Title

Micrococcin cysteine-to-thiazole conversion through transient interactions between the scaffolding protein TclI and the modification enzymes TclJ and TclN.

Permalink

Authors

Calvopina-Chavez, Diana G
Bursey, Devan M
Tseng, Yi-Jie
et al.

Publication Date

DOI

Copyright Information

Peer reviewed

1 **Micrococcin cysteine-to-thiazole conversion through transient interactions**

2 **between a scaffolding protein and two modification enzymes**

3

4 Diana G Calvopina-Chavez[1], Devan M Bursey[1], Yi-Jie Tseng[2], Leena M Patil[2], Kathryn D

5 Bewley[3, 4], Philip R Bennallack[1, 5], Josh M McPhie[2], Kimberly B Wagstaff[2], Anisha Daley[2],

6 Susan M Miller[3], James D Moody[2], John C Price[2], *Joel S Griffitts[1]

7

8 [1]Department of Microbiology and Molecular Biology, Brigham Young University, Provo, UT

9 84602

10

11 [2]Department of Chemistry and Biochemistry, Brigham Young University, Provo, UT 84602

12

13 [3]Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94143

14

15 [4]Currently at: Genentech Inc, San Francisco, CA 94080

16

17 [5]Currently at: Werfen North America, Bedford, MA 01730

18

19 *Correspondence should be addressed to J.S.G (email: joelg@byu.edu)

20

21 Running Title: A scaffold protein coordinates cysteine to thiazole conversion

22

23 Keywords: thiazole, thiopeptide, post-translationally modified peptides

24   ABSTRACT

25   Ribosomally synthesized and post-translationally modified peptides (RiPPs) are a broad group of

26   compounds mediating microbial competition in nature. Azole/azoline heterocycle formation in

27   the peptide backbone is a key step in the biosynthesis of many RiPPs. Heterocycle formation in

28   RiPP precursors is often carried out by a scaffold protein, an ATP-dependent cyclodehydratase,

29   and an FMN-dependent dehydrogenase. It has generally been assumed that the orchestration of

30   these modifications is carried out by a stable complex including the scaffold, cyclodehydratase

31   and dehydrogenase. The antimicrobial RiPP micrococcin begins as a precursor peptide (TclE)

32   with a 35-amino acid N-terminal leader and a 14-amino acid C-terminal core containing six Cys

33   residues that are converted to thiazoles. The putative scaffold protein (TclI) presumably presents

34   the TclE substrate to a cyclodehydratase (TclJ) and a dehydrogenase (TclN) to accomplish the

35   two-step installation of the six thiazoles. In this study, we identify a minimal TclE leader region

36   required for thiazole formation, we demonstrate complex formation between TclI, TclJ and

37   TclN, and further define regions of these proteins required for complex formation. Our results

38   point to a mechanism of thiazole installation in which TclI associates with the two enzymes in a

39   mutually exclusive fashion, such that each enzyme competes for access to the peptide substrate

40   in a dynamic equilibrium, thus ensuring complete modification of each Cys residue in the TclE

41   core.

42

43   IMPORTANCE

44   Thiopeptides are a family of antimicrobial peptides characterized for having sulfur-containing

45   heterocycles and for being highly post-translationally modified. Numerous thiopeptides have

46   been identified; almost all of which inhibit protein synthesis in gram-positive bacteria. These

47   intrinsic antimicrobial properties make thiopeptides promising candidates for the development of

48   new antibiotics. The thiopeptide micrococcin is synthesized by the ribosome and undergoes

49   several post-translational modifications (PTMs) to acquire its bioactivity. In this study, we

50   identify key interactions within the enzymatic complex that carries out cysteine to thiazole

51   conversion in the biosynthesis of micrococcin.

52

53   INTRODUCTION

54   Ribosomally synthesized and post-translationally modified peptides (RiPPs) are natural products

55   produced by many bacteria that exhibit diverse biological activities including antimicrobial

56   functions (1-6). RiPP biosynthesis starts with the ribosomal translation of a precursor peptide

57   that is then heavily modified by multiple enzymes. The precursor peptide consists of an N-

58   terminal leader sequence, also known as the recognition sequence, responsible for recruiting

59   enzymes that carry out post-translational modifications (PTMs) (7-9), and a C-terminal core

60   peptide sequence where PTMs occur (10-12). In most cases, RiPP biosynthetic gene clusters

61   encode an E1-ubiquitin activating-like (E1-like) protein that has been implicated in leader

62   peptide binding. This E1 homolog contains a RiPP recognition element (RRE) that adopts a

63   highly conserved winged helix-turn-helix (wHTH) structure with three α-helices and a three-

64   stranded β-sheet. RiPP leader peptides bind to RRE domains by interacting at the interface of the

65   3α/3β fold acting as a fourth β strand (13-15). After proper substrate recognition, numerous

66   possible modifications take place on the core peptide culminating with the proteolytic removal of

67   the leader from the core yielding a mature RiPP (7, 8).

68

69   Thiazole/oxazole-modified microcins (TOMMs) are a class of RiPPs that feature thiazol(in)e and

70   oxazol(in)e heterocycles resulting from intramolecular reactions of cysteine, serine or threonine

71   residues in the precursor peptide (16). Thiazole/oxazole biosynthesis is a two-step process in

72   which an ATP-dependent cyclodehydratase (member of the YcaO superfamily) yields

73   thiazoline/oxazoline heterocycles that are then oxidized into azoles by an FMN-dependent

74   dehydrogenase. In addition to the cyclodehydratase and optional dehydrogenase, TOMM clusters

75   encode proteins that facilitate coupling of the precursor peptide with these enzymes, but the

76   different TOMM systems are highly variable in this respect. Most include an E1-like scaffold

77   protein (mentioned above) and/or a second type of protein-protein interaction domain annotated

78   as "Ocin-ThiF-like". Either or both of these may contain an RRE, and the E1-like domain is

79   often fused to the cyclodehydratase (17-23). This structural variability in TOMM complexes is

80   illustrated in **Fig. 1**, which depicts four examples for which studies on the architecture of these

81   complexes have been carried out. During biosynthesis of the cyanobactin trunkamide, the

82   enzyme TruD catalyzes formation of azoline heterocycles. The crystal structure of TruD shows a

83   fused cyclodehydratase with an NTD that contains an E1-like domain with an RRE, while the

84   CTD comprises a YcaO domain responsible for catalyzing heterocycle formation (24, 25). The

85   biosynthetic gene cluster for the bacteriocin heterocycloanthracin (HCA) contains a single copy

86   of an Ocin-ThiF-like protein (HcaF), a fused E1-YcaO cyclodehydratase (HcaD), and a

87   dehydrogenase (HcaB). In this system, the Ocin-ThiF-like protein interacts in a 1:1 ratio with the

88   E1-like domain on the fused cyclodehydratase to yield azoline heterocycles. These azoline rings

89   are then oxidized into azoles by a dedicated dehydrogenase, however, studies to characterize

90   protein interactions with this enzyme have been unsuccessful (26, 27). For Microcin B17, the

91   E1-like scaffold protein (McbB) interacts with a discrete cyclodehydratase (McbD) and a

92   dehydrogenase (McbC) in a higher order octameric complex with a ratio of 4:2:2 (22, 28). In a

93   more complicated system, the biosynthetic gene cluster for the thiopeptide sulfomycin (Sul)

94   encodes multiple copies of RREs (SulB, SulF), E1/Ocin-ThiF-like proteins (SulB, SulC, SulE,

95   SulF), cyclodehydratases (SulC, SulD) and dehydrogenases (SulF, SulG). Combinations of these

96   proteins form three complexes (SulBC, SulEFG, SulDEFG) that achieve Cys, Thr and Ser

97   conversion into their corresponding thiazole, methyloxazole, and oxazole (23, 29). All these

98   TOMM systems share certain biochemical features across a vast evolutionary distance, but they

99   vary in their intersubunit architectures.

100

101   Micrococcin is a thiopeptide produced by several Gram-positive bacteria, including *Bacillus*

102   *cereus* and *Macrococcus caseolyticus* (30-32). Its biosynthesis involves several PTMs, including

103   thiazole formation, C-terminal decarboxylation, dehydroamino acid formation, and the creation

104   of a pyridine-anchored macrocycle (33). The gene cluster responsible for micrococcin production

105   in *M. caseolyticus* (**Fig. 2A**) is located on a plasmid and consists of 12 *tcl* genes, which is

106   simpler than the 24-gene cluster found in *Bacillus cereus* (34, 35). Out of these 12 *tcl* genes, 8

107   are essential for micrococcin production (31, 33). The roles of these genes are illustrated in **Fig.**

108   **2B**. The precursor peptide for micrococcin, TclE, has an N-terminal leader of 35 amino acids and

109   a C-terminal core of 14 amino acids. Its biosynthesis begins with the conversion of all six

110   cysteine residues in the core to thiazoles (31). Thiazole installation is required for all subsequent

111   modifications (**Fig. 2B**). The work presented here focuses on the thiazole installation step in

112   micrococcin biosynthesis. Each thiazole conversion is a two-step process requiring three

113   proteins: a putative scaffold (TclI), a cyclodehydratase (TclJ), and a dehydrogenase (TclN) (**Fig.**

114   **2C**). We have previously shown that in the absence of TclI, no formation of thiazolines or

115   thiazoles occurs, suggesting that this putative scaffold protein is essential for cys-to-thiazole

116   conversion. When the TclJ cyclodehydratase is absent, there are no detectable thiazolines or

117   thiazoles. In the absence of the TclN dehydrogenase, all six thiazoline heterocycles accumulate,

118   suggesting that each Cys residue does not require complete modification before the next one is

119   processed (31).

