Nonribosomal Peptide Synthetases

Bastiaan Dijkstra

supervisor: prof. dr. Arnold Driessen

Bachelor thesis
Major Molecular Life Sciences
11-06-2015
studentnr. s1766473

Abstract

Nonribosomal peptide synthetases (NRPS) are multimodular mega-proteins. They are able to use nonproteinogenic amino acids, and are used by bacteria and fungi to produce secondary metabolites. This includes many of the known antibacterial and anti-fungal compounds known today.

In this thesis the function and mechanisms of NRPSs are discussed. We look at the function of the many different domains used by these proteins, and give an overview of the different catalyzed reactions. Also the organization of the modules and how they interact on a structural level is shown. Finally a number of different methods for NRPS manipulation and obtaining novel peptides are discussed.

Table of contents

1. Introduction

In nature enzymes are the chemical factories that are involved in almost all biological processes. In terms of size the family of nonribosomal peptide synthetases (NRPS) are one of the largest. For example the NRPS involved in Cyclosporin A synthesis has a mass of around 1,6MDa (Weber, Schorgendorfer et al. 1994). As the name suggests they are able to synthesize peptides separate from the ribosomal system. The range of peptides that can be synthesized by NRPSs is much larger than the range of regular ribosomes. Ribosomes are limited to the proteinogenic amino acids, while NRPSs can use proteinogenic-, non-proteinogenic amino acids (L-and D- variants), amines and carboxylic acids for synthesis. (Caboche, Leclere et al. 2010) This wide range of substrates brings a wide variety of compounds that can be synthesized by NRPSs. This gives the products many different applications, most notably NRPSs are responsible for the production of a large portion of known antibiotics. This makes NRPSs very interesting as a target for research and engineering in order to develop novel antibiotics.

1.1 Secondary metabolites

Products formed by NRPSs fall into the group of secondary metabolites. This group of small compounds are used by bacteria and fungi to perform non essential tasks in and around the cell. They are known to have antibacterial, anti-fungal, cytostatic and immunosuppresive functions, and have been linked to cell development, spore formation, communication and cell defense. (Calvo, Wilson et al. 2002) Secondary metabolites, or natural products, are used in a responsive way by the cell, produced when the cell receives the proper external stimuli. For example to secure the environment and remove competition by use of antibiotics, or to combat iron deficiency in the environment by producing siderophores. (Winkelmann 2007)

1.2 Ribosome vs NRPS

Compared to the ribosome, NRPSs use a wildly different way of recognizing substrates. Ribosomes are versatile enzymes that can produce many different peptides using just one ribosome. They use messenger RNA to recognize and produce the proper peptides. NRPS use a different recognizing method. The substrate recognizing function is a built-in part of an NRPS, so a single NRPS complex is only able to produce one product, whereas a ribosome can produce many different products. However, NRPSs can utilize a much larger pool of possible substrates, giving the overall group of NRPSs a wider variety of products than a ribosome.

NRPSs are multimodular systems with many different domains. They use a multi template model (Stein, Vater et al. 1994) and the synthesis resembles a factory assembly line. Domains are structured in modules, and each module handles one substrate (amino acid, amine etc). Modules produce intermediates which are bound together in a number of ways: In linear synthesis the module sequence resembles the sequence of the product. In iterative synthesis a single substrate is incorporated multiple times, often to produce cyclic compounds. Finally, in non-linear synthesis the order of the modules does not correspond with the order of substrates incorporated in the final peptide. (Ali 2013)

2. Main NRPS domains

2.1 A domain

The A domain, or adenylation domain is the first domain in a module and is responsible for recognition of the proper substrate. It also activates the recognized substrate for the next step of its incorporation in the growing peptide chain. This is a 2 step process, the first step is the formation of an aminoacyl adenyl intermediate. This reaction expends an ATP and links the amino acid to an AMP while releasing pyrophosphate, and is catalyzed by the presence of a Mg^{2+}. This is comparable to the function of the aminoacyl-tRNA-synthetases in ribosomal synthesis (Eriani, Dirheimer et al. 1990). The second step binds this aminoacyl adenyl intermediate to the 4'-phosphopantetheine arm of the PCP domain next to the A domain.

The structural mechanics of these two steps have been shown to include a ~140^0 change in the C-terminal domain. (Hur, Vickery et al. 2012) (figure 5) The mechanism of recognition in the A domain has long been elusive. The first step was the elucidation of the structure of an isolated A domain of the first module of the gramicidin S synthetase (PheA)(Conti, Franks et al. 1996) This structure, with AMP and phenylalanine bound, showed that the binding pocket of the domain selects for the correct substrate. The crystal structure also showed that the A domain is compiled out of a large N-terminal and a small C-terminal domain, with the active site located at the intersection of these 2 domains. The shape and available functional groups in the binding pocket is called the 'nonribosomal code' of the A domain. By analyzing the binding pocket residues and comparisons with other A domain sequences, 8-10 residues were found to be determining the binding pocket specificity(Stachelhaus, Mootz et al. 1999). These residues are referred to as the 'codons' in nonribosomal synthesis. (Stachelhaus, Mootz et al. 1999) more recently *in silico* analysis resulted in predictive systems that can give researchers an idea of the product of an unknown NRPS on the basis of the A domain sequences. (Rausch, Weber et al. 2005) This method is not completely failsafe, because only the product of linear NRPSs can be predicted this way. The system cannot predict iterative and non-linear products by amino acid sequence, so there is still room for improvement in this area. Another problem with the known A domain code, is the fact that none of the known prediction methods developed for prokaryotes seem to work on eukaryotes. This is most likely due to different evolution in prokaryotes and eukaryotes, with different mechanisms producing the same specificity.(von Dohren 2004) The crystal structure of an A domain in the fungus *Neotyphodium lolii* has been resolved. (Lee, Johnson et al. 2010) This structure might be the first step in elucidating the nonribosomal code in fungi.

