

Kristallstruktur einer molekularen Synthesemaschine**

Kira J. Weissman und Rolf Müller*

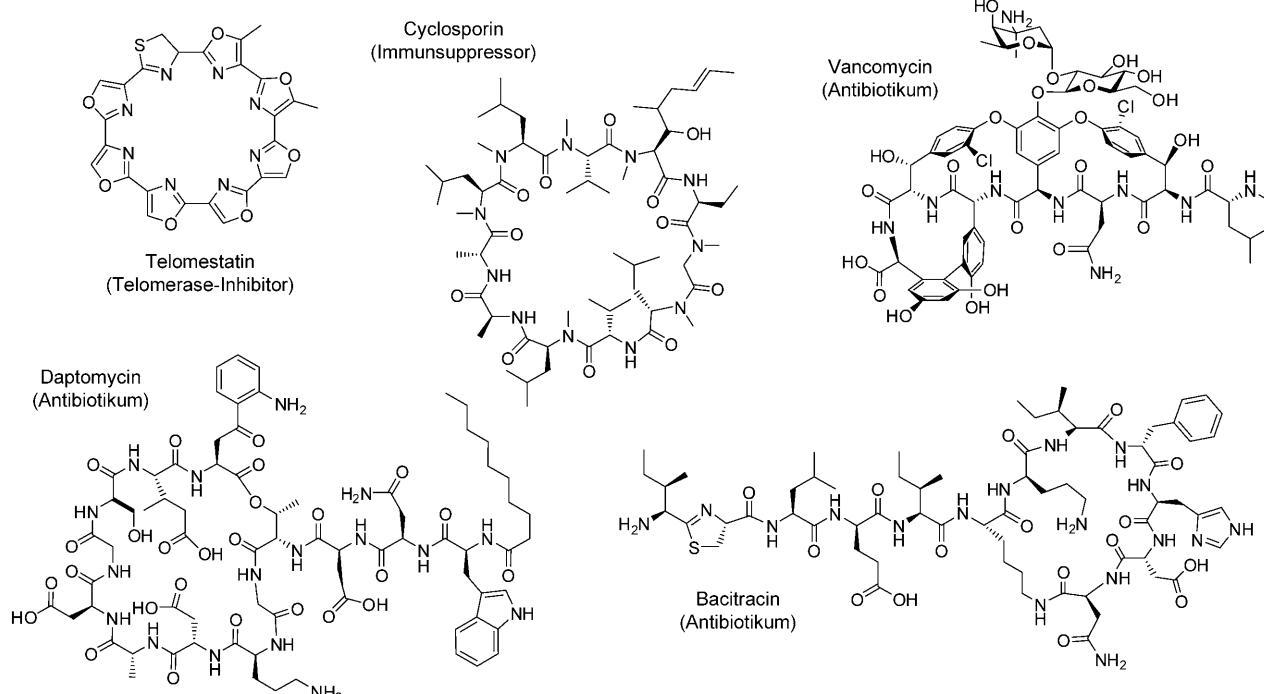
Biosynthese · Enzyme · Naturstoffe ·

Strukturaufklärung

Nie aufgeben! Nach jahrelangen vergeblichen Bemühungen beschrieben die Arbeitsgruppen um Essen und Marahiel in einer kürzlich erschienenen Veröffentlichung in *Science*^[1] die erste Kristallstruktur eines vollständigen Moduls mit mehreren Domänen einer nicht-ribosomalen Peptidsynthetase (NRPS). Obwohl Marahiel und Mitarbeiter bereits 1997^[2] die erste Struktur einer einzelnen Domäne einer NRPS publizierten, vergingen lange elf Jahre, bis nun geeignete Kristalle eines Moduls erzeugt und analysiert werden konnten.

