Chemical Science

EDGE ARTICLE

View Article Online
View Journal | View Issue



Cite this: *Chem. Sci.*, 2019, **10**, 9466

All publication charges for this article have been paid for by the Royal Society of Chemistry

# A proof-reading mechanism for non-proteinogenic amino acid incorporation into glycopeptide antibiotics†

Milda Kaniusaite, [ab] Julien Tailhades, [ab] Edward A. Marschall, [ab] Robert J. A. Goode, [ac] Ralf B. Schittenhelm [ac] and Max J. Cryle *[ab]

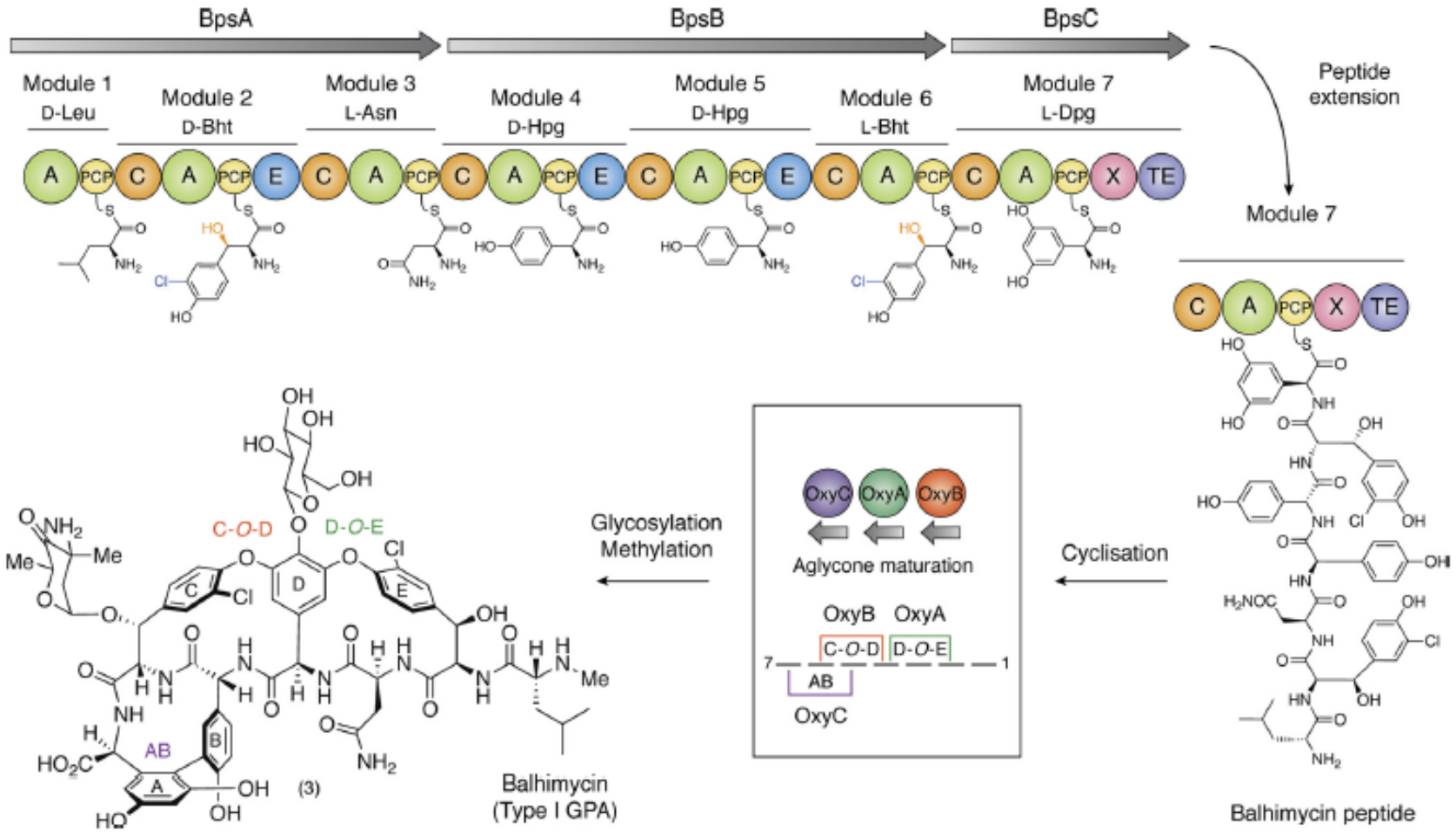



**Figure 3. Balhimycin biosynthesis.** Peptide extension is achieved by the NRPS multimodular machinery, which comprises seven modules spread across three polypeptide chains. Chlorine atoms (*blue*) are added when the tyrosine is linked to the PCP domain in the NRPS by a halogenase, and the β-hydroxyl groups (*orange*) are added prior to activation of β-hydroxytyrosine by the main NRPS assembly line (*A,* adenylation domain; *C,* condensation domain; *PCP,* peptidyl carrier protein domain; *E,* epimerization domain; *X,* Oxy recruitment domain; *TE,* thioesterase domain; *Oxy enzymes,* cytochromes P450). *Bht,* β-hydroxytyrosine; *Hpg,* 4-hydrophenylglycine; *Dpg,* 3,5-dihydroxyphenylglycine.