120

121   In this study, we investigate how the thiazole installation proteins in the micrococcin

122   biosynthetic pathway interact with the substrate peptide, and with each other, to carry out these

123   modifications. By conducting a truncation analysis on the TclE leader peptide sequence, we

124   determined a 20-amino acid minimal recognition region required for thiazole installation.

125   Furthermore, by using computational modeling and an *E. coli*-based expression system for

126   mutagenesis and copurification experiments, we demonstrate complex formation between TclI,

127   TclJ and TclN, and we propose a mechanism for cysteine to thiazole conversion in which the

128   scaffold protein TclI coordinates thiazole installation by presenting the TclE substrate to each

129   modifying enzyme in dynamic equilibrium.

130

131   RESULTS

132   **TclE, TclI, TclJ, and TclN can be functionally expressed in *E. coli***

133   We engineered a system in which *E. coli* would express codon-optimized *tcl* genes encoding

134   TclE, TclI, TclJ, and TclN. Each of these was engineered with affinity tags in a manner which

135   was previously shown to preserve functionality (31). To test the functionality of *E. coli*-

136   expressed Tcl proteins, we evaluated the *in vitro* conversion of the six TclE Cys residues to

137   heterocycles by mass spectrometry (**Fig. 3**). TclE was purified with an N-terminal cleavable GST

138   tag, and the three other Tcl proteins were purified as complexes using N-terminally His-tagged

139   TclI (these complexes are shown in figures below). In the absence of modifying enzymes, TclE

140   purification and proteolytic removal from the GST tag yields a leader-plus-core fragment of the

141   expected molecular weight (m/z = 5373). When treated with *E. coli*-produced TclI and TclJ

142   (TclN excluded), TclE resolved to two major peaks of m/z=5285 and 5266, consistent with the

143   appearance of 5 and 6 thiazolines, respectively. When treated with *E. coli*-produced TclI, TclJ

144   and TclN, the prominent TclE product has m/z = 5253, consistent with the complete 6-thiazole

145   product with an expected -120 Da change ($-6 \times (H_2O+2H)$) compared to the unmodified peptide.

146

TclI$_{NTD}$ directly interacts with TclE

148   We hypothesized that TclI plays a role in TclE recognition. Bioinformatic analysis using the

149   program HHpred (36-38) indicates that TclI is structurally similar to Ocin-ThiF proteins. We

150   obtained a structural model of the TclI:TclE dimer using AlphaFold2 (**Fig. 4A**). The TclI model

151   contains two distinct domains, the N-terminal domain (NTD) is a wHTH structure comprised of

152   three α helices and three β strands, consistent with an RRE. This model places TclE at the

153   interface between the α helices and the 3-stranded β sheet with the leader sequence acting as a

154   fourth β strand, similar to what has been shown for crystallography solved structures in other

155   TOMM systems (20, 22, 24). Key TclE residues in this interaction start with F17 occupying a

156   hydrophobic pocket at the interface of TclI helix 3 and β3. Other predicted key interactions

157   between the TclE leader and the TclI RRE involve three salt bridges: TclE(E21)-TclI(K31),

158   TclE(E22)-TclI(K78), and TclE(E28)-TclI(K22) (**Fig. 4A**). To further investigate whether the

159   TclI$_{NTD}$ is an RRE, we tested for TclI$_{NTD}$ binding to TclE by copurification. We co-expressed

160   His-tagged TclI$_{NTD}$ (residues 1-85) with GST-tagged TclE in *E. coli* and carried out

161   copurification with nickel-NTA beads. As shown in **Figure 4B**, $TclI_{NTD}$ pulls down TclE (**Fig.**

162   **4B, Lane 3**), indicating a non-covalent interaction between the two proteins. Furthermore,

163    $TclI_{NTD}$ copurifies the TclE leader region when the TclE core region is absent (**Fig. S1**). We then

164   attempted to determine a minimal TclE leader that interacts with $TclI_{NTD}$, so we generated a

165   series of leader truncations in which 3 residues were consecutively removed from the N-terminus

166   of TclE (**Fig. 5A**) for a total of six TclE truncation variants: Δ3, Δ6, Δ9, Δ12, Δ15, and Δ18. Each

167   truncated variant was co-expressed with His-tagged $TclI_{NTD}$, and nickel pull-down experiments

168   were carried out. As shown in **Figure 5B**, some copurification could be detected for all TclE

169   truncations, but the interaction appears weakened beyond the Δ9 truncation. We observe that, as

170   more of TclE is removed, the amount of $TclI_{NTD}$ purified becomes less. This suggests that the

171   stability of this putative RRE domain is enhanced by the presence of a fully functional leader

172   peptide.

173

174   **Determination of a functionally minimal TclE leader peptide**

175   We then used the TclE truncation series to investigate leader sequence requirements for thiazole

176   installation in *E. coli* cells co-expressing TclI, TclJ, and TclN. We used four of the TclE

177   truncated variants (Δ9, Δ12, Δ15, and Δ18) and assessed thiazole installation by Orbitrap liquid

178   chromatography mass spectrometry (LC-MS) after GST-TclE purification from co-expressing

179   cells (**Fig. 6**). For this analysis, we define "fully modified peptides" as those in which all Cys

180   residues are converted to thiazoles. For the Δ9 truncation 100% of detected TclE peptides were

181   fully modified. The Δ12 variant also yielded products consistent with having a fully modified

182   core, while the Δ15 variant produced a slightly reduced yield of fully modified product (mean=

183    $99.8 \pm 0.2\%$; n=3), with some detected peptides containing intermediates with 4-5 thiazoles. The

184   Δ18 leader truncation yielded multiple intermediates containing a combination of thiazoles,

185   thiazolines and cysteines. The mean of fully modified TclE with the Δ18 truncation is $71.6 \pm$

186   4.8% (n=3). We conclude from this that the first 12 N-terminal amino acids of TclE are not

187   required for thiazole installation, and amino acids 12-15 have minimal impact. When amino

188   acids 15-18 are removed, it significantly impairs thiazole installation, potentially due to loss of

189   key TclE-RRE non-covalent interactions (see Fig. 4A). Recall that TclE(F17) was already

190   predicted to mediate an important interaction in a hydrophobic pocket of the modeled $TclI_{NTD}$.

191   Previous studies with TOMM systems such as streptolysin and microcin B17 have shown that

192   the leader peptide is primarily engaged through a conserved FXXXB (B= V, I, or L) motif (39,

193   40). The TclE leader contains a similar motif (FXXXXB) in residues 17-22. The convergence of

194   these empirical findings with the AlphaFold2 structural prediction and these observations in

195   other TOMM systems gives credibility to a model in which the $TclI_{NTD}$ functions as a genuine

196   RRE in micrococcin biosynthesis.

197

198   **$TclI_{CTD}$ binds to TclJ and TclN**

199   Given that $TclI_{NTD}$ engages with TclE, we hypothesized that the C-terminal domain (CTD) of

200   TclI may be primarily involved in recruiting the enzymatic proteins TclJ and TclN. Ocin-ThiF

201   proteins like TclI have previously been shown to mediate interactions with TOMM enzymes (3,

202   23, 26). To test interactions of full-length TclI to the modifying enzymes, we co-expressed His-

203   tagged TclI with TclJ or with TclN in *E. coli* and carried out nickel copurification experiments.

204   As shown in **Figure 7** (Lane 3), TclI interacts with both TclJ and TclN when all three proteins

205   are expressed together. When co-expressed with each individual enzyme, TclI also copurifies

206   them (**Fig. 7, Lanes 4-5**). His-tagged $TclI_{CTD}$ (residues 85-242) also copurifies TclJ and TclN

207   (**Fig. 7, lane 7-8**), though TclI$_{CTD}$:TclN shows a weaker interaction than TclI$_{CTD}$:TclJ. These

208   results point to TclI$_{CTD}$ as being sufficient for binding to both TclJ and TclN.

209

210   Initially, we interpreted these results to mean that TclI has two unique interaction surfaces on its

211   CTD, simultaneously recruiting TclJ and TclN as a stable enzymatic complex. To test whether

212   the TclIJN proteins form a ternary stable complex, we carried out copurification experiments

213   with varying tagging arrangements. We created nine strains expressing different combinations of

214   His-tagged TclI, TclJ and TclN. These experiments show that TclI copurifies TclJ and TclN

215   regardless of whether TclI is N-terminally or C-terminally His-tagged (**Fig. 8, Lanes 2-3**),

216   although TclI may be more poorly expressed or less stable when C-terminally tagged.

217   Furthermore, TclI binding to TclJ appears to be favored when TclI is C-terminally tagged. N-

218   terminally tagged TclJ does not copurify any detectable amounts of TclI or TclN (**Fig. 8, Lane**

219   **4**); however, when the tag is moved to the TclJ C-terminus, TclJ robustly pulls down TclI, with

220   little evidence of TclN copurifying (**Fig. 8, Lane 5**). When an N-terminal region of

221   TclJ (residues 1-115) is C-terminally tagged, it very robustly copurifies TclI, and a minor band

222   consistent with TclN can be observed (**Fig. 8, Lane 6**). TclN strongly copurifies TclI regardless

223   of the location of the tag (**Fig. 8, Lanes 8-9**), with possibly a minor TclJ copurification product.

