Isolation, purification and structure of exochelin MS, the extracellular siderophore from *Mycobacterium smegmatis*

Gary J. SHARMAN,* Dudley H. WILLIAMS,* David F. EWING† and Colin RATLEDGE‡§

*Cambridge Centre for Molecular Recognition, Department of Chemistry, Cambridge University, Lensfield Road, Cambridge CB2 1EW, U.K., †Department of Chemistry, University of Hull, Hull HU6 7RX, U.K. and ‡Department of Applied Biology, University of Hull, Hull HU6 7RX, U.K.

The extracellular siderophore from *Mycobacterium smegmatis*, exochelin MS, was isolated from iron-deficiently grown cultures and purified to > 98 % by a combination of ion-exchange chromatography and h.p.l.c. The material is unextractable into organic solvents, is basic (pI = 9.3–9.5), has a λ_{max} at 420 nm and a probable K_s for Fe^{3+} of between 10^{25} and 10^{30} . Its structure has been determined by examination of desferri- and ferri-exochelin and its gallium complex. The methods used were electrospray-m.s. and one- and two-dimensional (NOESY, DQF-COSY and TOCSY) ^1H n.m.r. The constituent amino acids were examined by chiral g.l.c analysis of *N*-trifluoroacetyl isopropyl and *N*-pentafluoropropionyl methyl esters after hydrolysis, and reductive HI hydrolysis, of the siderophore. The exochelin is a formylated pentapeptide: *N*-(δ -*N*-formyl, δ -*N*-hydroxy-*R*-ornithyl)- β -alaninyl- δ -*N*-hydroxy-*R*-ornithinyl-*R*-allo-threoninyl-

δ -*N*-hydroxy-*S*-ornithine. The linkages involving the three ornithine residues are via their δ -N(OH) and α -CO groups leaving three free α -NH₂ groups. Although there are two peptide bonds, these involve the three *R* (D)-amino acids. Thus the molecule has no conventional peptide bond, and this suggests that it will be resistant to peptidase hydrolysis. The co-ordination centre with Fe^{3+} is hexadenate in an octahedral structure involving the three hydroxamic acid groups. Molecular modelling shows it to have similar features to other ferric trihydroxamate siderophores whose three-dimensional structures have been established. The molecule is shown to have little flexibility around the iron chelation centre, although the terminal (Orn-3) residue, which is not involved in iron binding except at its δ -N atom, has more motional freedom.

INTRODUCTION

Macham and Ratledge (1975) reported the occurrence of a new group of extracellular siderophores from the mycobacteria which previously had been known to produce mycobactin as a putative siderophore (Snow, 1970). Mycobactin was located wholly intracellularly and the mechanism by which it solubilized iron from the external environment was never apparent (Macham et al., 1975). A prior suggestion (Ratledge and Marshall, 1972) had been made for the participation of salicylic acid as an extracellular iron solubilizing and transport agent, but this was subsequently discounted (Ratledge et al., 1974) because of the inability of salicylic acid to function in the presence of phosphate ions and other complexing agents. The occurrence of the exochelins in a number of mycobacterial species, including the pathogens, *Mycobacterium tuberculosis*, *M. africanum*, *M. avium* and *M. paratuberculosis* (Barclay and Ratledge, 1983, 1988), then established the likely role of these compounds as the means of iron solubilization and transport in the mycobacteria (Stephenson and Ratledge, 1979, 1980) for both *in vitro* and *in vivo* growth. The more recent finding of possible receptor proteins for the uptake of exochelins in *M. avium* and *M. leprae* recovered from animal infections (Sritharan and Ratledge, 1990) strengthens the case for exochelins fulfilling a key role in iron uptake and in mycobacterial pathogenicity (Wheeler and Ratledge, 1994).

Although considerable work was carried out to purify the exochelins using *M. smegmatis* as a suitable non-pathogen (Macham et al., 1977), the low yields of exochelin (about

2–3 $\mu\text{g}/\text{ml}$) even from iron-deficient cultures where the yield is considerably increased over iron-sufficient cultures, coupled with the pervasiveness of short peptides during purification, prevented the isolation of sufficient purified material to permit an unequivocal structural determination to be carried out at this time. Now, with improved h.p.l.c. facilities and the increased analytical power of n.m.r. and m.s. techniques, we have succeeded in purifying the exochelin from *M. smegmatis* and in determining its structure. Although the exochelin is similar to siderophores from other micro-organisms, in being a trihydroxamate derivative, it shows some unusual features of linkage between the component amino acids of the molecule.

EXPERIMENTAL

Organism and iron-deficient growth

M. smegmatis N.C.I.M.B. 8548 was grown in 250 ml conical flasks, pretreated to remove trace metals (see Ratledge and Hall, 1971), containing 100 ml of medium. The medium consisted of 5 g of KH_2PO_4 , 5 g *S*-(L)asparagine adjusted to pH 7.0 with KOH in 900 ml of ultrapure (membrane-filtered) water. After autoclaving at 121 °C for 15 min in the presence of 2 % (w/v) aluminium oxide (chromatography grade), the medium was filtered through Whatman no. 541 paper, dispensed in 90 ml aliquots into the flasks and re-sterilized. To each flask, immediately before inoculation, 1 g of glucose in 10 ml of ultrapure water was added (having been treated for iron removal as above) and sterilized by autoclaving (121 °C for 15 min). Flasks were

Abbreviations used: COSY, correlation spectroscopy; DQF-COSY, double-quantum-filtered COSY; DMSO, dimethylsulphoxide; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement spectroscopy; TOCSY, total correlation spectroscopy.

