and difficult to obtain or modify through ordinary synthetic means, but with re-engineered or evolved NRPSs even more diverse groups of non-natural compounds could be obtained [3,6]. Expansion of the scope of such systems of combinatorial biosynthesis by directed evolution would ultimately result in powerful systems that can make novel "natural" products or easily produced structural analogs of present natural products. Several groups have attempted to harness this power for combinatorial biosynthesis through rational design, directed evolution, and module swapping [12–16]. These attempts have met varying degrees of success, but most revealed that there are still selectivity obstacles to be overcome.

The Tyrocidine A biosynthetic pathway has been studied as a model NRPS system for decades [17,18]. It was one of the first fully sequenced NRPS [19] and was used for one of the first cell-free syntheses of a number of Tyrocidine derivatives (A–E). Those

* Corresponding author. Present address: Physiology and Pharmacology, Oregon Health & Science University, 3181 SW Sam Jackson Park Road, L334, Portland, OR 97239, USA. Fax: +1 5034944352.
** Corresponding author. Present address: Biocatalysis & Organic Chemistry, Delft University of Technology, Julianalaan 136, 2628 BL Delft, The Netherlands. Fax: +31 15 2781415.
E-mail addresses: schaffem@ohsu.edu (M.L. Schaffer), l.g.otten@tudelft.nl (L.G. Otten).

1381-1177/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.molcatb.2009.02.004

