A Widely Distributed Bacterial Pathway for Siderophore Biosynthesis Independent of Nonribosomal Peptide Synthetases

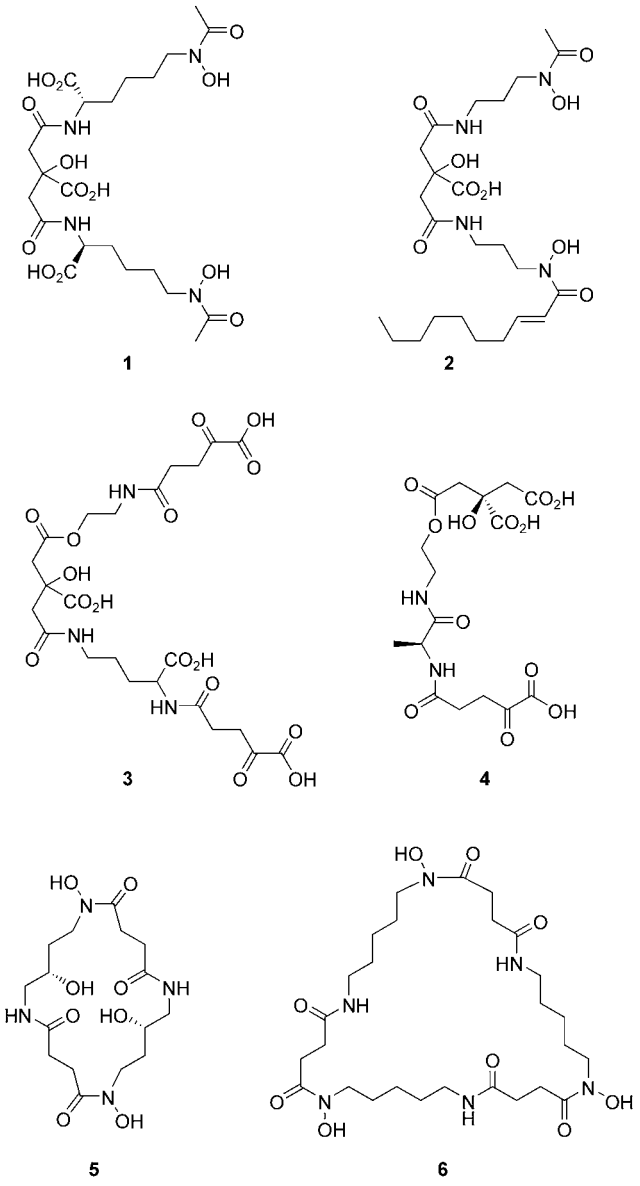
Gregory L. Challis*[a]

1. Introduction

Iron is an essential nutrient for virtually all microorganisms because it is a cofactor for several electron-transport proteins involved in vital life processes like aerobic and anaerobic ATP biosynthesis. However, the bioavailability of iron, which exists predominantly in its ferric form in aerobic environments, such as soil, is very low despite the fact that iron is the fourth most abundant element in the Earth's crust. This is because, at neutral and alkaline pH, ferric iron forms insoluble, polymeric oxyhydroxide complexes that cannot be assimilated by microorganisms. Consequently, iron acquisition from the environment poses a significant challenge to saprophytic microorganisms. Similar bioavailability problems exist in the intercellular matrices of higher eucaryotes, where ferric iron is tightly bound to solubilising transport and storage glycoproteins, such as transferrin and lactoferrin. Thus, iron assimilation by invading pathogens, which is considered essential for establishing infection, also poses a significant challenge.[1]

A common strategy used by many pathogenic and saprophytic microorganisms to tackle the problem of low iron bioavailability is the biosynthesis and excretion of high-affinity iron chelators known as siderophores.[1,2] Once an excreted siderophore has scavenged ferric iron from the environment or host, the resulting iron–siderophore complex is readsorbed by bacterial cells by a membrane-associated ATP-dependent transport system that often exhibits high substrate selectivity.[3] In fungi, the readsorption of iron–siderophore complexes is mediated by the siderophore iron transport (SIT) family of the major facilitator protein superfamily. Several different mechanisms have been proposed for the recovery of ferric iron from the siderophore complex and reduction to the ferrous form for storage and utilisation.[4,5]

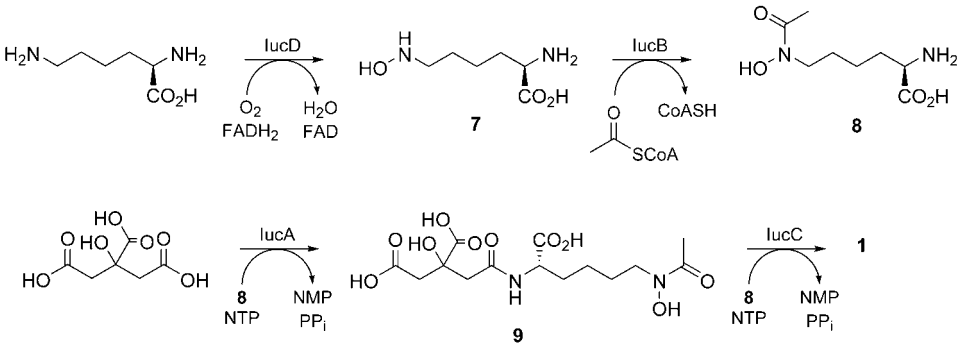
Many siderophores are polypeptides that are biosynthesised by members of the nonribosomal peptide synthetase (NRPS) multienzyme family, which is also responsible for the biosynthesis of the majority of microbial peptide antibiotics. The enzymology of NRPS-catalysed siderophore biosynthesis has been intensively studied over the last decade, and the biosynthetic mechanisms for several types of structurally diverse peptide siderophore are now well understood.[6] On the other hand, several bacterial siderophores are not polypeptides, but are assembled instead from alternating dicarboxylic acid and diamine or amino alcohol building blocks (which are nevertheless derived from amino acids) linked by amide or ester bonds. Examples include aerobactin (**1**),[7] rhizobactin 1021 (**2**),[8] achromobactin (**3**),[9] vibrioferrin (**4**),[10,11] alcaligin (**5**),[12–14] and desferrioxamine E (**6**).[15] Pioneering biochemical genetic studies in the 1980s by Neilands and co-workers established that aero-



[a] *Dr. G. L. Challis*
Department of Chemistry, Warwick University
Coventry CV4 7AL (UK)
Fax: (+44) 24-7652-4112
E-mail: g.l.challis@warwick.ac.uk

 DOI: 10.1002/cbic.200400283

bactin is biosynthesised in *Escherichia coli* via an NRPS-independent pathway that utilises two siderophore synthetase enzymes, IucA and IucC, to catalyse the formation of the key amide bonds that link the dicarboxylic acid and diamine units containing the iron-chelating functional groups together.[16,17] For nearly a decade, the aerobactin system remained as an isolated example of an NRPS-independent pathway for siderophore biosynthesis. However, over the last eight years, genetic studies have established that seven further siderophores (rhizobactin 1021, alcaligin, desferrioxamines, vibrioferrin, staphylobactin, anthrachelin and achromobactin) are biosynthesised by NRPS-independent pathways, all utilising at least one enzyme with significant sequence similarity to the aerobactin siderophore synthetases.[18–28] Several of these pathways are found in pathogens, and the siderophore products of most of these have been shown to be either required for growth in the host or important in pathogenesis. This under-investigated NRPS-independent siderophore (NIS) biosynthesis pathway, which appears to be present in over 40 species of bacteria, is the subject of this article.