2.2 PCP/T domain

The next domain in the main modules is the Peptidyl Carrier Protein (PCP), also called the thiolation (T) domain. This domain is used in conjunction with a cofactor, a 4' phosphopantetheine (4'PP) arm. This 4'PP is covalently linked to a conserved serine on the PCP domain. This cofactor is post-translationally transferred from a Coenzyme A by 4-phosphopantetheinyl transferases that are associated with the corresponding NRPS, thus converting the *apo* form without the 4'PP arm into the *holo* form with the 4'PP arm connected.(Quadri, Weinreb et al. 1998) The activated aminoacyl adenyl intermediate from the A domain is covalently linked to the 4'PP arm at the -SH end.
The main function of this domain is the transfer of recognized and activated substrate to the next domain of the main module (C domain). This carrier function is not unique to NRPSs: other large multi modular synthesis systems like polyketide synthetases (PKS) and fatty acid synthetases (FAS) share this function. (Hopwood 1997)

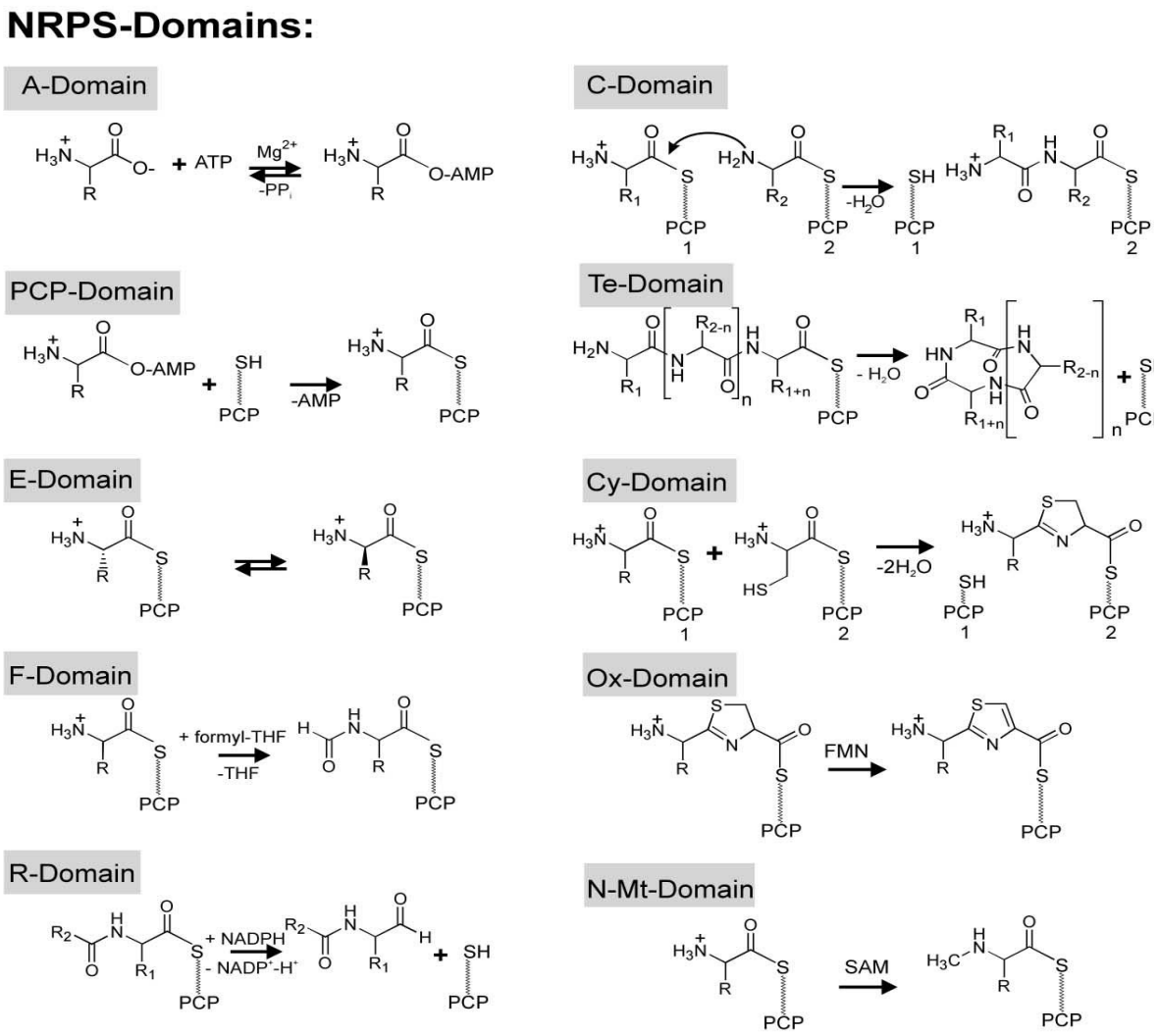

figure 1. Reactions of NRPS domains. *Source: (Schwarzer, Finking et al. 2003)*

2.3 C domain

The last domain in the main NRPS module is the condensation (C) domain. Its main function is to form the peptide bonds between the two bound and activated intermediates of the upstream and downstream PCP domains. This is a large (~450aa) monomeric domain which is compiled from a C- and N-terminal subdomain, arranged in a V shape with the active site at the junction of the 2 subdomains. (Stachelhaus, Mootz et al. 1998) (figure 4) Activated intermediates are transferred by the PCP domain to this domain, where the peptide bond is formed. This reaction is carried out between the aminoacyl intermediate bound to the 4'PP arm of the adjacent PCP domain and the growing peptide chain, bound to the 4'PP arm of the PCP domain of the previous module in the NRPS chain. The thioester group of the first module is attacked by the nucleophilic amino group of the substrate of the next module. This forms an amide bond and the newly formed peptide is now only bound to the PCP of the second module.