224   In all cases where TclI is pulled down by a tagged version of TclJ or TclN, the band intensities

225   indicate a stoichiometric excess of TclI ranging from 2 to 8 based on densitometry that accounts

226   for staining intensity and molecular weight. In all cases where TclI is pulled down by either

227   tagged enzyme, copurification of the untagged other enzyme is absent or barely evident. From

228   these overall results, we hypothesize that TclI engages with TclJ and TclN in a competitive

229   fashion, and that there may be a weak interaction between the two enzymes.

230

231   **One surface of TclI facilitates binding to TclJ and TclN**

232   We generated structural models of TclI:TclJ and TclI:TclN dimers using AlphaFold2 (**Fig. 9A,**

233   **9B**). According to these two models, TclI$_{NTD}$ folds in a similar conformation for both, while

234   TclI$_{CTD}$ takes on a slightly different structure in each predicted complex, with Helix 6 (TclI

235   residues 188-213) being a central structure for binding to both enzymes. These models reinforce

236   the notion that the TclI:TclJ and TclI:TclN complexes are alternative and mutually exclusive

237   structures. To investigate whether Helix 6 is the primary interaction surface of TclI for binding to

238   the enzymes, we genetically dissected the region corresponding to Helix 6 by substituting each

239   of its 15 surface-exposed residues with an arginine residue and tested for binding to TclJ and

240   TclN under the same conditions used for the copurification shown in **Fig. 8** Lane 2. For this, 15

241   *E. coli* strains were constructed that express TclJ, TclN and His-tagged TclI with its

242   corresponding Helix 6 substitutions (**Fig. S2**). Copurification experiments show that TclI

243   residues Y189, I198, I202, and T212 are important for the TclI:TclJ interaction since these Arg

244   substitutions abolish binding interactions between the two proteins (**Fig. 10A, Lanes 2, 4, 6, 10**).

245   TclI residues H194, N201, S205, F208, and L209 are critical for interaction with TclN, as when

246   these residues are substituted, TclI:TclN interactions are disrupted (**Fig. 10A, Lanes 3, 5, 7, 8,**

247   **9**). Substitutions T188, L197, C199, E206, Y210, and S213 had no effect in TclI interactions

248   with either TclJ or TclN (**see Fig. 10B; Fig. S2**). Substitutions T212R and F208R resulted in a

249   weaker TclI band, suggesting that these changes may also have a negative effect on TclI folding

250   or stability. These findings indicate that Helix 6 is a pivotal structure within TclI responsible for

251   recruiting both TclJ and TclN. Key residues of Helix 6 facilitating binding to each enzyme are

252 interspersed within this postulated surface of TclI$_{CTD}$, further supporting the notion of two

253 enzymes competing for one TclI docking site.

254

255 **An independent domain of TclJ facilitates binding to TclI**

256 We wanted to further investigate how the N-terminal region of TclJ (when C-terminally tagged),

257 is able to pull down TclI, while the N-terminally tagged full-length TclJ is unable to (see **Fig. 8,**

258 **Lanes 4-5**). We reasoned that the N-terminus of full-length TclJ is important for TclI interaction.

259 We obtained AlphaFold2 models of TclJ and TclN (**Fig. S3A, and S3B**). The TclJ model

260 features two distinct domains, corresponding to a small NTD (residues 1-105) with the enzyme

261 active site in a larger YcaO-like CTD (residues 115-242). HHpred analysis showed that the NTD

262 of TclJ is a short E1-like domain likely involved in protein-protein interactions. The TclJ C-

263 terminus is embedded in its putative active site, while its N-terminus is predicted to be surface

264 exposed on this E1-like NTD. We hypothesized that TclJ$_{NTD}$ folds independently and facilitates

265 binding to TclI. To test this, two versions of the TclJ$_{NTD}$ (resides 1-105 and residues 1-115) were

266 co-expressed with His-tagged TclI for nickel copurification. Both versions of TclJ$_{NTD}$ copurified

267 with TclI (**Fig. 11, Lanes 3-4**) indicating that the E1-like domain of TclJ (residues 1-105) folds

268 independently and forms a TclI-binding domain. Tags on the N-terminus of this domain appear

269 to obstruct TclI binding (see **Fig. 8, Lane 4**).

270

271 **Analysis of TclIJN complexes by size-exclusion chromatography**

272 To investigate this model, we evaluated the TclIJN complexes using size-exclusion

273 chromatography. We first established the oligomeric state of purified TclI, which was consistent

274 with a monomer (**Fig. 12**). However, when TclI, TclJ and TclN were analyzed as a mixture, a

275 new peak arose with a retention time consistent with a molecular weight of ~120 kDa, suggesting

276 molecular units that contain two copies of TclI and one copy of either enzyme (2TclI:1TclJ,

277 expected: 121 kDa or 2TclI:1TclN, expected: 111 kDa) (**Fig. 12**).

278

279 DISCUSSION

280 The Tcl biosynthetic pathway from *M. caseolyticus* synthesizes micrococcin from a precursor

281 peptide (TclE) by orchestrating the following post-translational modifications: Cys-to-thiazole

282 conversion, C-terminal decarboxylation, Ser/Thr dehydration, and macrocyclization. The work

283 provided in this study focuses on the thiazole installation step to provide a better understanding

284 of the complex protein interactions of thiazole modifying enzymatic proteins. Thiazole

285 installation occurs in a two-step process in which a cyclodehydratase (TclJ) converts six

286 cysteines in the core peptide of TclE into thiazolines, and then a dehydrogenase (TclN) converts

287 the thiazolines into thiazoles. In this study, we determined a minimal region on the leader peptide

288 that is required for thiazole installation, we show complex formation between TclI, TclJ, and

289 TclN, and further define conserved domains within these proteins that play a critical role

290 mediating protein interactions for cysteine to thiazole conversion.

291

292 Thiazole/oxazole modified microcins (TOMMs) are derived from a precursor peptide that

293 contains an N-terminal leader and a C-terminal core. Modifications in TOMM precursor peptides

294 are orchestrated by an E1-like scaffold protein, a YcaO cyclodehydratase, and a dehydrogenase;

295 in some cases, a highly divergent Ocin-ThiF-like protein is also present (16, 21, 26, 41, 42). The

296 E1-like and Ocin-ThiF-like proteins are involved with leader peptide binding through an RRE

297 domain that features a winged helix-turn-helix (wHTH) structure comprised of three α-helices

298   and a three-stranded β-sheet (15). The purpose of the leader peptide is to recruit the biosynthetic

299   enzymes, but they have shown variations in their roles in the biosynthesis of diverse TOMMs

300   (20, 43-45). For instance, studies in Lantibiotics such as Nisin have shown that the core can be

301   processed in the absence of the leader peptide, albeit with reduced efficiency (46, 47). Reports

302   for cyanobactin processing by LynD have shown that in the absence of the leader, the core does

303   not get modified; however, if the leader is provided as a separate peptide (*in trans*) or appended

304   to the N-terminus of the biosynthetic enzyme, the substrate is fully modified (20, 43, 48). In this

305   study, we define the NTD of TclI as the RRE that directly interacts with the TclE leader, and we

306   define a minimal leader region that is essential for binding to the RRE and allows for complete

307   thiazole installation by the biosynthetic enzymes. Reports on streptolysin and microcin B17 have

308   determined that the leader peptide is recognized by a conserved FXXXB (B= V, I, or L) motif

309   (39, 40). Our studies show that truncation of a similar motif (FXXXXB) in the TclE leader

310   significantly reduces RRE binding and thiazole installation. This suggest a shared strategy in the

311   leader of TOMM precursors that has been conserved through evolutionary timescale.

312

313   TOMM biosynthetic clusters have shown a wide variety of structural variations in how the

314   biosynthetic enzymes interact with each other to install azole heterocycles onto a precursor

315   peptide. The solved structures of the TruD and LynD cyclodehydratases provide an example of

316   fused cyclodehydratases that contain an E1-like and a YcaO domain (49, 50). TOMM proteins

317   for the Heterocycloanthracin (Hca) biosynthetic cluster feature a single copy of an Ocin-ThiF-

318   like protein (HcaF), a fused E1-YcaO cyclodehydratase (HcaD), and a dehydrogenase (HcaB),

319   with the Ocin-ThiF-like protein comprising of an RRE and interacting in a 1:1 ratio with the

320   fused cyclodehydratase; attempts to elucidate interactions between the scafold protein and the

321   dehydrogenase for the Hca pathway have not been possible because the dehydrogenase is heavily

322   proteolyzed when expressed in *E. coli* (26). The Microcin B17 synthetase features a discrete

323   cyclodehydratase (McbD) and dehydrogenase (McbC) that are dependent on the presence of an

324   E1-like protein (McbB) for substrate recognition forming a higher order active complex with a

325   ratio of 4McbB:2McbD:2McbC (22). The TOMM proteins for the biosynthesis of sulfomycin

326   feature combinations of RREs (SulB, SulF), E1/Ocin-ThiF-like domains (SulB, SulC, SulE,

327   SulF), cyclodehydratases (SulC, SulD) and dehydrogenases (SulF, SulG). Combinations of these

328   proteins form three active complexes (SulBC, SulEFG, SulDEFG) that achieve azole formation

329   in the substrate peptide (23). Our findings show that the TOMM proteins from the Tcl

330   biosynthetic cluster comprise of an Ocin-ThiF-like protein (TclI) that includes an RRE, a fused

331   cyclodehydratase (TclJ) that includes an E1-like domain and a YcaO domain, and a

332   dehydrogenase (TclN). TclI binds to the TclE leader through its RRE and presents the core

333   peptide to each enzyme in a mutually exclusive manner, forming two distinct complexes

334   (TclI:TclJ, and TclI:TclN). Our results also suggest that the enzymes may interact weakly with

335   each other creating spatial constraints, so that both enzymes are readily available to bind to TclI

336   when taking turns to modify TclE in a coordinated manner (**Fig. 13**).