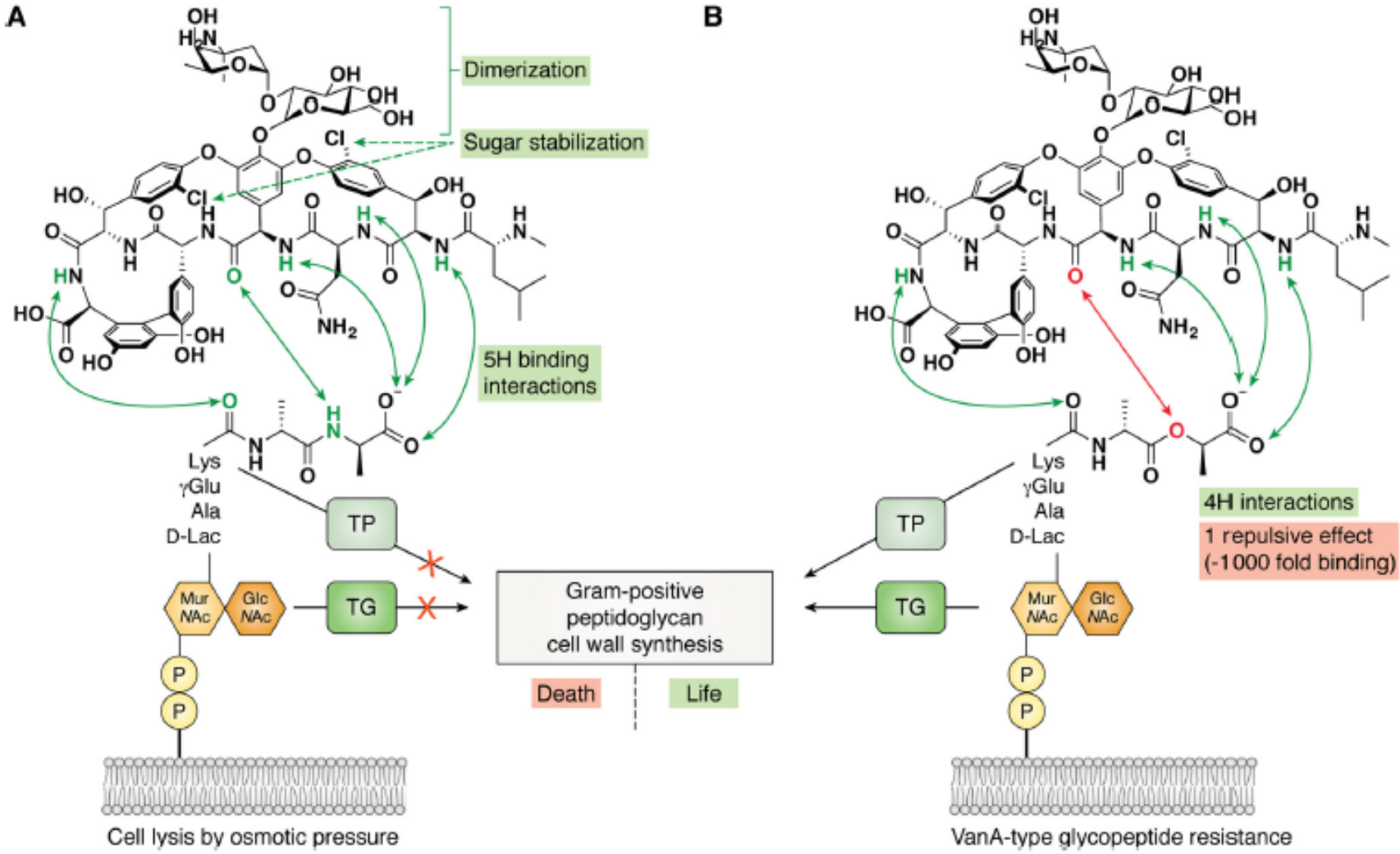



**Figure 1. Comparison of the interactions between GPAs and their lipid II target shown for vancomycin against sensitive (*A*) and resistant (*B*) bacteria.** In the case of sensitive bacteria, the interaction between the GPA and lipid II is centered on five hydrogen bonds between the peptide backbones of both compounds (*green arrows*), which result in complex formation that then inhibits the actions of cell wall synthesis enzymes (*TP*, transpeptidase; *TG*, transglycosylase) and leads to eventual cell lysis. Resistance mediated via the exchange of the final D-Ala moiety for a D-Lac leads to the loss of one hydrogen bond and replaces this with lone pair/lone pair repulsion (*red arrow*), leading to a loss of GPA-binding affinity of 3 orders of magnitude and rendering GPAs ineffective against such bacterial strains.

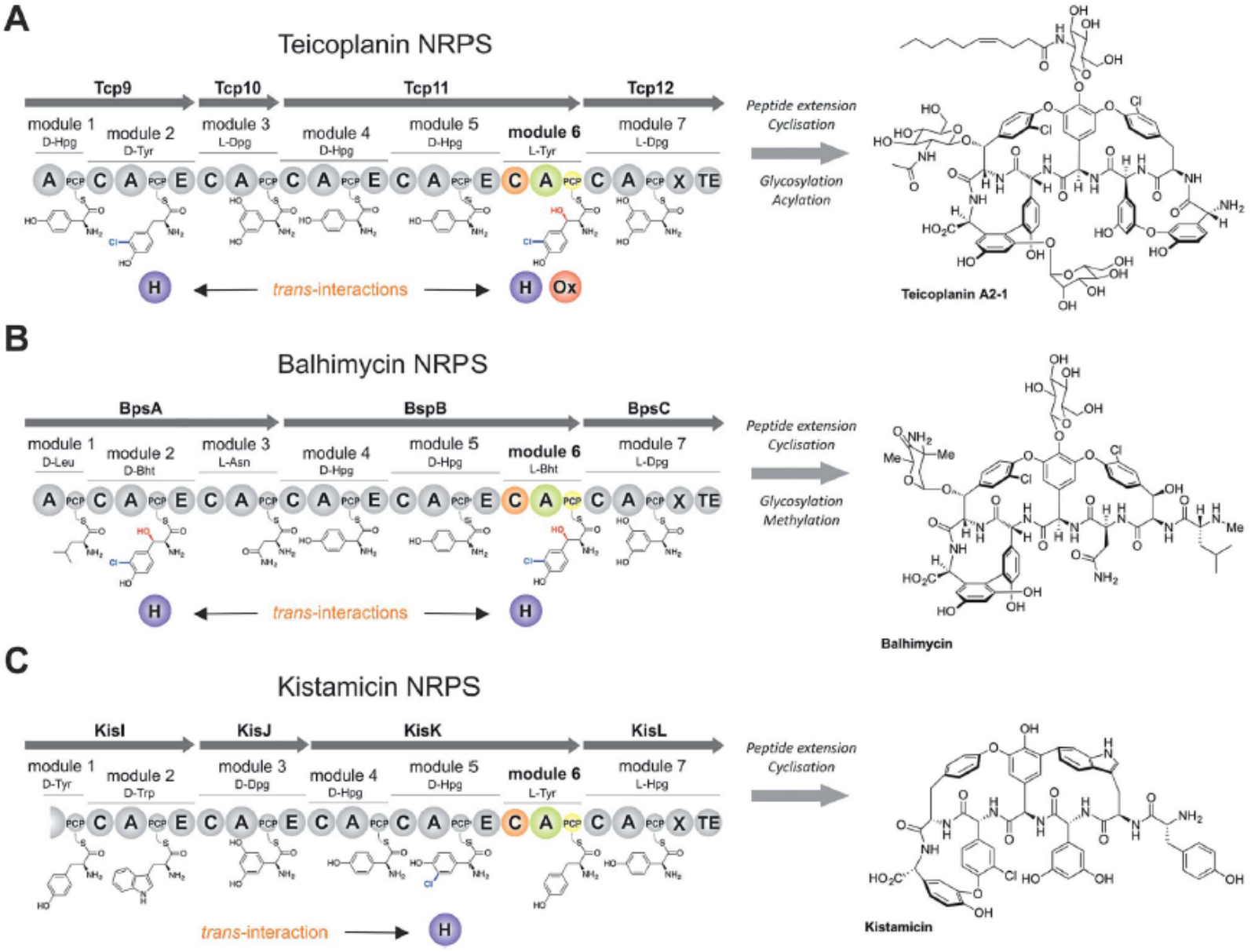



**Fig. 1** Non-ribosomal peptide biosynthesis of the glycopeptide antibiotics teicoplanin (A), the vancomycin-type GPA balhimycin (B) and kistamicin (C), concentrating on the NRPS proteins and *trans*-interacting enzymes. Module 6 (M6) of the NRPS machineries are shown in colour, the rest of the NRPS is indicated in grey. A – adenylation domain, C – condensation domain, PCP – peptidyl carrier protein domain, E – epimerisation domain, X – cytochrome P450 recruitment domain, TE – type I thioesterase domain H – flavin-dependent halogenase, Ox – non-heme iron oxygenase, Hpg – 4-hydroxyphenylglycine, Dpg – 3,5-dihydroxyphenylglycine, Tyr – tyrosine (**1**), Bht – β-hydroxytyrosine (**3**), Trp – tryptophan, Leu – leucine, Asn – asparagine.