Scheme 1. *Pathway for aerobactin biosynthesis.*[16,17]

2. Genetic Studies of NRPS-Independent Siderophore Biosynthesis

2.1. The aerobactin pathway

Aerobactin (**1**) mediates iron acquisition in many virulent *E. coli* strains and is known to be important for pathogenesis.[29] Genes encoding the aerobactin biosynthesis enzymes and an outer-membrane receptor protein involved in the uptake of the ferric aerobactin complex can be either plasmid- or chromosome-borne, but they were first identified, cloned and characterised from the large plasmid pColV-K30.[29,30] The aerobactin cluster consists of five genes. Four (*iucABCD*) direct aerobactin biosynthesis, while the fifth (*iutA*) encodes the receptor protein (Figure 1).[16] The role of each of the IucABCD enzymes in aerobactin biosynthesis was established by Neilands and coworkers using a plasmid containing the *iucABCD* and *iutA* genes in a series of gene deletion and subcloning experiments.[16] Thus, expression of *iucD* in *E. coli* resulted in accumulation of *N*-6-hydroxylysine (**7**); this suggested that IucD catalyses N6-hydroxylation of lysine.[16] Subsequent in vitro experiments on a purified recombinant derivative of IucD with a modified N terminus confirmed that it is an FAD-dependent L-lysine N6-hydroxylase (Scheme 1).[31] Extracts of *E. coli* expressing only *iucA* and *iucB* exhibited *N*-6-hydroxylysine *N*-6-acetyltransferase activity on addition of **7** and acetyl CoA, whereas extracts from *E. coli* expressing only *iucA* did not.[16] Thus, IucB is implicated in the N6-acetylation of **7** with acetyl CoA to form *N*6-acetyl-*N*6-hydroxylysine (**8;** Scheme 1). This was confirmed by purification and biochemical characterisation of IucB.[32] The hydroxamic acid **9**, but not aerobactin (**1**), accumulated in *E. coli* transformed with the plasmid containing *iucABC-DiutA*, but with a 0.3 kb deletion in *iucC*; this indicated that IucA catalyses condensation of **8** with one of the prochiral carboxyl groups of citric acid to form **9** and that IucC condenses **8** with **9** to form aerobactin (Scheme 1).[16] The roles of IucA and IucC were confirmed by further deletion and complementation experiments in conjunction with analyses for the accumulation of **8** and **9**.[17] Subsequently, the *iucABCD* gene cluster was sequenced, thus confirming that *iucD* codes for a flavin-dependent monooxygenase and *iucB* codes for an acetyl-CoA-dependent acyltransferase.[33] Surprisingly, given that IucA and IucC are implicated as amide synthetases, the deduced sequences of these enzymes do not contain any of the known highly conserved sequence motifs for nucleotide triphosphate (NTP) binding.[34] Nevertheless, the condensation of **8** with both citric acid and **9** would require some means of activating the carboxyl groups participating in the reactions to drive the equilibrium towards the products. Such activation reactions usually require an NTP, most often ATP, but the requirement for NTPs as cosubstrates remains to be established for IucA and IucC. In addition to *E. coli*, the aerobactin pathway has been detected in several other bacteria including *Aerobacter aerogenes*,[35] *Salmonella arizona*,[36] *Vibrio mimcus*,[37] *Erwinia carotovora* subsp. *carotovora*,[38] *Enterobacter cloacae*,[39] *Escherichia fergusonii*,[40] *Shigella boydii*,[41] *Shigella flexneri*[42] and *Shigella sonnei*.[43] In many of these organisms the aerobactin pathway enhances virulence.

The pioneering studies of Neilands and colleagues led to a model for aerobactin biosynthesis (Scheme 1)[16] that serves as a rational basis for suggesting plausible biosynthetic pathways to other siderophores assembled by the NRPS-independent pathway, in which the biosynthetic genes have been identified and sequenced.

Although the aerobactin pathway remained the isolated example of an NIS biosynthetic pathway for nearly a decade, seven further gene clusters that direct NRPS-independent biosynthesis of siderophores have been identified and sequenced in the last eight years. Four of these seven clusters have been identified in the last year; this suggests that interest in this area is rapidly expanding, partly as a result of the recent dramatic increase in microbial sequence data made available

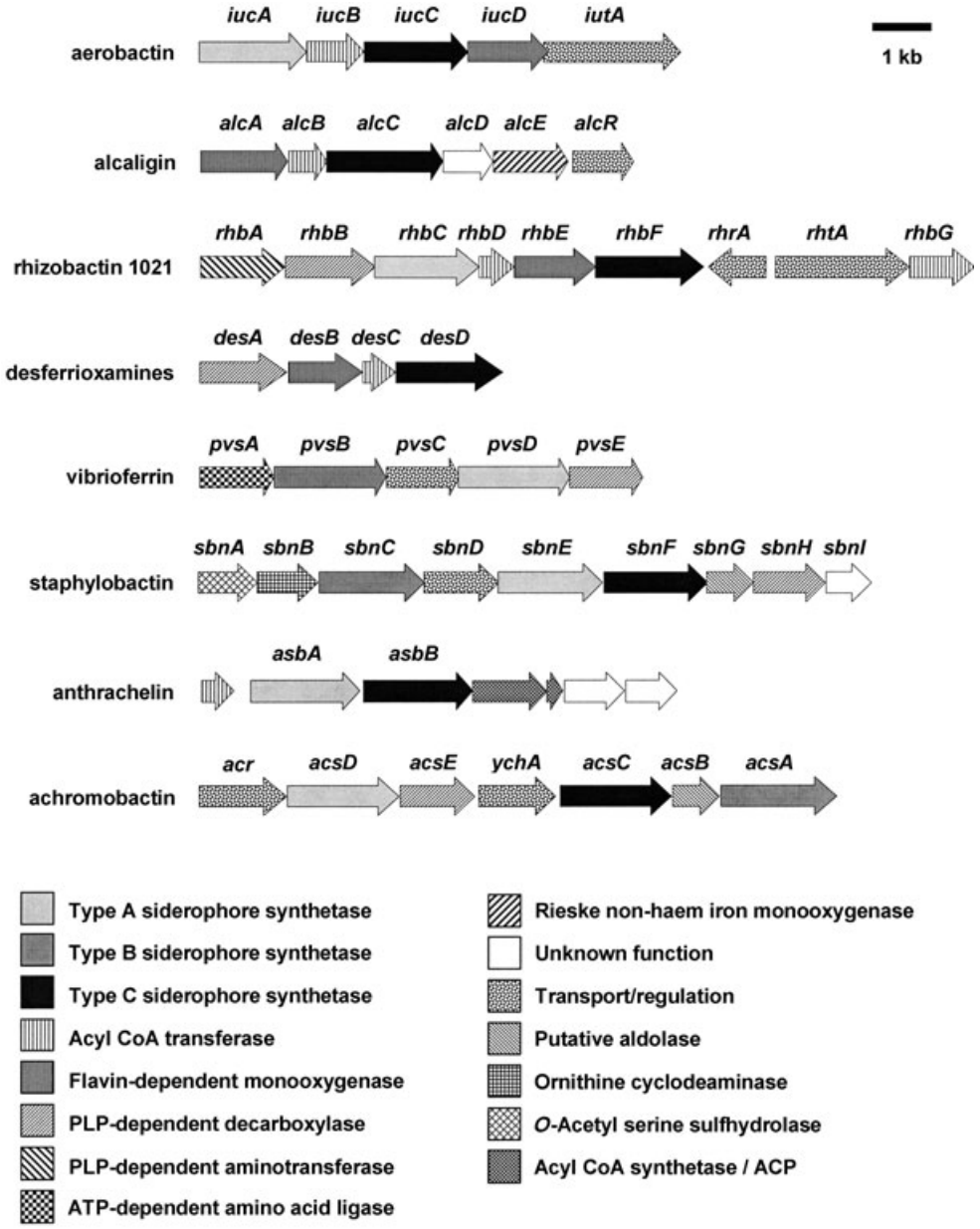


Figure 1. *Gene clusters for NIS biosynthesis. The name of the siderophore whose production is directed by each cluster is listed. Each cluster contains at least one member of the siderophore synthetase superfamily shaded light grey, dark grey or black depending on whether it is a type A, B or C enzyme.*

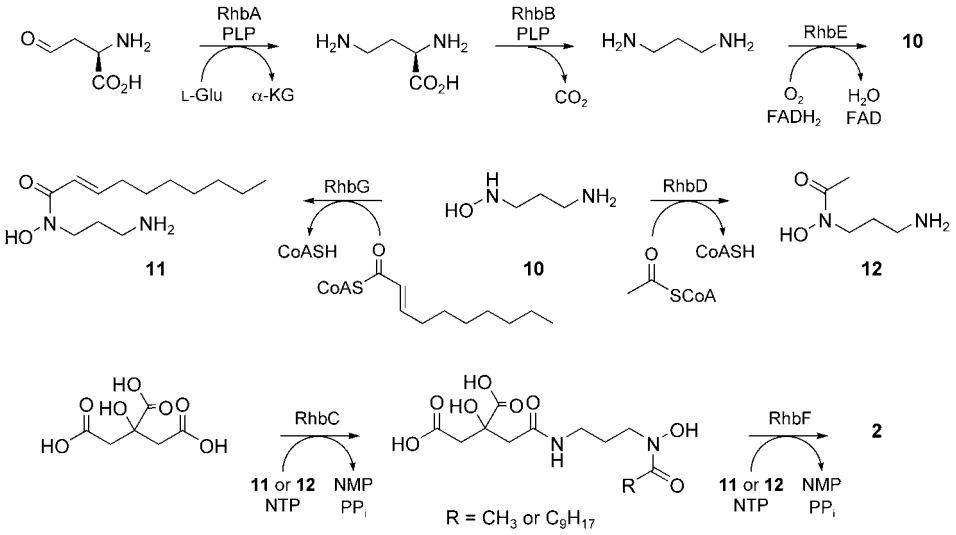
through genome projects. The following sections summarise the published work to date on each of these pathways.