C domain analysis has shown that the C domain also exhibits a selective function alongside the A domain. (Belshaw, Walsh et al. 1999) This specificity is oriented towards the incoming aminoacyl intermediate, and functions like a filter that selects for cognate amino acids. (Belshaw, Walsh et al. 1999)

3. Modifying domains

3.1 TE domain

Peptide formation is terminated and released from the last PCP by the thioesterase (TE) domain. This 280aa domain is only found in the last module of the NRPS.(Marahiel 2009) The structure of the domain resembles a α/β hydrolase and the active site is capped off with a lid that is highly variable between different TE domains. This lid is responsible for selecting against presented substrate and the promiscuity of this lid differs between TE domains. Most NRPSs have a quite low selectivity for substrate in their TE domain, but in some exceptions the TE domain is used as a logic gate. An example of this is the nocarcidin synthesis, here a TE domain selects for the proper formation of a β-lactam ring.(Gaudelli and Townsend 2014) When the ring has been properly formed, the TE domain continues to properly hydrolyse the peptide, but this reaction will not take place if the ring is not there.

The reaction carried out by the TE domain is a two-step process in which a conserved serine in the active side of the TE performs a nucleophilic attack on the formed peptide chain bound by a thioester bond to the PCP of the last module. This forms an acyl-O-TE intermediate. After this reaction one of two different second half reactions can take place. The fully formed peptide can be hydrolyzed off the PCP and form a linear peptide. The other option is macrocyclisation by an internal nucleophile in the peptide. (Kohli and Walsh 2003) Many of the known secondary metabolites use some sort of cyclic structures, probably to promote structural toughness and enhance activity. (Hur, Vickery et al. 2012)

3.2 E domain

One of the interesting characteristics of NRPSs, is their ability to use D enantiomers of amino acids for synthesis. These enantiomers are not used by the ribosomal system, and serve as a way to produce the special structural conformations found in natural products. The NRPS system has a number of ways to use and activate these unnatural amino acids. One way is to directly activate them by use of the A domain, thus using a D-amino acid produced by an external system.(Li and Jensen 2008) The other way uses the epimerization (E) domain, and transforms a normal L-amino acid into its D form.

This racemic change is performed on the L-amino acid bound to the PCP. This means that the A domain selects for the natural L-form, and during the catalytic process the amino acid is flipped to its D-form. (Stachelhaus and Walsh 2000) Only recently the first crystal structure of an epimerization domain has been published, this structure may be able to lead to new insights into the mechanisms and selectivity of the E domain (Samel, Czodrowski et al. 2014)

3.3 Cy domain

A distinctive property of many of the products formed by NRPSs is the presence of heterocyclic rings. These are rings of 5 atoms that are derived from either cysteine, called a thiazoline ring or a serine/threonine called an oxazoline ring. These rings are incorporated into the peptide by the cyclization (Cy) domain using a three step process. First, the amino acid containing a nucleophilic side chain attacks the activated amino acyl intermediate of the upstream module. This forms an amide bond between the 2 amino acids that is internally attacked by the nucleophillic thiol (in cysteine) or hydroxyl (in serine and threonine) forming the 5 ring in the second step. The last step is dehydration using a proton provided by the Cy domain and the hydroxy group connected to the heterocyclic ring. (Hur, Vickery et al. 2012) The Cy domain is used in duplo in the NRPS chain, as demonstrated by the the vibriobactin producing VibF protein.(Marshall, Burkart et al. 2001) The first Cy domain is responsible for the heterocyclic ring formation and the second domain performs the condensation reaction usually carried out by the C domain. This results in the absence of a C domain when a Cy domain is present, although the two domains share structural homology.

3.4 Ox and R domain

After the heterocyclation reaction performed by the Cy domain, there is an option for the NRPS to oxidize or reduce the formed 5 ring. This is facilitated by the oxidation (Ox) domain, which produces thiazole/oxyzole, and the reduction (R) domain, which produces thiazolidine/oxazolidine. The Ox domain is incorporated into the NRPS structure and needs a cofactor for its function, called flavin mononucleotide (FMN). The R domain however, acts more like a free enzyme and uses NADPH to reduce its substrates. (Hur, Vickery et al. 2012, Grünewald and Marahiel 2013)

3.5 F domain

In prokaryotic ribosomal peptide synthesis the addition of a formyl group is used to initiate the synthesis process.(Laursen, Sorensen et al. 2005) Formylation is also used in some nonribosomally formed peptides, catalysed by the formylation (F) domain. Mutational analysis on this domain has showed that its function is essential in gramicidin synthesis, because the A domain cannot accept the formylated version of its substrate, while formylation is necessary to be accepted by the C domain. (Schoenafinger, Schracke et al. 2006)

3.6 MT domain

Many of the natural products synthesized by NRPSs have methylated amino acids in their structure. The addition of these methyl groups is catalyzed by the methyltransferase (MT) domain. This domain uses (S)-adenosyl methionine (SAM) as a cofactor, and transfers a methyl group from this cofactor to the amino acid.(Hur, Vickery et al. 2012) These methyl groups can be attached to a carbon or nitrogen atom in the amino acid, and are subsequently called C-methyl and N-methyl groups. N-methyls are usually produced by methylating the PCP-bound amino acid(Strieker, Tanovic et al. 2010), but the

methylation of external substrates is also observed.(Hornbogen, Riechers et al. 2007, Hur, Vickery et al. 2012)
C-MT domains do not directly methylate the substrate, but instead perform their catalytic function on their precursors. (Milne, Powell et al. 2006) This requires that the associated A domain has to be able to recognize the methylated form of the substrate.