# Obiettivo del lavoro

- Determinare i meccanismi di selettività alla base della capacità di incorporare in modo specifico diversi derivati della Tyr nel modulo 6 di NRPS che producono GPA
- Era già noto che:
  - Teicoplanina M6-A riconosce Tyr che poi viene modificata da idrossilasi e alogenasi per ottenere Cl-Bht che viene incorporata nel peptide
  - Balhimicina M6-A riconosce Bht che poi viene modificata da alogenasi per ottenere Cl-Bht che viene incorporata nel peptide
  - Kistamicina M6-A riconosce Tyr che viene incorporata nel peptide

# Come fare?

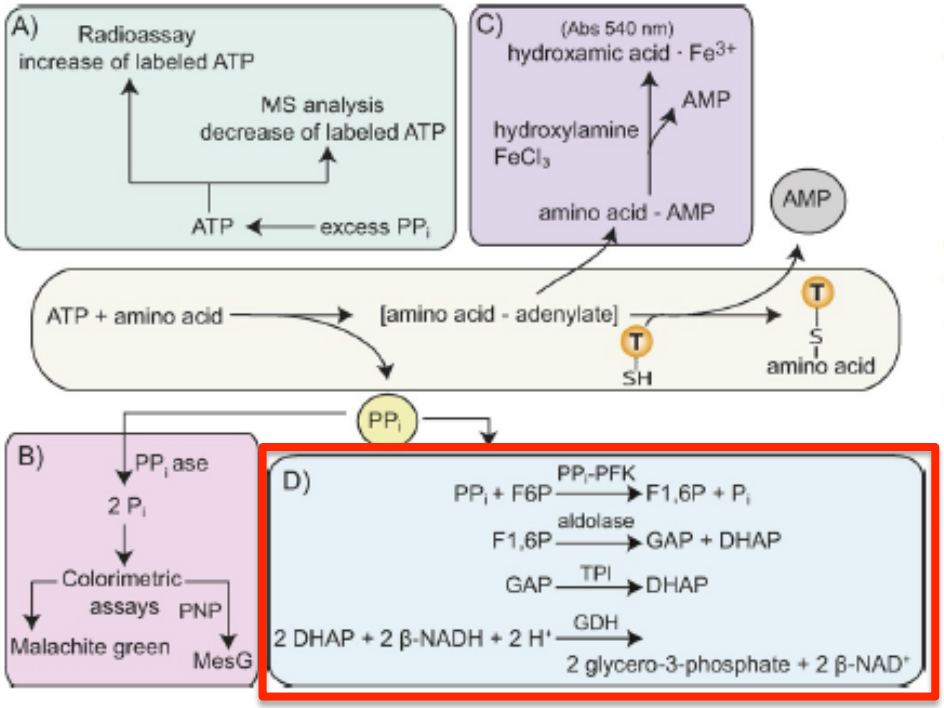
- Ottenere moduli M4, M5 e M6 separati e funzionali per ricostruire il processo di sintesi *in vitro*
- Analizzare la specificità di riconoscimento del dominio M6-A
  - attivazione di Tyr e analoghi della Tyr
- Analizzare la specificità di riconoscimento del dominio M5/6-C
  - formazione di un esapeptide dopo aver fornito precursori (Hpg, substrato di M4-A e M5-A) e Tyr o analoghi della Tyr (substrato M6-A) e aver attivato M3 con CoA-tripeptide
- Analizzare il ruolo di tioesterasi
  - capacità di ripristinare la sintesi dell'espapeptide corretto in presenza di una miscela di Tyr e analoghi della Tyr

# Metodologie

- Clonaggi per ottenere singoli moduli delle NRPS
- Espressione in E. coli e purificazione delle proteine ricombinanti
- Attivazione e saggi di attività enzimatica delle NRPS
- Analisi dei prodotti delle reazioni (LC-MS)

# Online Pyrophosphate Assay for Analyzing Adenylation Domains of Nonribosomal Peptide Synthetases

Tiia Kittilä,[a] Melanie Schoppet,[a] and Max J. Cryle*[a, b, c]




**Scheme 2.** A domain activity assays. A domains catalyze amino acid activation in two steps. The reaction uses ATP and produces $PP_i$ and AMP. A) ATP–$PP_i$ exchange assays are commonly used because of their sensitivity. B) Colorimetric phosphate assays detect $PP_i$ formation by cleaving $PP_i$ into two monophosphates. $PP_i$ase: pyrophosphatase; PNP: purine nucleoside phosphorylase: MesG: 2-amino-6-mercapto-7-methylpurine ribonucleoside. C) The hydroxylamine-trapping assay is useful for A domains that bind $PP_i$ tightly. The AMP intermediate is cleaved by using hydroxylamine and yields a very specific signal. D) The NADH/$PP_i$ coupled assay is suitable for analyzing A domain activity online, simply by mixing enzymes and substrates in assay buffer. F6P: D-fructose-6- phosphate; F1,6P:fructose-1,6-diphosphate; $PP_i$-PFK: phosphofructokinase (pyrophosphate dependent); GAP: glyceraldehyde-3-phosphate; DHAP: dihydroxyacetone phosphate; TPI: triosephosphate isomerase; GDH: glycererophosphate dehydrogenase.

Separazione E/C
Il modulo 5 è prodotto a livelli molto bassi!

I domini A sono funzionali
I domini PCP vengono attivati da Sfp con CoA o peptidyl-CoA

