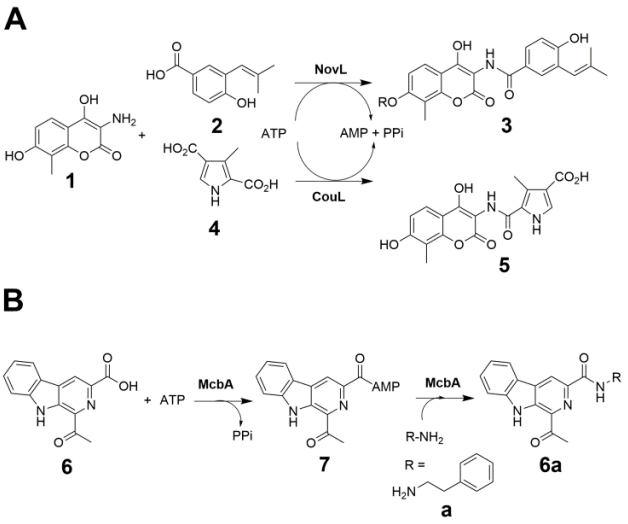
# Aryl acid specificity of the amide bond synthetase McbA suggests broad potential for the biocatalytic synthesis of amides

*Mark Petchey,[a] Anibal Cuetos,[a] Benjamin Rowlinson,[a] Stephanie Dannevald,[a] Amina Frese,[a] Peter W. Sutton,[b]† Sarah Lovelock,[b]‡ Richard C. Lloyd,[b] Ian J.S. Fairlamb[a]* and Gideon Grogan[a]**

Abstract: Amide bond formation is one of the most important reactions in pharmaceutical synthetic chemistry. The development of sustainable methods of amide bond formation, including those that are catalyzed by enzymes, is therefore of significant interest. The ATP-dependent amide bond synthetase (ABS) enzyme McbA, from *Marinactinospora thermotolerans*, catalyzes the formation of amides as part of the biosynthetic pathway towards the marinacarboline secondary metabolites. The reaction proceeds via an adenylate intermediate, with both adenylation and amidation steps catalyzed within one active site. In this study, McbA has been applied to the synthesis of pharmaceutical-type amides from a range of aryl carboxylic acids with partner amines provided at 1-5 molar equivalents. The structure of McbA has revealed the structural determinants of aryl acid substrate tolerance and differences in conformation associated with the two half-reactions catalyzed. The catalytic performance of McbA, coupled with the structure, suggest that this and other ABS enzymes may be engineered for applications in the sustainable synthesis of pharmaceutically relevant (chiral) amides.

The formation of the amide bond is one of the most important reactions in pharmaceutical synthetic chemistry, accounting for up to 16% of all synthetic steps in medicinal chemistry laboratories.[1] Amide bond synthesis typically involves the activation of the carboxylic acid, using either toxic chlorinating agents, or atom-inefficient coupling reagents,[2] and so there is an urgent need to investigate methods of amide bond formation that are sustainable in terms of reagent safety and atom efficiency. Enzymatic methods of amide bond synthesis that display these advantages are appealing therefore.[3,4] The amide bond is encountered in secondary metabolites that are synthesised by non-ribosomal peptide synthases (NRPSs).[5] In this case, the formation of the amide is complex, requiring ATP-dependent formation of an intermediate adenylate of the carboxylic acid substrate, followed by transfer to a phosphopantetheine thiol enabled by an acyl/peptidyl carrier protein (ACP or PCP), followed by condensation with the amine substrate, each process taking place in a separate domain of a large, multi-domain NRPS enzyme. While NRPS-mediated amide bond formation may have a role in synthetic biology pathways towards amides,[6] their structural and catalytic complexity militates against their application in preparative *in vitro* biocatalysis. Hydrolases, such as *N*-acylases[7] and lipases,[8] have been employed for amide bond formation, and are attractive in terms of their simplicity and efficiency, but often require ester substrates, and suffer from poor substrate scope. More recently, Flitsch and co-workers have recruited the adenylate-forming ability of a carboxylate reductase (CAR) adenylation-domain-plus-peptidyl carrier protein (CAR-A-PCP) to the formation of amides, although a 100-fold excess of amine was required to drive reactions to high conversions.[9]

