Iron acquisition in plague: modular logic in enzymatic biogenesis of yersiniabactin by *Yersinia pestis*

Amy M Gehring[1]*, Edward DeMoll[2], Jacqueline D Fetherston[2], Ichiro Mori[1†], George F Mayhew[3], Frederick R Blattner[3], Christopher T Walsh[1] and Robert D Perry[2]

Background: Virulence in the pathogenic bacterium *Yersinia pestis*, causative agent of bubonic plague, has been correlated with the biosynthesis and transport of an iron-chelating siderophore, yersiniabactin, which is induced under iron-starvation conditions. Initial DNA sequencing suggested that this system is highly conserved among the pathogenic *Yersinia*. Yersiniabactin contains a phenolic group and three five-membered thiazole heterocycles that serve as iron ligands.

Results: The entire *Y. pestis* yersiniabactin region has been sequenced. Sequence analysis of yersiniabactin biosynthetic regions (*irp2–ybtE* and *ybtS*) reveals a strategy for siderophore production using a mixed polyketide synthase/nonribosomal peptide synthetase complex formed between HMWP1 and HMWP2 (encoded by *irp1* and *irp2*). The complex contains 16 domains, five of them variants of phosphopantetheine-modified peptidyl carrier protein or acyl carrier protein domains. HMWP1 and HMWP2 also contain methyltransferase and heterocyclization domains. Mutating *ybtS* revealed that this gene encodes a protein essential for yersiniabactin synthesis.

Conclusions: The HMWP1 and HMWP2 domain organization suggests that the yersiniabactin siderophore is assembled in a modular fashion, in which a series of covalent intermediates are passed from the amino terminus of HMWP2 to the carboxyl terminus of HMWP1. Biosynthetic labeling studies indicate that the three yersiniabactin methyl moieties are donated by *S*-adenosylmethionine and that the linker between the thiazoline and thiazolidine rings is derived from malonyl-CoA. The salicylate moiety is probably synthesized using the aromatic amino-acid biosynthetic pathway, the final step of which converts chorismate to salicylate. YbtS might be necessary for converting chorismate to salicylate.

Addresses: [1]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115, USA. [2]Department of Microbiology and Immunology, MS415 Medical Center, University of Kentucky, Lexington, KY 40536-0084, USA. [3]Department of Genetics, University of Wisconsin, 445 Henry Hall, Madison, WI 53706, USA.

Present addresses: *Department of Microbiology and Cellular Biology, Harvard University, Cambridge, MA 02115, USA. †Nippon Glaxo Ltd, Tsukuba Research Laboratories, Chemistry Department, Ibaraki, Japan.

Correspondence: Christopher T Walsh
E-mail: walsh@walsh.med.harvard.edu

Key words: heterocyclization, nonribosomal peptide synthetase, polyketide synthase, siderophore, yersiniabactin

Received: **6 August 1998**
Accepted: **28 August 1998**

Published: **21 September 1998**

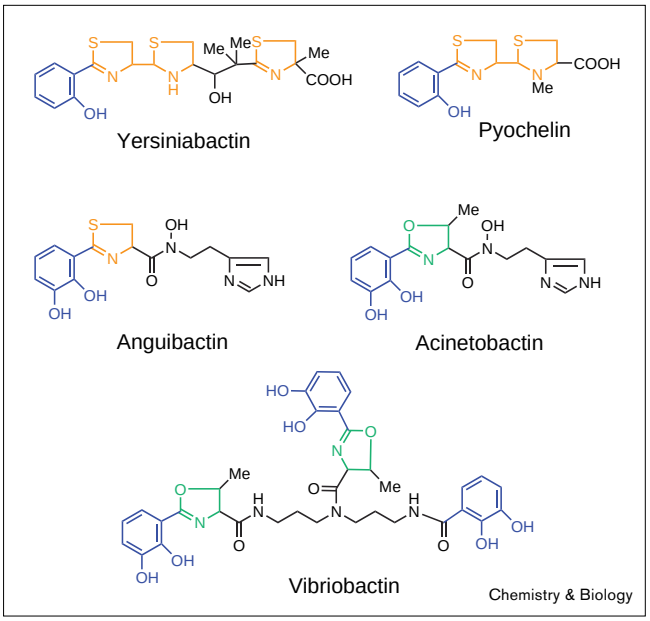
Chemistry & Biology October 1998, **5**:573–586
http://biomednet.com/elecref/1074552100500573

Introduction

Nearly all free living organisms need iron for growth. Many microbes have carefully orchestrated strategies for the regulated expression of genes that code for the biosynthesis of siderophores, small molecular mass compounds with enormous avidity for ferric iron (Fe^{3+}), and for proteins involved in the specific uptake of the iron-loaded siderophore [1–3]. In vertebrate hosts where iron is scarce due to chelation by host proteins [4,5], microbial siderophore genes are turned on and contribute to the virulence of numerous bacteria including *Yersinia pestis* [6,7], *Vibrio cholerae* [8] and *Pseudomonas aeruginosa* [9,10], the causative agents of plague, cholerae and cystic fibrosis-associated respiratory disease, respectively. Siderophores are typically small water-soluble molecules that present a constellation of functional groups with avidity for ferric iron coordination such that the aggregate affinity makes the ferric siderophore thermodynamically favored over any host protein-coordination sites. For example, the *Escherichia coli* siderophore enterobactin is particularly successful in this design: with an estimated K_D of 10^{-52} M for iron binding [11] it can wrest ferric ions from any locus.

Siderophores are often fashioned from small peptides. The peptides can be *N*-hydroxylated to yield NHOH ligands for iron (e.g., the *N*-hydroxyornithines in the *E. coli* siderophore aerobactin) and/or *N*-acylated with salicyl or 2,3-dihydroxbenzoyl groups to provide acidic phenolic oxygen ligands (e.g., the *E. coli* siderophore enterobactin) [12]. A third class of iron-coordination sites in peptidic ligands is created by the cyclization of serine, threonine or cysteine sidechains on the upstream amide carbonyl to yield oxazoline and thiazoline five-membered rings with a now basic imine nitrogen to coordinate the metal cation.

Figure 1



Structures of yersiniabactin and other microbial thiazoline/oxazoline-containing siderophores. Hydroxyphenyl, thiazoline/thiazolidine and oxazoline moieties of each siderophore are shown in blue, orange and green, respectively.

Examples include anguibactin from the fish pathogen *Vibrio anguillarum* [13], acinetobactin from *Acinetobacter baumanni* [14], pyochelin from *P. aeruginosa* [15], vibriobactin from *V. cholerae* [16] and yersiniabactin (Ybt) from *Y. pestis* and *Y. enterocolitica* ([17,18], R.D.P., P.B. Balbo, H.A. Jones, J.D.F. and E.D., unpublished observations; Figure 1). Ybt, pyochelin and vibriobactin utilize the phenol and thiazoline/oxazoline strategies, whereas acinetobactin and anguibactin use all three (catechol, thiazoline/oxazoline, *N*-hydroxylation) to create Fe^{3+}-binding sites.

Ybt is of particular interest both for its relationship to *Yersinia* virulence and for its intriguing structure. The high virulence of *Y. pestis* and *Y. enterocolitica* strains in infected animals requires the high-pathogenicity island that encodes the Ybt biogenesis and transport genes [6,19–25]. Ybt, an intriguing small-molecule natural product of mixed polyketide/polypeptide origin [17,18], has a high affinity for ferric ions (K_D of 4×10^{-36}) with no significant binding of ferrous ions (R.D.P., P.B. Balbo, H.A. Jones, J.D.F. and E.D., unpublished observations). Ybt is a peptide incorporating three cysteines that have been cyclized to five-membered ring heterocycles, the first and third at the dihydrooxidation state and the second at the tetrahydrooxidation state. The enzymatic logic for cyclization of cysteine and serine moieties into thiazolines and oxazolines in both siderophore and nonribosomal peptide antibiotic biosynthesis has been opaque until now. The Ybt peptide chain is initiated with an aryl *N*-cap, derived biosynthetically from salicylate. The second and third cysteine moieties, the thiazolidine and second thiazoline, are interrupted by a branched four-carbon insert, which suggests a polyketide synthase strategy has been at work. We have recently validated that YbtE, expressed in and purified from *E. coli*, initiates siderophore construction by activating salicylate as the salicyl-AMP mixed anhydride and transferring the salicyl group covalently onto the amino terminus of HMWP2 [26]. In addition, we have demonstrated that an amino-terminal fragment of HMWP2 adenylates cysteine and forms a heterocyclic hydroxyphenyl thiazoline product [26].

The sequences of several genes encoding putative and proven Ybt biosynthetic genes have been reported for *Y. pestis* and *Y. enterocolitica* [6,27,28]. We have now completed sequencing of the Ybt region in the plague organism *Y. pestis* KIM6+ (Figure 2). As anticipated from previous sequencing results and studies indicating cross-functionality of the Ybt system among the highly pathogenic *Yersinia* [6,21,23,24,27,28], there is a very high degree of similarity between the *irp2*, *irp1*, *ybtU*, *ybtT*, *ybtE* and *psn* sequences of *Y. pestis* and *Y. enterocolitica*. We focus our attention here on the *Y. pestis* Ybt biosynthetic regions — *irp2–ybtT* and *ybtS*. Genetic and sequence evidence suggests that YbtS is required for salicylate synthesis. Sequencing of *irp1*, encoding a 3,163 residue, 349 kDa HMWP1 polypeptide, and *irp2*, encoding a 2,035 residue, 229 kDa HMWP2 polypeptide, permits a dissection of the logic for the initiation, elongation and termination steps in the multidomainal enzymes of Ybt assembly. HMWP2 and HMWP1 together provide the modular assembly line for Ybt elongation, with seven and nine domains respectively, discernible by homology to domains of known function in other nonribosomal peptide synthetases and polyketide synthases. These 16 functions give considerable insight into the timing of biochemical transformations, the location of intermediates attached to and elongated along the length of these two enzymes, as well as the mechanism of both C–C bond-forming and heterocyclic-ring-forming reactions.