Separazione C/A
Tutte le proteine sono prodotte con rese adeguate

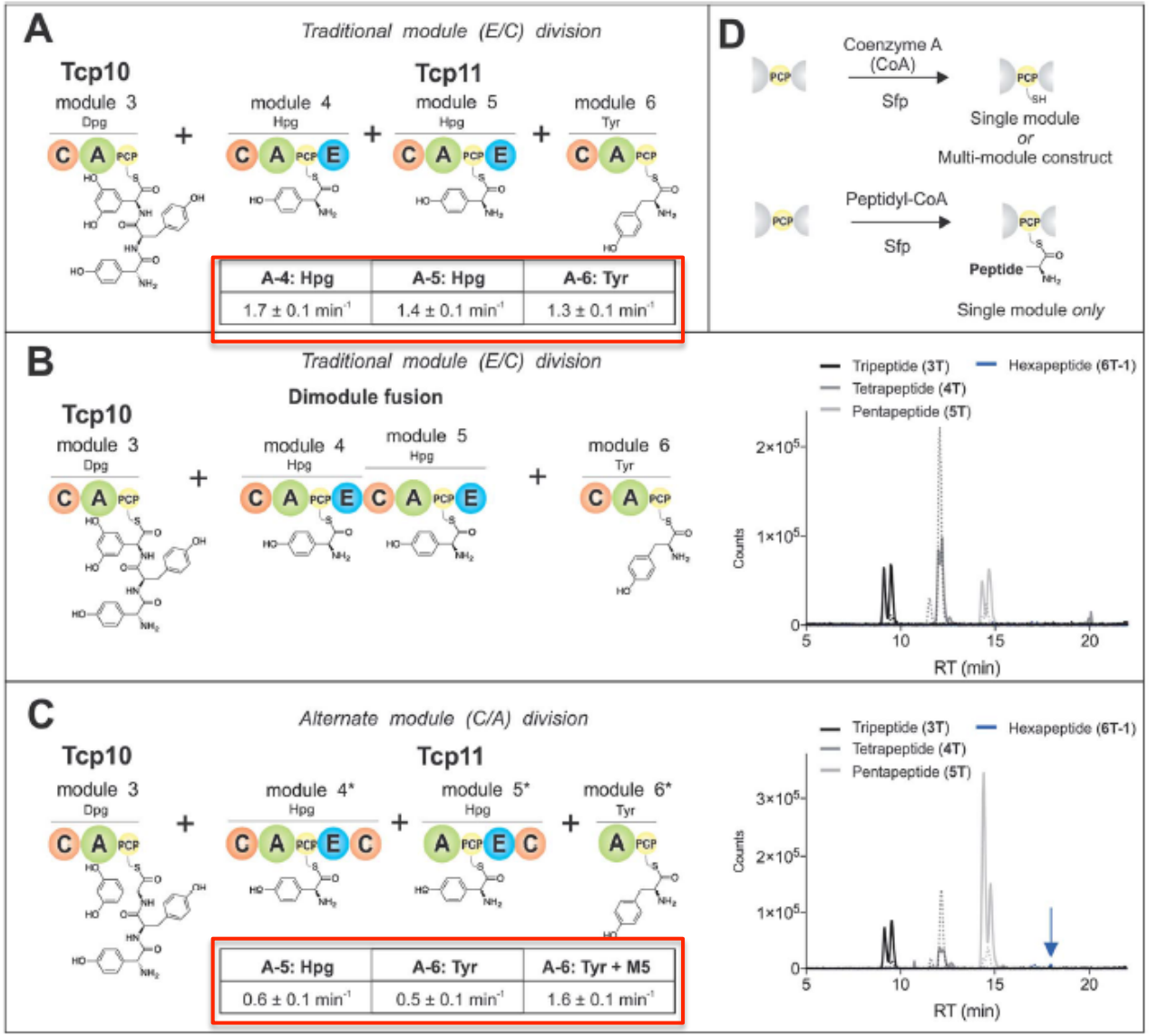



**Fig. 2** Reconstitution of peptide biosynthesis from the teicoplanin NRPS proteins Tcp10 and Tcp11, utilising two strategies to isolate individual modules 4–6 from Tcp11: either a C-A-PCP-E module architecture (A and B) or A-PCP-E-C architecture (C), together with the rationale behind the need for modularisation of the NRPS – the ability to load individual modules with peptide substrates using phosphopantetheinyl transferases (D). Rate of activation of the natural A-domain substrates for (A) and (C) were determined using a continuous, enzyme-coupled pyrophosphate detection assay; experiments performed in triplicate and standard deviation indicated. Peptide biosynthesis was reconstituted from tripeptide **3T** loaded on M3, together with ATP, 4-Hpg and Tyr (**1**) using both the C-A-PCP-E module architecture and an M4–M5 fusion (B) or the A-PCP-E-C architecture (C). Peptide products were determined by LCMS analysis (ESI, positive mode), with solid lines indicating methylamide peptides (PCP-bound) and dashed lines indicating hydrolysed peptides (tripeptide **3T**: black line; tetrapeptide **4T**: dark grey line; pentapeptide **5T**: light grey line; hexapeptide **6T-1**: blue line). A – adenylation domain, C – condensation domain, PCP – peptidyl carrier protein domain, E – epimerisation domain, Hpg – 4-hydroxyphenylglycine, Tyr – tyrosine (**1**).

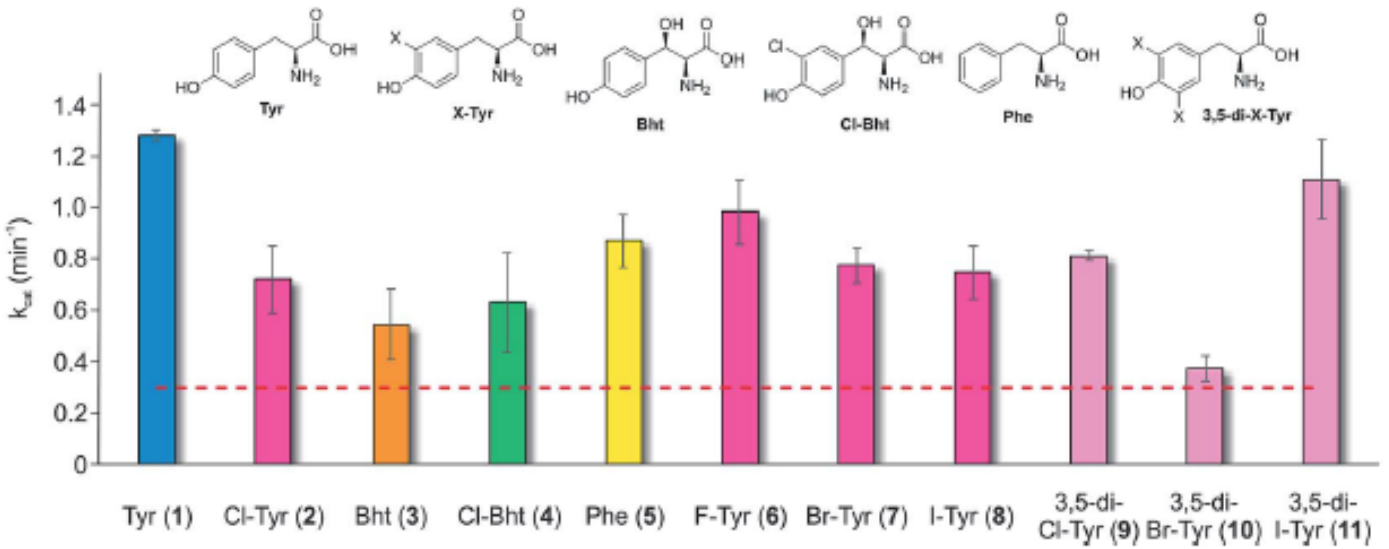



**Fig. 3** Module 6 A-domain amino acid activation rates determined for the teicoplanin NRPS. Amino acids tested as substrates include tyrosine (**1**, blue), halogenated tyrosine residues (**2** and **6–8**, magenta), di-halogenated tyrosine residues (**9–11**, pale pink), Bht (orange, **3**), Cl-Bht (green, **4**) and phenylalanine (yellow, **5**). Triplicate experiments, standard deviations indicated. Dotted line indicates the method detection limit.