3.7 H domain

Halogens in secondary metabolites are reasonably common with about 4000 halogenated compounds known.(Gribble 2004) Two types of halogenating enzymes are found that are active in NRPSs, flavin-dependent halogenases and non-heme iron-dependent halogenases. Flavin-dependent halogenases are responsible for the binding of halogens to aromatic rings of growing peptide chains. As the name suggest these enzymes need flavin ($FADH_2$) as a cofactor in its catalytic function. Non-heme iron-dependent halogenases work on unactivated aliphatic carbon centers in the growing peptide.(Hur, Vickery et al. 2012) The presence of halogens in the final peptide has a significant effect, and it has been shown that some antibiotics lose 8-16 fold in activity when the halogen is removed. (Dorrestein, Yeh et al. 2005)

3.8. COM-linker domains

In fungi NRPSs are usually one large protein, that is located on the genome as a single gene. In bacteria this is different, as they are located on the genome in a single open reading frame but are synthesized as separate proteins. These proteins do however have to function together as a single system and have to be bound together in some way. This function is performed by the COM domains, small regions (15-25aa) on either side of the protein, that act like a lock-and-key mechanism. The upstream NRPS contains a c-terminal donor COM domain that is recognized by the n-terminal COM domain of the downstream NRPS. These two domains are complementary to each other, ensuring that the proper peptide is formed even when multiple NRPSs have to work together.(Hur, Vickery et al. 2012)

3.9 X domain

Very recently a new domain has been discovered with a relatively unknown function. It has been suggested this domain is involved with the synthesis of glycopeptide antibiotics. These peptides need crosslinking of their aromatic side chains by oxygenases to reach their bioactive conformation. The recruitment of the necessary oxygenases is thought to be performed by the X domain.
The structure of this domain has been revealed by x-ray crystallography and resembles the C domain in structure, but its amino acid composition is different and appears to be noncatalytic. (Haslinger, Peschke et al. 2015)

4. Associated proteins

4.1 4'PP transferases

The enzymes responsible for the addition of the 4'PP arm to the PCP domain are called 4' phosphopantetheinyl transferases. This reaction is essential for NRPS function, because the individual substrates cannot be bound together if this reaction doesn't take place. The 4'PP arm is taken from a free Coenzyme A, and linked to a serine on the PCP domain, thus activating the PCP domain and preparing it for peptide synthesis.

4.2 Type 2 Thioesterase domain (TEII)

The serine group on the PCP domain receiving the 4'PP arm has a relative low specificity for Coenzyme A (CoA), so other CoA-like compounds are able to bind to this serine (eg. acyl-CoA). This is called mispriming of the 4'PP arm, and it is essential for the proper binding of the aminoacyl adenyl intermediate to the 4'PP arm that this mispriming is resolved. The improper recognition and binding of the wrong amino acid to the 4'PP arm is another form of mispriming. Thioesterase type II (TEII) is used to hydrolyze the misprimed 4'PP arm and produces an -SH end on the 4'PP arm.(figure 5) This stand-alone enzyme has to be able to recognize all the PCPs and be able to cleave off the acyl or peptidyl groups that are misprimed, without hydrolysing the correctly forming peptide chain. Research showed that the the active side of this enzyme consists of 3 functional amino acids and one amino acid essential for integrity of the structure.(Linne, Schwarzer et al. 2004) This structure of 3 catalytic amino acids is comparable to that found in the thioesterase type I domain.

5. NRPS function

5.1 Modules

The general layout of the domains of a NRPS follows a regular pattern. The domains are ordered in modules that each are responsible for recognition, binding and activation of a single building block of the final product of the NRPS. Modules comprise of a single A, PCP and C domain, other domains are integrated into modules when that specific activity is necessary for the correct formation of the peptide. The first module of a NRPS, called the initiation module contains only an A and PCP domain. The following modules follow a C-A-PCP motive, and the last module often contains a terminating TE domain. The substrate bound by the A domain is transferred by the PCP domain using the 4'PP arm. This arm is a flexible carbon chain, that is used to shuttle the growing peptide chain from one module to the next. This shuttling activity is often described as a 'swinging arm'. The modules can recognize and bind substrates at the same time, and the swinging arm is used to transfer the first substrate to the second and the subsequently formed peptide to the third substrate ect. (Hur, Vickery et al. 2012) (Ali 2013, Samol 2015) (figure 2)

5.2 Conformational changes in domain reactions

The PCP domain has three different conformations, A(apo), H (holo) and A/H. The A/H form is shared with both the apo and holo form. The 4'PP transferases are known to only be able to attack the A conformation, and thereby triggering the conformational change to the H form. The H conformation can transform into the A/H form, which generates the 100^{o} swing of the 4'PP arm. (see figure 3)
When the PCP is in the holo form, and any mispriming is resolved by TEII activity the PCP is ready to

receive an activated substrate from the upstream A domain. This A domain is known to have 2 different conformations which are performed by a small domain that acts like a lid. This lid domain is able to move 140° and starts in the open position. In this conformation the recognizing binding pocket of the A domain is open to the medium and substrate can bind in this state. When substrate has been recognized and bound, the first half reaction takes place and the aminoacyl adenyl intermediate is formed. Next the lid domain makes the 140° swing and goes into closed conformation. The pyrophosphate that remains from the previous reaction is released and the intermediate is transferred to the 4' PP arm of the downstream PCP domain in the second half reaction. (Koglin, Mofid et al. 2006, Strieker, Tanovic et al. 2010)(figure 5)

The structure of the C domain has been revealed and shows a V shaped structure with the active side at the tip of the V. The two PCP domains are located to either side of the structure, with the 4'PP arms extending into the middle towards the active side. A conserved histidine is thought to be the catalytic center and is located in a conserved His motive. (Hur, Vickery et al. 2012) (figure 4)