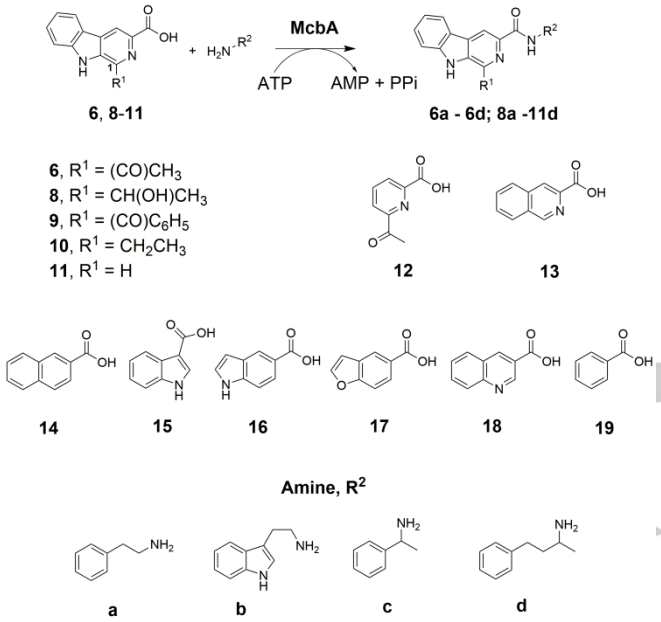
ATP-dependent amide bond synthetases (ABSs) have been discovered in secondary metabolite pathways leading to compounds such as novobiocin,[10] cloromycin[11] and coumermycin.[12] The enzymes NovL and CouL, for example, couple the aminoflavone **1** with carboxylic acids **2** and **4** in their respective biosynthetic pathways to give amide products **3** and **5** (Scheme 1A).




**Scheme 1. A**: Amide bond synthesis by ATP-dependent amide bond synthetases (ABSs) NovL and CouL involved in secondary metabolism in actinomycetes; **B**: Formation of marinacarbolines such as **6a** from β-carboline acid **6** and amine **a** catalyzed by McbA from *Marinactinospora thermotolerans*.

These enzymes are apparently capable of catalyzing formation of the adenylate, but also of binding the amine for the coupling reaction within the same active site, with no involvement of separate acyl carrier protein or condensation domains. A further enzyme, McbA, has recently been reported in the biosynthetic pathway towards the marinacarbolines in *Marinactinospora thermotolerans* by Ji and co-workers.[13,14] McbA catalyzes the ATP-dependent coupling of the β-carboline derivative **6** with 2-phenylethylamine **a** to form **6a** *via* adenylate **7** (Scheme 1B), but, unlike other ABSs, McbA also accepted other amines, including substituted 2-phenylethylamines and tryptamines, indicating some promise for the application of this enzyme to reactions with a wider substrate spectrum.[14] These studies prompted us to further examine the activity of McbA to include a range of β-carboline, and other aryl carboxylic acid acceptors. We have also determined the structure of McbA in complex with AMP and **6**, which has revealed the substrate binding interactions within the active site, and may serve as a platform for protein engineering to expand the substrate specificity of this and related ABS enzymes.

McbA was expressed from a synthetic gene and the protein purified using IMAC followed by gel filtration (Figure S1, S2). A sample of the pure enzyme (1 mg $mL^{-1}$; 9.4 nmol) was applied to the coupling of 0.4 mM acid **6** and 0.6 mM 2-phenylethylamine **a**, in the presence of 2 mM ATP on an analytical scale, and full conversion to product amide **6a** was observed within 1 h using HPLC. Standard amide product was prepared by reaction of the β-carboline ester with the amine in DMSO at 90°C (Supporting Information). We then further explored the constraints on substrate structure in the McbA-catalyzed reaction, but, in contrast to previous work,[14] we focused on the structure of the carboxylic acid acceptor (Scheme 2).