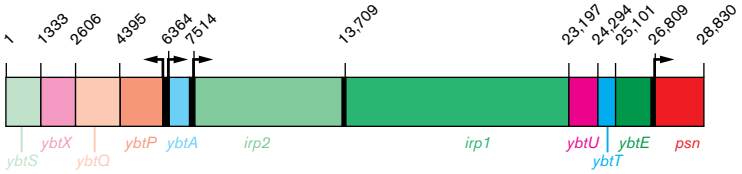
Results and discussion

Sequence analysis of the Ybt region

We have sequenced a region of the *pgm* locus of *Y. pestis* KIM6+ [29,30] that encompasses the entire high-pathogenicity island, which encodes the Ybt system (GenBank accession # YBT AF091251). Features of the high-pathogenicity island unrelated to the Ybt system are noted in the database entry; analyses presented here include only Ybt-related genes. Although we present the entire gene sequence from *ybtS* to *psn* (Figure 2), several of these *Y. pestis* KIM6+ genes have been sequenced and described previously. YbtA is an AraC-like regulator that activates expression from promoters for *psn*, *irp2* and *ybtP* and represses expression of its own promoter. In addition, all

Figure 2

Genetic organization of the Ybt biosynthetic, regulatory, and transport operons. Numbering corresponds to nucleotide numbers with 1 being the last nucleotide in the open reading frame of *ybtS*. Numbering associated with *ybtA, irp2, irp1, ybtU, ybtT, ybtE* and *psn* correspond to the start of each gene whereas that with *ybtS, ybtX, ybtQ, ybtP* and *psn* correspond to the stops in each gene. For Psn, predicted masses and pI values were calculated for the unprocessed and processed forms of the polypeptide. The table gives relevant features of each protein encoded within this region.



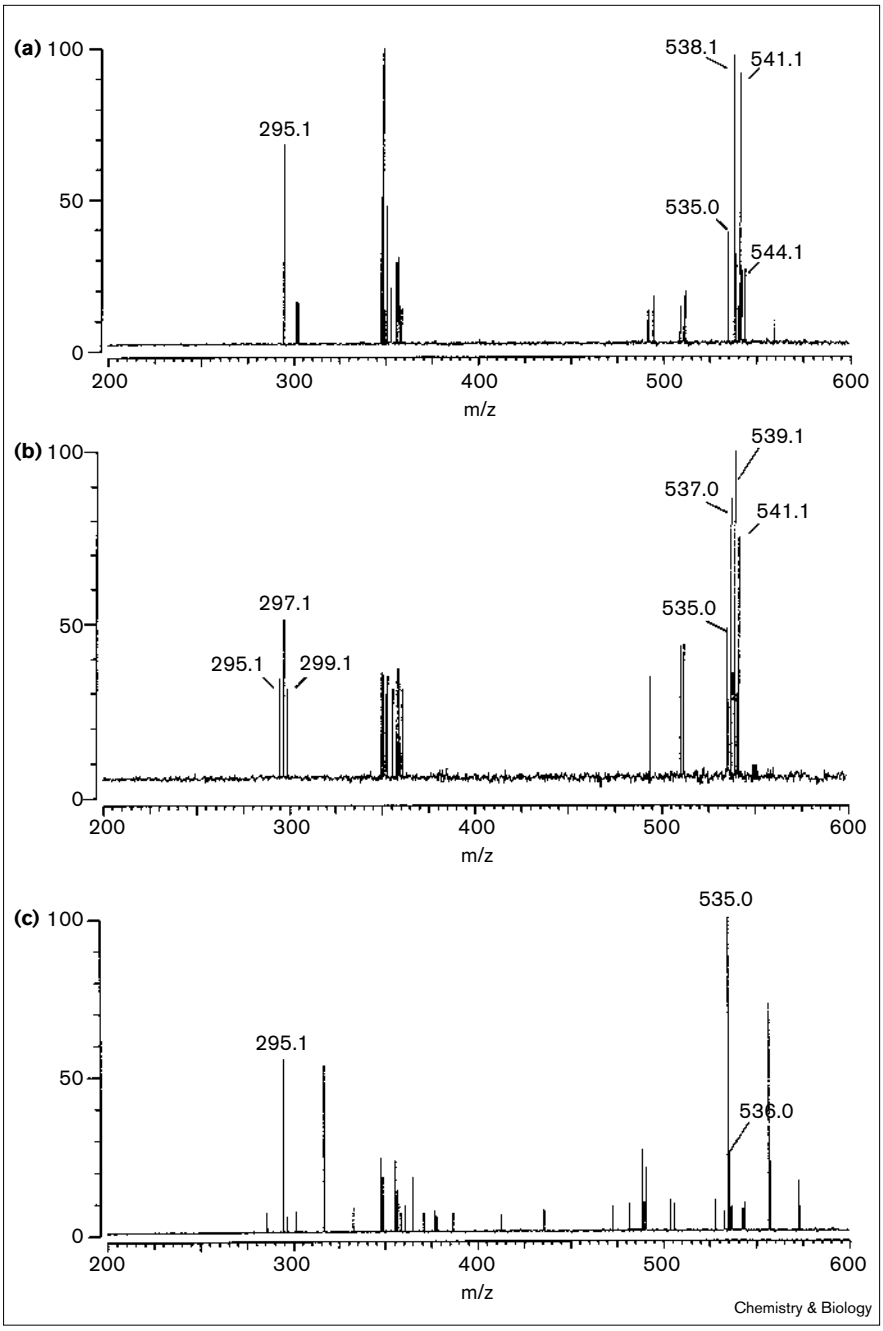
Protein (gene)	Predicted mass (Da)	pI	Function
YbtS	47,990	5.54	Salicylate synthesis
YbtX	44,781	10.32	Unknown
YbtQ	66,468	6.75	Ybt uptake; IM permease
YbtP	66,289	10.12	Ybt uptake; IM permease
YbtA	35,865	8.11	AraC-like regulator
HMWP2 *(irp2)*	228,666	5.73	Ybt peptide synthetase
HMWP1 *(irp1)*	348,820	5.29	Ybt peptide/ polyketide synthetase
YbtU	40,915	6.35	Unknown
YbtT	29,752	5.80	Unknown; thioesterase
YbtE	57,227	5.16	Salicyl-AMP ligase
Psn	73,786/71,476	5.35/5.11	Ybt and pesticin receptor

Chemistry & Biology

four promoter regions possess ferric uptake regulation (Fur)-binding sites and iron and Fur regulation of each operon has been observed ([6,23,31] J.D.F., V.J. Bertolino and R.D.P., unpublished observations). Psn is the outer membrane receptor for Ybt and the bacteriocin pesticin that is required for utilization of Ybt [23]. YbtP and YbtQ are inner-membrane permeases that appear to be required for utilization of Ybt and transport of iron. YbtX is a highly hydrophobic protein with 9–11 potential transmembrane domains; its role in Ybt synthesis, transport, or regulation has not been established. We have used the second of two possible methionine start sites for YbtX, because the first overlaps *ybtQ* by 116 nucleotides, whereas the second methionine has an overlap, with *ybtQ*, of only eight nucleotides (J.D.F., V.J. Bertolino and R.D.P., unpublished observations). The role of YbtE, a homolog of EntE from *E. coli*, in Ybt synthesis has been demonstrated genetically and biochemically [6,26]. YbtT possesses a thioesterase-like domain; the two gene products most closely related to YbtT are AngT and NrpT. AngT is encoded in the anguibactin biosynthetic region of *V. anguillarum* [32] and shows 39% identity and 59% similarity at the predicted amino-acid level to YbtT. The second gene, *nrpT*, encoded within a *Proteus mirabilis* locus required for swarm-cell differentiation [33], has 44% identity and 57% similarity to YbtT. The relevance of YbtT to the Ybt system has not been demonstrated. Like YbtT, YbtU is encoded within the apparent *irp2–ybtE* operon (Figure 2) [6]. Although this suggests that YbtU is involved in Ybt synthesis, transport or regulation, its function is undetermined. It is intriguing that *nrpU*, another gene in the *P. mirabilis* swarm-cell locus [33], has 43% identity and 64% similarity to the predicted amino-acid sequence of YbtU. No other nonyersinial genes currently in the database show significant similarity to *ybtU*.

A number of the Ybt genes in *Y. enterocolitica* have been sequenced — *irp2*, *irp1*, *irp3* (*ybtU*), *irp4* (*ybtT*), *irp5*(*ybtE*) and *fyuA* (*psn*). Comparison of the deduced amino-acid sequences of these genes from *Y. enterocolitica* and *Y. pestis* KIM6+ shows ≥ 98% identity [6,23,24,27,28]. The Ybt systems distributed among the pathogenic yersiniae are therefore nearly identical. This report focuses on *ybtS*, *irp1* and *irp2*, which are considered in detail below.

Figure 3



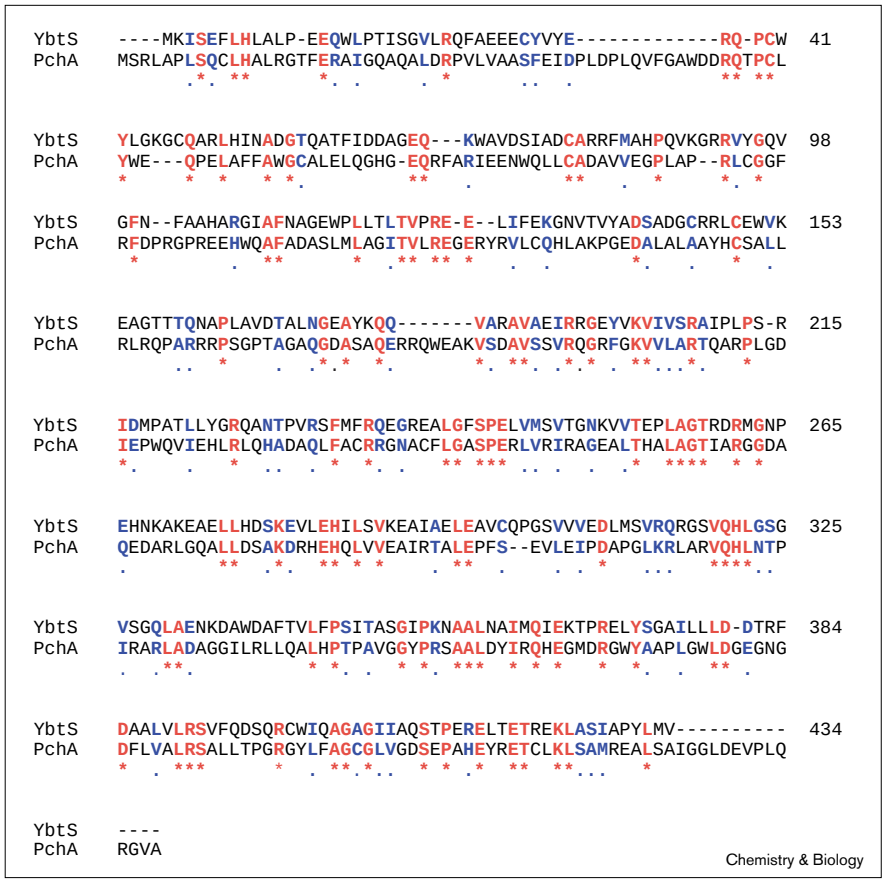
MALDI mass spectra of Fe–yersiniabactin isolated from cultures containing **(a)** L-[methyl-$^2H_3$]methionine (approximately 50 atom %), **(b)** L-[3,3,3′,3′-^{2}H]cystine (98 atom % ^{2}H), or **(c)** only natural abundance isotopes.