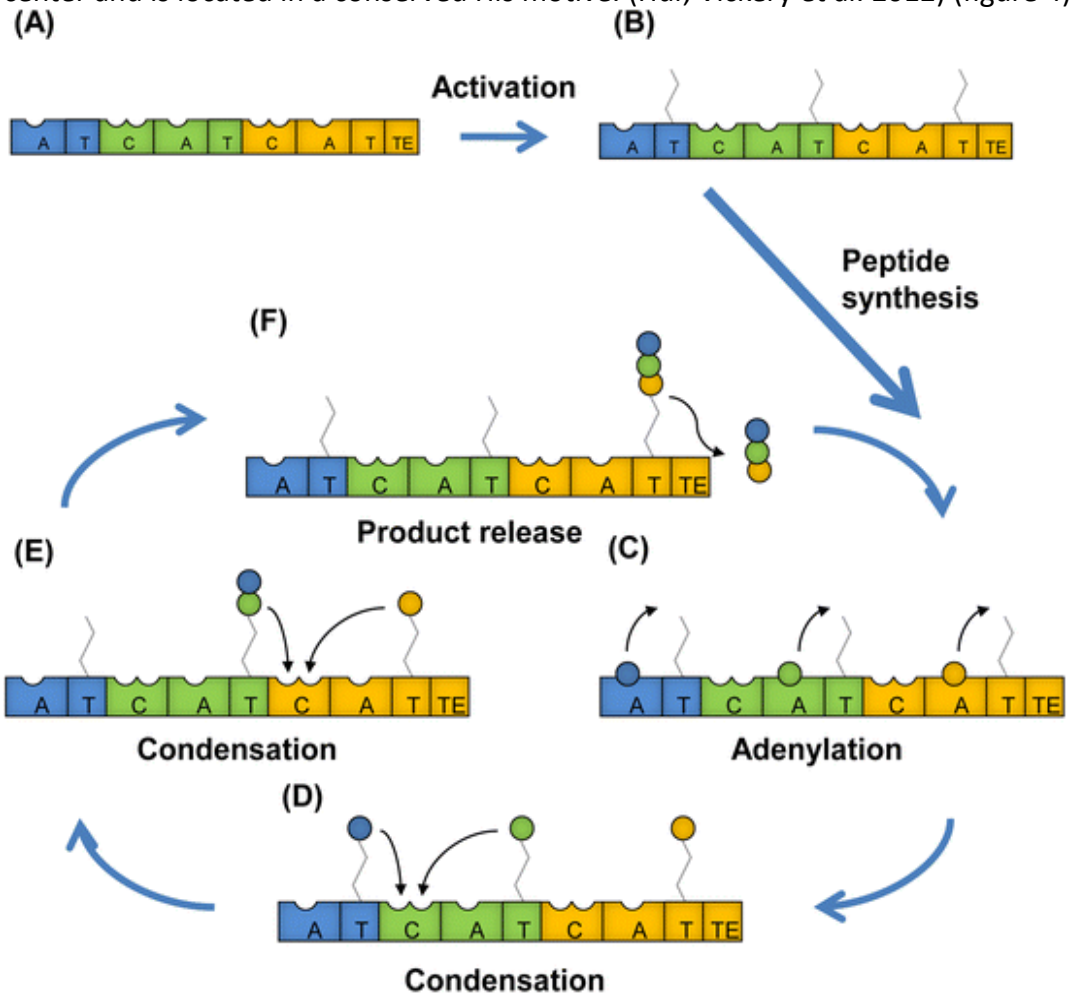

figure 2. Work flow of peptide synthesis in NRPSs. (A) shows the unactivated NRPS without the attached 4'PP arms that are present in (B). (C) shows the attached substrates to each A domain, and the attachment of the substrate to the 4'PP arm. In (D) these substrates are brought together at the C domain by the 'swinging arms' of the T/PCP domain. There they are bound together and transferred to the next C domain (E). This continues with all the modules and ultimately the final peptide is released by the TE domain (F). *source: (Calcott and Ackerley 2014)*

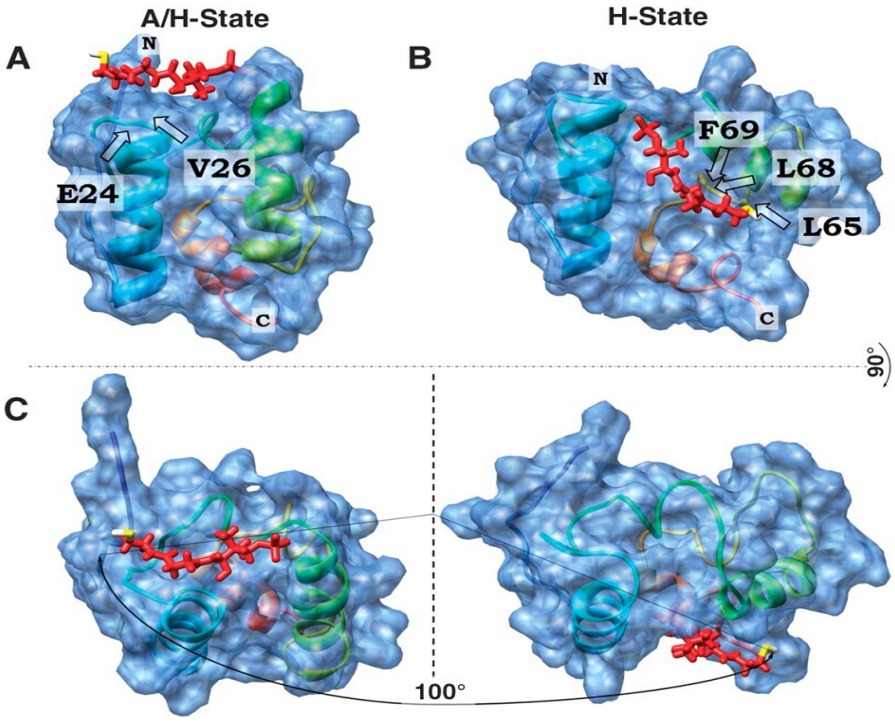

figure 3. Conformational changes in the PCP domain. (A) shows the A/H state with the 4'PP arm shown in red oriented towards the N-terminus of the domain. The H state (B) shows the 4'PP arm oriented towards the C terminus. (C) shows the difference in conformation and the swing of the 4'PP arm of 100°. *source: (Koglin, Mofid et al. 2006)*