**Scheme 2.** β-carboline carboxylic acid and amine partners for McbA-catalyzed amide bond formations.

Carboxylic acid substrates **6** and **8-11** were synthesized using Pictet-Spengler condensations of tryptophan-derived esters with relevant aldehydes, followed by hydrolysis of the ester products, as detailed in the Supporting Information. Substitution of the acyl group in position 1 of the pyridine ring with a secondary hydroxymethyl (**8**) reduced the conversion to 21% of product **8a** after 1 h, but 96% conversion was achieved after 24 h. An increase in the size of the substituent to benzoyl **9** gave conversions to product **9a** of 8% and 82% after 1 h and 24 h respectively. Replacement by an ethyl group **10** showed that an oxygen atom on this substituent was not necessary, with 100% conversion to product **10a** again achieved after 24 h. Even replacement with H (**11**) resulted in 6% conversion to **11a** after 24 h. The activity of the enzyme was then tested using carboxylic acids **6** and **8**-**11** with amines **a**-**d** (Scheme 2, Table 1).

**Table 1.** ATP-dependent amide couplings by McbA. Values represent % conversions to amide products after 24 h as determined by HPLC analysis

| | Amine | | | |
|---|---|---|---|---|
| **Acid** | **a** | **b** | **c** | **d** |
| **6** | 100 | 81 | 6 | 26 |
| **8** | 96 | 100 | 0 | 17 |
| **9** | 82 | 0 | 0 | 0 |
| **10** | 100 | 91 | 0 | 0 |
| **11** | 6 | 0 | 0 | 0 |

McbA catalysed the formation of **6a** and **6b** with 100% and 81% conversion respectively, confirming the findings of Ji and co-workers.[14] However, McbA also catalysed the coupling of **6** to the racemic chiral amines **c** and **d** to form **6c** and **6d**, although with lower conversions of 6% and 26%. Interestingly, analysis of the product **6d** by chiral HPLC revealed it to be predominantly of the (*S*)-configuration, and with an e.e. of 96% (Figure S3), indicating that McbA is capable of performing kinetic resolution reactions. Tryptamine **b** was also coupled to **10** to give **10b** in 91% yield. 1-Phenylethylamine **c** proved to be a poor donor overall, but 3-phenyl-1-methylpropylamine **d** was coupled with **8** to form amide **8d** with a conversion of 17%. Some of the β-carboline amides synthesized here have been shown to have biological activity, most notably against benzodiazepine receptors, first reported in the 1980s,[15] and are of interest as treatments against, for example, alcohol abuse.[16]

In order to investigate the broader applicability of McbA to amide bond synthesis, aryl acid substrates **12**-**19**, with structures more diverged from the β-carboline acid platform, were then tested with 2-phenylethylamine **a**, in an effort to produce amide products **12a** to **19a (**Table 2).

**Table 2.** ATP-dependent coupling of **12**-**19** with amine **a** by McbA. Values represent % conversions to amide products after 24 h as determined by HPLC analysis.

| **Acid** | **Product** | **Conversion (%)** |
|---|---|---|
| **12** | **12a** | 34 |
| **13** | **13a** | 23 |
| **14** | **14a** | 43 |
| **15** | **15a** | 40 |
| **16** | **16a** | 41 |
| **17** | **17a** | 39 |
| **18** | **18a** | 4 |
| **19** | **19a** | 24 |

Overall McbA displayed a surprisingly broad substrate range, including substrates such as naphthoic acid **14**, indole carboxylic acids **15** and **16** and benzofuran-5-carboxylic acid **17**, although quinoline substrate **18** was transformed with low conversion. Strikingly, McbA, was also able to couple benzoic acid **19** with **a** with a modest but significant conversion of 24%.