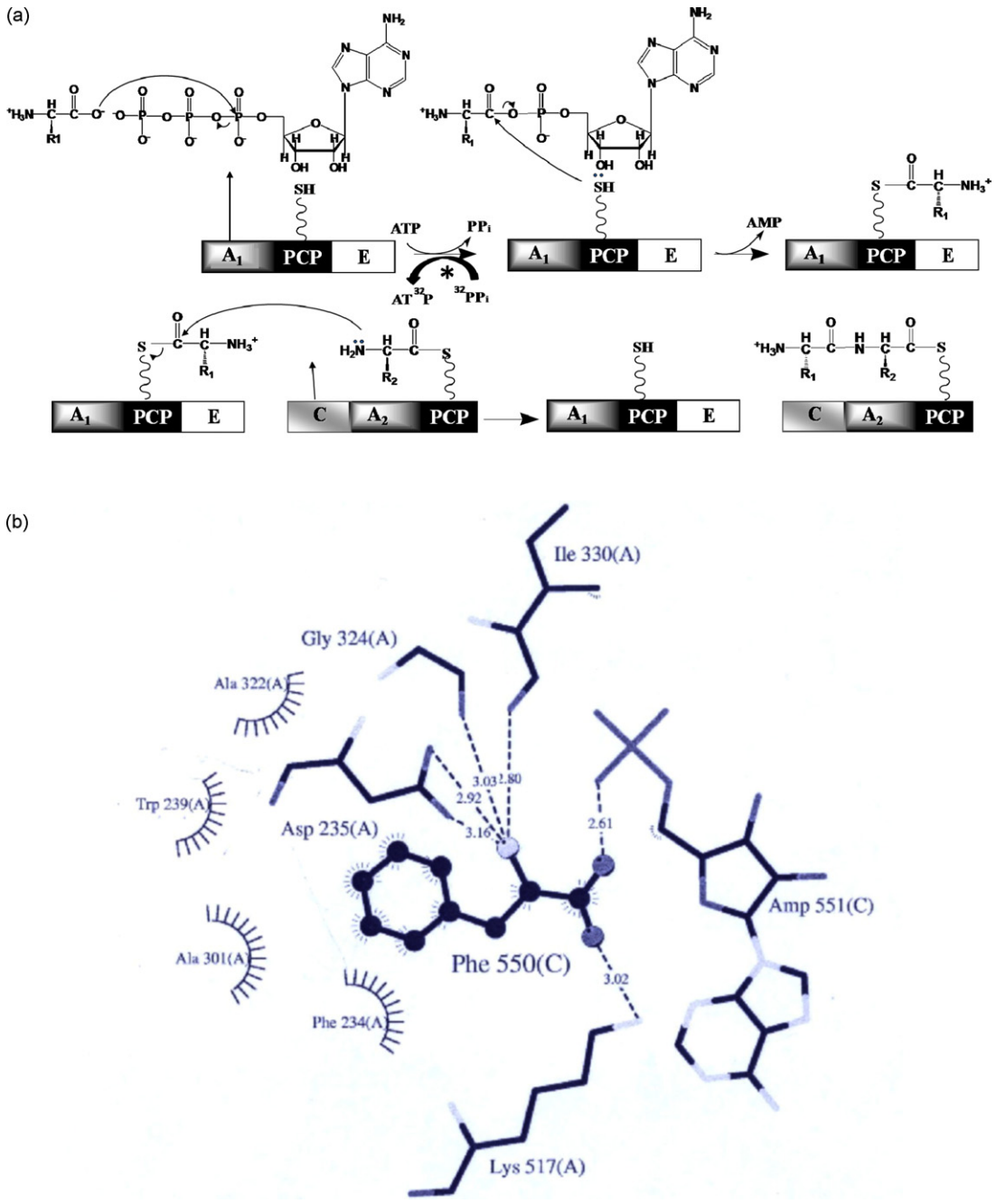

Fig. 1. (a) Core NRPS modules and reactions. The core domains are the condensation (C), adenylation (A), and the peptidyl carrier protein (PCP) with the Ppant arm (squiggly line). The reaction mechanism is shown by the curvy arrows. The R1 and R2 represent the amino acid side chain of the first and second adenylation reaction. The extra domain is an epimerization domain, one of the extra catalytic domains that perform tailoring reactions. *The reverse reaction forces the formation of radioactive ATP by addition of excess $^{32}PP_i$. This radioactive ATP is measured in the ATP-PP_i exchange assay. (b) Structural representation of the active site of GrsA-A1 showing residues within 5 Å of the bound L-Phe substrate. The illustration was generated with LIGPLOT [33].

experiments revealed that depletion of necessary substrates will stunt Tyrocidine production, but will not stop the NRPSs from incorporating replacement substrates to make a full length derivative [20,21]. The structures of a number of amino acid activating A domains have been solved, e.g., the initiating A domain of pore forming antibiotic Gramcidin A, GrsA A1, a L-Phe activating domain [22], the termination module Surfactin A–C, a L-Leu activating domain [23], and the initiation module DhbE, a dihydroxy benzoic acid activating domain [24]. The TycA A1 is similar to GrsA A1 (65% overall identity), whose structure was solved in complex with substrate L-Phe and adenosine mono-phosphate (AMP) revealing a binding pocket for L-Phe [22]. The main interactions between the enzyme and substrate are derived from interactions of the amino and carboxyl ends of the L-Phe (Fig. 1b) with residues Lys-517 and Asp-235 (GrsA numbering), respectively [22]. The phenyl side chain is bound by the selectivity-conferring residues specific for L-Phe (Fig. 1b) [9,11], which form strong hydrophobic interactions with the aromatic side chain. As altering the substrate specificity of a full NRPS by directed evolution or redesign is attempted, downstream editing mechanisms from other domains may hinder the progress of the new compounds. For example, for TycA A1 its epimerization (E) domain (Fig. 1a) is thought to play an important role in the correct formation of Tyrocidine [25]. However, results from TLC assays with GrsA [26] demonstrated that the epimerization domain

will convert noncognate L-amino acids to their corresponding D-configuration. Furthermore, HPLC assays with TycA [27] showed the correct formation of dipeptides with the non-native substrate nitro-tyrosine, in the place of L-Phe, in the TycA pathway. However, both showed reduced activity as compared to the conversion and peptide bond formation of L-Phe. This illustrates that while the E- and C-domain are strong gatekeepers the pathway is not completely halted when presented with a non-native substrate.