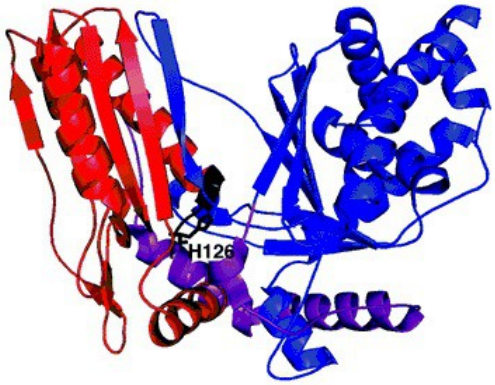

figure 4. structure of a free C domain of VibH, from *Vibrio cholerae*. The N-terminal (red) and C-terminal (blue) subdomains are oriented in a V shape with a linker region (purple). The catalytic site is located in the middle of the domain (black) and consists of a His motive with one identified catalytic histidine residue. The up- and downstream PCP domains are situated next to the N-terminal (upstream) and C-terminal (downstream) subdomains, with their 4'PP arms able to reach towards the catalytic site. *Source:* (Hur, Vickery et al. 2012)

6. Engineering of NRPS

NRPSs are prime candidates for modification in the search for new bioactive compounds, primarily new antibiotics. Since microorganisms are developing resistances to currently known antibiotics all the time, these novel antibiotics are highly sought after. NRPSs are useful vectors in this search because of their modular composition. And one way to develop novel compounds is the exchange of domains of NRPS to produce a new peptide.

6.1 A domain substitution

The A domain is the first and obvious target for substitution because of its recognizing function. On the genetic level the new A domain is inserted at the location of the old. This was performed on the surfactin synthetases in *Bacillus subtilus.* (Stachelhaus, Schneider et al. 1995) Different A domains were inserted in the SrfC gene, and the production of peptides by these chimeras was measured. But the yields of these modified NRPSs was much lower than the wild type SrfC. This is most likely due to the recognizing function of the C domain, which was not substituted in this study. So the A domain will recognize the new substrate, but the synthesis will be severely slowed by the C domain that cannot process this new substrate effectively. Also the introduction of a new A domain will most likely disrupt the interdomain interactions. The new domain basically doesn't 'fit' its surroundings. (Calcott and Ackerley 2014) A way to circumvent this C domain selectivity is to target the initiation module of a NRPS. These lack a C domain, and thus are not effected by its negative influence. This does however limit the possibilities of substitution significantly.

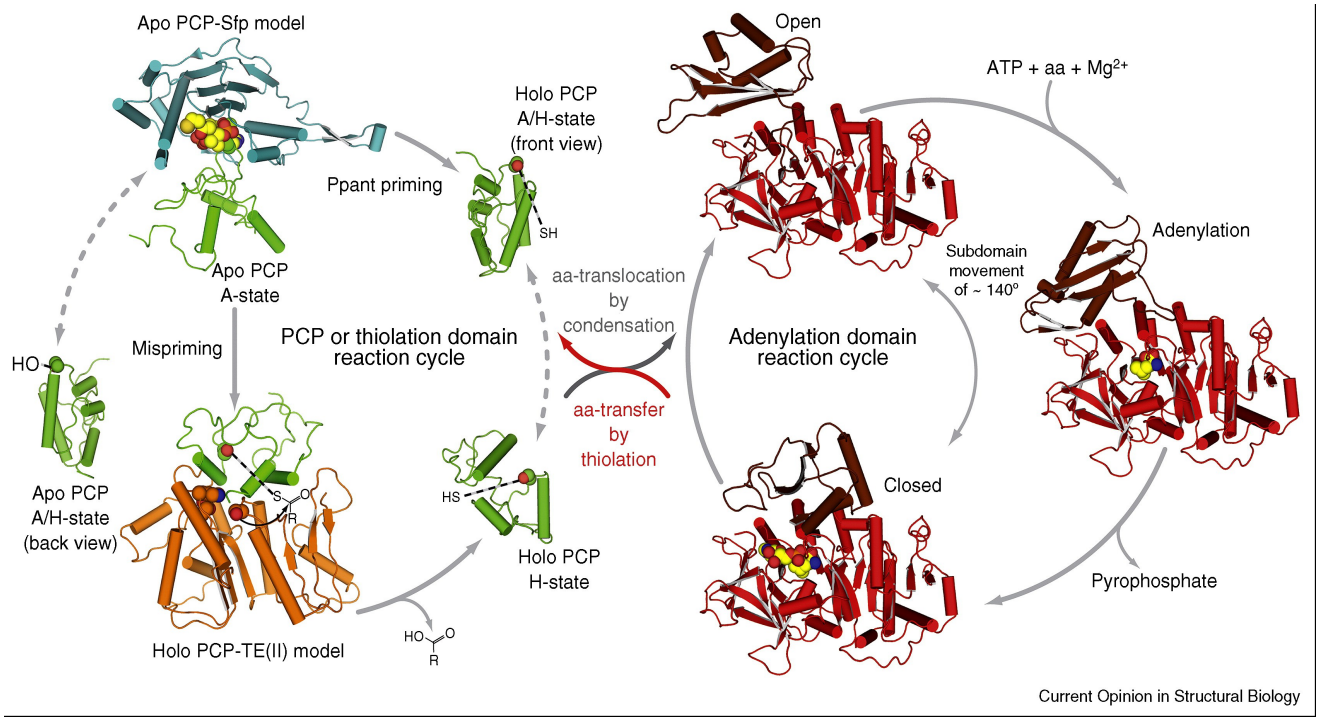

figure 5. Reactions in PCP and A domains. The reaction starts with the inactive apo state of the PCP domain (left hand, green) This state is activated by 4'PP transferases to the holo state by attaching the 4'PP arm to the PCP domain. When the domain is misprimed by the addition of an acylated 4'PP arm the TEII domain can hydrolyze this (left hand, orange) and form the proper holo PCP. The A domain (right hand, red) can be found in three conformations. The open conformation can take up a recognized substrate plus ATP and form an aminoacyl adenyl intermediate by adenylation. The pyrophosphate is released and the domain goes into the closed formation to protect the intermediate. Next the intermediate is transferred to the PCP when it is in the holo conformation. *(source: Strieker, 2010)*