2.2. The alcaligin pathway

The second NIS biosynthetic pathway to be identified and sequenced was that for alcaligin (**5**) in *Bordetella pertussis* and *Bordetella bronchiseptica*—the mammalian pathogens that cause whooping cough in humans, kennel cough in dogs and atrophic rhinitis in swine. Alcaligin was first isolated from the marine bacterium *Alcaligenes dentrificans* subsp. *xylosoxydans* and was subsequently identified as the siderophore of *B. pertussis* and *B. bronchiseptic*a.[12–14] The structure of **5** has been confirmed by total synthesis.[44] Reduced virulence in swine has been reported for *B. bronchiseptica* mutants unable to produce alcaligin.[45]

In 1995, Dyer and colleagues identified and sequenced a *B. bronchiseptica* gene (*alcA*) coding for an enzyme similar to flavin-dependent amino acid N-hydroxylases (e.g. IucD) involved in siderophore biosynthesis in other bacteria and required for alcaligin biosynthesis.[18] The following year, Armstrong and co-workers reported the identification and sequencing of three genes *alcABC* required for alcaligin biosynthesis in *B. pertussis*.[20] In the same year, this group reported that the *odc* gene of *B. pertussis* and *B. bronchiseptica*, which codes for a pyridoxal phosphate (PLP) dependent ornithine decarboxylase and is not clustered with *alcABC*, is also required for alcaligin biosynthesis.[19] Feeding of *odc*-knockout mutants

with 1,4-diaminobutane restored alcaligin biosynthesis; this suggests that 1,4-diaminobutane is a precursor of **5**. The sequences of the *alcA* genes from *B. bronchiseptica* and *B. pertussis* are virtually identical. Sequence comparisons with database proteins showed that AlcB and AlcC are similar to the IucB and IucC enzymes of aerobactin biosynthesis, respectively. This prompted Armstrong and colleagues to outline plausible early steps for alcaligin biosynthesis, but a detailed proposal for each step in alcaligin biosynthesis has not yet appeared in the literature (see Section 4.1 and Scheme 5, below).[20] Subsequently, Dyer and colleagues reported the identification and sequencing of the *alcABC* operon from *B. bronchisepti*ca.[21] Transcriptional analyses of the *alcABC* genes showed they were cotranscribed and that the transcript extended about 3.6 kb beyond the 3′-end of *alcC*; this suggested that the *alc* cluster might contain additional genes required for alcaligin biosynthesis.[22] This suggestion was confirmed by Locht and colleagues who showed that there are three further genes *alcD*, *alcE* and *alcR* downstream of *alcC* in *B. bronchiseptica* (Figure 1).[23] The *alcR* gene encodes a transcriptional regulator required for alcaligin biosynthesis and transport.[22,23] The question of whether AlcD and AlcE are involved in alcaligin biosynthesis remains open, and no suggestions for the possible role they might play have been put forward. Very recently, the presence of the *alcABCDER* operon in *B. pertussis* and *B. parapertussis* has been confirmed by whole-genome sequencing of these organisms.[46]



Scheme 2. *Modified proposal for the biosynthetic pathway to rhizobactin 1021 (**2**), based on ref. [24].*

2.3. The rhizobactin 1021 pathway

The structure of rhizobactin 1021 (**2**), a siderophore produced by the alfalfa symbiont *Synorhizobium meliloti*, was reported in 1993.[8] Biosynthetic genes for **2** were first identified in 1989, but the sequence of a region of the pSymA megaplasmid of *S. meliloti* directing regulation, biosynthesis and transport of **2** was reported only recently by O'Connell and colleagues.[24] The region contains an operon of six genes (*rhbABCDEF*) assigned a role in the biosynthesis of **2**, one gene (*rhrA*) assigned a role in the regulation of rhizobactin 1021 biosynthesis and one gene (*rhtA*) assigned a role in the transport of **2**, on the basis of gene disruption experiments and sequence comparisons (Figure 1). Disruption of the *rhbG* gene directly downstream of *rhtA* did not abolish siderophore production as measured by chrome azurol S (CAS) and siderophore-uptake assays.[24] However, since the authors did not determine whether **2** or a structurally related siderophore (e.g. in which the enoyl lipid moiety in **2** is replaced with an acetyl group) was being produced by the mutant, *rhbG* cannot be ruled out as a rhizobactin 1021 biosynthetic gene. The similarity of RhbC, RhbD, RhbF and RhbE, to IucA, IucB, IucC and IucD, respectively, the structural similarity between **1** and **2**, and the similarity of RhbA and RhbB to PLP-dependent enzymes, known to be involved in the production of 1,3-diaminopropane, led the authors to propose a largely plausible pathway for rhizobactin 1021 biosynthesis (Scheme 2).[24] The question of how the enoyl–lipid moiety is incorporated into **2** was left open. However, RhbG shows significant similarity to IucB and other acyl-CoA-dependent acyl transferases, and therefore it seems likely that this could catalyse the acylation of *N*-hydroxydiaminobutyrate (**10**)s with 2-decenoyl CoA (Scheme 2). Consequently, the possible role of RhbG in the biosynthesis of **2** warrants further investigation.

2.4. The desferrioxamine pathway

Desferrioxamines are a group of related siderophores composed of alternating *N*-hydroxycadaverine or *N*-hydroxyputrescine units and succinate units (e.g. desferrioxamine E (**6**)) that are characteristic siderophores of *Streptomyces* spp.[15] Genes that direct the biosynthesis of desferrioxamine B in *Streptomyces pilosus* were identified by Schupp and colleagues in 1987,[47,48] although the sequence of these genes was not reported. Subsequently, Expert and colleagues reported that a gene coding for a protein with similarity to *alcA* is required for desferrioxamine biosynthesis in *Erwinia amylovora*, although the sequence of this gene was apparently also not deposited in the databases.[49] In addition, they showed that desferrioxamine production plays a role in the plant pathogenicity of this bacterium. In 2002, Challis and co-workers identified a cluster of four genes *desABCD* in the genome sequence of *Streptomyces coelicolor* M145 that was proposed to direct the biosynthesis of desferrioxamines.[50] A very similar cluster was also identified in the genome sequence of *Streptomyces avermitilis*.[51]

Very recently, the Challis group reported that inactivation of *desD* abrogates production of **6**, the major desferrioxamine produced by *S. coelicolor*, along with three other yet to be unambiguously identified *tris*-hydroxamate siderophores, presumed also to belong to the desferrioxamine family.[25] Sequence comparisons identified that DesA is similar to several PLP-dependent decarboxylases, DesB is similar to flavin dependent N-oxygenases, DesC is similar to acyl CoA-dependent acyl transferases and DesD is similar to AlcC and IucC. This led to the proposal of a possible pathway for desferrioxamine biosynthesis from L-lysine and succinyl CoA that made use of these four enzymes (Scheme 3).[25] This pathway is consistent with early steps for desferrioxamine B biosynthesis previously proposed by Schupp and co-workers.[47–48]