337

338   Size exclusion chromatography data suggests that TclI exists in a monomeric conformation when

339   expressed by itself, but it adopts a homodimeric structure when interacting with the enzymes

340   (**Fig. 12**). Furthermore, preliminary copurification experiments (**Fig. S4**) suggest that TclI may

341   indeed form a homodimer, thus the ratio of the active complexes might consist of 2TclI:1TclJ,

342   and 2TclI:1TclN, but further investigation is required to validate this claim. Overall, the findings

343   from this study have not only elucidated key domains of TOMM proteins in the Tcl biosynthetic

344   cluster, but it provides evidence of the high functional and structural diversity of the TOMM

345   protein family.

346

347   MATERIALS AND METHODS

348   **Plasmids, strains, and culture conditions**

349   The bacterial strains and plasmids used in this study are summarized in Tables S1 and S2. For

350   full plasmid sequences, refer to the Supplemental Materials file. Plasmids were constructed and

351   maintained in *E. coli* strain DH5α. The *tcl* genes were synthesized with codon optimization for

352   expression in *E. coli*. The sequences of vectors and *tcl* gene inserts are given in the supplemental

353   information file. For protein expression, plasmids were transformed into strain BL21, Nico 21

354   (DE3) or DH5α. All bacterial cultures were grown in Luria broth (LB: per liter, 10 g Bacto

355   tryptone, 5 g Bacto yeast extract, 5 g NaCl, 1 ml 2N NaOH). Antibiotics used were kanamycin

356   (30 µg/ml), ampicillin (100 µg/ml), and chloramphenicol (30 µg/ml). Cultures were induced for

357   protein expression using 0.3 mM isopropyl β-D-1-thiogalactopyranoside (IPTG).

358

359   **Tcl protein expression and purification**

360   To prepare samples for SDS-PAGE analysis, overnight liquid cultures (4 ml) were grown from

361   single colonies in the presence of appropriate antibiotics. 50-ml cultures were inoculated with 2

362   ml of overnight culture, allowed to grow at 30°C for 1 h, followed by induction with IPTG for

363   another 6 h at 30°C. Cells were collected by centrifuging the culture, and cell pellets were frozen

364   at -80°C for a minimum of 1 h. Cell pellets were then processed for protein purification with

365   either Ni-NTA-linked (for the His$_6$ tag) or glutathione-linked (for the GST tag) resin, as detailed

366   below.

367

368   For His purification, cell pellets were thawed on ice and re-suspended in 1 ml of lysis buffer (50

369   mM HEPES pH 7.8, 300 mM NaCl, 0.2 % Triton X-100, 0.5 mg/ml lysozyme, 60

370   mM imidazole, 1 mM EDTA). Lysis took place for 1 h at 4°C. Cell lysates were then sonicated 4

371   × 20 sec using a probe sonicator to ensure complete lysis and fragmentation of DNA. Samples

372   were centrifuged at 13,000 rpm for 10 min (4°C) and approximately 1 ml of supernatant was

373   transferred to a new microcentrifuge tube. Supernatant was incubated end-over-end with 50 µl of

374   NTA-nickel agarose beads (Qiagen) at 4°C for 30 min. Nickel beads were pelleted at 13,000 rpm

375   for 30 seconds and washed 3 × 1 ml with wash buffer (60 mM imidazole, 300 mM NaCl, 50 mM

376   HEPES pH 7.8). Purified proteins were then eluted in 50 µl of 2x SDS sample buffer (20%

377   glycerol, 83 mM Tris pH 6.8, 40 mg/ml sodium dodecyl sulfate (SDS), 0.01% bromophenol

378   blue, 0.03 µl/ml 2-mercaptoethanol).

379   For GST purifications, cell pellets were thawed on ice and re-suspended in 1 ml lysis buffer (50

380   mM Tris 8.0, 150 mM NaCl, 0.5 mg/ml lysozyme, 2 mM EDTA, and 0.2% Triton X-100). Cells

381   were lysed for 1 h at 4°C then dithiothreitol (DTT) was added to a final concentration of 1.5

382   mM. Samples were sonicated 4 × 20 sec. Cell lysates were centrifuged at 13,000 rpm at 4°C for

383   10 min to pellet cell debris. Approximately 1 ml of supernatant was transferred to a new

384   microcentrifuge tube. 50 µl of unwashed glutathione-agarose beads was added and samples were

385   rotated end over end for 45 min at 4°C. Slurry was pelleted at 13,000 rpm for 30 seconds.

386   Supernatant was removed and beads were washed 3 × 1 ml GST buffer (50 mM Tris 8.0, 150

387   mM NaCl). GST buffer was completely removed, and proteins were eluted from resin in 50 µl of

388   2x SDS sample buffer.

389

390   Purified samples were heated at 100°C for 5 min. Unless stated otherwise, 8 μl of supernatant

391   was loaded onto a 12% resolving Laemmli gel with a 4% stacking gel. Gels were run using 1x

392   Laemmli running buffer, stained overnight in coomassie blue stain, followed by destaining and

393   soaking in water prior to imaging.

394

395   **Mass spectrometry analysis of TclE processing**

396   For purification of His$_6$-tagged enzymes for mass spectrometry analysis, 25-ml overnight

397   cultures were grown from single colonies. These overnight cultures were then used to inoculate 1

398   L induction cultures (30°C). After 1 h, IPTG was added and the cultures were grown for an

399   additional 6 hours. The cells were harvested by centrifugation and the cell pellets were frozen at

400   -80°C overnight. For copurification of His$_6$-TclIJ, His$_6$-TclIN, or His$_6$-TclIJN the cells were then

401   thawed on ice with the addition of lysis buffer (50 mM HEPES, 150 mM NaCl, pH 7.8). A

402   protease inhibitor tablet (Roche), 0.2% Triton X-100 and 0.5 mg/ml lysozyme were added and

403   the cells were incubated on ice for 1 h. Complete lysis was achieved by sonication for 2 min on

404   ice using a Branson Sonifier 450, followed by centrifugation for 20 min at 32,539 × g. The

405   supernatant was incubated with 1 ml of Talon resin for 30 min at 4°C. Resin was washed with 3

406   × 10 ml lysis buffer, followed by elution with lysis buffer plus 75 mM imidazole (4 × 1 ml). The

407   elution fractions containing protein were buffer exchanged back into lysis buffer, concentrated,

408   flash frozen with 10% glycerol, and stored at -80°C.

409

410   For purification of GST-tagged TclE, 30-ml cultures were inoculated with 1 ml overnight

411   culture, grown at 37°C until an OD$_{600}$ = 0.6, then IPTG was added and the cells were grown for

412   an additional 20 h at 25°C. The cells were harvested by centrifugation and the cell pellets were

413   frozen at -80°C for at least 30 min. The cells were then thawed and resuspended in 1 ml lysis

414   buffer (50 mM Tris pH 8, 150 mM NaCl, 0.5 mg/ml lysozyme, 2 mM EDTA and one Roche

415   protease inhibitor tablet per 10 ml). Complete lysis was achieved after a 15-min incubation at

416   room temperature, followed by addition of DTT to 1.5 mM. Lysate was processed with several

417   short sonication pulses with a microtip. Insoluble material was centrifuged at 7,000 × g, and the

418   supernatant was combined with 30 µl of glutathione-agarose resin (slurry) at 4°C for 45 min

419   (rotating). The resin was pelleted and the beads were washed with GST buffer three times and

420   the peptide was eluted with 40 µl GST buffer plus 10 mM reduced glutathione. The eluant was

421   either frozen at -80°C for later use, or directly treated with tobacco etch virus (TEV) protease

422   and ZipTipped (using the manufacturer's instructions).

423   Activity of Tcl enzymes was tested *in vitro*. 20-µl reactions containing 20 µM GST-TclE, 5 mM

424   DTT, 2 mM ATP, 20 mM $MgCl_2$, 1 µM enzymes, and 1 µg TEV protease, were allowed to react

425   for 40 min at RT. Reactions were zip-tipped (using the manufacturer's instructions) and analyzed

426   by matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometry.

427

428   **Orbitrap Liquid Chromatography Mass Spectrometry (LC-MS) for TclE truncation series**

429   GST-TclE leader-truncated samples were co-expressed with TclIJN as described in the Tcl

430   protein expression and purification section. GST-TclE leader truncations were purified using

431   glutathione beads as described above and TclE peptides were removed from the GST tag through

432   TEV protease cleavage. Truncated TclE peptide samples were alkylated to cap any reduced

433   cysteines using chloroacetamide at 20 mM. The peptides were then separated and measured via

434   liquid chromatography-mass spectrometry (LC-MS) on an Easy nLC 1200 in connection with a

435   Thermo Easy-spray source and an Orbitrap Fusion Lumos. Peptides were pre-concentrated with

436　buffer A (3% acetonitrile, 0.1% formic acid) onto a PepMap Neo Trap Cartridge (particle size 5

437　μm, inner diameter 300 μm, length 5 mm) and separated with an EASY-Spray™ HPLC Column

438　(particle size 2 μm, inner diameter 75 μm, length 25 mm) with increasing buffer B (80%

439　acetonitrile, 0.1% formic acid) gradient:

440　Samples were eluted using a gradient of 5% B to 22% B over 85 minutes (128 minutes for

441　muscle), 22% to 32% B over 15 minutes (22 minutes for muscle), with a wash of 32% to 95% B

442　over 15 minutes, which was held at 95% B for 15 minutes followed by a wash step consisting of

443　two washes going from 95% B to 2% B over 3 minutes, holding at 2% B for 3 minutes, returning

444　to 95% over 3 minutes and holding for 3 minutes were performed. Sample loading and

445　equilibration were performed using the HPLC's built in methods. LC-MS only runs were

446　performed using 2400 V in the ion source, scan range of 375-1700 m/z, 30% RF Lens,

447　Quadrupole Isolation, 8 *105 AGC Target and a maximum injection time of 50 ms. The MS-

448　based data-dependent acquisition method was set to a 3 second cycle time. MS1 scans were

449　acquired by the orbitrap at a resolution of 120,000. Precursors with a charge > 1 and < 6 were

450　selected for MS2 fragmentation. MS2 scans of CID precursor fragments were detected with the

451　linear ion trap at a scan rate of 33.333 Da/sec with a Dynamic injection time. CID collisions

452　were set to 30% for 10ms. After 3 selections a 60 second dynamic exclusion window was

453　enabled; isotopes and unassigned charge states were excluded.