To investigate the promiscuity in aryl acid recognition by McbA, the structure of the enzyme was determined using X-ray crystallography. A combination of a K483A mutant, co-crystallised with AMP, magnesium ions and the β-carboline acid substrate **6**, yielded crystals of diffraction quality. The structure was solved using molecular replacement, and refined to a resolution of 2.80 Å (Data collection and refinement statistics can be found in the Supporting Information, Table S2). McbA adopts a structure within the ANL superfamily of adenylating enzymes defined by Gulick,[17] and which includes firefly luciferase,[18] acetyl-CoA synthase,[19] and standalone adenylation domains of NRPSs, including the dihydroxybenzoate adenylating enzyme enzyme DhbE[20] from the bacillibactin biosynthetic pathway. Each ANL enzyme is characterised by a two-domain structure in which a large N-terminal domain is coupled to a smaller C-terminal one, of 4-500 and 100 amino acid residues respectively. Crucially, the structures of many ANL enzymes have revealed substantial relative domain rotation during the reaction coordinate, with one conformation ('adenylation') assumed for adenylate formation from the carboxylate substrate and ATP, and a second, ('thiolation') in which the C-terminal domain rotates to enable the ligation reaction between the phosphopantetheinate thiol and the adenylate intermediate.

The McbA structure features the two predicted domains: the large N-terminal domain of residues 1-394 ($McbA_N$) and the smaller C-terminal one featuring residues 395-494 ($McbA_C$), with the hinge residue at Q394. The McbA structure is unusual however in that, of five molecules within the asymmetric unit, four (A-D) are in the 'adenylation' conformation (Figure 1a) and one (E) is in the 'thiolation' state described above, which we have termed 'amidation' (Figure 1b) for McbA. The adenylation conformer of McbA superimposes well with other ANL enzymes crystallized in the same state, including PheA (1AMU, 2.1 Å over 431 Cα atoms),[21] DhbE (1MD9, 2.3 Å over 445)[20] and the recently determined structure of the relevant 'A' domain from CAR from *Nocardia iowenis* (*Ni*CAR) in complex with benzoic acid and AMP (5MSD, 2.8 Å over 406).[22]

The substrate **6** is bound within a hydrophobic pocket stacked between the loop between G295 and F301 on one face and L202 on the other (Figure 2). Few specific interactions between active site residues and the substrate are made, except that the carbonyl oxygen and pyridine nitrogen of **6** interact with the backbone N-H and C=O of G295 at distances of approximately 3.9 and 3.4 Å respectively. The relative lack of specific interactions may help to explain the relaxed specificity of McbA for the range of aryl acids used in this study. The exception was quinoline derivative **18**, in which the N-atom in the 1-position has had an adverse effect on conversion, perhaps disrupting the hydrophobic interaction made by the native ligand with F301. Residue D201, which points away from **6** in the active site, is replaced in CARs and DhbE by H300, and is close to the benzoic acid carboxylate in the CAR-AMP complex 5MSD. This histidine has been implicated in the catalysis of both adenylation and thiolation reactions in 4-chlorobenzoyl-CoA ligase (4CBL, 3CW8, 3CW9).[23,24]

The amidation conformer of McbA superimposes well with the thiolation conformer of 4-CBL (3CW9, 2.0 Å over 430 Cα atoms) and also the 'A' domain in the structure of CAR-A-PCP (5MSS, 2.5 Å over 402), the construct used in amide-bond forming reactions described previously.[9] The structure of 4-CBL in the thiolation state, in complex with 4-chlorophenacyl- CoA (4-CPA-CoA, 3CW9, Figure S3a)[23] locates the binding site for the thiolation product, and permits a comparison with the equivalent region of McbA (Figure S3b) The surfaces of the enzymes show that positively charged residues in 4-CBL, such as R87, which interact with the phosphates in 4-CPA-CoA, are absent in McbA; indeed the entrance to the active site features three residues with carboxylate side chains, D463, E221 and E400 which may favour interaction with incoming amine substrates.