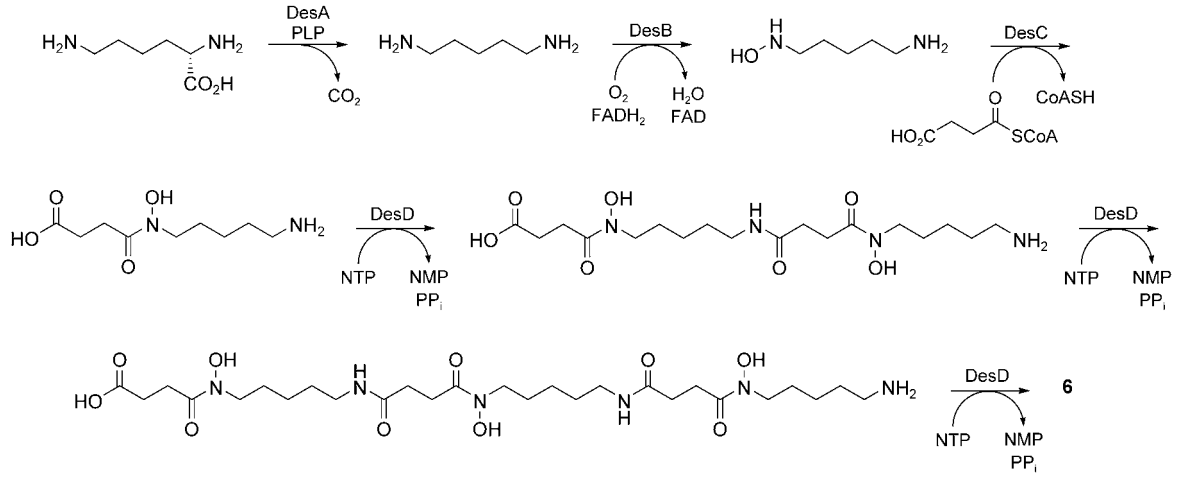
2.5. The vibrioferrin pathway

Vibrio parahaemolyticus is a halophilic Gram-negative marine/estuarine bacterial pathogen that causes seafood-related gastroenteritis throughout Asia. The isolation of the unusual siderophore vibrioferrin (**4**) from *V. parahaemolyticus* was reported in 1992.[11,12] The structure and absolute stereochemistry of vibrioferrin were subsequently confirmed by chemical synthesis.[52] In December 2003, Yamamoto and co-workers reported that an operon of five genes *pvsABCDE* directs vibrioferrin biosynthesis and transport in *V. parahaemolyticus*, on the basis of gene-knockout experiments.[26] Thus, disruption of *pvsA*, *pvsD* and *pvsE* all resulted in abrogation of vibrioferrin biosynthesis, as judged by the CAS plate assay.[26] PvsA returns a hit to PFAM01820, a family of proteins exemplified by the D-Ala-D-Ala ligase enzyme of bacterial cell-wall biosynthesis, in a conserved-domain search,[53] whereas PvsB and PvsD both return hits to PFAM04183, exemplified by IucA and IucC, in a similar search. PvsC is similar to membrane-spanning transport proteins; this suggests that it might play a role in vibrioferrin transport rather than biosynthesis. PvsE is similar to several PLP-dependent decarboxylases involved in siderophore biosynthesis. Although Yamamoto and colleagues did not propose a detailed pathway for vibrioferrin biosynthesis on the basis of their work,[26] a plausible pathway, consistent with the results of the above sequence similarities, can be hypothesised (see Section 4.2 and Scheme 6, below).

2.6. Pathways to staphylobactin and anthrachelin

In 2004, two gene clusters that direct the biosynthesis of siderophores of as yet undetermined structure were identified in the genome sequences of the notorious pathogens *Staphylococcus aureus* and *Bacillus anthracis*.[27,28] The *S. aureus* cluster contains an operon of nine genes (*snbABCDEFGHI*) that, on the basis of insertional inactivation of *snbE* (which abolishes siderophore production) and sequence comparisons, has been proposed to code for enzymes involved in the biosynthesis of the unknown siderophore dubbed "staphylobactin" (Figure 1).[28] In addition, the production of staphylobactin was shown to enhance the virulence of *S. aureus* in a mouse model.[28] The *B. anthracis* cluster contains six genes in an apparent operon, one of which (*asbA*) abolishes production of the unknown siderophore dubbed "anthrachelin".[27] This cluster might also encompass the gene upstream of *absA*, because this gene shows sequence similarity to other acyl transferases involved in siderophore biosynthesis. The *absA* mutant exhibited severely attenuated growth in macrophages and attenuated virulence in mice.[27] In contrast, inactivation of a second independent NRPS pathway in *B. anthracis*, which is believed to direct biosynthesis of siderophore similar to bacillibactin (the catecholate siderophore of *Bacillus subtilis*), did not affect virulence or growth in macrophages.[27] Blast searches reveal that the *abs* cluster is also present in *Bacillus cereus* and *Bacillus thuringiensis* serovar *konkukian*.[54]

Both the *sbn* and the *abs* clusters contain genes (*sbnC*, *sbnE*, *sbnF* and *absA*, *absB*) that code for proteins showing similarity to IucA and IucC. It is therefore suggested that staphylobactin and anthrachelin are likely to be siderophores with similar structural features and iron-chelating functional groups to compounds **1**–**6**.



Scheme 3. *Proposed pathway to desferrioxamine E (**6**).*[25]

2.7. The achromobactin pathway

Achromobactin (**3**) is a siderophore that was recently isolated from *Erwinia* (now *Pectinobacterium*) *chrysanthemi*.[9] Although no reports regarding genes that direct the biosynthesis of **3** have yet appeared in the literature, the sequence of a cluster of genes involved in achromobactin biosynthesis was recently deposited in the Genbank database (accession numbers AF416739 and AF416740, Figure 1). This cluster shows striking similarity to the *sbn* cluster that directs staphylobactin biosynthesis in *S. aureus*.[28] Like the other seven clusters that direct siderophore biosynthesis through the NRPS-independent pathway, the achromobactin cluster contains genes that encode proteins with sequence similarity to IucA and IucC. Sequence analysis of the enzymes encoded by the achromobactin cluster provides a rational basis for proposing plausible pathways for achromobactin biosynthesis (see Section 4.3). It should be noted, however, that the achromobactin cluster might encompass more than the seven genes reported in the two Genbank files, which could alter these proposals.

3. Sequence Analysis of the Siderophore Synthetase Superfamily

All eight of the NIS biosynthetic pathways characterised by genetic studies to date contain at least one and often two or three genes coding for an enzyme with similarity to IucA and IucC from the aerobactin pathway, which is the only pathway where the roles of these enzymes have been examined experimentally.[16,17] By using IucA and IucC as probes, a BLAST search retrieves over 80 sequences from more than 40 bacterial species with significant (37–77%) sequence similarity.[54] On the basis of the experimentally determined functions of IucA and IucC and likely roles deduced for orthologous enzymes from the other NIS-biosynthesis clusters discussed above, these proteins are proposed to form a new siderophore synthetase superfamily that defines the NIS pathway.

A multiple sequence alignment of 88 members of this siderophore synthetase superfamily reveals several new insights into the probable functions and catalytic mechanisms of theses enzymes. Thus, phylogenetic analysis shows that the superfamily splits into three types termed A, B and C (Figure 2). IucA is representative of the type A group, which includes RhbC, PvsD, AcsD, SbnE and AsbA, whereas IucC is representative of the type C group, which includes RhbF, AcsC, SbnF, AsbB, AlcC and DesD. The type B group does not currently contain any proteins of experimentally investigated function, but is nevertheless represented by PvsB, AcsA and SbnC. By considering the known roles for IucA and IucC in aerobactin biosynthesis in conjunction with the relationship between the structures of the products (where known) for the seven other genetically characterised pathways, two plausible models for the role each type of siderophore synthetase enzyme plays in siderophore biosynthesis can be suggested. In the first model (Scheme 4, top), each type of siderophore synthetase is specific for a particular type of carboxylic acid sub-