Nicht-ribosomal gebildete Peptide (NRPs) sind eine diverse Gruppe zumeist cyclischer Naturstoffe (Schema 1), die

nicht nur proteinogene Aminosäuren, sondern auch hunderte anderer Biosynthesebausteine enthalten.^[3] Diese Metabolite werden in Mikroorganismen durch die nicht-ribosomalen Peptidsynthetasen (NRPSs) gebildet, die in der Natur eine nucleinsäureunabhängige Peptidsynthese ermöglichen. Dabei wird die Sequenz der inkorporierten Aminosäuren im Produkt durch die Reihenfolge der autonom gefalteten enzymatischen Domänen innerhalb gigantischer Multienzyme bestimmt. Jede Domäne ist für einen spezifischen Schritt der Assemblierung zuständig, sodass NRPSs als eine Art molekulares Fließband betrachtet werden können. Domänen werden zu Funktionseinheiten gruppiert, die „Module“ ge-



Schema 1. Strukturen und biologische Aktivitäten von nicht-ribosomal gebildeten Peptiden.

[*] Dr. K. J. Weissman, Prof. R. Müller
Pharmazeutische Biotechnologie, Universität des Saarlandes
Postfach 151150, 66041 Saarbrücken (Deutschland)
Fax: (+49) 681-3027-0201
E-Mail: rom@mx.uni-saarland.de

[**] Diese Arbeit wurde vom Bundesministerium für Bildung und Forschung und der Deutschen Forschungsgemeinschaft unterstützt.

nannt werden, von denen jeweils ein Kettenverlängerungsschritt durchgeführt wird. Im Minimalmodul selektiert und aktiviert die Adenylierungsdomäne (A) eine Aminosäure als Adenylat, und eine Kondensationsdomäne (C) verknüpft diesen Baustein mit der wachsenden Biosynthesekette. Zwischen diesen Domänen befinden sich katalytisch nicht aktive Domänen, die Peptidyl-Carrier-Proteine (PCPs) genannt

werden und mit einer prothetischen Gruppe – dem Phosphopantetheinyl-Arm (Ppant-Arm) – ausgestattet sind. An die terminale Thiolgruppe des Ppant-Arms wird die jeweilige Aminosäure gebunden, und nachfolgend findet hier auch die Kettenverlängerung statt, bei der die wachsende Peptidylkette durch die Biosynthesemaschinerie weitergeleitet wird. Die Thioester-Verknüpfung aktiviert die Aminosäure für die Kondensationsreaktion und ermöglicht eine dirigierte Weiterleitung der Intermediate zu den verschiedenen aktiven Zentren der NRPS, wodurch die Gesamteffizienz des Prozesses steigt. Zusätzlich zu den oben beschriebenen minimal vorhandenen Domänen eines Moduls können weitere vorkommen, welche die Aminosäuren modifizieren, so z.B. Domänen, die Epimerisierungen, *N*-Methylierungen oder Cyclodehydratisierungen durchführen. In der Regel wird die finale Kettenfreisetzung von einer Thioesterasedomäne (TE) durch Makrocyclisierung katalysiert, die an das Ende des letzten Moduls fusioniert vorliegt. Durch diese Verteilung der Biosyntheseschritte ermöglicht die Natur die Assemblierung hochkomplexer Strukturen aus einfachen Bausteinen. Die Polyketide – eine zweite große Klasse von mikrobiellen Naturstoffen – werden durch den Einsatz einer sehr ähnlichen Biosynthesestrategie gebildet.^[4]

Die Entdeckung des modularen Aufbaus von NRPSs hat große Hoffnungen für die Herstellung neuer Varianten der Produkte geweckt, weil es möglich schien, dies durch eine Reprogrammierung der Synthetasen zu erreichen. Diese Idee ist nicht nur von akademischem Interesse, weil die NRP-Metabolite vielfach therapeutisch nutzbare Aktivitäten aufweisen (z.B. als Krebsmittel, als Antibiotika oder als Immunsuppressiva), weshalb Derivate für die Wirkstoffentwicklung von großer Bedeutung sind.