The McbA-catalyzed synthesis of selected amides was scaled up to utilise 50 mg (0.20 mmol) of aryl acids **6**, **10**, **14**, **16**, **17** and **19** in a 50 mL reaction volume, with 1-5 molar equivalents 2-phenylethylamine **a** and two equivalents of ATP (198 mg; 0.40 mmol). Amide products **6a**, **10a**, **14a, 17a** and

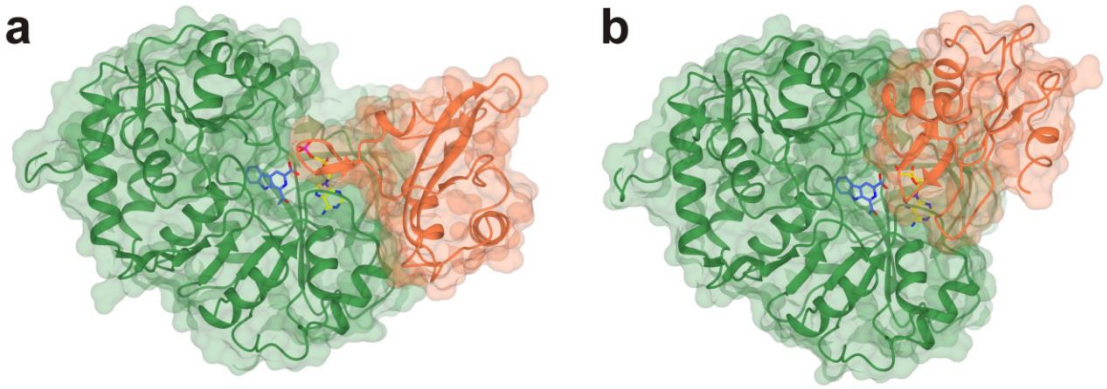



**Figure 1.** Structures of McbA in the adenylation (**a**) and amidation (**b**) conformations. In each case the protein is coloured in green from residues 1-394 ($McbA_N$) and coral from 395-494 ($McbA_C$). AMP (yellow) is bound at the domain interface and **6** (blue) within a binding pocket in the N-terminal domain. The $McbA_C$ domain in **b** is rotated 149° relative to the larger $McbA_N$ domain compared with their orientations in **a**.

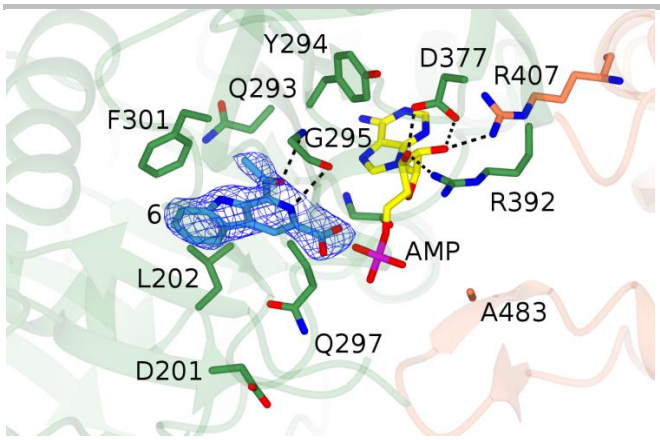



**Figure 2.** Active site of adenylation conformer of McbA showing residues involved in the binding of substrate **6** and AMP. Electron density is the $F_o$-$F_c$ (omit) map at a level of 3$\sigma$ before inclusion of the ligand atoms in refinement. Ligand atoms have been added for clarity. Selected interactions are indicated by black dashed lines.