Attivazione di Tyr e Tyr modificate da parte di M6-A

➔ Tyr è il substrato naturale ma la specificità di riconoscimento è piuttosto bassa

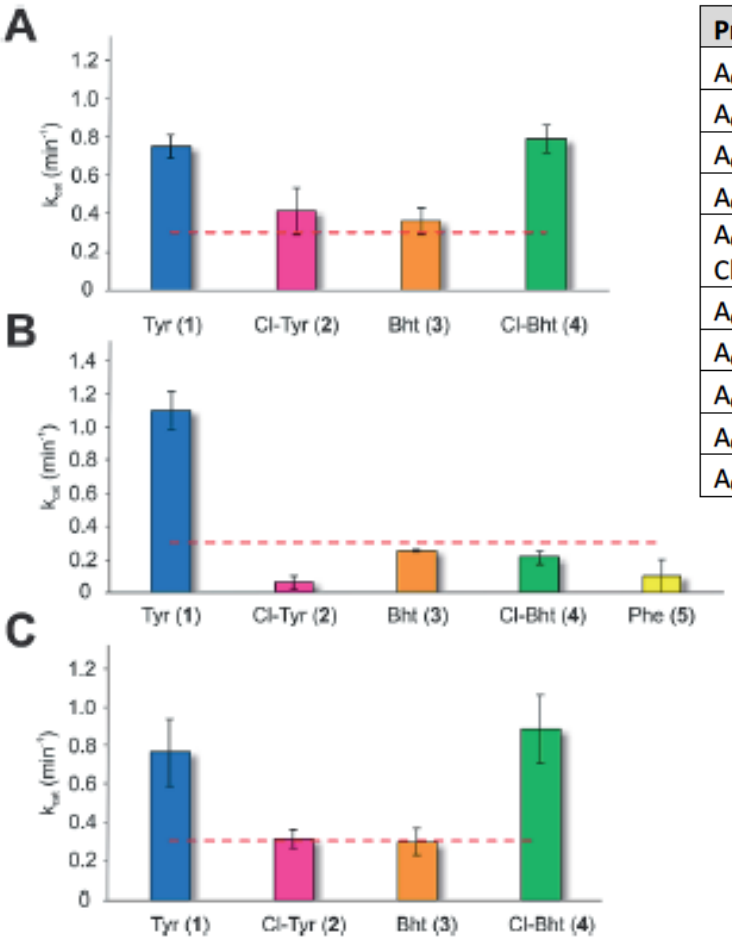



Fig. 4 Module 6 A-domain amino acid activation rates determined for the balhimycin (A) and kistamicin (B) NRPS together with the kistamicin double mutant modelled on the teicoplanin A-domain pocket (C). Aromatic amino acids tested include tyrosine (blue, **1**), 3-chlorotyrosine (magenta, **2**), Bht (orange, **3**), Cl-Bht (green, **4**) and phenylalanine (yellow, **5**). Triplicate experiments, standard deviations indicated. Dotted line indicates the method detection limit.

| Protein[ref] | $A_6$ selectivity code[1] | | | | | | | | | | Description |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $A_6$ Teicoplanin[2] | D | A | S | T | I | A | G | V | C | K | Tyr permissive pocket (tei-type) |
| $A_6$ UK-68597[3] | D | A | S | T | I | A | G | V | C | K | |
| $A_6$ Balhimycin[4] | D | A | S | T | L | G | A | I | C | K | Tyr permissive pocket (bal-type) |
| $A_6$ Vancomycin[5] | D | A | S | T | L | G | A | I | C | K | |
| $A_6$ Chloroerymomycin[6] | D | A | S | T | L | G | A | I | C | K | |
| $A_6$ Arylomycin[7] | D | A | S | T | V | A | A | V | C | K | Tyr specific pocket |
| $A_6$ A40926[8] | D | A | S | T | V | A | A | V | C | K | |
| $A_6$ Complestatin[9] | D | A | S | T | V | A | A | V | C | K | |
| $A_6$ Kistamicin[10] | D | A | S | T | V | A | A | V | C | K | |
| $A_6$ Kistamicin-mut* | D | A | S | T | I | A | G | V | C | K | Tyr permissive pocket |

Attivazione di Tyr e Tyr modificate da parte di M6-A balhimicina (A)

➔ specificità di riconoscimento piuttosto bassa

M6-A kistamicina (B)

➔ specificità di riconoscimento elevata

M6-A kistamicina mutato (C)

➔ conversione della specificità di riconoscimento

Ricostituzione dell'estensione del tripeptide caricato su M3

L'esapeptide si forma pochissimo fornendo a M6 Tyr (B) e progressivamente meglio con Cl-Tyr (C), Bht (D) e Cl-Bht (E)

Il dominio M5/6-C è selettivo per Cl-Bht l'amminoacido correttamente modificato

NOTA: l'esperimento è possibile perché M6-A è poco selettivo e può accettare le Tyr modificate. *In vivo* le modifiche sono catalizzate da enzimi accessori (idrossilasi e alogenasi) che non è stato possibile utilizzare *in vitro* perché non si producevano in forma attiva.