454

455　**Data Processing for Label-free Quantitation**

456　Raw files were searched against a FASTA data base for the TCL operon (containing I, J, N, and

457　E entries) with the *E. coli* proteome as a contaminant (Uniprot Reference UP000000625) using

458　Peaks Studio analysis. The parent mass error tolerance was set to 10 ppm and the fragment mass

459   error tolerance was set to 0.5 Da. Cysteine carbamidomethylation, thiazoline, and thiazole were

460   set as a variable modification, as well as methionine oxidation and pyro-glu from glutamine were

461   set as variable modifications in the search. Digest mode was set to unspecific, and the peptide

462   length range was set to 6 – 55 amino acids. The false discovery rate (FDR) for peptide matches

463   was set to 1%, and protein ID significance was set to -10log(P-value) ≥ 15. Label-free data was

464   normalized using the TIC option in PEAKS then the total signal for the modified form was

465   compared as a fraction of the total signal for the peptide of interest (**File S1**).

466

467

468

469   **Modeling and optimization of Tcl protein structures**

470   The sequences of TclI, TclJ, TclN, as well as the protein hetero-dimers such as TclE:TclI,

471   TclI:TclJ and TclI:TclN were submitted online to the AlphaFold2 Google colab (51, 52) for

472   structural predictions. The top-ranked models were selected for each Tcl protein and dimer. Each

473   model underwent optimization using the FastRelax algorithm (53) within PyRosetta. Energy

474   calculations for each model were carried out using the ref2015 score function (54, 55) in

475   PyRosetta. Binding energies were determined by subtracting the energies of the bound and

476   unbound state models. Additionally, to assess the binding modes, we calculated the shape

477   complementarities (56) of the complex models using Python logic and PyRosetta. These models

478   were visualized using PyMOL, and a list of interacting residues (L187R, T188R, Y189R,

479   H194R, I198R, C199R, N201R, I202R, S205R, E206R, F208R, L209R, Y210R, T212R, S213R)

480   was identified for use in mutagenesis experiments.

481

Size Exclusion Chromatography of TclI and TclIJN complex

Purified complex was filtered through 0.2 μm cellulose filter (14000 g, 2 min). All the samples were vacuum dried, and then resuspended in SEC buffer (100 mM sodium phosphate with pH 6.8, 0.023% NaN$_3$) to make the final concentration to 1 μg/μL. The Agilent 1260 Infinity HPLC System (Agilent, Santa Clara, CA) equipped with quaternary pump, manual injector, thermostatted column compartment, diode array detector (DAD) was used to carry out the analytical size exclusion chromatography. A Yarra-1.8 μm X 150Å ,150 x 4.6 mm HPLC column (00F-4631-E0, Phenomenex, USA) was used for separation of molecules. The Agilent system and column were equilibrated with 100 mM Sodium Phosphate with pH 6.8, 0.023% NaN$_3$ at a flow rate of 0.3 mL/min at 25 °C. The molecular weight calibration curve for SEC was obtained by running a protein standard mix containing bovine thyroglobulin (670 kDA), IgA (300 kDa), IgG (150 kDa), ovalbumin (44 kDa), myoglobin (17 kDa) and uridine (0.244 kDa), (AL0-3042, Phenomenex, CA, USA) to relate the molecule's size to elution volume. The injection volume for the protein standard and the protein samples was 5 μL, and the elution volume was measured using the UV detector. All the data was collected at 280 nm, with a reference wavelength of 600 nm.

Bioinformatics analysis of E1/YcaO, and Ocin-ThiF-like domains

Protein sequences of TclI and TclJ were submitted to HHpred (36-38) in FASTA format using standard parameters, and PDB70 and TIGRFAMs databases to analyze for the presence of E1/YcaO/Ocin-ThiF-like domains using pairwise comparison of profile HMM (hidden Markov model). The results of this analysis provided a list of known homologs with well-defined

504   domains, and multiple sequence alignments that were used to define each domain on the TclI and

505   TclJ proteins.

506

512

513

514   REFERENCES

515   1.   Kelly WL, Pan L, Li C. Thiostrepton biosynthesis: prototype for a new family of

516   bacteriocins. J Am Chem Soc. 2009;131(12):4327-34.

517   2.   Kutscher AH, Seguin L, Zegarelli EV, Piro JD. Antimicrobial activity of thiostrepton:

518   tube dilution studies. J Am Dent Assoc. 1959;59:715-20.

519   3.   Benazet F, Cartier M, Florent J, Godard C, Jung G, Lunel J, et al. Nosiheptide, a sulfur-

520   containing peptide antibiotic isolated from Streptomyces actuosus 40037. Experientia.

521   1980;36(4):414-6.

522   4.   Procopio RE, Silva IR, Martins MK, Azevedo JL, Araujo JM. Antibiotics produced by

523   Streptomyces. Braz J Infect Dis. 2012;16(5):466-71.

524   5.   Nicolaou KC, Zak M, Rahimipour S, Estrada AA, Lee SH, O'Brate A, et al. Discovery of

525   a biologically active thiostrepton fragment. J Am Chem Soc. 2005;127(43):15042-4.

526   6.   Aminake MN, Schoof S, Sologub L, Leubner M, Kirschner M, Arndt HD, Pradel G.

527   Thiostrepton and derivatives exhibit antimalarial and gametocytocidal activity by dually

528   targeting parasite proteasome and apicoplast. Antimicrob Agents Chemother. 2011;55(4):1338-

529   48.

530   7.   Oman TJ, van der Donk WA. Follow the leader: the use of leader peptides to guide

531   natural product biosynthesis. Nat Chem Biol. 2010;6(1):9-18.

532   8.   Chekan JR, Ongpipattanakul C, Nair SK. Steric complementarity directs sequence

533   promiscuous leader binding in RiPP biosynthesis. Proc Natl Acad Sci U S A.

534   2019;116(48):24049-55.

535   9.   Burkhart BJ, Kakkar N, Hudson GA, van der Donk WA, Mitchell DA. Chimeric Leader

536   Peptides for the Generation of Non-Natural Hybrid RiPP Products. ACS Cent Sci.

537   2017;3(6):629-38.

538   10.   Burkhart BJ, Hudson GA, Dunbar KL, Mitchell DA. A prevalent peptide-binding domain

539   guides ribosomal natural product biosynthesis. Nat Chem Biol. 2015;11(8):564-70.

540   11.   Barr I, Latham JA, Iavarone AT, Chantarojsiri T, Hwang JD, Klinman JP. Demonstration

541   That the Radical S-Adenosylmethionine (SAM) Enzyme PqqE Catalyzes de Novo Carbon-

542   Carbon Cross-linking within a Peptide Substrate PqqA in the Presence of the Peptide Chaperone

543   PqqD. J Biol Chem. 2016;291(17):8877-84.

544   12.   Wieckowski BM, Hegemann JD, Mielcarek A, Boss L, Burghaus O, Marahiel MA. The

545   PqqD homologous domain of the radical SAM enzyme ThnB is required for thioether bond

546   formation during thurincin H maturation. FEBS Lett. 2015;589(15):1802-6.

547   13.   Sumida T, Dubiley S, Wilcox B, Severinov K, Tagami S. Structural Basis of Leader

548   Peptide Recognition in Lasso Peptide Biosynthesis Pathway. ACS Chem Biol. 2019;14(7):1619-

549   27.

550   14.   Grove TL, Himes PM, Hwang S, Yumerefendi H, Bonanno JB, Kuhlman B, et al.

551   Structural Insights into Thioether Bond Formation in the Biosynthesis of Sactipeptides. J Am

552   Chem Soc. 2017;139(34):11734-44.

553   15.   Kretsch AM, Gadgil MG, DiCaprio AJ, Barrett SE, Kille BL, Si Y, et al. Peptidase

554   Activation by a Leader Peptide-Bound RiPP Recognition Element. Biochemistry.

555   2023;62(4):956-67.

556   16.   Melby JO, Nard NJ, Mitchell DA. Thiazole/oxazole-modified microcins: complex natural

557   products from ribosomal templates. Curr Opin Chem Biol. 2011;15(3):369-78.

558   17.   Burkhart BJ, Schwalen CJ, Mann G, Naismith JH, Mitchell DA. YcaO-Dependent

559   Posttranslational Amide Activation: Biosynthesis, Structure, and Function. Chem Rev.

560   2017;117(8):5389-456.

561   18.   Koehnke J, Bent AF, Zollman D, Smith K, Houssen WE, Zhu X, et al. The cyanobactin

562   heterocyclase enzyme: a processive adenylase that operates with a defined order of reaction.