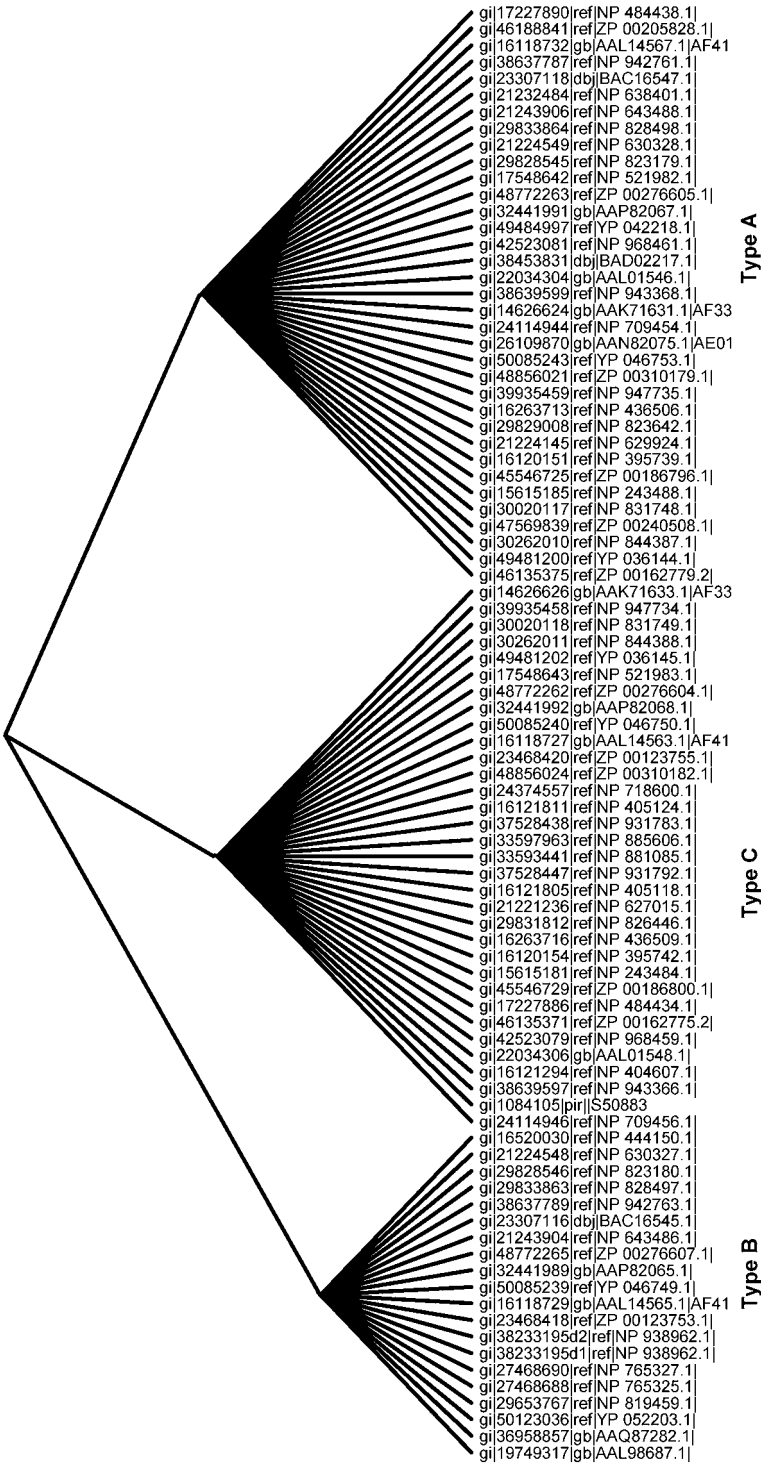
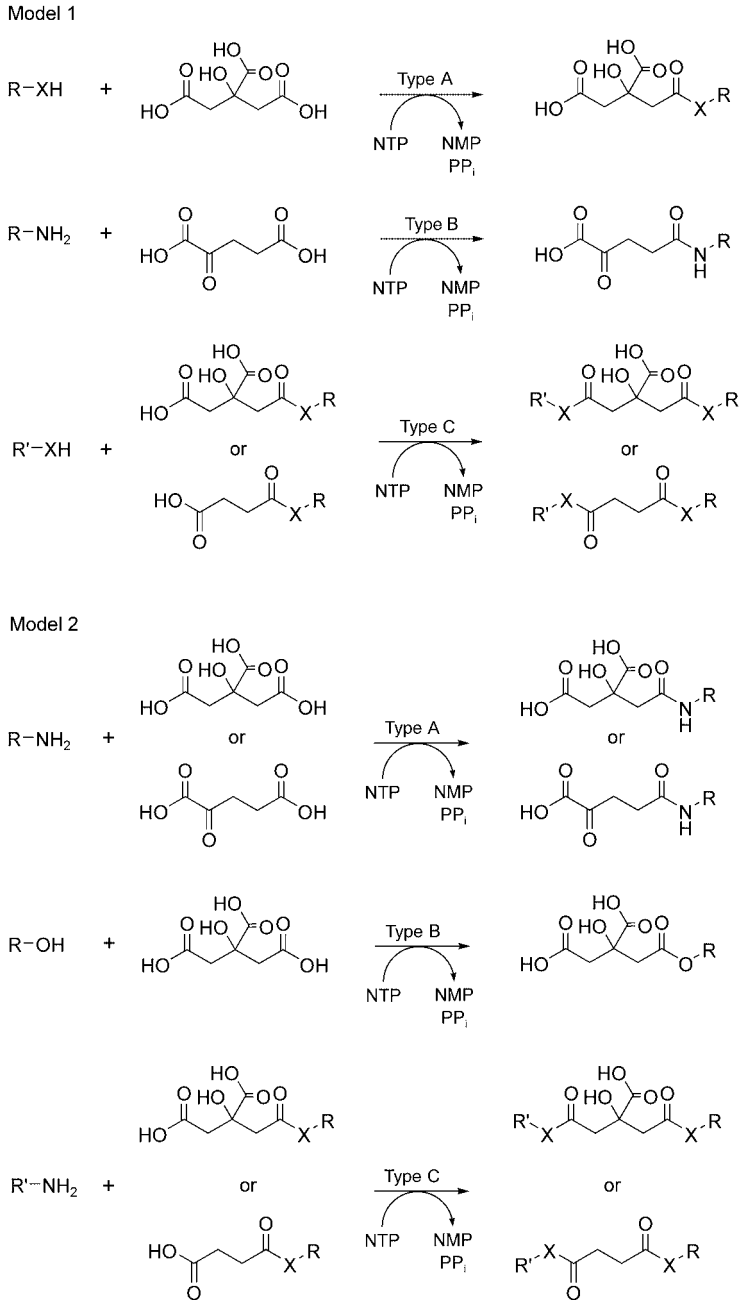


Figure 2. *Cladogram illustrating how the siderophore synthetase enzyme family splits into three subfamilies based on a multiple sequence alignment of 88 IucA and IucC homologues in the NCBI database.*



Scheme 4. *Two models for the reactions catalysed by type A, B and C siderophore synthetases (X = NH_2 or OH).*

strate and can utilise either amines or alcohols as substrates in amide- and ester-bond-forming reactions, respectively. Thus, the type A enzymes catalyse the formation of an amide or ester bond between an amino or hydroxyl group in a variety of substrates and one of the two prochiral carboxyl groups of citric acid; the type B enzymes catalyse amide (and possibly also ester) bond formation between an amino or hydroxyl group in a variety of substrates and the C5 carboxyl group of α-ketoglutaric acid; and the type C enzymes catalyse the formation of an amide or ester bond between an amine or alcohol substrate and a carboxyl group in monoamide or monoester derivatives of citric or succinic acid. In the second model (Scheme 4, bottom), each type of enzyme exhibits high preference for either alcohol or amine nucleophilic substrates, yet can tolerate different carboxylic acid substrates in the ester and amide synthetase reactions. Thus, the type A enzymes catalyse the formation of an amide bond between an amine substrate and either one of the prochiral carboxyl groups in citrate or the C5 carboxyl group of α-ketoglutarate (α-KG); the type B enzymes catalyse ester-bond formation between an alcohol substrate and one of the prochiral carboxyl groups of citric acid; and the type C enzymes catalyse amide-bond formation between an amino group in diverse substrates and a monoamide or monoester derivative of citric or succinic acid. Both of the models are consistent with the relationship between the structures of the products and the genes that direct their biosynthesis discussed in Section 2. Although further biochemical and/or genetic experiments will be required to discriminate between the two models, at this stage Model 1 is favoured for two main reasons. Firstly, in either model, each of the enzyme types has to tolerate considerable structural variation in the nucleophilic substrate regardless of whether it is an amine or an alcohol. Secondly, the amine nucleophiles are very likely to be protonated at physiological pH. Thus, regardless of whether an amino or hydroxyl group is the nucleophile, a general base would be required to deprotonate it, that is, the same kind of catalytic chemistry would be employed. Thus, it seems logical to conclude that the division of the siderophore synthetase superfamily into three types reflects specificity towards the different types of carboxylic acid intermediates in siderophore biosynthesis.

The second illuminating feature of the siderophore synthetase superfamily revealed by the multiple sequence alignment is several amino acid residues that are conserved in all or virtually all members of the family. Thus, His242 in DesD is universally conserved in all of the 87 other aligned sequences. This might act as a general base required to activate the hydroxyl or protonated amino group of the incoming nucleophile by deprotonation as part of the catalytic cycle of the enzymes. Likewise, Arg280, Lys294, Arg306, Arg373 and Asn445 (DesD numbering) are also universally conserved. Although a conserved Walker A- or B-like motif cannot be detected within these aligned sequences,[34] the reaction catalysed by the siderophore synthetase superfamily would be expected to utilise an NTP to activate the carboxyl group of the carboxylic acid substrates as an acyl adenylate or acyl phosphate intermediate as part of the catalytic mechanism. The universally conserved Lys residue (or perhaps one of the three universally conserved Arg residues) might interact with the negatively charged oligophosphate of an NTP and possibly also the carboxylate group of the carboxylic substrate, as does a universally conserved Lys residue in other NTP-utilising enzymes.[34] Within the conserved motif S/A-X-R-

S/T, the first residue is Ser in all but two of the 88 sequences aligned and the fourth residue is either Ser or Thr in all of the sequences. Similarly, the last residue within the conserved sequence motif K/N/R-X_{14}-P/A-D/E is Asp or Glu in all of the sequences aligned. Together with the universally conserved His, these highly conserved Ser/Thr and Asp/Glu residues might form part of an α,β-hydrolase-like His-Ser-Asp catalytic triad that covalently tethers the carboxylic acid substrate through an ester linkage to Ser or Thr by reaction with an acyl adenylate or acyl phosphate intermediate.[55] Thus, this multiple sequence alignment provides significant insights and suggests possible lines of future enquiry regarding the catalytic mechanism of the siderophore synthetase superfamily.