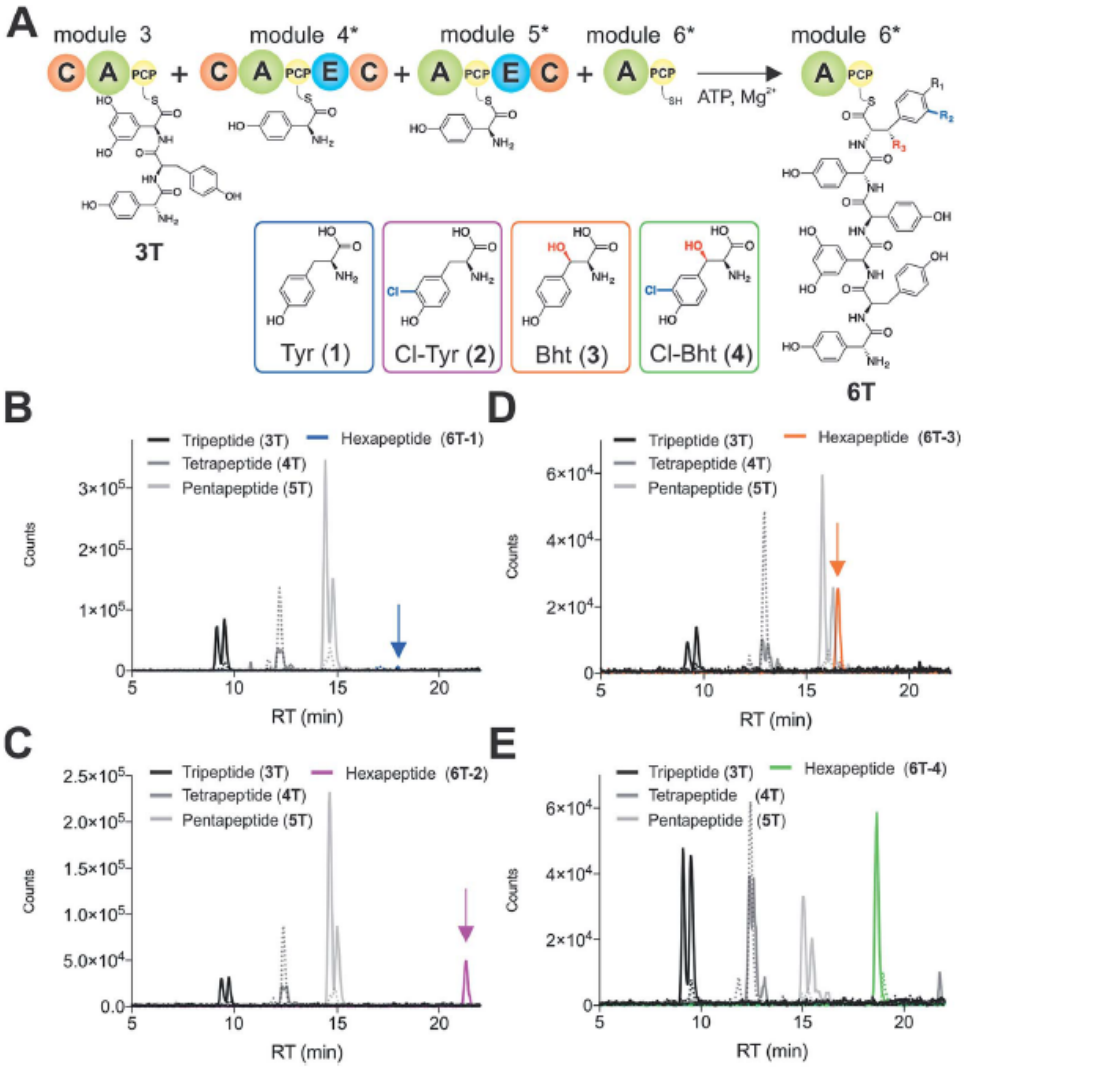



Fig. 5 Reconstitution of tripeptide extension using the separated modules (A-PCP-E-C architecture) from teicoplanin biosynthesis M3–M6 (Tcp10/Tcp11) using different substrates for module 6 (A). Peptide biosynthesis reconstituted using ATP, tripeptide (**3T**)-loaded M3, 4-Hpg, plus tyrosine (**1**) (B), Cl-Tyr (**2**) (C), Bht (**3**) (D) and Cl-Bht (**4**) (E) as M6 substrates, and determined by LCMS analysis (ESI, positive mode) with solid lines indicating methylamide peptides (PCP-bound) and dashed lines indicating hydrolysed peptides (tripeptide **3T**: black line; tetrapeptide **4T**: dark grey line; pentapeptide **5T**: light grey line; hexapeptides **6T**: blue line (Tyr, **6T-1**), magenta line (Cl-Tyr, **6T-2**), orange line (Bht, **6T-3**) or green line (Cl-Bht, **6T-4**)). A – adenylation domain, C – condensation domain, PCP – peptidyl carrier protein domain, E – epimerisation domain, Tyr – tyrosine (**1**), Cl-Tyr – 3-chlorotyrosine (**2**), Bht – β-hydroxytyrosine (**3**), Cl-Bht – 3-chloro-β-hydroxytyrosine (**4**).

Ricostruzione dell'estensione del pentapeptide caricato su M5
L'ordine di preferenza dei substrati di M6-A da parte di M5/6-C è:
Cl-Bht
Bht
Cl-Tyr

*In vivo* quale enzima deve agire per primo su M6-PCP-Tyr: l'alogenasi o l'idrossilasi?

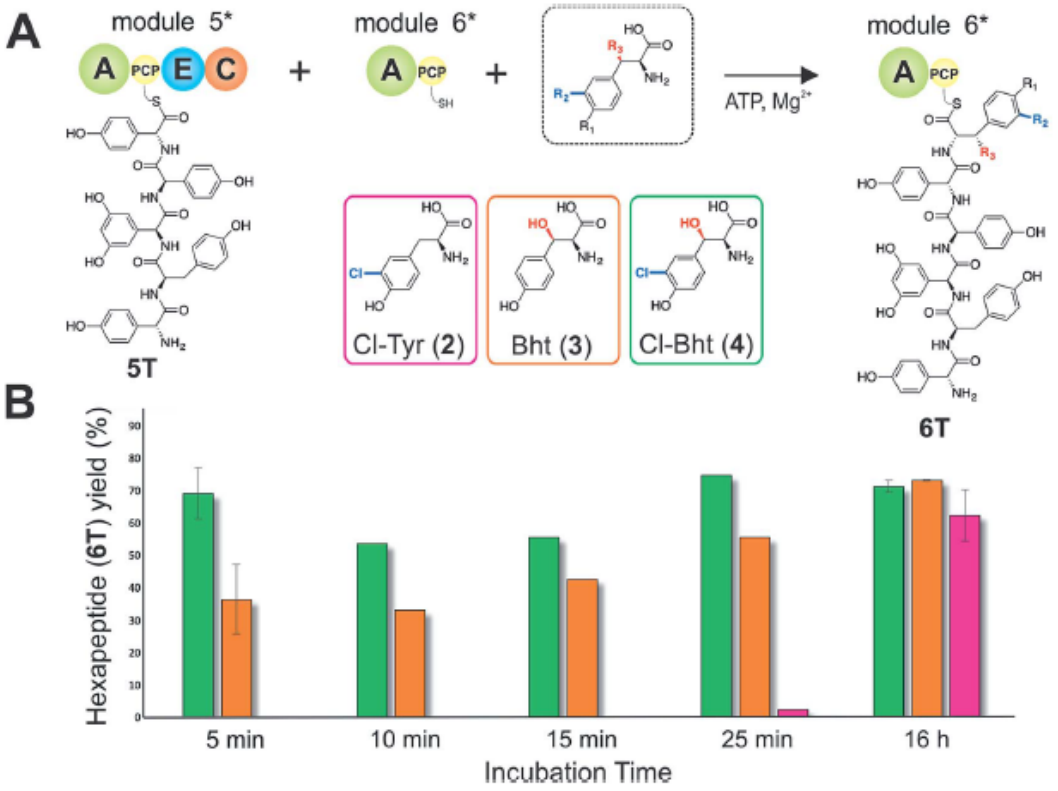
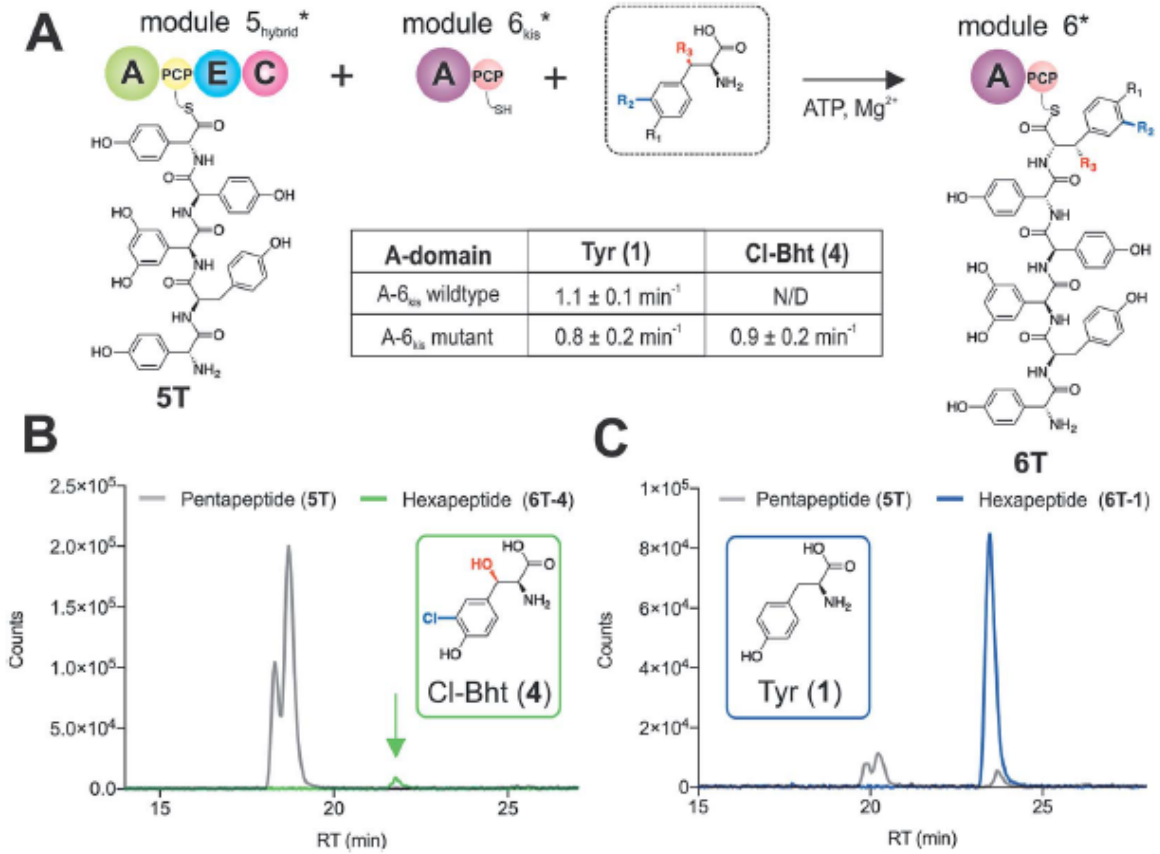