563   Angew Chem Int Ed Engl. 2013;52(52):13991-6.

564   19.   Koehnke J, Bent AF, Houssen WE, Mann G, Jaspars M, Naismith JH. The structural

565   biology of patellamide biosynthesis. Curr Opin Struct Biol. 2014;29:112-21.

566   20.   Koehnke J, Mann G, Bent AF, Ludewig H, Shirran S, Botting C, et al. Structural analysis

567   of leader peptide binding enables leader-free cyanobactin processing. Nat Chem Biol.

568   2015;11(8):558-63.

569   21.   Dunbar KL, Melby JO, Mitchell DA. YcaO domains use ATP to activate amide

570   backbones during peptide cyclodehydrations. Nat Chem Biol. 2012;8(6):569-75.

571   22.   Ghilarov D, Stevenson CEM, Travin DY, Piskunova J, Serebryakova M, Maxwell A, et

572   al. Architecture of Microcin B17 Synthetase: An Octameric Protein Complex Converting a

573   Ribosomally Synthesized Peptide into a DNA Gyrase Poison. Mol Cell. 2019;73(4):749-62 e5.

574   23.   Du Y, Qiu Y, Meng X, Feng J, Tao J, Liu W. A Heterotrimeric Dehydrogenase Complex

575   Functions with 2 Distinct YcaO Proteins to Install 5 Azole Heterocycles into 35-Membered

576   Sulfomycin Thiopeptides. J Am Chem Soc. 2020;142(18):8454-63.

577   24.   McIntosh JA, Donia MS, Schmidt EW. Insights into heterocyclization from two highly

578   similar enzymes. J Am Chem Soc. 2010;132(12):4089-91.

579   25.   McIntosh JA, Schmidt EW. Marine molecular machines: heterocyclization in

580   cyanobactin biosynthesis. Chembiochem. 2010;11(10):1413-21.

581   26.   Dunbar KL, Tietz JI, Cox CL, Burkhart BJ, Mitchell DA. Identification of an Auxiliary

582   Leader Peptide-Binding Protein Required for Azoline Formation in Ribosomal Natural Products.

583   J Am Chem Soc. 2015;137(24):7672-7.

584   27.   Melby JO, Dunbar KL, Trinh NQ, Mitchell DA. Selectivity, directionality, and

585   promiscuity in peptide processing from a Bacillus sp. Al Hakam cyclodehydratase. J Am Chem

586   Soc. 2012;134(11):5309-16.

587   28.   Regni CA, Roush RF, Miller DJ, Nourse A, Walsh CT, Schulman BA. How the MccB

588   bacterial ancestor of ubiquitin E1 initiates biosynthesis of the microcin C7 antibiotic. EMBO J.

589   2009;28(13):1953-64.

590   29.   Egawa Y, Umino K, Tamura Y, Shimizu M, Kaneko K. Sulfomycins, a series of new

591   sulfur-containing antibiotics. I. Isolation, purification and properties. J Antibiot (Tokyo).

592   1969;22(1):12-7.

593   30.   Bennallack PR, Burt SR, Heder MJ, Robison RA, Griffitts JS. Characterization of a novel

594   plasmid-borne thiopeptide gene cluster in Staphylococcus epidermidis strain 115. J Bacteriol.

595   2014;196(24):4344-50.

596   31.   Bennallack PR, Bewley KD, Burlingame MA, Robison RA, Miller SM, Griffitts JS.

597   Reconstitution and Minimization of a Micrococcin Biosynthetic Pathway in Bacillus subtilis. J

598   Bacteriol. 2016;198(18):2431-8.

599   32.   Bennallack PR, Griffitts JS. Elucidating and engineering thiopeptide biosynthesis. World

600   J Microbiol Biotechnol. 2017;33(6):119.

601   33.   Bewley KD, Bennallack PR, Burlingame MA, Robison RA, Griffitts JS, Miller SM.

602   Capture of micrococcin biosynthetic intermediates reveals C-terminal processing as an

603   obligatory step for in vivo maturation. Proc Natl Acad Sci U S A. 2016;113(44):12450-5.

604   34.   Liao R, Duan L, Lei C, Pan H, Ding Y, Zhang Q, et al. Thiopeptide biosynthesis

605   featuring ribosomally synthesized precursor peptides and conserved posttranslational

606   modifications. Chem Biol. 2009;16(2):141-7.

607   35.   Wieland Brown LC, Acker MG, Clardy J, Walsh CT, Fischbach MA. Thirteen

608   posttranslational modifications convert a 14-residue peptide into the antibiotic thiocillin. Proc

609   Natl Acad Sci U S A. 2009;106(8):2549-53.

610   36.   Soding J, Biegert A, Lupas AN. The HHpred interactive server for protein homology

611   detection and structure prediction. Nucleic Acids Res. 2005;33(Web Server issue):W244-8.

612   37.   Zimmermann L, Stephens A, Nam SZ, Rau D, Kubler J, Lozajic M, et al. A Completely

613   Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. J Mol Biol.

614   2018;430(15):2237-43.

615   38.   Gabler F, Nam SZ, Till S, Mirdita M, Steinegger M, Soding J, et al. Protein Sequence

616   Analysis Using the MPI Bioinformatics Toolkit. Curr Protoc Bioinformatics. 2020;72(1):e108.

617   39.   Mitchell DA, Lee SW, Pence MA, Markley AL, Limm JD, Nizet V, Dixon JE. Structural

618   and functional dissection of the heterocyclic peptide cytotoxin streptolysin S. J Biol Chem.

619   2009;284(19):13004-12.

620   40.   Roy RS, Kim S, Baleja JD, Walsh CT. Role of the microcin B17 propeptide in substrate

621   recognition: solution structure and mutational analysis of McbA1-26. Chem Biol. 1998;5(4):217-

622   28.

623   41.   Melby JO, Li X, Mitchell DA. Orchestration of enzymatic processing by

624   thiazole/oxazole-modified microcin dehydrogenases. Biochemistry. 2014;53(2):413-22.

625   42.   Dunbar KL, Chekan JR, Cox CL, Burkhart BJ, Nair SK, Mitchell DA. Discovery of a

626   new ATP-binding motif involved in peptidic azoline biosynthesis. Nat Chem Biol.

627   2014;10(10):823-9.

628   43.   Khusainov R, Kuipers OP. When the leader gets loose: in vivo biosynthesis of a

629   leaderless prenisin is stimulated by a trans-acting leader peptide. Chembiochem.

630   2012;13(16):2433-8.

631   44.   Oman TJ, Knerr PJ, Bindman NA, Velasquez JE, van der Donk WA. An engineered

632   lantibiotic synthetase that does not require a leader peptide on its substrate. J Am Chem Soc.

633   2012;134(16):6952-5.

634   45.   Thibodeaux CJ, Wagoner J, Yu Y, van der Donk WA. Leader Peptide Establishes

635   Dehydration Order, Promotes Efficiency, and Ensures Fidelity During Lacticin 481 Biosynthesis.

636   J Am Chem Soc. 2016;138(20):6436-44.

637   46.   Goto Y, Ito Y, Kato Y, Tsunoda S, Suga H. One-pot synthesis of azoline-containing

638   peptides in a cell-free translation system integrated with a posttranslational cyclodehydratase.

639   Chem Biol. 2014;21(6):766-74.

640   47.   Dunbar KL, Mitchell DA. Insights into the mechanism of peptide cyclodehydrations

641   achieved through the chemoenzymatic generation of amide derivatives. J Am Chem Soc.

642   2013;135(23):8692-701.

643   48.   Mavaro A, Abts A, Bakkes PJ, Moll GN, Driessen AJM, Smits SHJ, Schmitt L. Substrate

644   recognition and specificity of the NisB protein, the lantibiotic dehydratase involved in nisin

645   biosynthesis. J Biol Chem. 2011;286(35):30552-60.

646   49.   Ortega MA, Hao Y, Zhang Q, Walker MC, van der Donk WA, Nair SK. Structure and

647   mechanism of the tRNA-dependent lantibiotic dehydratase NisB. Nature. 2015;517(7535):509-

648   12.

649   50.   Schmidt EW, Nelson JT, Rasko DA, Sudek S, Eisen JA, Haygood MG, Ravel J.

650   Patellamide A and C biosynthesis by a microcin-like pathway in Prochloron didemni, the

651   cyanobacterial symbiont of Lissoclinum patella. Proc Natl Acad Sci U S A. 2005;102(20):7315-

652   20.

653   51.   Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly

654   accurate protein structure prediction with AlphaFold. Nature. 2021;596(7873):583-9.

655   52.   Mirdita M, Schutze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold:

656   making protein folding accessible to all. Nat Methods. 2022;19(6):679-82.

657   53.   Tyka MD, Keedy DA, Andre I, Dimaio F, Song Y, Richardson DC, et al. Alternate states

658   of proteins revealed by detailed energy landscape mapping. J Mol Biol. 2011;405(2):607-18.

659   54.   Park H, Bradley P, Greisen P, Jr., Liu Y, Mulligan VK, Kim DE, et al. Simultaneous

660   Optimization of Biomolecular Energy Functions on Features from Small Molecules and

661   Macromolecules. J Chem Theory Comput. 2016;12(12):6201-12.

662   55.   Alford RF, Leaver-Fay A, Jeliazkov JR, O'Meara MJ, DiMaio FP, Park H, et al. The

663   Rosetta All-Atom Energy Function for Macromolecular Modeling and Design. J Chem Theory

664   Comput. 2017;13(6):3031-48.

665   56.   Lawrence MC, Colman PM. Shape complementarity at protein/protein interfaces. J Mol

666   Biol. 1993;234(4):946-50.

667   57.   Benjamini Y, Krieger AM, Yekutieli D. Adaptive linear step-up procedures that control

668   the false discovery rate. Biometrika. 2006;93(3):491-507.