Leider sind biologische Systeme aber selten so einfach aufgebaut wie wir sie uns wünschen, was dazu führte, dass die meisten der ersten moderaten Versuche der gezielten Rekonfiguration von NRPSs durch den Austausch von Domänen sehr enttäuschende Resultate zeigten.^[5-7] Konsequenterweise wurde dann verstärkt an der Aufklärung der detaillierten molekularen Enzymologie und der Strukturbioologie von NRPSs gearbeitet. Obwohl die Kristallstruktur eines ganzen Moduls von Anfang an das Ziel war, wurde zunächst eine diskrete A-Domäne kristallographisch beschrieben,^[2] gefolgt von verschiedenen TEs^[8,9] und einer C-Domäne.^[10] Nachfolgend wurde die Struktur einer PCP-C-Didomäne gelöst, die auch die Übergangsregion zwischen zwei NRPS-Modulen enthält.^[11] Ein repräsentatives PCP wurde mittels NMR-Spektroskopie beschrieben.^[12,13] Die Kristallstrukturen und ihre Analyse ermöglichen wichtige Einblicke in die katalytischen Mechanismen der Domänenaktivitäten. Zudem konnte der allgemeine Spezifitätscode von A-Domänen definiert werden (der zehn Aminosäuren umfassende nicht-ribosomale Code), der mittlerweile weitverbreiteten Einsatz für die Vorhersage hypothetischer Produkte von NRPSs unbekannter Funktion findet.^[14,15] Es konnte gezeigt werden, dass das PCP drei alternative Konformationen einnehmen kann,^[13] was darauf schließen lässt, dass die Änderung der Konformation zur Programmierung der alternativen Interaktionen mit den verschiedenen Partnern der Domäne genutzt wird. Bislang fehlten allerdings jegliche Strukturdaten

zur räumlichen Verknüpfung der Domänen innerhalb eines typischen Moduls, was als entscheidende Information angesehen werden kann, um zu verstehen, wie das PCP-gebundene Substrat zwischen den verschiedenen katalytischen Zentren bewegt wird und wie diese Bewegungen kontrolliert werden.^[16]

Die Kristallstruktur von SrfA-C als typisches Modul der Form C-A-PCP-TE aus der Surfaktinsynthetase (Abbildung 1) bei 2.6 Å Auflösung^[1] ermöglicht es nun zum ersten Mal, diese Fragen anzugehen. Um lichtbrechende Kristalle zu erzeugen, mussten die Autoren das konservierte Serin des katalytischen Zentrums des PCP durch Alanin substituieren. Diese Mutation eliminiert die Möglichkeit der post-transkriptionalen Modifikation mit dem Ppant-Arm, und es wurde gezeigt, dass das PCP dadurch in einer der drei möglichen Konformationen festgehalten wird.^[13] Dadurch gelang es, die gesamte konformativ Heterogenität des Moduls einzuschränken, welche vermutlich bisherige Kristallisierungsversuche verhinderte.

Das Zentrum des Moduls ist eine stabile und rechteckige katalytische Plattform, die von der C-Domäne und der N-terminalen Zentraleinheit der A-Domäne (A_{core}) gebildet wird, wobei beide aktiven Zentren auf derselben Seite liegen. Der Komplex wird durch eine große Interaktionsfläche zwischen der C- und der A_{core} -Domäne, eine enge Interaktion der katalytischen Regionen und einen gut definierten dazwischenliegenden „Linker“ zusammengehalten. Der Rest der A-Domäne und das PCP sind mit der Plattform durch kurze flexible Verbindungsstücke verbunden. Diese Beobachtung führt zu einem Modell, in dem sich die beiden Domänen während jedes katalytischen Zyklus relativ zu der statischen C- A_{core} -Plattform bewegen. Diese Mobilität ist definitiv notwendig, weil die katalytischen Zentren der C- und A_{core} -Domänen 63 Å voneinander entfernt liegen, was durch den Ppant-Arm (ca. 20 Å) nicht überbrückt werden kann. Die Struktur der TE ist im Prinzip identisch mit derjenigen der separat exprimierten Domäne^[8,9] und nimmt eine unabhängige Region im Modul ein.