Cysteine incorporation into Ybt

Both the structure of Ybt ([17,18], R.D.P., P.B. Balbo, H.A. Jones, J.D.F. and E.D., unpublished observations) and the organization of catalytic domains within HMWP1 and HMWP2 (see below) suggest that three cysteines are incorporated into Ybt. We used matrix-assisted laser desorption/ionization (MALDI) mass analysis of purified Ybt to experimentally demonstrate this. MALDI mass analysis of purified Ybt from cultures without stable isotope supplementation yielded a spectrum with the major ion intensity at m/z 535.035, consistent with the formula $C_{21}H_{25}N_3O_4S_3Fe^+$ (Figure 3c). This ion is that of the intact molecular ion form of iron-bound Ybt. Because of its high relative intensity, the isotope cluster derived from this ion was used for stable-isotope incorporation determinations. When cultures were grown in a deferrated, defined medium, PMH, containing 75 μM L-[3,3,3′,3′-^{2}H]cystine (98 atom % ^{2}H), the mass spectrum of Ybt–Fe^+ showed a

Figure 4

Alignment of predicted amino-acid sequences of YbtS and PchA. Numbering corresponds to YbtS amino-acid residues. Identical amino acids (in red) and conserved residues (in blue) are shown as asterisks and dots, respectively, on a consensus line below the aligned sequences.



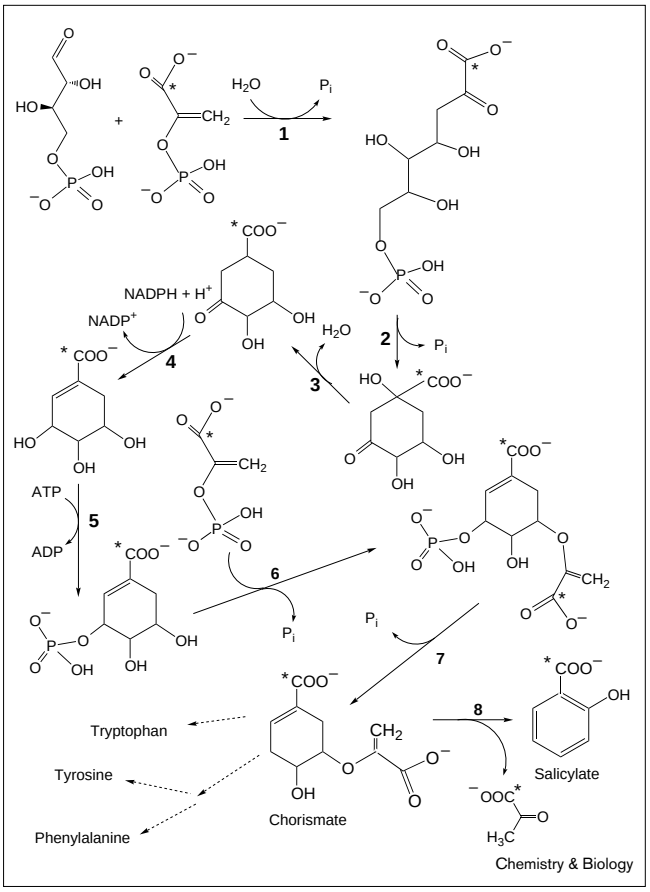
clearly significant enhancement of the M^++2, M^++4, and M^++6 ion intensities (Figure 3b) compared to those seen in the natural abundance mass spectrum of Ybt. These results confirm that three cysteine molecules are incorporated into the Ybt molecule.

Biosynthesis of salicylic acid

YbtS encodes a 434 amino-acid residue protein with a pI of 5.54 and a molecular mass of 47.99 kDa. The intergenic region between the start codon of *ybtS* and the stop codon of *ybtX* is 27 nucleotides. Sequence analysis suggests that *ybtS* is transcribed as part of a putative *ybtPQXS* operon. A BLAST search [34] indicates a high degree of similarity of YbtS to anthranilate synthase component I from a number of different organisms. To a lesser, but still highly significant degree, YbtS also shows similarities to para-aminobenzoate synthases and isochorismate synthases. In *E. coli*, anthranilate synthase is part of the aromatic amino-acid biosynthetic pathway required for tryptophan synthesis. Because *Y. pestis Δpgm* mutants, which have deleted 102 kb of chromosomal DNA in which the Ybt region is encoded [29,30], are not tryptophan auxotrophs (R.D.P., unpublished observations), however, YbtS is not required for tryptophan synthesis. A portion of the aromatic amino-acid biosynthetic pathway appears to be required for synthesis of Ybt because an *aroA* mutation in *Y. enterocolitica* did not secrete a chrome azurol S-reactive compound [35]. PchA, an isochorismate synthase of *P. aeruginosa*, is involved in biosynthesis of the siderophore pyochelin [36,37] and has the highest similarity to YbtS of the isochorismate synthases in the database (Figure 4). Serino *et al.* [36] concluded that, in *P. aeruginosa*, PchA converts chorismate to isochorismate whereas PchB converts isochorismate to salicylate. If YbtS catalyzes the same biochemical reaction as PchA, a PchB homolog should be essential for Ybt synthesis. Sequence analysis failed to identify a *Y. pestis* homolog of PchB anywhere within the high pathogenicity island encoding the *ybt* region (GenBank accession # YBT AF091251. Consequently, either a PchB homolog is encoded outside the pathogenicity island that encodes all other biosynthetic genes or YbtS catalyzes the conversion of chorismate directly to salicylate.

To support the hypothesis that the salicylate moiety in Ybt (Figure 1) is from the aromatic amino-acid biosynthetic pathway, we followed the incorporation of ^{13}C from D-[1-^{13}C]glucose into the Ybt molecule. In this experiment, 75 μM natural abundance cystine was also included

Figure 5



Proposed origin of salicylate as a branching of the aromatic amino-acid biosynthetic pathway. An asterisk indicates the ^{13}C label that originated in D-[1-^{13}C]glucose. Enzymes are indicated as follows: **1**, 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase; **2**, 5-dehydroquinate synthase; **3**, 5-dehydroquinate dehydratase; **4**, shikimate dehydrogenase; **5**, shikimate kinase; **6**, 3-enolpyruvylshikimate-5-phosphate synthase; **7**, chorismate synthase; **8**, *Y. pestis* YbtS enzyme or enzymes.

in deferrated PMH medium; the presence of the cystine and the 1 mM methionine already in PMH medium precluded nonspecific incorporation of ^{13}C into any of the atoms derived from those two amino acids. The results show that essentially no general incorporation of ^{13}C was seen (data not shown). The amount of ^{13}C incorporation into Ybt corresponded with labeling of a single carbon atom of salicylate. This is consistent with the 1-carbon of glucose being incorporated into the carboxyl carbon of salicylate, as one would expect if salicylate is derived from chorismate in a branching off of the aromatic amino-acid biosynthetic pathway (Figure 5).

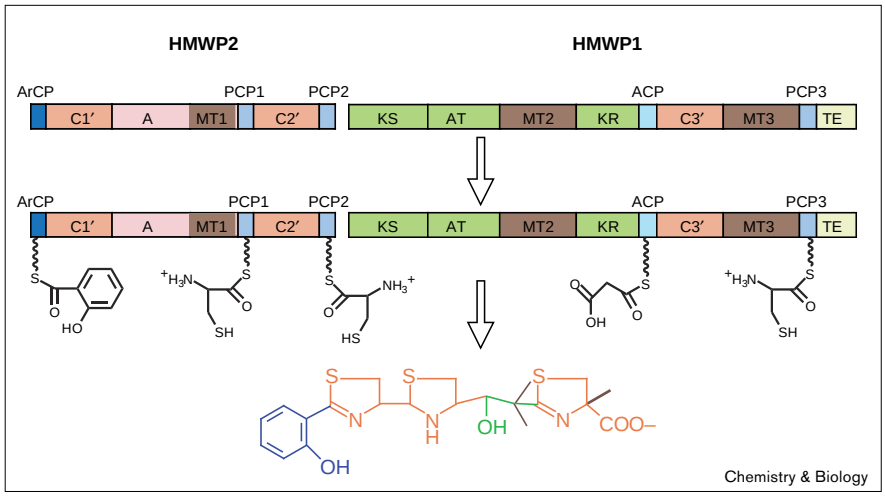
To test whether YbtS is required for Ybt synthesis, we constructed a mutation in the *ybtS* gene. *Y. pestis* KIM6-2070.1 (*ybtS*::*kan2070.1*) was unable to grow on PHM-S plates, indicating a defect in the Ybt system. Cell-free culture supernatants from iron-starved KIM6-2070.1 cells did not support the growth of *Y. pestis* KIM6-2046.1, a strain unable to synthesize Ybt due to a mutation in *irp2*. Thus KIM6-2070.1 is defective in synthesis and/or secretion of Ybt. These results and the sequence of YbtS suggest it is either an isochorismate synthase or a salicylate synthase.

HMWP2: a nonribosomal peptide synthetase with two heterocyclization domains

HMWP2, a 2035 amino-acid protein with a molecular mass of 229 kDa, shows a size and organization typical of nonribosomal peptide synthetases [38] with some intriguing variations. Seven domains are detectable in HMWP2 (Figure 6): an aryl carrier protein (ArCP) domain, residues 1–100; a bond-forming condensation domain with heterocyclization capacity C1′, residues 101–544; an adenylation (A) domain, residues 545–1382 interrupted starting at residue 1011 by the fourth domain (337 amino acids) with homology to cyclosporin synthetase methyltransferase (MT) domains; a second carrier protein domain at residues 1383–1491, designated peptidyl carrier protein 1 (PCP1); a second condensation/heterocyclization (C′) domain C2′, residues 1492–1926 and, finally, a third carrier protein domain designated PCP2, at the carboxyl terminus of HMWP2, residues 1927–2035.