6.2 C/A domain and whole module substitution

The problem of C domain selectivity and domain interface problems can be countered by substitution of the combined C and A domain, or by substituting the whole module. Especially in bacteria this is effective because of the COM-linkers. A whole module with COM-linkers designed to match the other NRPS proteins has been shown to influence the success of incorporation. (Calcott and Ackerley 2014)

An example of module exchange is the work done on the daptomycin producing NRPS from *Streptomyces roseosporus*. (Nguyen, Ritz et al. 2006) Single and multiple substitutions were made, and 15 different new NRPS sequences were obtained. Some of the new peptides were shown to be as active as wild type daptomycin, and one peptide even showed increased antibacterial activity.

Whole module substitution seems to be the most effective method to obtain novel peptides by substituting domains, but it is important to note that the yield of these experiments was mostly lower than wild type. So multiple substitutions will become increasingly difficult, because of the compounding effect of these lower yields. (Calcott and Ackerley 2014)

6.3 A domain code modification

Another way to manipulate the NPRS into the production of novel peptides is the modification of the recognizing residues in the A domain, the A domain code. This circumvents the problem of the sterical strain by introducing a new domain, but still has the problem of C domain recognition. The software for product prediction discussed earlier can also be used to develop the necessary mutations to obtain a new compound. Another method is the mimicking of the binding pocket of a known A domain. The third way is to introduce random mutations to the binding pocket residues, and scan for bioactive peptides. (Calcott and Ackerley 2014) These methods have been effective in producing new products while maintaining the structure of the NRPS, and thus do not have the problems related to breaking the interdomain connections.

6.4 Silent gene cluster activation

Many of the NRPSs found in the genomes of sequenced bacteria and fungi are not expressed under standard conditions, and therefore have not been seen in the laboratory. Efforts have been made to activate these silent NRPSs, and to test them for novel peptide products.(Brakhage and Schroeckh 2011), Samol 2015) Methods used to express these genes are manipulating histone regulating genes to release the NRPS gene from heterochromatin, and using transcription factors associated with the target NRPS and overexpressing them. These methods revealed many new proteins and peptides, and research on these new proteins is ongoing.

7. Conclusion

Nonribosomal peptide synthetases are large multidomain proteins responsible for many bioactive compounds. They are found in many bacteria and fungi and are responsible for a large portion of these organisms secondary metabolites. These compounds are useful for the organism and are linked to many biological functions. Like cell division, chemical signaling, cell defense and many others. But for humans these compounds are also very interesting because of their many anti-fungal, anti-bacterial and anti-cancer activities. NRPSs utilize many different domains, and use these in modular fashion to recognize, activate and bind the different building blocks of the produced peptide together. The main module consists of an Adenylation, Peptide Carrier Protein and a Condensation domain. Many other tailoring domains are used to produce the proper peptide. Recognition is handled internally by the A domains, and uses ~10 conserved recognizing residues called the nonribosomal code. Much work has been done to obtain novel peptides from NRPSs, mainly to find new antibiotics. Future research into these new peptides and NRPSs may reveal new antibiotic compounds that can be used in medicine and hopefully save lives.

8. References

Ali, H. (2013). "Nonribosomal peptide synthestases and secondary metabolite production in *Penicillium chrysogenum*."

Belshaw, P. J., C. T. Walsh and T. Stachelhaus (1999). "Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis." Science **284**(5413): 486-489.

Caboche, S., V. Leclere, M. Pupin, G. Kucherov and P. Jacques (2010). "Diversity of monomers in nonribosomal peptides: towards the prediction of origin and biological activity." J Bacteriol **192**(19): 5143-5150.

Calvo, A. M., R. A. Wilson, J. W. Bok and N. P. Keller (2002). "Relationship between secondary metabolism and fungal development." Microbiol Mol Biol Rev **66**(3): 447-459, table of contents.

Conti, E., N. P. Franks and P. Brick (1996). "Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes." Structure **4**(3): 287-298.

Dorrestein, P. C., E. Yeh, S. Garneau-Tsodikova, N. L. Kelleher and C. T. Walsh (2005). "Dichlorination of a pyrrolyl-S-carrier protein by FADH2-dependent halogenase PltA during pyoluteorin biosynthesis." Proceedings of the National Academy of Sciences of the United States of America **102**(39): 13843-13848.

Eriani, G., G. Dirheimer and J. Gangloff (1990). "Aspartyl-tRNA synthetase from Escherichia coli: cloning and characterisation of the gene, homologies of its translated amino acid sequence with asparaginyl- and lysyl-tRNA synthetases." Nucleic Acids Res **18**(23): 7109-7118.

Gaudelli, N. M. and C. A. Townsend (2014). "Epimerization and substrate gating by a TE domain in beta-lactam antibiotic biosynthesis." Nat Chem Biol **10**(4): 251-258.

Gribble, G. W. (2004). "Natural Organohalogens: A New Frontier for Medicinal Agents?" Journal of Chemical Education **81**(10): 1441.

Grünewald, J. and M. A. Marahiel (2013). Chapter 21 - Nonribosomal Peptide Synthesis. Handbook of Biologically Active Peptides (Second Edition). A. J. Kastin. Boston, Academic Press**:** 138-149.

Haslinger, K., M. Peschke, C. Brieke, E. Maximowitsch and M. J. Cryle (2015). "X-domain of peptide synthetases recruits oxygenases crucial for glycopeptide biosynthesis." Nature **521**(7550): 105-109.

Hopwood, D. A. (1997). "Genetic Contributions to Understanding Polyketide Synthases." Chemical Reviews **97**(7): 2465-2498.