The challenge of engineering these enzymatic assembly lines towards expanding the production of novel peptides with increasing structural diversity relies on further elucidating the substrate specificity for each reaction in NRPSs. For this study we started at the assembly line initiation by seeking to determine the flexibility of the main structural interactions of the adenylation reaction; firstly to explore the side chain interactions, and secondly to explore the flexibility of the main amino and carboxyl interactions. Previously panels of substrates have been shown in the literature to be activated by the TycA adenylation domain in single time point analysis [8,28], but few have been fully characterized (K_m and k_{cat} calculations). An ATP-PP_i exchange screening assay [29] was used to determine the full range of activities from a diverse panel of substrates. Following the initial screen we were able to provide detailed kinetic information on a select group of substrates. The results indicate a previously unknown flexibility in the adenylation reaction and reveal future possibilities for the incorporation of novel substrates into "natural" products.

2. Experimental

2.1. Materials

All amino acids were purchased at the commonly known chemical distributors with greater than 99% purity (ACS Grade).

2.2. Enzyme purification and ATP-PP_i exchange assay

The purification of the TycA A1 protein was performed as described previously [29]. For the activity tests we used Procedure B as described in Otten et al. [29]. An amino acid concentration of 2 mM was used for the screening assay. For determination of the kinetic parameters, the amino acid concentrations were varied (0.005–6 mM), the ATP and [^{32}P]PP_i concentrations were held constant at 2 mM and 0.0007 μCi/μL, respectively, and the TycA protein concentration was either 0.1 or 0.2 μM. The time scales of the experiments ranged from 1 to 60 min. The Michaelis–Menten parameters (K_m and k_{cat}) were calculated from the initial velocity time curves for each substrate concentration by non-linear least square fit to the Michaelis–Menten equation (Kaleidagraph, Synergy Software, Reading, PA) as described previously [29].

3. Results and discussion

3.1. Screening a diverse panel of compounds for TycA activity in the ATP-PP_i exchange assay

In this study 25 substrates were tested representing a panel of 15 L-Phe derivatives (structurally similar compounds including β-, homo-, amino- and hydroxyl- L-Phe derivatives) and nine completely unrelated amino acid-like derivatives (Fig. 2a). In a 96-well plate format these 25 substrates were screened at a single concentration and a single time point with the TycA enzyme in the ATP-PP_i exchange assay (mechanism in Fig. 1a). The time point was taken at 30 min to ensure that reactions could be detected for even the lowest-activity substrates.

The results of the single point screening assay are reported (Fig. 2b) as a percentage of the total activity of the native substrate L-Phe (**1** 100%). While the substrates showed varied activities relative to the cognate substrate activity, a number of compounds were more active than predicted by the strongest interactions seen in the crystal structure of GrsA. A feature required to preserve activity (>30%) is the presence of an aromatic group. In addition the relative positioning of the aromatic group to the amino group is important. The aromatic substrates without a primary amine (**24**) or with a β-shifted amino group (**17**) as compared to L-Phe showed very little activity. This demonstrates the catalytic importance of the positioning of the amino and carboxylic groups of L-Phe. Compounds capped at the amino-group (**4**, **5**, **6**, **11**, **12**) maintained sufficient activity as compared to the cognate substrate showing that a doubly substituted nitrogen group is still capable of making the necessary interactions with the enzyme, but with variable loss of activity. The most structurally analogous compound **3** gave very robust signals. The experiments demonstrate both the utility of the 96-well format of the assay and the flexibility of the TycA enzyme.

Table 1
Kinetic values for TycA substrates. The Michaelis–Menten parameters were calculated from at least 5 different substrate concentrations and 6 time points. Substrates are listed according to their catalytic efficiency (k_{cat}/K_m).