**19a** were isolated by simple liquid-liquid extraction into ethyl acetate, giving isolated yields of between 50% and 85% of the amide products without further purification (Table 3). **16a** was also

**Table 3.** ATP-dependent coupling of carboxylic acids with amine **a** by McbA on 50 mg scale (*carried out on 30 mg scale).

| Acid | Equivalents of amine | Isolated Yield (%) |
|---|---|---|
| **6** | 2 | 85 (**6a**) |
| **10** | 2 | 70 (**10a**)* |
| **14** | 2 | 50 (**14a**) |
| **16** | 5 | 15 (**16a**) |
| **17** | 5 | 51 (**17a**) |
| **19** | 5 | 21 (**19a**) |

isolated, although only in 15% yield owing to complications in extraction of this indole product. These results compare favourably with the application of CAR-A-PCP to preparative amide bond formation,[9] as 100 equivalents of amine donor were required for those reactions, against the low ratios reported here. The difference may reflect of the active participation of McbA in amine binding and catalysis of amidation, which was not yet established for the CAR-A-PCP reaction. The results offer promise for the application of McbA-type enzymes in preparative biocatalyzed reactions, particularly if ATP recycling methods, currently a subject of significant interest in biocatalysis,[25] were to be employed.

The efficient biocatalytic formation of amide bonds for the preparation of pharmaceutical-type amides, is one of the most significant reactions for which biocatalytic alternatives are being sought. The studies presented here suggest that McbA and related ABS enzymes may provide starting points for the evolution of activity towards a wider range of carboxylic acid and amine partners, using the structure of McbA as a basis for structure-guided evolution.

## Experimental Section

Details of target selection, gene cloning and expression, enzyme purification and assay, synthesis of substrates and product standards, HPLC analyses, biotransformation protocols, crystallisation and data collection and refinement can be found within the Supporting Information.

## Acknowledgements

M.P. received an iCASE studentship award from GSK in collaboration with the Biotechnology and Biological Sciences Research Council (BBSRC). A.C. thanks the Principado de Asturias for a Clarín postdoctoral fellowship. We thank Dr Johan P. Turkenburg and Mr Sam Hart for X-ray data collection, and the Diamond Light Source for access to beamline i03 under proposal number mx-9948.