Die Struktur gibt auch zwei wichtige Hinweise darauf, wie aufeinanderfolgende Module miteinander interagieren und Kettenverlängerungen durchführen. In der Struktur ist das strangabwärts liegende PCP (Akzeptor) an der Akzeptorstelle der C-Domäne festgemacht. Die Bindestelle des strangaufwärts liegenden PCP (Donor) liegt direkt auf der anderen Seite eines Canyons, der durch das Zentrum der C-Domäne verläuft. Dadurch liegen beide PCPs gleichzeitig innerhalb der Reichweite des katalytischen Histidins. Allerdings sind die zusammengehörenden Donor- und Akzeptor-PCPs in vielen Fällen auf separaten Multienzymen zu finden. Deshalb ist es notwendig, dass die beiden Polypeptide spezifisch miteinander assoziieren, um das richtige Produkt aufzubauen, während gleichzeitig unerwünschte Kontakte mit anderen Multienzymen der Assemblierungslinie verhindert werden. Dieser auch als „Docking“ bekannte Prozess wurde in vorhergehenden Arbeiten zusammenpassenden Sequenzregionen („communication-mediating“(COM)-Domänen) zugeschrieben, die sich an den C- und N-terminalen Enden der Proteine befinden und die Selektivität für den Partner entscheidend beeinflussen.^[17] Nach diesem „Docking-Mo-