Hornbogen, T., S.-P. Riechers, B. Prinz, J. Schultchen, C. Lang, S. Schmidt, C. Mügge, S. Turkanovic, R. D. Süssmuth, E. Tauberger and R. Zocher (2007). "Functional Characterization of the Recombinant N-Methyltransferase Domain from the Multienzyme Enniatin Synthetase." ChemBioChem **8**(9): 1048-1054.

Hur, G. H., C. R. Vickery and M. D. Burkart (2012). "Explorations of catalytic domains in non-ribosomal peptide synthetase enzymology." Nat Prod Rep **29**(10): 1074-1098.

Kohli, R. M. and C. T. Walsh (2003). "Enzymology of acyl chain macrocyclization in natural product biosynthesis." Chem Commun (Camb)(3): 297-307.

Laursen, B. S., H. P. Sorensen, K. K. Mortensen and H. U. Sperling-Petersen (2005). "Initiation of protein synthesis in bacteria." Microbiol Mol Biol Rev **69**(1): 101-123.

Lee, T. V., L. J. Johnson, R. D. Johnson, A. Koulman, G. A. Lane, J. S. Lott and V. L. Arcus (2010). "Structure of a eukaryotic nonribosomal peptide synthetase adenylation domain that activates a large hydroxamate amino acid in siderophore biosynthesis." J Biol Chem **285**(4): 2415-2427.

Li, J. and S. E. Jensen (2008). "Nonribosomal Biosynthesis of Fusaricidins by Paenibacillus polymyxa PKB1 Involves Direct Activation of a d-Amino Acid." Chemistry & Biology **15**(2): 118-127.

Linne, U., D. Schwarzer, G. N. Schroeder and M. A. Marahiel (2004). "Mutational analysis of a type II thioesterase associated with nonribosomal peptide synthesis." Eur J Biochem **271**(8): 1536-1545.
Marahiel, M. A. (2009). "Working outside the protein-synthesis rules: insights into non-ribosomal peptide synthesis." J Pept Sci **15**(12): 799-807.
Marshall, C. G., M. D. Burkart, T. A. Keating and C. T. Walsh (2001). "Heterocycle Formation in Vibriobactin Biosynthesis: Alternative Substrate Utilization and Identification of a Condensed Intermediate†." Biochemistry **40**(35): 10655-10663.
Milne, C., A. Powell, J. Jim, M. Al Nakeeb, C. P. Smith and J. Micklefield (2006). "Biosynthesis of the (2S,3R)-3-Methyl Glutamate Residue of Nonribosomal Lipopeptides." Journal of the American Chemical Society **128**(34): 11250-11259.
Quadri, L. E. N., P. H. Weinreb, M. Lei, M. M. Nakano, P. Zuber and C. T. Walsh (1998). "Characterization of Sfp, a Bacillus subtilis Phosphopantetheinyl Transferase for Peptidyl Carrier Protein Domains in Peptide Synthetases." Biochemistry **37**(6): 1585-1595.
Rausch, C., T. Weber, O. Kohlbacher, W. Wohlleben and D. H. Huson (2005). "Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs)." Nucleic Acids Res **33**(18): 5799-5808.
Samel, S. A., P. Czodrowski and L. O. Essen (2014). "Structure of the epimerization domain of tyrocidine synthetase A." Acta Crystallogr D Biol Crystallogr **70**(Pt 5): 1442-1452.
Schoenafinger, G., N. Schracke, U. Linne and M. A. Marahiel (2006). "Formylation Domain: An Essential Modifying Enzyme for the Nonribosomal Biosynthesis of Linear Gramicidin." Journal of the American Chemical Society **128**(23): 7406-7407.
Stachelhaus, T., H. D. Mootz, V. Bergendahl and M. A. Marahiel (1998). "Peptide bond formation in nonribosomal peptide biosynthesis. Catalytic role of the condensation domain." J Biol Chem **273**(35): 22773-22781.
Stachelhaus, T., H. D. Mootz and M. A. Marahiel (1999). "The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases." Chem Biol **6**(8): 493-505.
Stachelhaus, T. and C. T. Walsh (2000). "Mutational Analysis of the Epimerization Domain in the Initiation Module PheATE of Gramicidin S Synthetase†." Biochemistry **39**(19): 5775-5787.
Stein, T., J. Vater, V. Kruft, B. Wittmann-Liebold, P. Franke, M. Panico, R. Mc Dowell and H. R. Morris (1994). "Detection of 4'-phosphopantetheine at the thioester binding site for L-valine of gramicidinS synthetase 2." FEBS Lett **340**(1-2): 39-44.
Strieker, M., A. Tanovic and M. A. Marahiel (2010). "Nonribosomal peptide synthetases: structures and dynamics." Curr Opin Struct Biol **20**(2): 234-240.
von Dohren, H. (2004). "Biochemistry and general genetics of nonribosomal peptide synthetases in fungi." Adv Biochem Eng Biotechnol **88**: 217-264.
Weber, G., K. Schorgendorfer, E. Schneider-Scherzer and E. Leitner (1994). "The peptide synthetase catalyzing cyclosporine production in Tolypocladium niveum is encoded by a giant 45.8-kilobase open reading frame." Curr Genet **26**(2): 120-125.
Winkelmann, G. (2007). "Ecology of siderophores with special reference to the fungi." Biometals **20**(3-4): 379-392.

Ali, H. (2013). "Nonribosomal peptide synthestases and secondary metabolite production in *Penicillium chrysogenum*."

Brakhage, A. A. and V. Schroeckh (2011). "Fungal secondary metabolites - strategies to activate silent gene clusters." Fungal Genet Biol **48**(1): 15-22.
Hur, G. H., C. R. Vickery and M. D. Burkart (2012). "Explorations of catalytic domains in non-ribosomal peptide synthetase enzymology." Nat Prod Rep **29**(10): 1074-1098.
Samol, M. (2015). "Genomic Wake-Up Call: Activating Silent Biosynthetic Pathways for Novel Metabolites in Penicillium chrysogenum."