**Fig. 6** The rate of pentapeptide extension by teicoplanin module 6 using differently modified tyrosine residues as substrates (A). Hexapeptide biosynthesis reconstituted using ATP, pentapeptide (**5T**)-loaded M5, plus Cl-Tyr (**2**), Bht (**3**) and Cl-Bht (**4**) as M6 substrates, analysed by LCMS analysis (ESI, positive mode) at various time intervals (B). 5 min and 16 h experiments were performed in triplicate with the standard deviation indicated; other time points are the result of single experiments. A – adenylation domain, C – condensation domain, PCP – peptidyl carrier protein domain, E – epimerisation domain, Cl-Tyr – 3-chlorotyrosine (**2**), Bht – β-hydroxytyrosine (**3**), Cl-Bht – 3-chloro-β-hydroxytyrosine (**4**).

Riprogrammazione M5 ibrido teicoplanina (A-PCP-E)-kistamicina (C) e M6 kistamicina

L'esapeptide si forma solo con Tyr sia con M6-A wild type che con M6-A mutato per alterare la specificità di riconoscimento (fig. 4)

Il dominio M5/6-$C_{kis}$ è selettivo per Tyr




| A-domain | Tyr (1) | Cl-Bht (4) |
|---|---|---|
| A-$6_{kis}$ wildtype | 1.1 ± 0.1 $min^{-1}$ | N/D |
| A-$6_{kis}$ mutant | 0.8 ± 0.2 $min^{-1}$ | 0.9 ± 0.2 $min^{-1}$ |

**Fig. 7** Reconstitution of pentapeptide extension by exchanging the teicoplanin Cl-Bht specific M5/6 C-domain with the Tyr-specific M5/6 C-domain from kistamicin biosynthesis (A). Peptide biosynthesis reconstituted using ATP, pentapeptide (**5T**)-loaded M5 hybrid, plus Cl-Bht (**4**) (B) and Tyr (**1**) (C) as $M6_{kis}$ substrates, and determined by LCMS analysis (ESI, positive mode) with solid lines indicating methylamide peptides (PCP-bound) peptides (pentapeptide **5T**: light grey line; hexapeptides **6T**: blue line (Tyr, **6T-1**) or green line (Cl-Bht, **6T-4**)). A – adenylation domain, C – condensation domain, PCP – peptidyl carrier protein domain, E – epimerisation domain, Tyr – tyrosine (**1**), Cl-Tyr – 3-chlorotyrosine (**2**), Bht – β-hydroxytyrosine (**3**), Cl-Bht – 3-chloro-β-hydroxytyrosine (**4**).

Teicoplanina M6-A è specifico per Tyr ma è scarsamente selettivo mentre M5/6-C è selettivo per Cl-Bht

Come fa la sintesi di teicoplanina ad essere efficiente *in vivo* dove sono presenti diversi potenziali substrati alternativi per M6-A?

Ricostruzione dell'estensione del pentapeptide caricato su M5
L'esapeptide si forma in maniera ridotta se vengono fornite miscele di **Tyr**/Cl-Bht o **Phe**/Cl-Bht