Key to an understanding of the mechanism of Ybt synthesis is the proof that the three 80–100 amino-acid carrier protein domains predicted in HMWP2 are sites of covalent post-translational phosphopantetheinylation (Ser52 in ArCP, Ser1439 in PCP1 and Ser1977 in PCP2). This phosphopantetheinylation was recently demonstrated by the expression and purification of approximately 100-residue fragments of ArCP and PCP1 in *E. coli*, followed by their successful covalent modification with tritiated CoASH in the presence of the serine-residue-specific phosphopantetheinyl transferase (PPTase) EntD [39] from *E. coli* [26]. As demonstrated in Figure 7, the purified PCP2 fragment (residues 1927–2035) of HMWP2 is likewise modified with [^{3}H]-phosphopantetheine (Ppant) in the presence of EntD. This establishes these three modified serine residues, with their 20 Å long phosphopantetheinyl terminal thiols, as 'way stations' for tethering the growing siderophore chain as it progresses down HMWP2. The ArCP and the two PCP domains are indeed differentiated functionally. The ArCP is efficiently acylated with salicyl-AMP by YbtE to initiate siderophore assembly whereas ArCP is only weakly aminoacylated by cysteine and the HMWP2 A domain (Figure 8) [26]. Instead, the A domain of HMWP2 (assayed as the HMWP2 1–1491 fragment) shows specificity for cysteine activation and then will transfer the cysteine group from cysteine-AMP to both PCP1 [26] and, as demonstrated here, PCP2 when presented as separate fragments in *trans* (Figure 8). Thus as shown in Figure 9, the salicyl group and the two cysteinyl groups are covalently deployed as thiolesters to the Ppant

Figure 6



Domain organization of the HMWP2 and HMWP1 subunits of the Ybt synthetase. The seven-domain HMWP2 is organized similarly to other nonribosomal peptide synthetase enzymes, whereas the nine-domain HMWP1 contains both polyketide synthase (green) and nonribosomal peptide synthetase-like domains. Distributed along the Ybt synthetase subunits are five carrier protein domains (shades of blue), each containing a conserved serine residue destined for phosphopantetheinylation (~~S), thereby allowing subsequent covalent acylation of these domains with substrates (salicyl, cysteinyl, malonyl groups) via the attached Ppant cofactor. In the final Ybt product, the hydroxyphenyl moiety derived from salicylate is shown in blue, the cysteine residues that have been heterocyclized to thiazoline/thiazolidine rings (presumably by action of the C′ domains) are shown in orange, the groups incorporated or modified by action of the polyketide synthase active sites are shown in green, and the methyl groups (presumably incorporated by the action of the MT domains) are shown in brown. ArCP, aryl carrier protein; C′, condensation/cyclization; A, adenylation; MT, methyltransferase; PCP, peptidyl carrier protein; KS, β-ketoacylsynthase; AT, acyltransferase; KR, β-ketoreductase; ACP, acyl carrier protein; TE, thioesterase.

termini with the regioselectivity expressed in the buildup of hydroxyphenyl–thiazoline–thiazolidine to yield the 'left-hand' side of Ybt.

It has been demonstrated previously that the condensation/heterocyclization domain (C1′) can fulfil its expected role in the 1–1491 fragment of HMWP2 — the hydroxyphenylthiazoline carboxylic acid was found after hydrolysis of the presumed acyl-S-Ppant–Ser1439 enzyme species [26]. This result established that the C1′ domain is likely a thiazoline-forming catalyst (Figure 9). The condensation/heterocyclization domain C2′ of HMWP2 is presumed to take the hydroxyphenylthiazoline-S-PCP1, **1**, condense it with Cys–S–PCP2 and then cyclize it to yield the hydroxyphenylthiazoline–thiazoline–S–PCP2 acyl enzyme intermediate, **2**, at the carboxyl terminus (Ser1977) of HMWP2 (Figure 9). The growing siderophore chain starts as the thioester on the initiating Ppant of the amino-terminal ArCP (Ser52) domain and then moves (via PCP1) with two bond-forming/thiazoline-forming steps to the carboxyl terminus of HMWP2, tethered on PCP2 (Figure 9), where **2** presumably will be handed off to a partner HMWP1 subunit in a heterooligomeric complex. The C1′ and C2′ domains in HMWP2 cluster in a subset of condensation domains in the nonribosomal peptide synthetase families, a subset found in the *Bacillus licheniformis* bacitracin synthetase that also conducts cysteine heterocyclization attendant to peptide-bond formation [40]. Residues shared by these heterocyclizing condensation domains are good candidates for structure–function mutagenesis studies to attempt uncoupling of the condensation and heterocyclization process, and for domain swaps in combinatorial biosynthesis to attempt designed heterocyclization of peptide moieties at specific cysteine or serine residues.

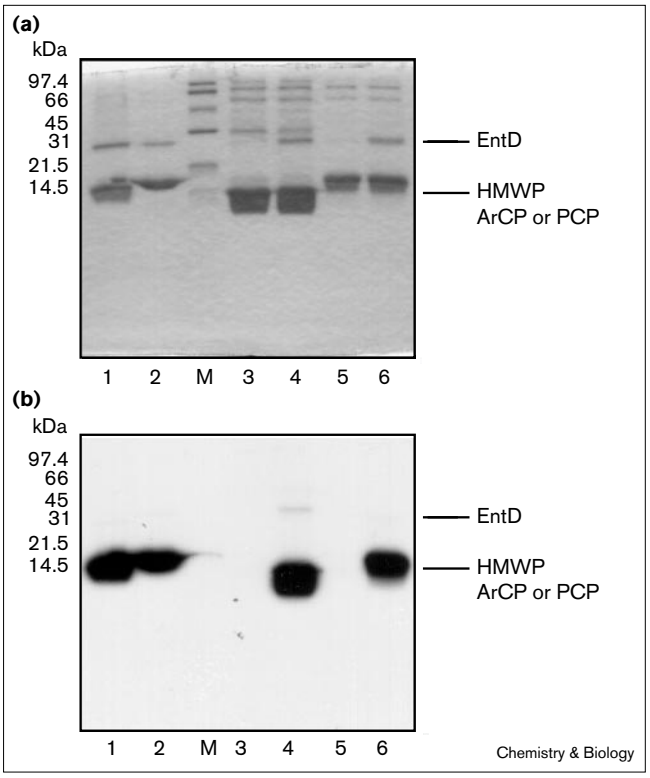
In addition to the remarkable cysteine-heterocyclization activity of HMWP2, the arrangement of the PCP1 and PCP2 domains is novel. Heretofore PCP domains covalently loaded with an aminoacyl fragment have been immediately downstream in the primary sequence of the aminoacyl-AMP-generating adenylation domains [38]. The three-dimensional architecture of HMWP2 must allow a close approach of both PCP1 and PCP2 to the A domain, as there is catalytic specificity imposed — the cysteine–AMP is not just a chemically reactive aminoacylating species fired off into the microenvirons by the A domain.

HMWP1: a mixed polyketide synthase/nonribosomal polypeptide synthetase with heterocyclization capacity

Primary sequence analysis of the *Y. pestis irp1* gene encoding HMWP1 reveals nine discernible domains in its 3163 amino-acid residues. The *Y. pestis* HMWP1 and the recently reported *Y. enterocolitica* 3161-residue HMWP1 [28] are highly related (97.9% identical). In their analysis of *Y. enterocolitica irp1*, Pelludat *et al.* [28] noted only homology to ketosynthase active sites of fatty acid synthases and polyketide synthases, as well as some general homology in the carboxy-terminal half of HMWP1 to regions of HMWP2.

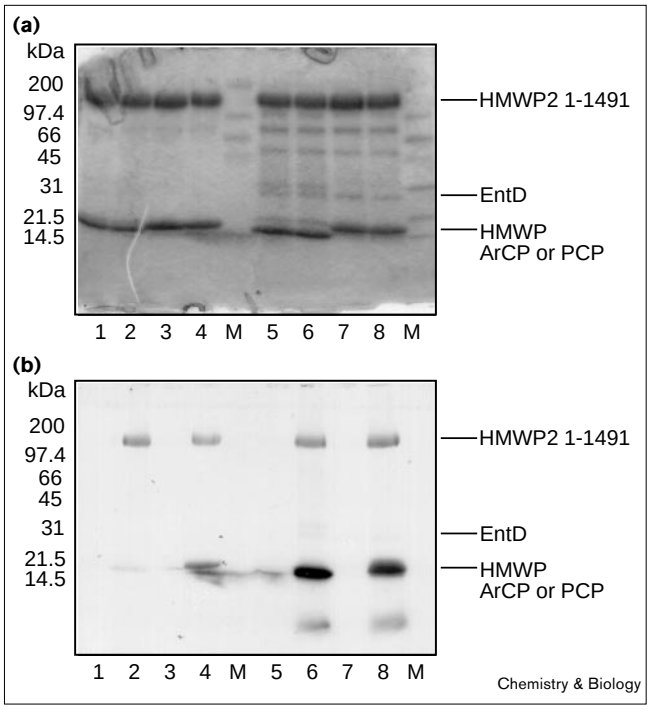
Here we analyze the putative functions of all nine of the domains that account for the full sequence of *Y. pestis* HMWP1. This megasynthetase, probably acting to

Figure 7



Demonstration of the phosphopantetheinylation of the Ybt synthetase ArCP and PCP domains dependent on *E. coli* EntD. **(a)** Coomassie-stained gel (15%) of TCA-precipitated reaction mixtures. **(b)** Autoradiograph of this gel. All reaction mixtures contained [^{3}H]CoASH (150 μM, 192 Ci/mol) and an apo-carrier protein domain substrate (6 μM). The PPTase enzyme EntD (800 nM) was added to reaction mixtures as indicated. Lane 1, ArCP and EntD; lane 2, PCP1 and EntD; lane M, molecular weight markers; lane 3, PCP2; lane 4, PCP2 and EntD; lane 5, PCP3; lane 6, PCP3 and EntD.

Figure 8



Demonstration of the ATP-dependent covalent loading of the Ybt synthetase PCP domains with [^{35}S]-L-cysteine. **(a)** Coomassie-stained gel of TCA-precipitated reaction mixtures (4–20%). **(b)** Autoradiograph of this gel. All reaction mixtures contained [^{35}S]-L-cysteine (0.7 mM, 285 Ci/mol), apo-HMWP2 1–1491 (1 μM) and a carrier protein domain substrate (15 μM). In the reactions shown in the odd-numbered lanes, ATP was omitted, whereas ATP was included in the even-numbered lanes (3 mM). Lanes 1 and 2, holo-ArCP; lanes 3 and 4, holo-PCP1; lanes 5 and 6, apo-PCP2; lanes 7 and 8, apo-PCP3; lane M, molecular weight markers. When apo-proteins were used as cysteinylation substrates (PCP2 and PCP3), CoASH and the PPTase EntD were included in the reaction mixture to allow generation of the holo-protein.

introduce the branched isobutyryl-alcohol linker and the last thiazoline moiety in Ybt assembly, looks to be organized as a polyketide/fatty acid synthase in its first 1895 residues and then as a modified peptide synthetase for residues 1896–3163. It notably lacks any semblance of an A domain (Figure 6). There are two discernible consensus serine phosphopantetheinylation sites, at Ser1853 and Ser2858, bringing the total of carrier protein domains in HMWP2 and HMWP1 together to five. The sequence at Ser1853 looks like a canonical acyl carrier protein (ACP) domain and is positioned as one would find it in a fatty acid or polyketide synthase [41]. The Ser2858 sequence fits the PCP motif [38] and so is designated PCP3, consistent with the need to activate a third cysteine molecule in the late stages of Ybt formation.