**Keywords:** amides • biocatalysis • adenylation • ligase • ATP

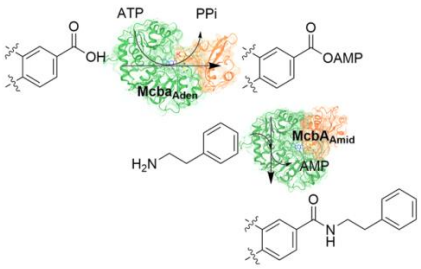
[1] J.R. Dunetz, J. Magano and G.A. Weisineburger, *Org. Proc. Res. Devel.*, **2016**, *20*, 140-177.
[2] J.S. Carey, D. Laffan, C. Thomson and M.T. Williams, *Org. Biomol. Chem.*, **2006**, *4*, 2337-2347.
[3] A. Goswami and S.G. van Lanen, *Mol. Biosyst.*, **2015**, *11*, 338-353.
[4] J. Pitzer and K. Steiner, *J. Biotechnol.*, **2016**, *235*, 32-46.
[5] R. Finking, *Annu. Rev. Microbiol.*, **2004**, *58*, 453-488.
[6] M. Winn, J.K. Fyans, Y. Zhuo and J. Micklefield, *Nat. Prod. Rep.*, **2016**, *33*, 317-347.
[7] E. Busto, V. Gotor-Fernández and V. Gotor, *Chem. Rev.*, **2011**, *111*, 3998-4035.
[8] V. Gotor, *Bioorg. Med. Chem.*, **1999**, *7*, 2189-2197.
[9] A.J.L. Wood, N.J. Weise, J.D. Frampton, M.S. Dunstan, M.A. Hollas, S.R. Derrington, R.C. Lloyd, D. Quaglia, F. Parmeggiani, D. Leys, N.J. Turner and S.L. Flitsch *Angew. Chem.*, **2017**, *129*, 14690-14693; *.Angew. Chem. Int. Ed.*, **2017,** *56*, 14498–14501.
[10] M. Steffensky, A. Műhlenweg, Z.X. Wang, S-M. Li and L. Heide, *Antimicrob. Agents. Chemother.*, **2000**, *44*, 1214-1222.
[11] A.S. Eustáquio, B. Gust, S-M. Li, K.F. Chater and L. Heide, *Chem. Biol.*, **2003**, *10*, 279-288.
[12] Z.X. Wan, S-M. Li and L. Heide, *Antimicrob. Agents. Chemother.*, **2000**, *44,* 3040-3048.
[13] Q. Chen, C. Ji, H. Huang, J. Ma, X. Tian and J. Ju, *Angew. Chem.* **2013**, *125*, 10164-10168; *Angew. Chem. Int. Ed. Engl.*, **2013**, *16*, 9980-9984.
[14] C. Ji, Q. Chen, Q. Li, H. Huang, Y. Song, J. Ma and J. Ju, *Tetrahedron Lett.*, **2014**, *55*, 4901-4904.
[15] K.P. Lippke, W.E. Műller and W.G. Schunack, *J. Pharm. Sci.*, **1985**, *74*, 674-680.
[16] W. Yin,; S. Majumder, T. Clayton, S. Petrou, M.L. VanLinn, O.A. Namjoshi, C. Ma, B.A. Cromer, B.L. Roth, D.M. Platt and J.M. Cook *Bioorg. Med. Chem.,* **2010**, *18*, 7548-7564.
[17] A. M. Gulick, *ACS Chem. Biol.* **2009**, *4*, 811-827.
[18] T. Nakatsu, S. Ichiyama, J. Hiratake, A. Saldanha, N. Kobashi, K. Sakata, H. Kato, *Nature* **2006**, *440*, 372.
[19] G. Jogl, L. Tong, *Biochemistry* **2004**, *43*, 1425-1431.
[20] J. J. May, N. Kessler, M. A. Marahiel, M. T. Stubbs, *Proc. Nat. Acad. Sci.* **2002**, *99*, 12120-12125.
[21] E. Conti, T. Stachelhaus, M. A. Marahiel, P. Brick, *EMBO J.* **1997**, *16*, 4174-4183.
[22] D. Gahloth, M. S. Dunstan, D. Quaglia, E. Klumbys, M. P. Lockhart-Cairns, A. M. Hill, S. R. Derrington, N. S. Scrutton, N. J. Turner, D. Leys, *Nat. Chem. Biol.* **2017**, *13*, 975.
[23] A. S. Reger, R. Wu, D. Dunaway-Mariano, A. M. Gulick, *Biochemistry* **2008**, *47*, 8016-8025.
[24] R. Wu, J. Cao, X. Lu, A. S. Reger, A. M. Gulick, D. Dunaway-Mariano, *Biochemistry* **2008**, *47*, 8026-8039.
[25] J.N. Andexer and M. Richter *ChemBioChem*, **2015**, *16*, 380-386.

**Entry for the Table of Contents** (Please choose one layout)

Layout 1:

# COMMUNICATION

The amide bond synthetase McbA from *Marinactinospora thermotolerans* catalyzes the ATP dependent two-step synthesis of phenyethylamides using a surprisingly broad range of aryl acid donors and at amine molar equivalents of just 1 to 5. The structure of the enzyme reveals the basis for broad aryl acid specificity and serves as a platform for further engineering of this system for the biocatalytic synthesis of amides.




*Mark Petchey,[a] Anibal Cuetos,[a] Benjamin Rowlinson,[a] Stephanie Dannevald,[a] Amina Frese,[a] Peter W. Sutton,[b]† Sarah Lovelock,[b]‡ Richard C. Lloyd,[b] Ian J.S. Fairlamb[a]* and Gideon Grogan[a]**

***Page No. – Page No.***

*Aryl acid specificity of the amide bond synthetase McbA suggests broad potential for the biocatalytic synthesis of amides*