669

670

671

672

673

674   FIGURE LEGENDS

675   **Figure 1. Schematic models of TOMM complexes from four RiPP biosynthetic pathways**

676   **for which the structural architecture has been elucidated.** Color codes for each protein are

677   described on the box on the right. The cyclodehydratase enzyme from the YcaO superfamily is

678   labeled YcaO. Black lines represent fused proteins, and dotted rectangles denote non-covalent

679   interactions. A. Cartoon representation of the cyclodehydratase form the Trunkamide

680   biosynthetic pathway (PDB: 4BS9) (18). B. Interactions between TOMM proteins from the

681   microcin Heterocycloanthracin. The question mark denotes that the dehydrogenase is part of this

682   pathway but attempts to characterize their protein interactions have not been successful (26). C.

683   Cartoon diagram based on the crystal structure of Microcin B17 (PDB: 6GRI) (22) D. Paradigm

684   of interactions between TOMM proteins in the biosynthesis of the thiopeptide sulfomycin.

685   Heterocycle formation in this system is mediated by three complexes labeled 1, 2, and 3 (23).

686

687   **Figure 2. Genes and proteins controlling micrococcin biosynthesis.** A. Map of the native *tcl*

688   gene cluster from *M. caseolyticus*. Essential proteins for complete micrococcin production are

689   annotated by colored blocks at the bottom. The gene encoding the precursor peptide (TclE) is

690   colored black, and the genes encoding proteins for thiazole installation (TclI, TclJ, TclN) are

691   colored blue. B. Overview of the micrococcin biosynthetic pathway that converts the TclE core

692   peptide into micrococcin. Modifications and corresponding enzymes are color coded.

693   Abbreviations: Tz, thiazolyl; Dc, decarboxyl; Dh, dehydro. C. Two-step conversion of TclE Cys

694   residues to thiazole by the enzymes TclJ, and TclN.

695

696   **Figure 3. Characterization of functional expression of Tcl TOMM proteins from *E. coli*.**

697   MS analysis (MALDI-TOF) of thiazole installation on TclE using purified components from *E.*

698   *coli*. Reactions included TEV protease to cleave the leader-core region of TclE from the GST tag

699   prior to mass spectrometry analysis. Numbers 1-4 on the left side of each MALDI-TOF spectra

700   represent each of the four strains we used to express the Tcl proteins.

701

702 **Figure 4. RRE of TclI interacts with TclE.** A. AlphaFold2 model of the TclE:TclI dimer. TclE

703 is shown in yellow. $TclI_{NTD}$ is highlighted in red, while $TclI_{CTD}$ is shown in brown. Residues

704 featuring key protein interactions are labeled as follows: $TclI_{NTD}$ (residues F7, V30, V31, I74,

705 V74, K78, N27, K22) in red, and TclE (residues F17, E21, E22, E28) in yellow. B. SDS-PAGE

706 analysis with coomassie blue staining to detect $TclE:TclI_{NTD}$ interactions. Protein identities are

707 given on the right side of the gel, and information of samples loaded in each lane is given above

708 the gel.

709

710 **Figure 5. TclE leader truncation analysis.** A. Schematic diagram of TclE leader truncations.

711 Native TclE from *M. caseolyticus* is shown on top followed by the mutagenesis analysis on the

712 N-terminus of the TclE leader. B. Nickel copurification experiment to detect $TclI_{NTD}$ interactions

713 with truncated variants of the TclE leader. The top panel shows the Tcl proteins that were co-

714 expressed for this experiment. His-tagged $TclI_{NTD}$ was co-expressed with each TclE leader

715 truncated variant in *E. coli* and subjected to purification with nickel-NTA beads and SDS-PAGE.

716 Protein identities are given on the right side of the gel. For each GST-TclE truncation variant, the

717 molecular weights are as follows: Δ3 (31.4 kDa), Δ6 (31.0 kDa), Δ9 (30.7 kDa), Δ12 (30.4 kDa),

718 Δ15 (30.1 kDa), Δ18 (29.7 kDa).

719

720 **Figure 6. Effects of TclE leader truncations on the production of fully modified core**

721 **peptides.** Fully modified TclE is defined by the presence of six thiazoles in the core peptide.

722 Each bar shows percentage of TclE peptides detected by LC-MS that are fully modified. Each

723 TclE truncated variant (see Fig. 4A; Δ9, Δ12, Δ15, and Δ18) was co-expressed with TclI, TclJ,

724 and TclN in *E. coli* and purified for LC-MS. Reactions included TEV protease to cleave TclE

725   from the GST tag prior to LC-MS analysis. Data are shown as the mean of three independent

726   replicates with standard deviation. Asterisks show statically significant differences (P < 0.05)

727   according to a parametric t-test carried out with the Benjamini, Krieger, and Yekutieli method

728   (57).

729

730   **Figure 7. Domain analysis of TclI binding to TclJ and TclN.** SDS-PAGE analysis to detect

731   Tcl protein expression and copurification. Maps of His-tagged TclI, TclI$_{CTD,}$ TclJ and TclN used

732   in this study are shown above the gel. Asterisk shows a weak band corresponding to TclN that

733   does not appear in the vector only control (Lane 1) but consistently appears in independently

734   replicated gels. The weak nature of this band could be explained by TclN being degraded by

735   proteases or weak interactions with TclI$_{NTD}$.

736

737   **Figure 8. Analysis of TclI, TclJ and TclN interactions as a stable complex.** Coomassie-

738   stained SDS-PAGE of purified TclIJN proteins expressed together in *E. coli* to detect complex

739   formation. The upper panel shows the combinations of each his-tagged Tcl protein. The

740   difference between Lane 5 and Lane 7 is that, in Lane 7, the linker between TclJ and the histidine

741   tag is longer (6-Gly) compared to the normal linker (Gly-Gly-Ser) used in Lane 5 and in the

742   other Lanes. Asterisks denote weak protein bands that do not appear in the vector only control

743   (Lane 1) but consistently show up in replicated copurification experiments.

744   **Figure 9. Predicted AlphaFold2 models of the TclI:TclJ (A) and TclI:TclN (B) dimers.**

745   TclI$_{NTD}$ is shown in red, and TclI$_{CTD}$ is shown in brown with the central Helix 6 highlighted in

746   magenta. TclJ is depicted in cyan and TclN is shown in green.

747

748 **Figure 10. Dissection of Helix 6 on TclI to determine key interactions with TclJ and TclN.**

749 A. Nickel copurification experiment with His-tagged TclI with its corresponding Helix 6

750 substitution co-expressed with TclJ and TclN. Each substitution on TclI is labeled on each lane.

751 B. Depiction of Helix 6 highlighting key residues for interacting with TclJ in cyan (Y189, I198,

752 I202, T212) and TclN in green (H194, N201, S205, F208, L209). Residues highlighted in gray

753 (T188, L197, C199, E206, Y210, S213) had no effect in TclI interactions with either TclJ or

754 TclN (see Fig. S2).

755

756 **Figure 11. E1-like domain analysis of TclJ.** A. Map of TclJ highlighting its E1-like and YcaO

757 domains. Labeled are the two NTD fragments of TclJ used in this study. B. SDS-PAGE analysis

758 to detect interactions between the two TclJ$_{NTD}$ fragments (NTD1, and NTD2) and TclI.

759

760 **Figure 12. Size-exclusion chromatography of purified TclI and TclIJN.** Yellow dotted line

761 indicates chromatogram of purified TclI. Grey dotted line shows the chromatograph for the

762 TclIJN purified complex. The blue chromatogram represents the protein ladder, and the dotted

763 orange line shows the chromatogram for the buffer only control.

764

765 **Figure 13. Paradigm for cysteine to thiazole conversion by TclI, TclJ and TclN on the**

766 **precursor peptide TclE in the biosynthesis of micrococcin.** Schematic model of the two

767 functional complexes (TclI:TclJ, and TclI:TclN) that assemble to achieve two-step thiazole

768 installation on the core of TclE.