A mixed polyketide synthase/nonribosomal peptide synthetase organization is not unique to the yersiniabactin synthetase (HMWP1 and HMWP2). A strikingly similar example, NrpS, has recently been uncovered in a *P. mirabilis* locus required for swarm-cell differentiation [33]. The NrpS mixed synthetase is arranged like the carboxy-terminal portion of HMWP1 — containing β-ketoreductase (KR), ACP, C′, PCP, and thioesterase (TE) domains (Figure 6). The amino-terminal portion of NrpS lacks the β-ketoacylsynthase (KS), acyl transferase (AT), and MT domains of HMWP1 (Figure 6), however, and instead has a short sequence with similarity to a *B. subtilis* polyketide synthase [33]. We would therefore expect that the product of the NrpS synthetase, which is presumably the final component of a synthetase complex, to resemble the right-hand portion of Ybt with a polyketide linker joined to a methylated thiazoline or oxazoline ring. It is clear from the NrpS example that features of the Ybt synthetase which initially seemed unique (e.g., use of a single adenylation domain to load multiple, nonadjacent PCP domains) are likely to be more widespread and will be uncovered as more bacterial genes and genomes are sequenced.

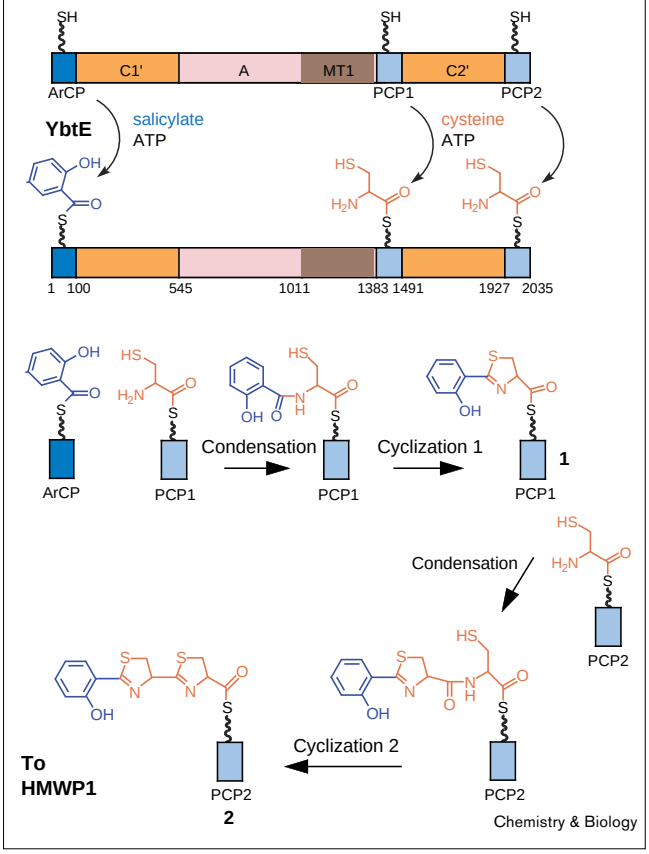
The provisional assignment of ACP and PCP3 function accords with the multidomain synthase logic that the placement of domains dictates the order of transformations in the growing natural product acyl chain [38]. To validate experimentally the function of the putative PCP3, we have overproduced residues 2808–2912 of HMWP1 as a carboxy-terminal hexahistidine-tagged fusion in *E. coli* and have purified this fragment. As shown in Figure 7, the apo-PCP3 fragment is converted to its phosphopantetheinylated holo form upon incubation with pure *E. coli* EntD PPTase and tritiated CoASH. Furthermore, when holo-PCP3 was then incubated with pure HMWP2 1–1491 fragment, ATP and [^{35}S]-cysteine, the radiolabel was specifically transferred to yield cysteinyl-S-PCP3 (Figure 8). These results validate two key predictions: that PCP3 is a way station for loading the third cysteinyl precursor of yersiniabactin and that the cysteine-specific A domain of HMWP2 can recognize PCP3 in *trans*.

The amino-terminal half of HMWP1 as a polyketide synthase array

The first 1895 of the 3163 residues of HMWP1, which contain a predicted KS domain (residues 1-490), AT domain (residues 491–940) and KR domain (residues 1430–1812) just upstream of the ACP domain (residues 1813–1895), are prototypic for a coordinated fatty acid synthase/polyketide synthase module (Figure 6) [41]. These catalysts could take an acyl CoA (e.g., malonyl CoA), transfer it onto a serine residue (e.g., Ser641) of the AT domain, and then carry out an acyl O→S shift onto the terminal thiol of the Ppant moiety of the ACP domain (e.g., Ser1853). Decarboxylation under the influence of the KS domain would yield a β-carbanion to act as nucleophile in a C–C bond-forming step. Indeed, the results of the D-[1-^{13}C]glucose labeling experiment that show incorporation of only a single atom of ^{13}C into both the intact Ybt–Fe molecule, as well as an m/z 295 fragment, are consistent with incorporation from a malonyl group because decarboxylation of the malonyl moiety would remove the ^{13}C-labeled malonyl carboxyl group. Additionally, the 2- and 3-carbons of the malonyl moiety would not be labeled with ^{13}C from D-[1-^{13}C]glucose. We would therefore not expect to see any ^{13}C incorporation from malonyl-CoA into Ybt, again consistent with our results.

The C2 acetyl-S-ACP carbanion formed by the polyketide synthase activities of HMWP1 could then attack the hydroxyphenyl bisthiazolinyl-S-PCP2, **2**, at the carboxyl terminus of HMWP2, as suggested in Figure 10, to transfer that bisthiazoline acyl group to the ACP domain of HMWP1 and produce the β-keto bisthiazolinyl-S-ACP acyl enzyme intermediate, **3**, on HMWP1. Such a species is a typical substrate for reduction of the β-keto to a β-hydroxyl group by NAD(P)H-utilizing KR domains of polyketide synthases, which would yield the secondary alcohol moiety found in Ybt. Additionally, the adduct **3** (Figure 10), from

Figure 9

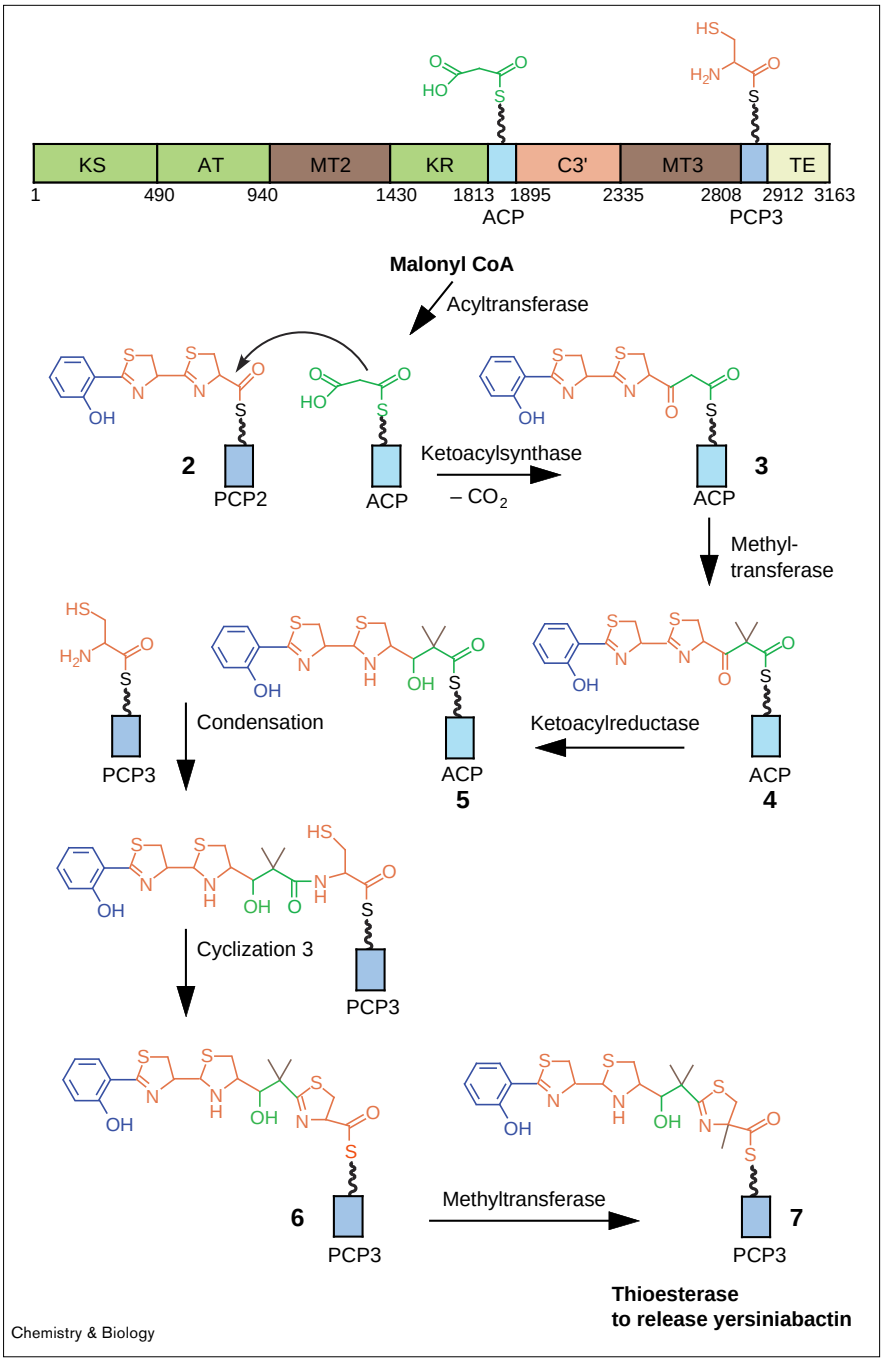


Proposed mechanism for Ybt biosynthesis by HMWP2. With its ArCP, two PCP and two C′ domains, HMWP2 should be able to condense and heterocyclize one salicylate and two cysteine moieties to give the enzyme-linked hydroxyphenyl-thiazolinyl-thiazoline species (**2**) thioesterified to the PCP2 domain. The loading of HMWP2 with salicylate and cysteine on the indicated carrier protein domains has been demonstrated in this and previous studies [26]. The release of hydroxyphenylthiazoline carboxylate from a HMWP2 1–1491 enzyme fragment [26] provides evidence for the formation of intermediate **1**. In this study, the single adenylation domain of HMWP2 (assayed as the 1–1491 fragment) has been shown to catalyze the covalent cysteinylation of all three of the Ybt synthetase PCP domains.

initial C–C bond formation that translocates the growing chain from the carboxyl terminus of HMWP2 to the ACP domain in HMWP1, contains the imine nitrogen in the first thiazoline that undergoes reduction. The C=N bond could isomerize from C5=N into a C2=N species that would be conjugated to the β-keto group. The KR domain could reduce both C=N and C=O bonds (this reduction predicts that the H* from NADPH* will end up at C5 of the thiazolidine) to produce the thiazoline–thiazolidine signature redox state of Ybt, **5**.