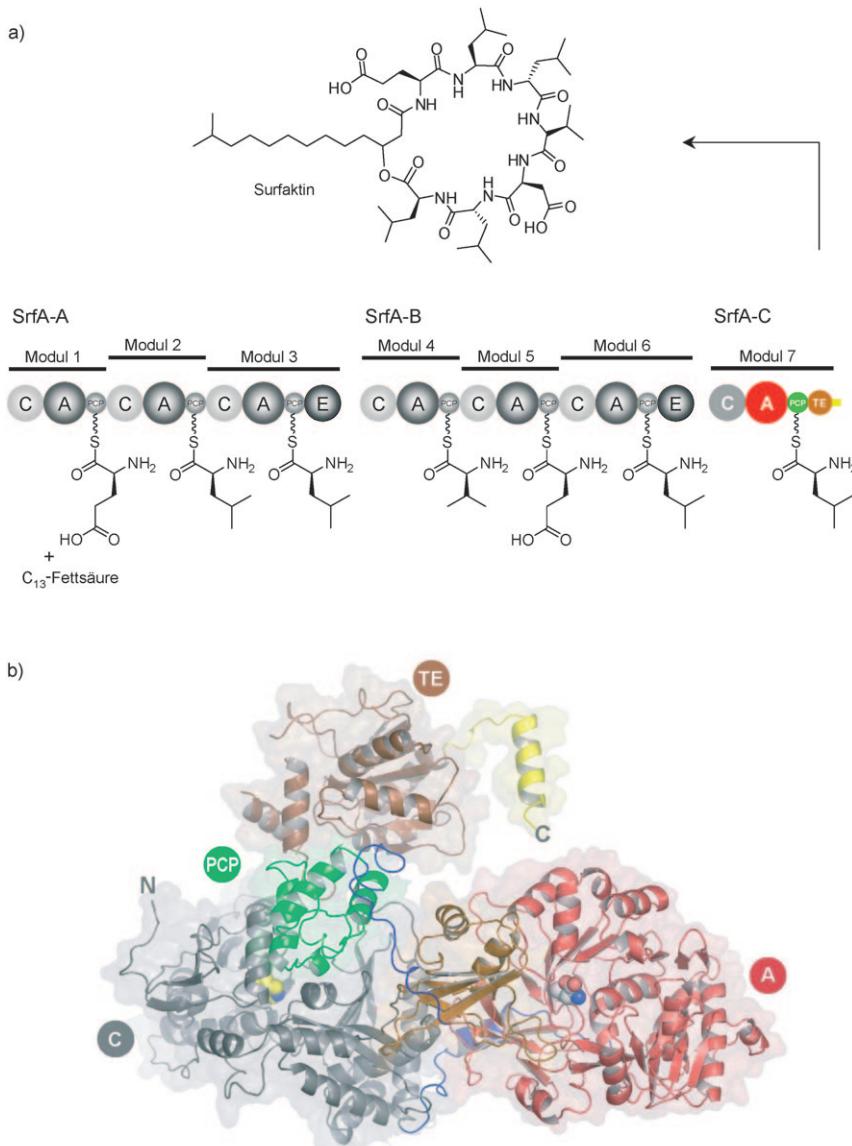


Abbildung 1. Biosynthese von Surfaktin in *Bacillus subtilis*. a) Surfaktin wird von einer NRPS gebildet, die die drei Untereinheiten SrfA-A, SrfA-B und SrfA-C aufweist. SrfA-A und SrfA-B enthalten je drei Module, während SrfA-C als Terminationsmodul fungiert. Die Aminosäuren, die vom jeweiligen Modul inkorporiert werden, sind dargestellt (die Epimerisierung erfolgt wahrscheinlich nach der Knüpfung der Peptidbindung). b) Die Struktur von SrfA-C bei einer Auflösung von 2.6 Å. Die farbliche Darstellung der Domänen ist so wie in (a). Das His des aktiven Zentrums der C-Domäne und eine Leucin-Bindung innerhalb des aktiven Zentrums der A-Domäne sind in der CPK-Darstellung zu sehen. Wiedergegeben mit Genehmigung nach Lit. [1]; Copyright 2008, American Association for the Advancement of Science.

dell“ interagieren fünf Aminosäuren jeder α -helicalen COM-Domäne mit komplementären Einheiten der Partnerdomäne und formen so eine antiparallele Doppelspirale.^[17] Zufälligerweise ähnelt der C-terminal an SrfA-C angebrachte Aufreinigungstag (*myc*-His₆) in seiner Sequenz und Hydrophobie den vermuteten COM-Spiralen. Dieser Sequenzbereich ist in der Struktur des Moduls in einem handförmigen Proteinmotiv der C-Domäne des angrenzenden Moduls eingewickelt, wobei eine Vielzahl von hauptsächlich hydrophoben Aminosäuren an den Kontaktan beteiligt sind. Diese COM-Hand besteht aus der bereits bekannten N-terminalen COM-Region und einer dreisträngigen β -Faltblattstruktur, die zur C-

Domäne gehört. Daraus wird klar, dass die strukturelle Basis für die spezifische Interaktion zwischen NRPS-Multienzymen komplexer ist als ursprünglich angenommen wurde. An dieser Stelle ist es interessant zu erwähnen, dass eine einzelne Punktmutation innerhalb der N-terminalen COM hinreichend war, um die Spezifität des Dockings umzukehren.^[17]

Insgesamt zeigt die Struktur, dass ausgiebige Umlagerungen der Domänen stattfinden müssen, während die NRPS arbeitet. Diese Feststellung führt direkt die Schwächen der Analyse der Kristallstruktur vor Augen, da diese nur eine Momentaufnahme einer wesentlich komplexeren Reaktionssequenz sein kann. Beispielsweise lässt sich nicht erkennen, wie die Gesamtstruktur sich verändert, um eine Interaktion zwischen dem PCP und der A-Domäne oder der TE zu ermöglichen oder wie die notwendigen Veränderungen herbeigeführt werden können. Die Abwesenheit des Ppent-Arms und damit auch des Substrates ist ebenfalls einen Schwachpunkt der Analyse, besonders da die Verknüpfung mit der prosthetischen Gruppe die Konformation des PCP moduliert.^[13] Offensichtlich werden in Zukunft weitere Strukturen benötigt, in denen ein aktiviertes PCP während der Interaktion mit seinen übrigen Partnern dargestellt werden kann. Kürzlich wurden Arbeiten publiziert, die hier die Richtung weisen könnten,^[18] da der Ppent-Arm des PCP durch einen Inhibitor des aktiven Zentrums der Zieldomäne ersetzt werden kann, wobei die breite Substratakzeptanz der aktivierenden Phosphopantetheinyltransferasen ausgenutzt wird. Diese Modifikation würde dem PCP alle Schritte bis hin zur Interaktion mit einer spezifischen Domäne erlauben und dann den Gesamtkomplex in genau einer seiner katalytisch relevanten Konformationen festhalten.

In Zukunft wird es hoffentlich möglich sein, eine Reihe solcher statischen Momentaufnahmen miteinander zu kombinieren und damit eine Art von Film eines kompletten katalytischen Zyklus einer NRPS zu erstellen.