Proposals for Other NRPS-Independent Siderophore Biosynthetic Pathways

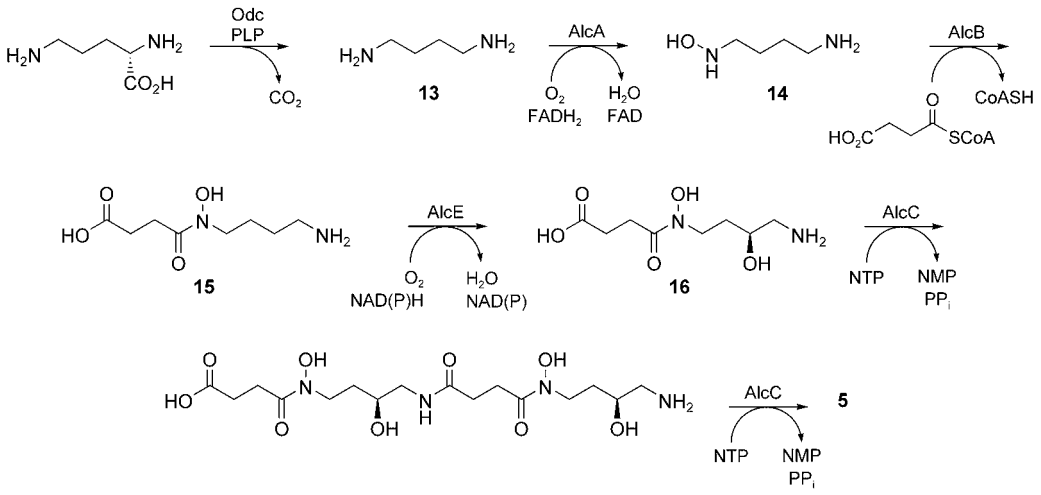
Plausible biosynthetic pathways for aerobactin, rhizobactin 1021 and desferrioxamine E have already been proposed in the primary literature (see Section 2). The sequence analysis presented in Section 3 is completely consistent with these proposals because both the aerobactin and the rhizobactin 1021 pathway utilise one type A and one type C siderophore synthetase,[16,17,24] whereas the desferrioxamine pathway utilises only one type C enzyme.[25] For alcaligin, vibrioferrin and achromobactin, either incomplete or no biosynthetic pathways have thus far been proposed. The analysis presented in Section 3, in combination with other sequence comparisons, allows plausible pathways to each of these siderophores to be suggested. In addition, it provides clues regarding the identity of some of the building blocks incorporated into siderophores of hitherto undetermined structure. Each of these is discussed in turn in the following sections.

4.1. Alcaligin

Armstrong and co-workers have provided compelling evidence that the first step in alcaligin biosynthesis in Bordetellae is decarboxylation of ornithine to yield putrescine (**13**), mediated by the PLP-dependent decarboxylase Odc (Scheme 5).[19] They propose that putrescine is converted by AlcA, which codes for a probable FAD-dependent monooxygenase, into its *N*-hydroxy derivative **14**.[20] This reaction is analogous to that shown to be catalysed by IucD in aerobactin biosynthesis, and it seems highly plausible that this is the second step in alcaligin biosynthesis. Armstrong and co-workers also suggested that **14** undergoes acylation with succinic acid, catalysed by AlcB.[20] This seems unlikely, however, because one of the carboxyl groups in succinic acid would need to be activated by, for example, reaction with ATP or as a thioester derivative in order for this reaction to be thermodynamically and kinetically favoured. It seems more likely that AlcB catalyses N-acylation of the hydroxylamine group in *N*-hydroxyputrescine with succinyl CoA—an activated monothioester derivative of succinic acid that is an intermediate in the Krebs cycle—to give **15** (Scheme 5). This reaction is very similar to the acylation of *N*6-hydroxylysine with acetyl CoA in the aerobactin pathway catalysed by IucB,[16] which shows significant sequence similarity to AlcB. Compound **15** could then undergo C-hydroxylation mediated by AlcE, which shows significant similarity to Rieske dioxygenases such as naphthalene dioxygenase, to yield **16**.[55] In this case AlcE would be acting as a monooxygenase. Finally, AlcC, which belongs to the type C siderophore synthetase family, could catalyse NTP-dependent dimerisation of **16** followed by NTP-dependent macrocyclisation to yield alcaligin (**5**; Scheme 5). The role played by AlcD, if any, in alcaligin biosynthesis remains uncertain. This hypothetical pathway closely resembles that recently proposed for the structurally related desferrioxamine E in *S. coelicolor* (see Section 2.2 and Scheme 3).[25]

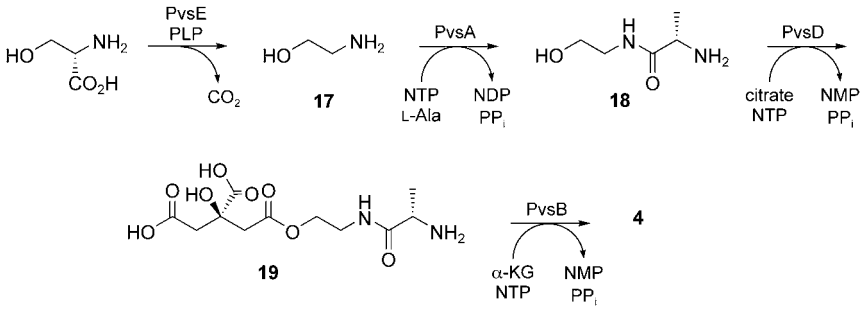
4.2. Vibrioferrin

Although Yamamoto and co-workers did not propose a detailed pathway for vibrioferrin biosynthesis in *Vibrio parahaemolyticus*, a plausible pathway can be suggested on the basis of sequence analysis of the proteins coded for by the



Scheme 5. *Hypothetical pathway for alcaligin biosynthesis, based partly on ref. [20].*

pvsABCDE cluster and the analysis presented in Section 3. Thus, PvsE is similar to PLP-dependent amino acid decarboxylases, and it seems likely that it would catalyse the decarboxylation of serine to yield ethanolamine (**17**; Scheme 6). PvsA shows similarity to ATP dependent amino acid ligases such as D-Ala-D-Ala ligase and therefore could catalyse initial activation of the carboxyl group in L-Ala through reaction with ATP to generate an acyl phosphate intermediate that could undergo nucleophilic attack by the amino group of ethanolamine to give **18** (Scheme 6).[56] PvsD belongs to the type A siderophore synthetases and, according to Model 1 (see Section 3, Scheme 4), could catalyse the formation of an ester bond between the hydroxyl group of **18** and one of the prochiral carboxyl groups of citrate to yield **19** (Scheme 6). Finally, PvsB belongs to the type B siderophore synthetases and, according to the same model, could catalyse the condensation of the amino group in **19** with the C5 carboxyl group of α-ketoglutarate (α-KG) to yield vibrioferrin (Scheme 6). If model 2 proves to be correct, the order and nature of the reactions catalysed by PvsB and PvsD would be different.

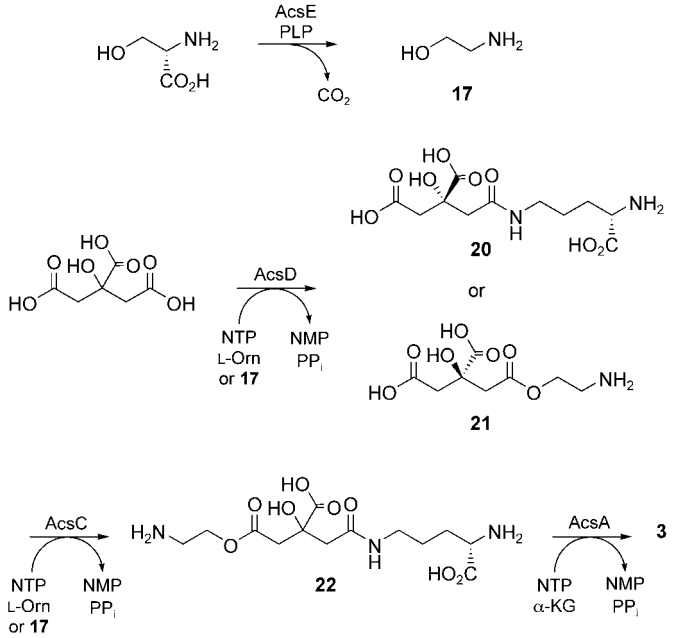


Scheme 6. *Hypothetical pathway for vibrioferrin biosynthesis.*

The vibrioferrin pathway clearly warrants further investigation either through determination of intermediates accumulated in *pvsA*, *pvsD* and *pvsB* mutants or through examination of the reactions catalysed by purified recombinant PvsA, PvsD and PvsB in vitro. In particular, it would be very interesting to learn the reaction catalysed by PvsD, which forms part of the hitherto uncharacterised type B siderophore synthetases.