The only unexpected domain of the first five HMWP1 domains is that of residues 941–1430, with all the hallmarks

Figure 10



Proposed mechanism for Ybt biosynthesis by HMWP1. The mixed polyketide synthase/nonribosomal peptide synthetase HMWP1 presumably completes the synthesis of Ybt by elongation of intermediate **2** obtained from HMWP2. The AT domain of HMWP1 could load the central ACP domain with a malonyl group derived from malonyl-CoA, and in a decarboxylative condensation reaction catalyzed by the KS domain, intermediate **3** could be formed with transfer of the elongating Ybt molecule from HMWP2 to the ACP domain of HMWP1. SAM-dependent methyltransfer followed by reduction would give, respectively, intermediates **4** and **5**. The C3′ domain of HMWP1 could then catalyze the condensation/heterocyclization of the final Cys residue, thioesterified to PCP3, with **5** to give the desmethyl intermediate **6**. SAM-dependent C-methylation of the last thiazoline ring would yield the completed Ybt molecule (**7**) still covalently tethered to HMWP1. Hydrolysis of the tethering thioester linkage by the carboxy-terminal TE domain would release Ybt from the synthetase, completing its biosynthesis.

of an MT domain [42], MT2, interposed between the AT and KR domains of the polyketide synthase architecture (Figure 6). The location of this MT domain suggests the origin of the two methyl groups of the isobutyryl moiety in Ybt. Although it remains to be shown experimentally whether the HMWP1 AT domain is specific for malonyl CoA or methylmalonyl CoA (homologies suggest malonyl CoA [43]), both of the branching C-methyl groups, as well as the methyl group of the α-carbon of the last thiazoline should come from *S*-adenosylmethionine (SAM). To experimentally determine the origin of the Ybt methyl groups, biosynthetic labeling studies of Ybt with [methyl-$^2H_3$]methionine were performed. MALDI analysis of purified Ybt (Figure 3a) clearly showed enhancement or appearance of the M^++3, M^++6, and M^++9 ions derived from the m/z 535 molecular ion. These ions arise from the

incorporation of three methyl groups from the [methyl-$^2H_3$]methionine precursor. All three of the methyl groups (at approximately 50 atom % ^2H) of Ybt (Figure 1) are therefore derived from the methyl group of methionine, indicating that these three methyl groups all originate via methylation from SAM and that malonyl CoA is the substrate of the AT domain.

C-methylation requires a carbanion to attack the C1-methyl fragment of SAM, and a carbanion of the β-ketobisthiazolinyl-S-ACP, **3**, (before reduction by the KR domain) is the probable nucleophile. It seems reasonable to propose methylation prior to reduction because methylation should prevent enol formation and thus stabilize the keto substrate for reduction. Bismethylation, perhaps catalyzed by MT2, could occur here to produce **4** as the substrate for the double reduction to **5** by the KR domain (Figure 10).

The carboxy-terminal half of HMWP1 for heterocyclization and chain release

Four domains are discernible in the remaining 1268 residues (1896–3163) of HMWP1 (Figure 6); the sequence motifs and the attendant chemical transformations show a switch from polyketide synthase back to polypeptide synthetase logic. The sixth domain, residues 1896–2335, is another condensation/cyclization domain, the third of the two-enzyme complex, designated C3′, which suggests another bond forming/heterocyclization step. In this step, the carbonyl tethered to the S-ACP domain is thought to be linked to the incoming nucleophile of the third cysteine (previously loaded onto PCP3, the eighth HMWP1 domain). Condensation of **5** with this PCP3 tethered cysteinyl group concomitant with heterocyclization would yield the desmethyl version **6** of Ybt with the acyl chain now tethered at the fifth and last Ppant way station of the two enzyme complex (Figure 10). The most surprising attribute of HMWP1 is that it has no A domain. It must rely on the A domain in HMWP2. Indeed the data shown in Figure 8 demonstrate that a purified PCP3 fragment, converted from apo to phosphopantetheinylated holo form by EntD, will serve in *trans* as a substrate for cysteinylation by the A domain of the 1–1491 fragment of HMWP2. This demonstrates that the A domain can act intermolecularly as well as intramolecularly, to load all three PCP domains with cysteine.

The seventh domain of HMWP1, between C3′ and PCP3, has homology to MT1 and MT2 and is designated MT3 (residues 2236–2807). It is most probably the C-methylation catalyst at the α-carbon of the last thiazoline. At the thioester stage of **6** the α-carbon is acidic, and a carbanion would be stabilized by the acyl-S-PCP3 linkage enough to form and attack an equivalent of SAM to produce the C-methyl species **7** that is the last substituent to be added to complete the Ybt skeleton (Figure 10). The purpose of this C-methylation has yet to be examined, but would prevent oxidation up to the aromatic thiazole level where the ring nitrogen would have lost basicity and metal-coordination affinity. Thus the α-methyl thiazoline, seen in other natural products such as thiangazole [44], could optimize iron-coordination sites in an oxidizing extracellular environment for these thiazoline siderophores.

The last domain of HMWP1 (residues 2913–3163), the 16th of the two-protein system, is designated TE because of its homology to the thioesterase motifs typically found at the carboxyl terminus of nonribosomal peptide synthetases. The TE domain most likely catalyzes acyl transfer of **7** to water (Figure 10), liberating the PCP3 of its fully formed Ybt substituent as it clears the acyl chain from covalent tethering.

Significance

Delineation of the genes for siderophore biosynthesis and definition of the function of specific open reading frames constitutes a first set of steps in the design of strategies for inhibition of yersiniabactin (Ybt) synthesis and induction of avirulence in the pathogenic species of *Yersinia*. Determination of the sequence of the *Y. pestis irp1* gene product, the 3163 residue HMWP1, has allowed primary sequence analysis, which suggests nine functional domains. Comparable analysis of the *irp2* gene product, the 2035 residue HMWP2, previously suggested seven domains [26]. The 16-domain organization of HMWP2 followed by HMWP1 suggests a two-enzyme complex that can convert salicylate, three cysteines, malonyl CoA and three methyl groups (from three S-adenosylmethionine molecules) into the tetracyclic siderophore Ybt, a known virulence factor of the yersiniae.

This analysis not only illuminates the molecular logic for Ybt assembly, using a mixed polyketide synthase/nonribosomal peptide synthetase strategy, but also reveals some novel features of domain construction and interaction in these multidomainal megasynthetases. The two-enzyme yersiniabactin synthetase (HMWP2 and HMWP1) therefore uses all three variants of phosphopantetheinyl carrier domain way stations, an ArCP domain for aryl *N*-cap chain initiation, three PCP domains for the amino-acid incorporation (three cysteines), as well as an ACP, presumably for malonyl tethering. In addition, there is but a single cysteine-activating adenylation domain to service the three PCP domains (the first instance of this layout in a nonribosomal peptide synthetase), which demands architectural flexibility within the HMWP1–HMWP2 complex to bring all three holo-PCP sites to the microenvirons of the cysteine-AMP loaded A domain. Finally, all three condensation domains, C1′ and C2′ in HMWP2 and C3′ in HMWP1, are likely to be heterocyclizing rather than simple peptide-bond forming condensation domains and have distinctive signature motifs. The modular logic as

revealed by the Ybt synthetase will probably be extended to synthetases in the biogenesis of related thiazoline and oxazoline siderophores, such as pyochelin in pseudomonads, vibriobactin in *Vibrio cholerae* and mycobactin in *Mycobacterium tuberculosis* and might also facilitate delineation of autonomous folding domains, which could be used for combinatorial swapping strategies.

Materials and methods

Plasmid, bacterial strains and cultivation

For plasmid isolation, cells of *E. coli* strains DH5α or DH5α(λ*pir*) (obtained from S.C. Straley) or *Y. pestis* were grown overnight at 37°C in Terrific broth [45] or Heart Infusion broth (Difco), respectively. Plasmids were isolated using alkaline lysis [46] and further purified, when necessary, using polyethylene glycol precipitation [47]. Recombinant plasmids pSDR498.3 and pSDR498.4 are subclones of pSDR498 [29] that together encompass the entire Ybt region. Suicide plasmid pSUCYbtS was constructed as follows. An ~4.6 kb *Bam*HI–*Eco*RI fragment from pSDR498 was ligated into pHC79 [48] and designated pSDR498.1. An ~2 kb *Sma*I–*Eco*RI fragment from pSDR498.1 was ligated into pBluescriptKS. The resulting plasmid pBSYbtS was digested with *Hpa*I and ligated with an ~1.3 kb *Hinc*II fragment containing the *kan* gene cassette from pUC-4K (Pharmacia Biotech) to yield pBSYbtSKan. An ~3.3 kb *Eco*RV–*Bam*HI fragment from pBSYbtSKan was ligated into the *Sma*I and *Bam*HI restriction sites in suicide plasmid pSUC1 (J.D.F. V.J. Bertolino and R.D.P., unpublished observations). The resulting suicide plasmid, pSUCYbtS, which contains a *kan* insert in *ybtS*, was used for allelic exchange. pPSN3 contains a 10 kb *Sal*I insert in pBGL2. The *Sal*I insert was cloned from pSDR498.4 [29] and encompasses most of *irp2* through to the *psn* gene.

All of the *Y. pestis* strains used in this study were derived from strain KIM6+, which contains the endogenous plasmids pMT1 and pPCP1. KIM6+ lacks the ~70 kb pCD1 plasmid that encodes essential calcium- and temperature-regulated virulence genes that are part of the low-calcium response stimulon [7]. *Y. pestis* KIM6-2045.1 contains an in-frame deletion in *psn*, which encodes the receptor for Ybt and the bacteriocin pesticin, and is unable to utilize Ybt. KIM6-2046.1 possesses an *irp2*::*kan2046.1* mutation that prevents synthesis of Ybt. KIM6-2046.1 cells can utilize exogenously supplied Ybt [23]. The *ybtS*::*kan* mutation was constructed by electroporation [23] of pSUCYbtS into KIM6+. Merodiploids were selected for growth on tryptose blood agar base plates containing kanamycin and carbenicillin at 50 μg/ml. Plasmid profiles and polymerase chain reaction (PCR) analysis with *ybtS*-specific oligonucleotides demonstrated that pSUCYbtS had integrated properly. One cointegrant, KIM6-2070, was grown overnight in Heart Infusion broth, diluted to an OD_{620nm} of 0.01, and 50 μl were spread onto a Tryptose Blood Agar Base (Difco) plate containing kanamycin (50 μg/ml) and 5% (w/v) sucrose. PCR analysis with *ybtS*-specific oligonucleotides of selected Suc^r colonies, identified clones that had lost the pSUC1 sequences and exchanged the mutated *ybtS* for the wild-type *ybtS* gene. One isolate was selected and designated KIM6-2070.1. For analysis of iron-deficient cells or culture supernatants, *Y. pestis* cells were grown for 6–9 generations in the defined, deferrated medium PMH as described previously [23,49].