Welche Lehren können bis dahin für die Manipulation von NRPS-Systemen gezogen werden? Die enge Assoziation der C- und A-Domänen legt die Strategie nahe, dass diese Domänen als ein funktionales Paar in einen neuen modularen Kontext eingebracht werden sollten. Damit wird gleichzeitig der Spezifität der C-Domäne an ihrer Akzeptorstelle Rechnung getragen.^[19] Zudem zeigt die Struktur bislang unbekannte Aminosäurepositionen auf, durch deren Mutation

neue und produktive C-PCP-Interaktionen herbeigeführt werden können.^[20] Die gezielte Veränderung von COM-Domänen wurde schon eingesetzt, um künstliche Kombinationen aus NRPS-Untereinheiten zu erzeugen.^[17] Durch den Einbezug der nun erkannten wichtigen β -Faltblattstruktur in der COM-Hand sollten sich nicht-native Interaktionsbereiche weiter optimieren lassen, um letztlich die Effizienz von künstlichen Hybridsynthetasen zu erhöhen. Die Ergebnisse solcher strukturbasierter „Engineering“-Experimente werden gespannt erwartet.

Online veröffentlicht am 9. September 2008

- [1] A. Tanovic, S. A. Samel, L.-O. Essen, M. A. Marahiel, *Science* **2008**, *321*, 659.
- [2] E. Conti, T. Stachelhaus, M. A. Marahiel, P. Brick, *EMBO J.* **1997**, *16*, 4174.
- [3] J. Grünwald, M. A. Marahiel, *Microbiol. Mol. Biol. Rev.* **2006**, *70*, 121.
- [4] J. Staunton, K. J. Weissman, *Nat. Prod. Rep.* **2001**, *18*, 380.
- [5] S. Doekel, M. A. Marahiel, *Chem. Biol.* **2000**, *7*, 373.
- [6] H. D. Mootz, D. Schwarzer, M. A. Marahiel, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 5848.
- [7] A. Schneider, T. Stachelhaus, M. A. Marahiel, *Mol. Gen. Genetics* **1998**, *257*, 308.
- [8] S. D. Bruner, T. Weber, R. M. Kohli, D. Schwarzer, M. A. Marahiel, C. T. Walsh, M. T. Stubbs, *Structure* **2002**, *10*, 301.
- [9] S. A. Samel, B. Wagner, M. A. Marahiel, L. O. Essen, *J. Mol. Biol.* **2006**, *359*, 876.
- [10] T. A. Keating, C. G. Marshall, C. T. Walsh, A. E. Keating, *Nat. Struct. Biol.* **2002**, *9*, 522.
- [11] S. A. Samel, G. Schoenafinger, T. A. Knappe, M. A. Marahiel, L. O. Essen, *Structure* **2007**, *15*, 781.
- [12] T. Weber, R. Baumgartner, C. Renner, M. A. Marahiel, T. A. Holak, *Structure* **2000**, *8*, 407.
- [13] A. Koglin, M. R. Mofid, F. Löhr, B. Schäfer, V. V. Rogov, M.-M. Blum, T. Mittag, M. A. Marahiel, F. Bernhard, V. Dötsch, *Science* **2006**, *312*, 273.
- [14] T. Stachelhaus, H. D. Mootz, M. A. Marahiel, *Chem. Biol.* **1999**, *6*, 493.
- [15] G. L. Challis, J. Ravel, C. Townsend, *Chem. Biol.* **2000**, *7*, 211.
- [16] K. J. Weissman, R. Müller, *ChemBioChem* **2008**, *9*, 826.
- [17] M. Hahn, T. Stachelhaus, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 275.
- [18] Y. Liu, S. D. Bruner, *ChemBioChem* **2007**, *8*, 617.
- [19] G. L. Challis, J. H. Naismith, *Curr. Opin. Struct. Biol.* **2004**, *14*, 748.
- [20] J. R. Lai, A. Koglin, C. T. Walsh, *Biochemistry* **2006**, *45*, 14869.