4.3. Achromobactin

Although no publications have yet appeared regarding the pathway to achromobactin (**3**) in *Pectobacterium chrysanthemi*, a plausible pathway can be proposed on the basis of the data deposited in Genbank (accession numbers AF416739 and AF416740) and the analysis presented above. Thus, AcsE shows similarity to PLP-dependent amino acid decarboxylases and probably catalyses the same reaction as PvsE in vibrioferrin biosynthesis, that is, decarboxylation of serine to yield ethanolamine. AcsD belongs to the type A siderophore synthetases and, according to Model 1, could therefore catalyse the condensation of either the hydroxyl group of ethanolamine (**17**) or the N5-amino group of ornithine with one of the prochiral carboxyl groups of citric acid to form **20** or **21** (Scheme 7). AcsC, which is a type C siderophore synthetase could then catalyse condensation of either ethanolamine or ornithine (depending on which of these was utilised by AcsD) with the other (originally prochiral) carboxyl group of the citrate unit in **20** or **21** to yield **22**. Finally, AcsA, which groups with the type B siderophore synthetases, could catalyse acylation of both of the amino groups in **22** with α-ketoglutarate (α-KG) to form achromobactin (**3**; Scheme 7). A plausible role for AcsB, which shows similarity to aldolases, in achromobactin biosynthesis cannot be proposed on the basis of the preceding analysis. It is important to note that it is impossible to rule out that further genes, beyond those that have been deposited in the Genbank files, could be required for achromobactin biosynthesis. It is also worth bearing in mind that the order of steps could be different from that proposed above for alcaligin, vibrioferrin and achromobactin biosynthesis.



Scheme 7. *Hypothetical pathway for achromobactin biosynthesis.*

4.6. Other siderophores

Even though the structures of staphylobactin and anthrachelin have yet to be determined, some clues regarding possible structural elements of these siderophores can be derived from analysis of the genes within the *sbn* and *asb* clusters.[27,28] Thus, SbnA and SbnB show similarity to a putative *O*-acetyl serine

sulfhydrolase and a putative ornithnine cyclodeaminase, respectively, which have recently been proposed to catalyse the formation of 2,3-diaminopropionic acid from *O*-acetyl serine and ornithine in the biosynthesis of the antibiotic viomycin.[57] SbnH shows similarity to PLP-dependent amino acid decarboxylases and might catalyse either the decarboxylation of 2,3-diaminopropionate to form 1,2-diaminoethane or the decarboxylation of other diamino or aminohydroxy acids to form diamines or hydroxyamines, respectively, that get incorporated into the staphylobactin structure. SbnE, SbnC and SbnF belong to the type A, type B and type C siderophore synthetase families, respectively, and according to Model 1 (see Section 3) might be responsible for incorporating diamide or amidoester derivatives of citric acid and δ-amido or δ-ester derivatives of α-ketoglutarate into the staphylobactin structure. Overall the staphylobactin cluster shows a high degree of similarity to the achromobactin cluster, and it seems reasonable to suggest that the achromobactin and staphylobactin structures might share many common structural elements.

AsbA and AsbB group with the type A and type C siderophore synthetase families, respectively; this suggests that anthrachelin might contain a diamido or amido ester derivative of citric acid and share some structural features with aerobactin and rhizobactin.

5. Future Opportunities and Challenges

The identification of eight distinct but clearly related NRPS-independent pathways for siderophore biosynthesis provides a unique opportunity to study the biochemistry and mechanistic enzymology underpinning the assembly of the siderophores within this structurally diverse family. In particular, understanding the catalytic mechanism employed by the siderophore synthetase superfamily and the mechanisms used by the type A, B and C enzymes within this superfamily to discriminate between their different substrates will be important and challenging future goals. A multidisciplinary approach combining molecular genetics, analytical chemistry, organic synthesis, mechanistic enzymology and structural biology will be required to effectively tackle these challenges. It is already clear that in many pathogens the NRPS-independent pathways to siderophores are required for pathogenicity or enhance virulence. A better understanding of the mechanistic enzymology of NIS biosynthesis, especially those enzymes (such as the siderophore synthetase superfamily) common to all the pathways, might create the opportunity to design, synthesise and test inhibitors as potential antibacterial agents.

6. Conclusion

For many years following the initial discovery and characterisation of the aerobactin pathway, relatively slow progress was made in the discovery of other NRPS-independent pathways for siderophore biosynthesis. As a consequence, the explosion of interest in the mechanistic enzymology of nonribosomal peptide biosynthesis in the early 1990s led to the impression that most microbial siderophores are biosynthesised by NRPS-dependent pathways.[6] The recent increase of interest in siderophore biosynthesis pathways coupled with rapid recent progress in microbial genome sequencing has led to the discovery of a widely distributed NRPS-independent pathway for siderophore biosynthesis. This pathway utilises members of the siderophore synthetase superfamily as a common strategy for assembling siderophores containing α-hydroxy acid and hydroxamic acid chelating ligands for ferric iron. This pathway is distributed over more than forty diverse bacterial species, including plant and animal pathogens, saprophytes and plant symbionts. In contrast to the NRPS-dependent pathway, very little is known about the mechanistic enzymology of the rapidly emerging NRPS-independent pathway. This is likely to become a major focus for future research, and the hypothetical pathways for siderophore biosynthesis put forward in this article, while not intended to be firm predictions, should serve as useful models for developing our understanding of this enzymology. The knowledge gained from such studies might allow inhibitors of the siderophore synthetase superfamily to be developed as potential antibacterial agents for use in medicine and agriculture and is certain to improve our understanding of the mechanisms employed by Nature for the assembly of complex natural products.

Note added in proof

After this manuscript had been accepted for publication, a report on the achromobactin biosynthetic gene cluster appeared. See: T. Franza, B. Mahe. D. Expert, *Mol. Microbiol.* **2005**, *55*, 261–275.

Acknowledgements

Research in the authors laboratory is supported by BBSRC, EPSRC, the Royal Society and the EU. Drs. Francisco Barona-Gómez and Sylvie Lautru are thanked for helpful discussions.

Keywords: biosynthesis · enzymes · genomics · natural products · siderophores

[1] J. R. Telford, K. N. Raymond, *Compr. Supramol. Chem.* **1996**, *1*, 245–266.
[2] H. Drechsel, G. Jung, *J. Pept. Sci.* **1998**, *4*, 147–181.
[3] W. Koster, *Res. Microbiol.* **2001**, *152*, 291–301.
[4] B. F. Matzanke, S. Anemueller, V. Schuenemann, A. X. Trautwein, K. Hantke, *Biochemistry* **2004**, *43*, 1386–1392.
[5] T. J. Brickman, M. A. McIntosh, *J. Biol. Chem.* **1992**, *267*, 12350–12355.
[6] J. H. Crosa, C. T. Walsh, *Microbiol. Mol. Biol. Rev.* **2002**, *66*, 223–249.
[7] F. Gibson, D. I. Magrath, *Biochim. Biophys. Acta* **1969**, *192*, 175–184.
[8] M. Persmark, P. Pittman, J. S. Buyer, B. Schwyn, P. R. Gill, Jr., J. B. Neilands, *J. Am. Chem. Soc.* **1993**, *115*, 3950–3956.
[9] M. Munzinger, H. Budzikiewicz, D. Expert, C. Enard, J. M. Meyer, *Z. Naturforsch. C J. Biosci.* **2000**, *55*, 328–332.
[10] S. Yamamoto, N. Okujo, T. Yoshida, S. Matsuura, S. Shinoda, *J. Biochem.* **1994**, *115*, 868–874.
[11] S. Yamamoto, Y. Fujita, N. Okujo, C. Minami, S. Matsuura, S. Shinoda, *FEMS Microbiol. Lett.* **1992**, *94*, 181–186.
[12] T. Nishio, N. Tanaka, J. Hiratake, Y. Katsube, Y. Ishida, J. Oda, *J. Am. Chem. Soc.* **1988**, *110*, 8733–8734.
[13] C. H. Moore, L.-A. Foster, D. G. Gerbig, Jr., D. W. Dyer, B. W. Gibson, *J. Bacteriol.* **1995**, *177*, 1116–1118.