Ybt bioassay

Iron-starved *Y. pestis* cells were streaked onto PMH-S plates and incubated at 37°C; strains defective in Ybt synthesis or transport are unable to grow on PMH-S plates. To test for Ybt synthesis, ~25 μl of culture supernatants from iron-starved cells were added to wells of PMH-S plates overlayered with KIM6-2045.1 cells as described previously [23]. KIM6-2045.1 cells are unable to grow on PMH-S plates unless supplied with exogenous Ybt.

Stable-isotope labeling and analysis

Y. pestis KIM6-2045.1 (a Ybt receptor mutant that overproduces Ybt but cannot use it) [23] was grown for ~8 generations in deferrated PMH with addition of labeled compounds as indicated. Ybt was isolated as described previously (R.D.P., P.B. Balboa, H.A. Jones, J.D.F. and E.D. *et al.*, unpublished observations). Briefly, Ybt was extracted with ethyl acetate. The extract was reduced in volume by evaporation, brought to approximately 60% ethanol, then passed through a C-18 SEP-PAK cartridge. Final purification was performed on a semipreparative C18 high performance liquid chromatography (HPLC) column with a methanol gradient. Incorporation of labeled precursor molecules into the iron-saturated form of Ybt was analyzed by high resolution MALDI mass analysis (positive ion mode) on an IonSpec 4.7 Tesla FT–MS system at the University of Kentucky Mass Spectrometry Facility.

Three compounds labeled with stable isotopes were tested for their possible role in Ybt biosynthesis. Deferrated PMH was supplemented with a) 75 μM L-[3,3,3′,3′-^{2}H]cystine (98 atom % ^{2}H); b) 1 mM L-[*methyl*-$^2H_3$]methionine (98 atom % ^{2}H) in addition to the 1 mM natural abundance methionine already in PMH yielding an atom % ^{2}H of approximately 50; or c) 75 μM natural abundance cystine and ^{13}C from D-[1-^{13}C]glucose to yield a final atom % ^{13}C of approximately 50% due to the 10 mM natural abundance D-glucose already in PMH.

DNA sequencing and analysis

Libraries were prepared from nebulized, size-fractionated DNA [50] from pSDR498.3 and pSDR498.4 DNA in the M13 Janus vector [51] and DNA templates were purified from random library clones [52]. Sequences were collected using dye-terminator labeled fluorescent cycle sequencing Prism reagents and ABI377 automated sequencers (Applied Biosystem Division of Perkin-Elmer). Sequences were assembled into contigs by the SeqMan II program (DNASTAR), and clones were selected for sequencing from the opposite end to fill-in coverage, resolve ambiguities and close gaps [51]. The DNA sequence has been deposited in the GenBank database and assigned accession number YBT AF091251. Homology searches were performed using BLAST [34]. Pairwise predicted amino-acid sequence alignments of YbtS versus PchA, YbtT versus NrpT, and YbtU versus NrpU were performed using CLUSTALW [53].

Overproduction and purification of the HMWP2 PCP2 and HMWP1 PCP3 domains

PCP domain fragments of *Y. pestis* HMWP2 (PCP2, residues 1927–2035) and *Y. pestis* HMWP1 (PCP3, residues 2808–2912) were overproduced in *E. coli* as carboxy-terminal hexahistidine-tagged fusion proteins. Gene fragments corresponding to the desired PCP domains were amplified by PCR using the *Pfu* polymerase (Stratagene) and then cloned into the *Nde*I/*Xho*I sites of pET22b (Novagen) to give plasmids pET22b-irp2PCP2 and pET22b-irp1PCP3. The gene fragment corresponding to PCP2 was amplified from the pIRP2 template [2] using the primer pair 5′-GAATTC***CATATG***GCAGAGCGCTCCC CGCGCG-TATGC-3′ and 5′-TGACCG***CTCGAG***TATCCGCCGCTGACGACG G-3′ (Integrated DNA Technologies, restriction sites bold and italicized). The gene fragment corresponding to PCP3 was amplified from the pPSN3 template using the primer pair 5′-GAATTC***CATATG***GCTCCG TCTGATGCGCCGACTGAGC-3′ and 5′-TGACCG***CTCGAG***ACCG-TCGCCC TGGCAAAGGGGT-3′. Fidelity of the inserts cloned into pET22b was confirmed by DNA sequencing (Dana Farber Molecular Biology Core Facility, Boston, MA).

Cultures of *E. coli* strain BL21(DE3) transformed with pET22b-irp2PCP2 or pET22b-irp1PCP3, for the overproduction of HMWP2 PCP2 and HMWP1 PCP3 respectively, (2 l, 2× YT or LB media, 50 μg/ml ampicillin) were grown at 37°C to an optical density (600 nM) of approximately 0.8 and then induced with 1 mM isopropyl β-D-thiogalactopyranoside (IPTG). Cells were harvested after an additional 3.5 h of growth, and the cell paste was resuspended in 20 mM Tris-HCl, 0.5 M NaCl, 5 mM imidazole, pH 7.9. Cells were lysed by two passages through a French pressure cell at 15,000 psi, and the lysate was clarified by centrifugation (27,000× g).

The PCP proteins were purified from these lysates by nickel chelate affinity chromatography over His•Bind resin (5 ml column) according to the manufacturer's specifications (Novagen). Fractions containing PCP2 or PCP3 were pooled and dialyzed against 50 mM Tris-HCl, pH 8.0, 2 mM DTT, 5% glycerol, flash frozen in liquid nitrogen and stored at –80°C. The protein concentrations of the PCP2 and PCP3 fractions were estimated using the Bio-Rad Protein Assay Reagent.

Assays for phosphopantetheinylation and cysteinylation of carrier protein domains

Assays for detection of the covalent phosphopantetheinylation or cysteinylation of the *Y. pestis* carrier protein domains were carried out as described previously [26]. Briefly, for the detection of phosphopantetheinylation, the apo carrier protein substrate (6 µM HMWP2 ArCP, HMWP2 PCP1, HMWP2 PCP2 or HMWP1 PCP3) was incubated with 75 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 5 mM DTT, 150 µM [^{3}H]-Coenzyme A (CoASH; 192 Ci/mol, 70% label in Ppant) and in some cases the purified PPTase enzyme EntD (800 nM). Incubation of the reaction mixture (100 µl final volume) was at 37°C for 15 min prior to quenching with 10% trichloroacetic acid (TCA).

For the detection of covalent cysteinylation, the carrier protein substrate (15 µM HMWP2 holo-ArCP, HMWP2 holo-PCP1, HMWP2 apo-PCP2, or HMWP1 apo-PCP3) was incubated with 75 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$, 5 mM DTT, 700 µM [^{35}S]-L-cysteine (285 Ci/mol), 1 µM of the cysteinylation catalyst HMWP2 1–1491, and where indicated 3 mM ATP. The reactions with apo-PCP2 and apo-PCP3 additionally contained 500 µM CoASH and 400 nM EntD and were allowed to preincubate for 10 min at room temperature prior to the addition of ATP, cysteine or HMWP2 1–1491. Incubation of the complete reaction mixture (100 µl final volume) was at 37°C for 15 min prior to quenching with 10% TCA.

Subsequent to quenching with TCA, both the phosphopantetheinylation and cysteinylation reaction mixtures were subjected to SDS–PAGE and autoradiography. To prepare the sample for SDS–PAGE, the protein pellet from the TCA precipitation was dissolved in SDS sample buffer (20 µl) and 1 M Tris base (3 µl) prior to loading on the gel. For visualization following electrophoresis, the gel was stained with Coomassie blue solution, destained and soaked in Amplify (Amersham) for 25 min. The dried gels were exposed to film for 1 day and 4 days, respectively, prior to film development for the phosphopantetheinylation and cysteinylation autoradiographs.

The HMWP2 ArCP, PCP1, and 1–1491 fragments used in these experiments were prepared as described previously [26].

Acknowledgements

Mass spectrometry was performed at the University of Kentucky Mass Spectrometry Facility by Amy Harms and Jan Pyrek. This work was supported by National Institutes of Health grant AI042736 (C.T.W., R.D.P. and E.D.). A.M.G. is a Howard Hughes Medical Institute Predoctoral Fellow. F.R.B. and G.F.M. were supported by PHS grant PO1 HG01428. J.D.F. was supported by National Institutes of Health grants AI25098 and AI33481. I.M. was a visiting scientist from Novartis, Japan. We thank Guy Peyrot for assistance with sequence editing.