Substrate	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (M^{-1} s^{-1})
L-Phe **1**	3.6 ± 0.1	0.015 ± 0.002	240,000
D-Phe **2**	3.6 ± 0.1	0.040 ± 0.003	90,000
ph-Ser **3**	2.5 ± 0.2	0.18 ± 0.04	13,812
Boc-Phe **6**	3.0 ± 0.6	0.30 ± 0.04	10,000
N-formyl-Phe **4**	1.9 ± 0.1	0.88 ± 0.09	2,159
L-Tyr **7**	0.5 ± 0.1	0.62 ± 0.28	833
Asp-Phe **5**	1.4 ± 0.3	1.78 ± 0.23	787
L-homo-Phe **9**	0.4 ± 0.05	0.61 ± 0.28	655
L-Trp **10**	0.5 ± 0.003	1.72 ± 0.24	290
L-Met **13**	0.6 ± 0.1	2.30 ± 0.52	261
Nitro-Tyr **8**	0.27 ± 0.03	3.0 ± 0.7	90

3.2. Kinetic analysis of the more active TycA substrates

From the results of the initial screen, we were able to determine the most active substrates. However, in this single point format only the relative activities could be ascertained; structural subtleties can only be further elucidated by determination of the full kinetic parameters. Substrates with greater than 20% activity of the cognate substrate in the initial assay (i.e. **1**–**13**) were assayed in a discontinuous time course utilizing the ATP-PP_i-exchange assay as described previously [29]. For our experiments only concentrations of the substrates of interest were varied. ATP was held at a constant concentration above saturation, in order to compare directly the differences between the novel substrates and L-Phe. The ATP concentration had been varied in the past for different substrates and this did not change the assessment of specificity [30]. The Michaelis–Menten (MM) values determined for each substrate are listed in Table 1. The results revealed that the assay is very robust and the activity of an adenylation domain can be characterized quickly and easily. The second order rate constants span over 3 orders of magnitude, an observation which establishes the sensitivity of the assay and the range of activities from the various substrates. This is a valuable assay for detecting the promiscuous activity towards a range of substrates and for obtaining kinetic information about different structural aspects of the interactions between multiple substrates and the enzyme.

3.3. Relationship of side chain interactions to adenylation activity

The most closely related substrates to L-Phe lose catalytic efficiency due largely to the reduced apparent affinity (higher apparent K_m values) with only small reductions in k_{cat}. As expected