[14] T. J. Brickman, J.-G. Hansel, M. J. Miller, S. K. Armstrong, *BioMetals* **1996**, *9*, 191 – 203.
[15] M. B. Hossain, D. van der Helm, M. Poling, *Acta Crystallogr. Sect. B* **1983**, *39*, 258 – 263.
[16] V. De Lorenzo, A. Bindereif, B. H. Paw, J. B. Neilands, *J. Bacteriol.* **1986**, *165*, 570 – 578.
[17] V. De Lorenzo, J. B. Neilands, *J. Bacteriol.* **1986**, *167*, 350 – 355.
[18] P. C. Giardina, L.-A. Foster, S. I. Toth, B. A. Roe, D. W. Dyer, *Gene* **1995**, *167*, 133 – 136.
[19] T. J. Brickman, S. K. Armstrong, *J. Bacteriol.* **1996**, *178*, 54 – 60.
[20] H. Y. Kang, T. J. Brickman, F. C. Beaumont, S. K. Armstrong, *J. Bacteriol.* **1996**, *178*, 4877 – 4884.
[21] P. C. Giardina, L.-A. Foster, S. I. Toth, B. A. Roe, D. W. Dyer, *Gene* **1997**, *194*, 19 – 24.
[22] F. C. Beaumont, H. Y. Kang, T. J. Brickman, S. K. Armstrong, *J. Bacteriol.* **1998**, *180*, 862 – 870.
[23] E. Pradel, N. Guiso, C. Locht, *J. Bacteriol.* **1998**, *180*, 871 – 880.
[24] D. Lynch, J. O'Brien, T. Welch, P. Clarke, P. O. Cuiv, J. H. Crosa, M. O'Connell, *J. Bacteriol.* **2001**, *183*, 2576 – 2585.
[25] F. Barona-Gómez, U. Wong, A. Giannakopulos, P. J. Derrick, G. L. Challis, *J. Ac. Chem. Soc.* **2004**, *126*, 16 282 – 16 283..
[26] T. Tanabe, T. Funahashi, H. Nakao, S. Miyoshi, S. Shinoda, S. Yamamoto, *J. Bacteriol.* **2003**, *185*, 6938 – 6949.
[27] S. Cendrowski, W. MacArthur, P. Hanna, *Mol. Microbiol.* **2004**, *51*, 407 – 417.
[28] S. E. Dale, A. Doherty-Kirby, G. Lajoie, D. E. Heinrichs, *Infect. Immun.* **2004**, *72*, 29 – 37.
[29] P. J. Warner, P. H. Williams, A. Bindereif, J. B. Neilands, *Infect. Immun.* **1981**, *33*, 540 – 545.
[30] N. H. Carbonetti, P. H. Williams, *Infect. Immun.* **1984**, *46*, 7 – 12.
[31] A. Thariath, D. Socha, M. A. Valvano, T. Viswanatha, *J. Bacteriol.* **1993**, *175*, 589 – 596.
[32] M. Coy, B. H. Paw, A. Bindereif, J. B. Neilands, *Biochemistry* **1986**, *25*, 2485 – 2489.
[33] J. L. Martinez, M. Herrero, V. de Lorenzo, *J. Mol. Biol.* **1994**, *238*, 288 – 293.
[34] E. V. Koonin, *J. Mol. Biol.* **1993**, *229*, 1165 – 1174.
[35] S. McDougall, J. B. Neilands, *J. Bacteriol.* **1984**, *159*, 300 – 305.
[36] V. L. Waters, J. H. Crosa, *J. Bacteriol.* **1988**, *170*, 5153 – 5160.
[37] Y.-H. Moon, T. Tanabe, T. Funahashi, K. Shiuchi, H. Nakao, S. Yamamoto, *Microbiol. Immunol.* **2004**, *48*, 389 – 398.
[38] C. T. Bull, S. R. Carnegie, J. E. Loper, *Phytopathology* **1996**, *86*, 260 – 266.
[39] J. E. Loper, C. A. Ishimaru, S. R. Carnegie, A. Vanavichit, *Appl. Environ. Microbiol.* **1993**, *59*, 4189 – 4197.
[40] D. Smajs, J. Smarda, G. M. Weinstock, *Folia Microbiol.* **2003**, *48*, 139 – 147.
[41] G. E. Purdy, S. M. Payne, *J. Bacteriol.* **2001**, *183*, 4176 – 4182.
[42] S. A. Vokes, S. A. Reeves, A. G. Torres, S. M. Payne, *Mol. Microbiol.* **1999**, *33*, 63 – 73.
[43] K. M. Lawlor, S. M. Payne, *J. Bacteriol.* **1984**, *160*, 266 – 272.
[44] R. J. Bergeron, J. S. McManis, P. T. Perumal, S. E. Algee, *J. Org. Chem.* **1991**, *56*, 5560 – 5563.
[45] K. B. Register, T. F. Ducey, S. L. Brockmeier, D. W. Dyer, *Infect. Immun.* **2001**, *69*, 2137 – 2143.
[46] J. Parkhill, M. Sebaihia, A. Preston, L. D. Murphy, N. Thomson, D. E. Harris, M. T. G. Holden, C. M. Churcher, S. D. Bentley, K. L. Mungall, A. M. Cerdeño-Tárraga, L. Temple, K. James, B. Harris, M. A. Quail, M. Achtman, R. Atkin, S. Baker, D. Basham, N. Bason, I. Cherevach, T. Chillingworth, M. Collins, A. Cronin, P. Davis, J. Doggett, T. Feltwell, A. Goble, N. Hamlin, H. Hauser, S. Holroyd, K. Jagels, S. Leather, S. Moule, H. Norberczak, S. O'Neil, D. Ormond, C. Price, E. Rabbinowitsch, S. Rutter, M. Sanders, D. Saunders, K. Seeger, S. Sharp, M. Simmonds, J. Skelton, R. Squares, S. Squares, K. Stevens, L. Unwin, S. Whitehead, B. G. Barrell, D. J. Maskell, *Nat. Genet.* **2003**, *35*, 32 – 40.
[47] T. Schupp, U. Waldmeier, M. Divers, *FEMS Microbiol. Lett.* **1987**, *42*, 135 – 139.
[48] T. Schupp, C. Toupet, M. Divers, *Gene* **1988**, *64*, 179 – 188.
[49] A. Dellagi, M. N. Brisset, J. P. Paulin, D. Expert, *Mol. Plant-Microbe Interact.* **1998**, *11*, 734 – 742.
[50] S. D. Bentley, K. F. Chater, A.-M. Cerdeno-Tarraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C.-H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabbinowitsch, M.-A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barrell, J. Parkhill, D. A. Hopwood, *Nature* **2002**, *417*, 141 – 147.
[51] H. Ikeda, J. Ishikawa, A. Hanamoto, M. Shinose, H. Kikuchi, T. Shiba, Y. Sakaki, M. Hattori, S. Omura, *Nat. Biotechnol.* **2003**, *21*, 526 – 531.
[52] Y. Takeuchi, T. Akiyami, H. T. Harayama, *Chem. Pharm. Bull.* **1999**, *47*, 459 – 460.
[53] A. Marchler-Bauer, J. B. Anderson, C. DeWeese-Scott, N. D. Fedorova, L. Y. Geer, S. He, D. I. Hurwitz, J. D. Jackson, A. R. Jacobs, C. J. Lanczycki, C. A. Liebert, C. Liu, T. Madej, G. H. Marchler, R. Mazumder, A. N. Nikolskaya, A. R. Panchenko, B. S. Rao, B. A. Shoemaker, V. Simonyan, J. S. Song, P. A. Thiessen, S. Vasudevan, Y. Wang, R. A. Yamashita, J. J. Yin, S. H. Bryant, *Nucleic Acids Res.* **2003**, *31*, 383 – 387.
[54] S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D. J. Lipman, *Nucleic Acids Res.* **1997**, *25*, 3389 – 3402.
[55] R. E. Parales, *J. Ind. Microbiol. Biotechnol.* **2003**, *30*, 271 – 278.
[56] V. L. Healy, L. S. Mullins, X. Li, S. E. Hall, F. M. Raushel, C. T. Walsh, *Chem. Biol.* **2000**, *7*, 505 – 514.
[57] M. G. Thomas, Y. A. Chan, S. G. Ozanick, *Antimicrob. Agents Chemother.* **2003**, *47*, 2823 – 2830.

Received: August 6, 2004
Published online on February 18, 2005