References

1. Griffiths, E. (1987). Iron in biological systems. In *Iron and Infection: Molecular, Physiological and Clinical Aspects.* (Bullen, J.J. & Griffiths, E., eds) pp.1-25, John Wiley & Sons, New York, New York.
2. Griffiths, E. (1987). The iron-uptake systems of pathogenic bacteria. In *Iron and Infection: Molecular, Physiological and Clinical Aspects* (Bullen, J.J. & Griffiths, E., eds) pp. 69-137, John Wiley & Sons, New York, New York.
3. Guerinot, M.L. 1994. Microbial iron transport. *Annu. Rev. Microbiol.* **48**, 743-772.
4. Griffiths, E. & Bullen, J.J. (1987). Iron-binding proteins and host defence. In *Iron and Infection: Molecular, Physiological and Clinical Aspects* (Bullen, J.J. & Griffiths, E., eds) pp. 171-209, John Wiley & Sons, New York, New York.
5. Mietzner, T.A. & Morse, S.A. (1994). The role of iron-binding proteins in the survival of pathogenic bacteria. *Annu. Rev. Nutr.* **14**, 471-493.
6. Bearden, S.W., Fetherston, J.D. & Perry, R.D. (1997). Genetic organization of the yersiniabactin biosynthetic region and construction of avirulent mutants in *Yersinia pestis. Infect. Immunol.* **65**, 1659-1668.
7. Perry, R.D. & Fetherston, J.D. (1997). *Yersinia pestis* – etiologic agent of plague. *Clin. Microbiol. Rev.* **10**, 35-66.
8. Henderson, D.P. & Payne, S.M. (1994). *Vibrio cholerae* iron transport systems: roles of heme and siderophore iron transport in virulence and identification of a gene associated with multiple iron transport systems. *Infect. Immunol.* **62**, 5120-5125.
9. Cox, C.D. (1982). Effect of pyochelin on the virulence of *Pseudomonas aeruginosa. Infect. Immunol. 36*, 17-23.
10. Meyer, J.-M., Neely, A., Stintzi, A., Georges, C. & Holder, I.A. (1996). Pyoverdin is essential for virulence of *Pseudomonas aeruginosa. Infect. Immunol.* **64**, 518-523.
11. Harris, W.R., *et al.,* & Raymond, K.N. (1979). Coordination chemistry of microbial iron transport compounds. 19. Stability constants and electrochemical behavior of ferric enterobactin and model complexes. *J. Am. Chem. Soc.* **101**, 6097-6104.
12. Neilands, J.B. (1981). Microbial iron compounds. *Annu. Rev. Biochem.* **50**, 715-731.
13. Jalal, M.A.F., Hossain, M.B., van der Helm, D., Sanders-Loehr, J., Actis, L.A. & Crosa, J.H. (1989). Structure of anguibactin, a unique plasmid-related bacterial siderophore from the fish pathogen *Vibrio anguillarum. J. Am. Chem. Soc.* **111**, 292-296.
14. Yamamoto, S., Okujo, N. & Sakakibara, Y. (1994). Isolation and structure elucidation of acinetobactin, a novel siderophore from *Acinetobacter baumannii. Arch. Microbiol.* **162**, 249-254.
15. Cox, C.D., Rinehart, K.L., Jr., Moore, M.L. & Cook, J.C., Jr. (1981). Pyochelin: novel structure of an iron-chelating growth promoter for *Pseudomonas aeruginosa. Proc. Natl Acad. Sci. USA* **78**, 4256-4260.
16. Griffiths, G.L., Sigel, S.P., Payne, S.M. & Neilands, J.B. (1984). Vibriobactin, a siderophore from *Vibrio cholerae. J. Biol. Chem.* **259**, 383-385.
17. Chambers, C.E., McIntyre, D.D., Mouck, M. & Sokol, P.A. (1996). Physical and structural characterization of yersiniophore, a siderophore produced by clinical isolates of *Yersinia enterocolitica. BioMetals* **9**, 157-167.
18. Drechsel, H., *et al.,* & Jung, G. (1995). Structure elucidation of yersiniabctin, a siderophore form highly virulent *Yersinia* strains. *Liebigs. Ann.* **1995**, 1727-1733.
19. Buchrieser, C., Prentice, M. & Carniel, E. (1998). The 102-kilobase unstable region of *Yersinia pestis* comprises a high-pathogenicity island linked to a pigmentation segment which undergoes internal rearrangement. *J. Bacteriol.* **180,** 2321-2329.
20. Carniel, E., Builvout, I. & Prentice, M. (1996). Characterization of a large chromosomal high-pathogenicity island in biotype 1B *Yersinia enterocolitica. J. Bacteriol.* **178**, 6743-6751.
21. Carniel, E., Guiyoule, A., Guilvout, I., & Mercereau-Puijalon, O. (1992). Molecular cloning, iron-regulation and mutagenesis of the *irp2* gene encoding HMWP2, a protein specific for the highly pathogenic *Yersinia. Mol. Microbiol.* **6**, 379-388.
22. de Almeida, A.M.P., Guiyoule, A., Guilvout, I., Iteman, I., Baranton, G. & Carniel, E. (1993). Chromosomal *irp2* gene in *Yersinia*: distribution, expression, deletion and impact on virulence. *Microb. Pathog.* **14**, 9-21.
23. Fetherston, J.D., Lillard, J.W., Jr. & Perry, R.D. (1995). Analysis of the pesticin receptor from *Yersinia pestis*: role in iron-deficient growth and possible regulation by its siderophore. *J. Bacteriol.* **177**, 1824-1833.
24. Rakin, A., Saken, E., Harmsen, D. & Heesemann, J. (1994). The pesticin receptor for *Yersinia enterocolitica*: a novel virulence factor with dual function. *Mol. Microbiol.* **13**, 253-263.
25. Schubert, S., Rakin, A., Karch, H., Carniel, E. & Heesemann, J. (1998). Prevalence of the 'high-pathogenicity island' of *Yersinia* species among *Escherichia coli* strains that are pathogenic to humans. *Infect. Immunol.* **66**, 480-485.
26. Gehring, A.M., Mori, I., Perry, R.D. & Walsh, C.T. (1998). The nonribosomal peptide synthetase HMWP2 forms a thiazoline ring during biogenesis of yersiniabactin, an iron-chelating virulence factor of *Yersinia pestis. Biochemistry* **37**, 11637-11650.
27. Guilvout, I., Mercereau-Puijalon, O., Bonnefoy, S., Pugslay, A.P. & Carniel, E. (1993). High-molecular-weight protein 2 of *Yersinia enterocolitica* is homologous to AngR of *Vibrio anguillarum* and belongs to a family of proteins involved in nonribosomal peptide synthesis. *J. Bacteriol.* **175**, 5488-5504.

28. Pelludat, C., Rakin, A., Jacobi, C.A., Schubert, S. & Heesemann, J. (1998). The yersiniabactin biosynthetic gene cluster of *Yersinia enterocolitica*: organization and siderophore-dependent regulation. *J. Bacteriol.* **180**, 538-546.
29. Fetherston, J.D. & Perry, R.D. (1994). The pigmentation locus of *Yersinia pestis* KIM6+ is flanked by an insertion sequence and includes the structural genes for pesticin sensitivity and HMWP2. *Mol. Microbiol.* **13**, 697-708.
30. Fetherston, J.D., Schuetze, P. & Perry, R.D. (1992). Loss of the pigmentation phenotype of *Yersinia pestis* is due to the spontaneous deletion of 102 kb of chromosomal DNA which is flanked by a repetitive element. *Mol. Microbiol.* **6**, 2693-2704.
31. Staggs, T.M., Fetherston, J.D. & Perry, R.D. (1994). Pleiotropic effects of a *Yersinia pestis fur* mutation. *J. Bacteriol.* **176**, 7614-7624.
32. Farrell, D.H., Mikesell, P., Actis, L.A. & Crosa, J.H. (1990). A regulatory gene, *angR*, of the iron uptake system of *Vibrio anguillarum*: similarity with phage P22 *cro* and regulation by iron. *Gene* **86**, 45-51.
33. Gaisser, S. & Hughes, C. (1997). A locus coding for putative non-ribosomal peptide/polyketide synthase functions is mutated in a swarming-defective *Proteus mirabilis* strain. *Mol. Gen. Genet.* **253**, 415-427.
34. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410.
35. Heesemann, J., *et al.,* & Berner, R. (1993). Virulence of *Yersinia enterocolitica* is closely associated with siderophore production, expression of an iron-repressible outer membrane polypeptide of 65 000 Da and pesticin sensitivity. *Mol. Microbiol.* **8**, 397-348.
36. Serino, L., Reimmann, C., Baur, H., Beyeler, M., Visca, P. & Haas, D. (1995). Structural genes for salicylate biosynthesis from chorismate in *Pseudomonas aeruginosa. Mol. Gen. Genet.* **249**, 217-228.
37. Serino, L., Reimmann, C., Visca, P., Beyeler, M., Chiesa, V.D. & Haas, D. (1997). Biosynthesis of pyochelin and dihydroaeruginoic acid requires the iron-regulated *pchDCBA* operon in *Pseudomonas aeruginosa. J. Bacteriol.* **179**, 248-257.
38. Marahiel, M.A., Stachelhaus, T. & Mootz, H.D. (1997). Peptide synthetases involved in nonribosomal peptide synthesis. *Chem. Rev.* **97**, 2651-2673.
39. Lambalot, R.H., *et al.,* & Walsh, C.T. (1996). A new enzyme superfamily – the phosphopantetheinyl transferases. *Chem. Biol.* **3**, 923-936.
40. Konz, D., Klens, A., Schörgendorfer, K. & Marahiel, M.A. (1997). The bacitracin biosynthesis operon of *Bacillus licheniformis* ATCC10716: molecular characterization of three multi-modular peptide synthetases. *Chem. Biol.* **4**, 927-937.
41. Hopwood, D.A. (1997). Genetic contributions to understanding polyketides synthases. *Chem. Rev.* **97**, 2465-2497.
42. Kagan, R.M. & Clarke, S. (1994). Widespread occurrence of three sequence motifs in diverse S-adenosylmethionine-dependent methyltransferases suggests a common structure of these enzymes. *Arch. Biochem. Biophys.* **310**, 417-427.
43. Kakavas, S.J., Katz, L. & Stassi, D. (1997). Identification and characterization of the niddamycin polyketide synthase genes from *Streptomyces caelestis. J. Bacteriol.* **179**, 7515-7522.
44. Kunze, B., *et al.,* & Reichenbach, H. (1993). Thiangazole, a new thiazoline antibiotic from *Polyangium* sp. (myxobacteria): production, antimicrobial activity and mechanism of action. *J. Antibiot.* **46**:1752-1755.
45. Tartof, K.D. & C.A. Hobbs, C.A. (1987). Improved media for growing plasmid and cosmid clones. *Focus* **9**, 12.
46. Birnboim, H.C. & Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7,** 1513-1523.
47. Humphreys, G.O. Willshaw, G.A., & Anderson, S.E. (1975). A simple method for the preparation of large quantities of pure plasmid DNA. *Biochim. Biophys. Acta* **383**, 457-463.
48. Hohn, B. & Collins, J. (1980). A small cosmid for efficient cloning of large DNA fragments. *Gene* **11**, 291-298.
49. Staggs, T.M. & Perry, R.D. (1991). Identification and cloning of a *fur* regulatory gene in *Yersinia pestis. J. Bacteriol.* **173**, 417-425.
50. Mahillon, J., *et al.,* & Blattner, R.F. (1998). Subdivision of *Escherichia coli* K-12 genome for sequencing: manipulation and DNA sequence of transposable elements introducing unique restriction sites. *Gene*, in press.
51. Burland, V., Daniels, D.L., Plunkett, G., III & Blattner, F.R. (1993). Genome sequencing on both strands: the Janus strategy. *Nucleic Acids Res.* **21**, 3385-3390.
52. Olson, C.H., Blattner, F.R. & Daniels, D.L. (1991). Simultaneous preparation of up to 768 single-stranded DNAs for use as templates in DNA sequencing. *Methods* **3**, 27-32.
53. Thompson, J.D., Higgins, D.G., & Gibson, T.J. (1994). CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673-4680.

Because *Chemistry & Biology* operates a 'Continuous Publication System' for Research Papers, this paper has been published via the internet before being printed. The paper can be accessed from http://biomednet.com/cbiology/cmb – for further information, see the explanation on the contents pages.