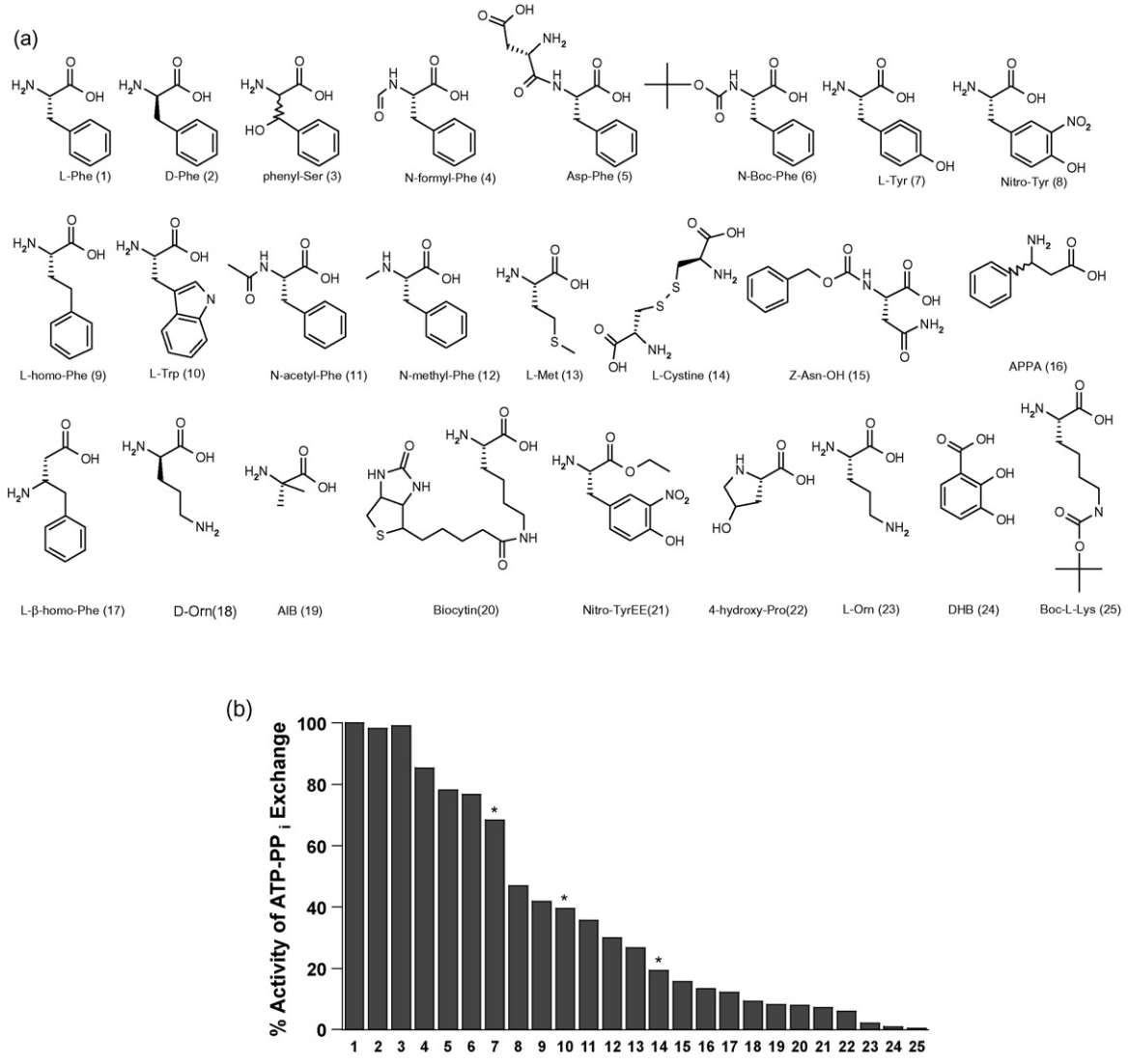

Fig. 2. (a) Amino acids and amino acid-like substrates. Chemical structures of the substrates tested for activation by TycA in the ATP-PP_i-exchange assay. L-Phenylalanine **1**, D-phenylalanine **2**, 3-phenyl-DL-serine **3**, *N*-formyl-L-phenylalanine **4**, L-aspartyl-L-phenylalanine **5**, *N*-boc-L-phenylalanine **6**, L-tyrosine **7**, 3-nitro-L-tyrosine **8**, L-homo-phenylalanine **9**, L-tryptophan **10**, *N*-acetyl-L-phenylalanine **11**, *N*-methyl-L-phenylalanine **12**, L-methionine **13**, L-cystine **14**, Z-L-asparagine **15**, DL-3-amino-3-phenyl-propanoic acid **16**, L-β-homo-phenylalanine **17**, D-ornithine **18**, 2-amino-isobutyric acid **19**, Nξ-(+)-biotinyl-L-lysine **20**, 3-nitro-L-tyrosine ethyl ester **21**, 4-hydroxy-L-proline **22**, L-ornithine **23**, 2,3-dihydroxy-benzoic acid **24**, boc-L-lysine **25**. (b) Single time point screen of active amino acids and other non-proteinogenic amino acids and their derivatives. *Three best natural amino acid substrates described in previous work [29].

the substrates with modifications in the aromatic region (**7**–**10**, **13**) showed the greatest loss in catalytic efficiency (Table 1, k_{cat}/K_m values). Interestingly, of the two amino acids that show the highest noncognate substrate activity in the very closely related L-Phe activating domain in GrsA (L-Tyr and L-Trp), greater catalytic activity towards L-Trp was seen in GrsA [25], while Tyc A preferentially activates L-Tyr. This difference shows that recognition and selectivity within the adenylation domain extend beyond the selectivity-conferring amino acids that have been described previously. In each adenylation domain there is a set of 8 amino acids that can be used to predict the specific substrate for that domain [9,12]. Our results show that the second-tier activity of the adenylation domain cannot be precisely predicted by the selectivity-conferring code of an A domain.

3.4. Exploring the kinetics of substrates containing modified amino or carboxyl groups

To determine the extent of the flexibility of the substrate binding pocket of the adenylation domain, we assayed substrates with the carboxyl group modified with an ethyl ester (**21**), or the amino group modified with an amide bond (**4**–**6**, **11**) or methyl group (**12**). A fascinating range of catalytic activities was observed.