L'efficienza di sintesi viene ripristinata dalla tioesterasi Tcp39

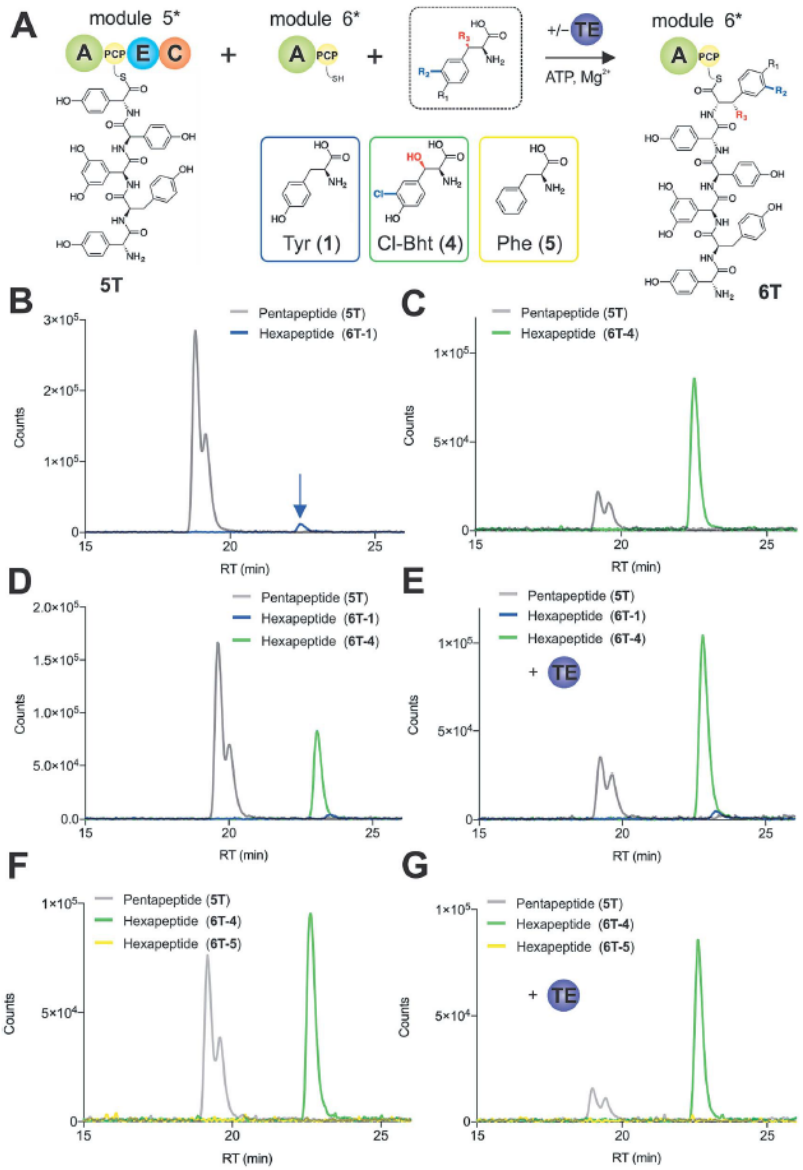



Fig. 8 Reconstitution of pentapeptide extension coupled with the actions of a type-II TE enzyme (A). Peptide biosynthesis reconstituted using ATP, pentapeptide (**5T**)-loaded M5, plus different combinations of possible amino acids as M6 substrates, as determined by LCMS analysis (ESI, positive mode) with solid lines indicating methylamide peptides (PCP-bound) peptides (pentapeptide **5T**: light grey line; hexapeptides **6T**: blue line (Tyr, **6T-1**), green line (Cl-Bht, **6T-4**) or yellow line (Phe, **6T-5**)). Results of peptide extension using tyrosine (**1**) alone (B), Cl-Bht (**4**) alone (C), an equimolar ratio of Tyr (**1**) and Cl-Bht (**4**) (D) as well as an equimolar ratio of Tyr (**1**) and Cl-Bht (**4**) together with the incorporation of the type-II TE enzyme Tcp39 (E). Results of peptide extension using an equimolar ratio of Phe (**5**) and Cl-Bht (**4**) (F) as well as an equimolar ratio of Phe (**5**) and Cl-Bht (**4**) together with the incorporation of the type-II TE enzyme Tcp39 shown (G). A – adenylation domain, C – condensation domain, PCP – peptidyl carrier protein domain, E – epimerisation domain, TE – type-II thioesterase, Tyr – tyrosine (**1**), Cl-Bht – 3-chloro-β-hydroxytyrosine (**4**), Phe – phenylalanine (**5**).

# In conclusione….

- Sistema composto da M3, M4, M5 e M6 isolati, purificati e attivati in vitro per analizzare i singoli passaggi biosintetici
- Il dominio M6-A non è selettivo (figura 3)
- Si forma l'esapeptide con elevata efficienza se M6 viene caricato con Cl-Bht (figura 5 e 6) ➔ il dominio M5/6-C è selettivo
- Saggi competitivi dimostrano l'efficacia della tioesterasi Tcp39 nel ripristinare la sintesi dell'esapeptide corretto in presenza di substrati M6-A alternativi (figura 8)

Schema dei filtri di selettività di M6-A e M5/6-C nei confronti degli enzimi di modificazione della Tyr (Ox e H)

➔ M5/6-C agisce solo dopo che Ox e H hanno prodotto Cl-Bht

➔ La tioesterasi Tcp39 rimuove substrati non corretti attivati da M6-A e non accettati da M5/6-C

Attività di proof-reading esercitata da M5-/6-C e Tcp39

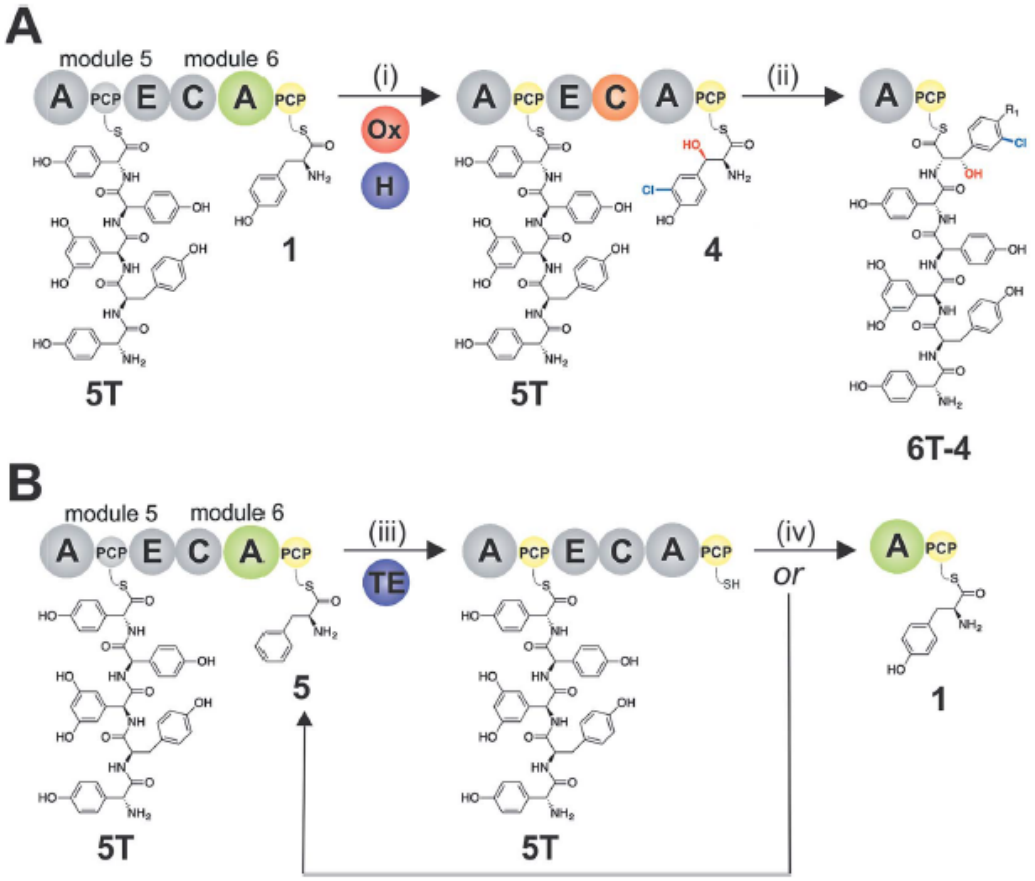



**Fig. 9** Schematic representation of A-domain and C-domain selectivity interfacing with the activity of *trans*-modifying enzymes and the type-II TE enzyme during teicoplanin biosynthesis. (A) Pentapeptide **5T** extension commences by the activation of Tyr (**1**) by the M6 A-domain, which is in turn modified by *trans* enzymes (**i**) to generate PCP-bound Cl-Bht (**4**), the recognised M5/6 C-domain acceptor substrate; this domain then catalyses peptide extension to generate the desired PCP-bound hexapeptide **6T-4** product (**ii**). (B) Incorrect amino acid activation (*e.g.* of Phe (**5**)) by the M6 A-domain leads to a PCP-bound intermediate that is not accepted by the halogenase or the M5/6 C-domain, which leads to a long-lived aminoacyl-PCP intermediate that is eventually cleaved by the type-II TE enzyme (**iii**). At this point, another round of amino acid activation catalysed by the M5/6 A-domain can lead to the loading of the correct amino acid (Tyr, **iv**) and entry into the productive pathway (A, **i** + **ii**), or the loading of an incorrect substrate that then will proceed to another round of TE-mediated substrate cleavage (**iii**). A – adenylation domain, C – condensation domain, PCP – peptidyl carrier protein domain, E – epimerisation domain, H – flavin-dependent halogenase, Ox – non-heme iron oxygenase, TE – type-II thioesterase.

# Inoltre…

- La sostituzione di M5/6-C teicoplanina con M5/6-C kistamicina converte la selettività da Cl-Bht a Tyr (figura 7)
- Nelle strategie di riprogrammazione delle NRPS è opportuno tenere in considerazione il filtro di selettività esercitato dai domini C sia nei confronti dei domini A-PCP accettori che nei confronti di enzimi di modificazione che agiscono *in trans*
- Ridisegnare/riassemblare unità C-A per ottenere riprogrammazione efficiente delle NRPS?