A

tclE *tclS* *orf18* *tclQ* *tclI* *tclJ* *tclK* *tclL* *tclM* *tclN* *tclP* *tclU*

■ Precursor peptide ■ Ser/Thr dehydration ■ Self-resistance □ Regulation
■ Thiazole formation ■ Pyridine formation ■ C-terminal dehydrogenases ■ Other/unknown

B

TclE precursor peptide

Core (14-aa)

LEADER (35-aa) **S C T T C V C T C S C C T T**

↓ TclI, TclJ, TclN (thiazole installation)

LEADER (35-aa) **S C T T C V C T C S C C T T**
     Tz   Tz Tz Tz TzTz

↓ TclP (C-terminal decarboxylation)

LEADER (35-aa) **S C T T C V C T C S C C T T**
     Tz   Tz Tz Tz TzTz **Dc**

↓ TclK, TclL (Ser/Thr dehydration)

        Dh   Dh   Dh   Dh
LEADER (35-aa) **S C T T C V C T C S C C T T**
     Tz   Tz Tz Tz TzTz **Dc**

↓ TclM (macrocyclization)

LEADER (35-aa)

+

Micrococcin P2

C_{Tz} =

T_{Dc} =

S_{Dh} =

T_{Dh} =

C

Cysteine → (TclJ, ATP) → Thiazoline → (TclN, FMN) → Thiazole

Strains	1	2	3	4
GST-TclE	+	+	+	+
TclI + TclN	-	+	-	-
TclI + TclJ	-	-	+	-
TclI + TclJ + TclN	-	-	-	+

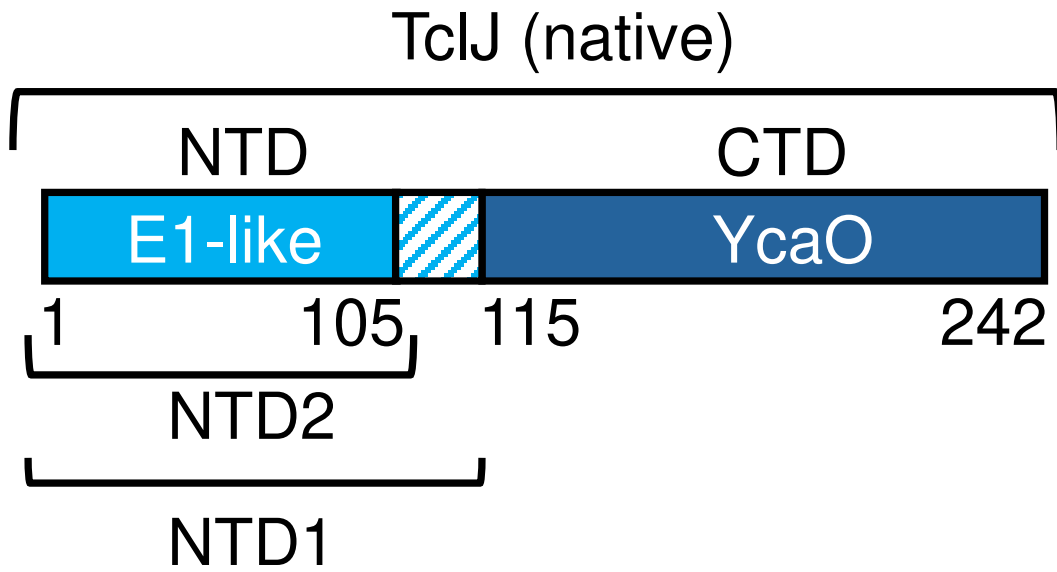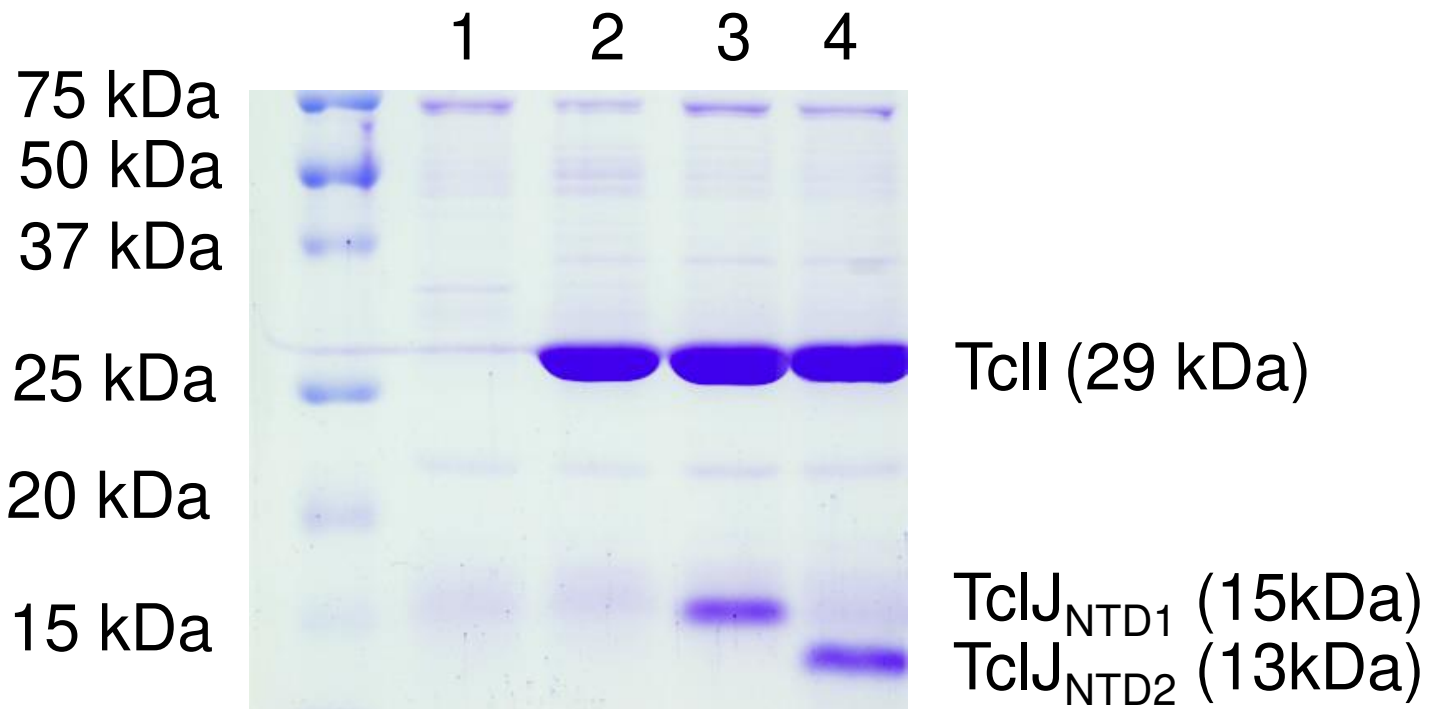| TclE peptide species	Calculated mass [M+H]$^+$
unmodified	5373.9
6 x thiazole	5253.7
6 x thiazoline	5265.8
5 x thiazoline	5283.8

A

B
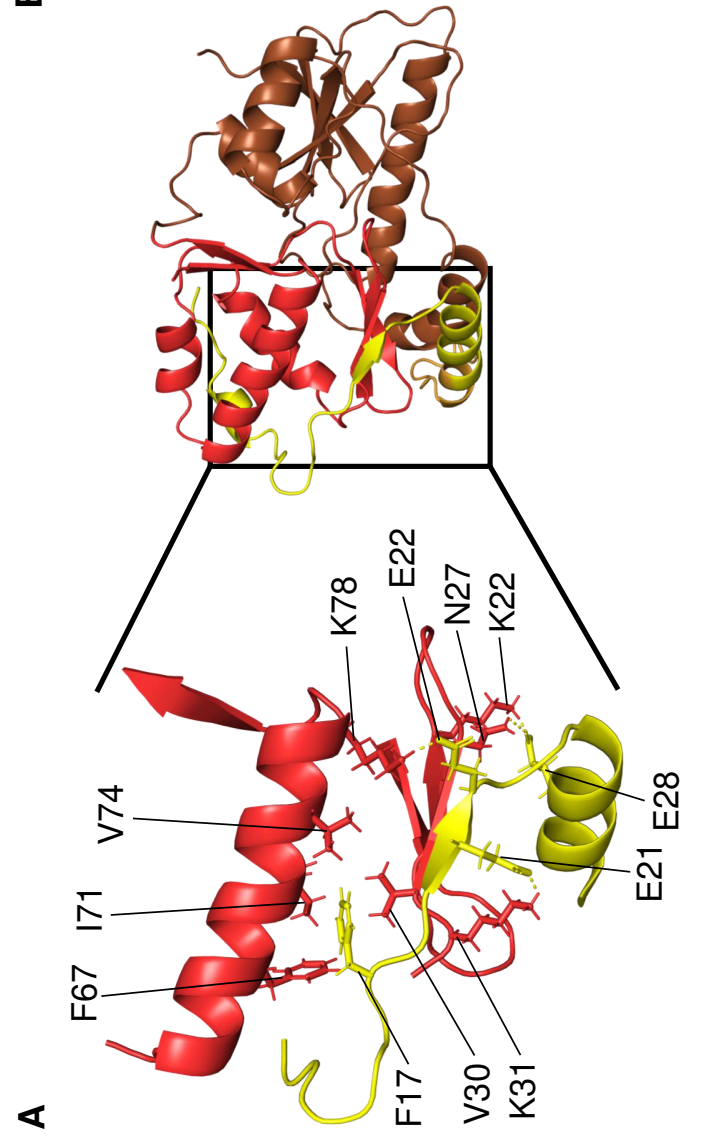
Lane 1: Vector only control

Lane 2: His₆ TcII_NTD

Lane 3: His₆ TcII_NTD   GST TclE

Lane 1: Vector only control

Lane 2: His$_6$ TclI

Lane 3: His$_6$ TclI   TclJ

Lane 4: His$_6$ TclI   TclJ   TclN

Lane 5: His$_6$ TclI   TclN

Lane 6: His$_6$ TclI$_{CTD}$

Lane 7: His$_6$ TclI$_{CTD}$ TclJ

Lane 8: His$_6$ TclI$_{CTD}$ TclN

TclJ (66 kDa)
TclN (53 kDa)
TclI (29 kDa)
TclI$_{CTD}$ (19 kDa)

75 kDa
63 kDa
48 kDa
35 kDa
25 kDa
20 kDa
17 kDa
10 kDa

A

TclJ (native)

NTD   CTD

E1-like   //////   YcaO

1   105   115   242

NTD2

NTD1

B

Lane 1: Vector only control

Lane 2: His$_6$ TclI

Lane 3: His$_6$ TclI   TclJ$_{NTD1}$

Lane 4: His$_6$ TclI   TclJ$_{NTD2}$

TclI (29 kDa)

TclJ$_{NTD1}$ (15kDa)
TclJ$_{NTD2}$ (13kDa)

B

A