The full kinetic characterization of the amino-substituted compounds **4** (*N*-formyl-Phe), **5** (Asp-Phe), and **6** (Boc-Phe) revealed a loss of catalytic efficiency as compared to **1** (100-fold, 300-fold, and 20-fold, respectively). This loss of activity is potentially due to the additional strain put on the formed aminoacyl-adenylate to keep an active conformation and remain bound to TycA. However, for certain phenylalanine derivatives that were modified at the amino-terminus, i.e. **12** (*N*-methyl-Phe) and **11** (*N*-acetyl-Phe), the kinetic parameters could not be accurately determined. This was due to the high concentrations of substrate needed to reach saturation. When comparing the results of compounds **5** and **6** with **11** and **12** it appears that TycA prefers a large group adjacent to the amide bond, i.e. Boc and L-Asp, rather than the smaller acetyl derivative or no amide bond at all. This is the first study in which the plasticity of binding pocket interactions has been extensively explored in TycA, and the substrates tested showed a much greater range of catalytic activity than expected.

4. Conclusion

4.1. Implications for combinatorial biosynthesis

Previously, close derivatives of L-Phe, i.e. phenyl ring modifications, were tested for adenylation activity in TycA A1, showing lower activity as compared to the cognate substrate, but the full kinetic parameters were not reported [31]. As the substrates of NRPSs are not restricted to the standard amino acids, non-proteinogenic amino acids, including D-, ω-, β- and α-hydroxy amino acids are used in the assembly of a number of natural products [6]. However, with the exception of hydro-cinnamic acid [32], compounds either derivatized at or lacking the amino group have not been screened for the adenylation reaction in NRPS modules.

This is the first instance that amino modified substrates, or even more significantly, a dipeptide (**5**), showed activity in an adenylation reaction for TycA A1 or any non-aryl adenylation domain. Modifications to the side chain affect both the apparent affinity and the turnover rate, indicating a degree of selectivity within the binding pocket. However, the activity retained by derivatized amino groups reveals the potential for the addition of numerous compounds to at least the first step of the biosynthetic reaction of NRPSs. Further obstructions down the assembly line may become apparent, but the newly discovered flexibility with the initial reaction could prove very useful. With the recent report of the crystal structure of the gatekeeping C-domain [23], its functionality may be more easily manipulated by mutagenesis to further elucidate the residues involved in its function. Numerous promising methods are being developed for the directed evolution and genetic engineering of NRPSs [13–15]. Manipulating and modulating the NPRS adenylation domain is the first step.

Presently, we have laid the foundation towards defining the critical factors of substrate binding in TycA. This A-domain is selective in terms of binding the native substrate, L-Phe. However, at higher concentrations of non-native substrates promiscuous activities can be detected and fully characterized. Taking these structural determinants and combining them with powerful genetic engineering techniques, we can make strides towards the ultimate goal of generating new and promising “unnatural natural” products.

Acknowledgements

We would like to thank Dr. F. Hollfelder of Cambridge University for the opportunity to carry out experiments in his lab. We would like to thank Prof. Jeffrey Karpen of Oregon Health & Science University for critical review of the manuscript. MLS received a scholarship from King’s College and the Sperling Foundation. LGO was supported by a Marie Curie Intra European Fellowship (MEIF-CT-2004-022448).

References

[1] H. von Dohren, U. Keller, J. Vater, R. Zocher, Chem. Rev. 97 (1997) 2675–2706.
[2] H.D. Mootz, D. Schwarzer, M.A. Marahiel, ChemBioChem 3 (2002) 490–504.
[3] S.A. Sieber, M.A. Marahiel, Chem. Rev. 105 (2005) 715–738.
[4] K.J. Weissman, R. Muller, ChemBioChem 9 (2008) 826–848.
[5] C.T. Walsh, Science 303 (2004) 1805–1810.
[6] D.A. Vosburg, C.T. Walsh, Ernst Schering Res. Found. Workshop (2005) 261–284.
[7] T. Stachelhaus, A. Huser, M.A. Marahiel, Chem. Biol. 3 (1996) 913–921.
[8] T. Stachelhaus, H.D. Mootz, V. Bergendahl, M.A. Marahiel, J. Biol. Chem. 273 (1998) 22773–22781.
[9] T. Stachelhaus, H.D. Mootz, M.A. Marahiel, Chem. Biol. 6 (1999) 493–505.
[10] H. von Dohren, R. Dieckmann, M. Pavela-Vrancic, Chem. Biol. 6 (1999) R273–R279.
[11] G.L. Challis, J. Ravel, C.A. Townsend, Chem. Biol. 7 (2000) 211–224.
[12] H.D. Mootz, D. Schwarzer, M.A. Marahiel, Proc. Natl. Acad. Sci. U.S.A. 97 (2000) 5848–5853.
[13] M.A. Fischbach, J.R. Lai, E.D. Roche, C.T. Walsh, D.R. Liu, Proc. Natl. Acad. Sci. U.S.A. 104 (2007) 11951–11956.
[14] D. Butz, T. Schmiederer, B. Hadatsch, W. Wohlleben, T. Weber, R.D. Sussmuth, ChemBioChem 9 (2008) 1195–1200.
[15] R.H. Baltz, Curr. Top. Med. Chem. 8 (2008) 618–638.
[16] K. Eppelmann, T. Stachelhaus, M.A. Marahiel, Biochemistry 41 (2002) 9718–9726.
[17] S.G. Lee, F. Lipmann, Methods Enzymol. 43 (1975) 585–602.
[18] M.A. Marahiel, M. Krause, H.J. Skarpeid, Mol. Gen. Genet. 201 (1985) 231–236.
[19] H.D. Mootz, M.A. Marahiel, J. Bacteriol. 179 (1997) 6843–6850.
[20] M.A. Ruttenberg, B. Mach, Biochemistry 5 (1966) 2864–2869.
[21] K. Fujikawa, Y. Sakamoto, T.J. Suzuki, K. Kurahashi, Biochim. Biophys. Acta (BBA)-Nucleic Acids Protein Synth. 169 (1968) 520–533.
[22] E. Conti, T. Stachelhaus, M.A. Marahiel, P. Brick, EMBO J. 16 (1997) 4174–4183.
[23] A. Tanovic, S.A. Samel, L.O. Essen, M.A. Marahiel, Science 321 (2008) 659–663.
[24] J.J. May, N. Kessler, M.A. Marahiel, M.T. Stubbs, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 12120–12125.
[25] L. Luo, C.T. Walsh, Biochemistry 40 (2001) 5329–5337.
[26] L. Luo, M.D. Burkart, T. Stachelhaus, C.T. Walsh, J. Am. Chem. Soc. 123 (2001) 11208–11218.
[27] S. Gruenewald, In vitro- und in vivo-Untersuchungen zur Evolution nichtribosomaler Peptidsynthetasen, Ph.D. Thesis, Philipps-Universität Marburg, Marburg DE, 2005.
[28] E. Pfeifer, M. Pavela-Vrancic, H. von Dohren, H. Kleinkauf, Biochemistry 34 (1995) 7450–7459.
[29] L.G. Otten, M.L. Schaffer, B.R.M. Villiers, T. Stachelhaus, F. Hollfelder, Biotechnol. J. 2 (2007) 232–240.
[30] H.J. Skarpeid, T.L. Zimmer, B. Shen, H. von Dohren, Eur. J. Biochem. 187 (1990) 627–633.
[31] R. Dieckmann, Y.O. Lee, H. van Liempt, H. von Dohren, H. Kleinkauf, FEBS Lett. 357 (1995) 212–216.
[32] J. Vater, H. Kleinkauf, Biochim. Biophys. Acta 429 (1976) 1062–1072.
[33] A.C. Wallace, R.A. Laskowski, J.M. Thornton, Protein Eng. 8 (1995) 127–134.