§ To whom correspondence should be addressed.

also supplemented before inoculation with 0.45 μg of Zn^{2+} /ml, 0.1 μg of Mn^{2+} /ml and 40 μg of Mg^{2+} /ml. For iron-deficient growth, 0.04 μg of Fe^{2+} /ml was added and for iron-sufficient growth 2.0 μg of Fe^{2+} /ml was added. The solutions containing all metal ions were sterilized separately.

For iron-deficient growth, flasks were inoculated with 1 ml of a 5- to 7-day-old culture of *M. smegmatis* grown with shaking in iron-deficient medium. Flasks were usually grown for 7 days before harvesting.

Isolation and purification of exochelin

Cultures were centrifuged (10000 *g* for 20 min), the supernatant solution adjusted to pH 7 with HCl, and a saturated solution of FeCl_3 added dropwise until a precipitate began to form. The solution was then stirred for 1 h and filtered through Whatman no. 1 paper in the cold, usually overnight. The filtrate, usually in batches of 750 ml, was passed through a cation-exchange column (Bio-Rad AG 50W-X8, 100–200 mesh in the NH_4^+ form; 8 cm $\times$ 3 cm) at a flow rate of 2 ml/min; the column was washed with 100 ml of water and the exochelin eluted with 1 M $\text{NH}_4\text{OH}/\text{NH}_4\text{Cl}$, pH 9.5, at 5 ml/min. An orange/deep yellow eluate was obtained in about 200 ml. This was evaporated in a rotary vacuum evaporator to about 10–15 ml. Any salt that precipitated was removed by filtration. The red solution was desalted by gel filtration through Sephadex G10 (48 cm $\times$ 3 cm), eluted with distilled water at 2 ml/min and 10 ml fractions collected as the red material began to elute. The conductivity of each fraction was monitored and as soon as this began to rise, indicating the elution of salts, the eluate was rejected. The collected, salt-free, fractions were pooled and evaporated to a volume of about 4 ml.

This partially purified material was re-fractionated using a high-resolving sulphonic acid ion-exchange resin (Bio-Rad AG 50W-X4, 200–400 mesh, NH_4^+ form, 40 cm $\times$ 2 cm). The column was eluted with a linear gradient of 0.1 M NH_4OH /acetic acid, pH 6.0, to 1 M NH_4OH /acetic acid, pH 9.0, and 3 or 4 ml fractions were collected. Each was monitored at 420 nm, the λ_{max} of the exochelin. The fractions comprising the main peak were pooled, evaporated to a volume of about 10 ml and desalted by gel filtration as before. The resulting exochelin solution was evaporated to 0.5 ml and further purified by h.p.l.c.

For h.p.l.c. samples, 0.2–0.3 ml of solution were passed through a semi-preparative column (25 cm $\times$ 1 cm) of Lichrosorb RP8, 10 μm , at room temperature with the following elution profile: solvent system (% v/v) at 0 min, 100% water; 15 min, 100% water; 20 min, 100% 0.1% trifluoroacetic acid; 45 min, 90% 0.1% trifluoroacetic acid in water/10% 0.1% trifluoroacetic acid in methanol; 60 min, as at 45 min.

The eluant was at 2 ml/min and the eluate was monitored continuously at 220 nm. The main exochelin peak, which was eluted after about 28 min, was freeze-dried, taken up in 0.5 ml of ultrapure water and a sample of 20 μl checked for purity using the above h.p.l.c. column with the following programme: solvent system (% v/v) at 0 min, 0.1% trifluoroacetic acid in water; 35 min, 90% 0.1% trifluoroacetic acid in water/10% 0.1% trifluoroacetic acid in methanol. The elution rate was 2 ml/min.

Preparation of desferri-exochelin

A sample of ferri-exochelin (> 98% pure) was diluted to 25 ml with ultrapure water, mixed with 145 mg (1 μmol) of 8-hydroxyquinoline in 25 ml of h.p.l.c.-grade methanol, and stirred overnight at room temperature. The solution, now dark green, was extracted with dichloromethane (5 $\times$ 20 ml). All glassware with which the desferri-exochelin solution subsequently came into

contact had to be iron-free to prevent reformation of the ferri-complex. Desferri-mycobactin was equally effective in producing desferri-exochelin, and could be readily removed by extraction into chloroform. However, the above procedure was the more convenient.

Preparation of gallium exochelin

This procedure was modified from that given by Stephan et al. (1993a). To a solution of desferri-exochelin (approx. 5 mg, about 10 μmol , in 25 ml of water) was added 15 mg (35 μmol) of GaSO_4 , dissolved in the minimum amount of 0.05 M H_2SO_4 (less than 0.5 ml), and after 30 min at room temperature, the pH was adjusted to 7.0 with NaOH and the solution was then freeze-dried. The freeze-dried material was dissolved in 1 ml of ultrapure water and any undissolved material removed by decantation. The solution was then separated from the residual gallium sulphate and also any 8-hydroxyquinoline still remaining from the preparation of the desferri-exochelin by semi-preparative h.p.l.c. using the above column with the following programme: solvent system (% v/v) at 0 min, 100% water; 5 min, 100% water; 9 min, 100% 0.1% trifluoroacetic acid in water; 24 min, 95% 0.1% trifluoroacetic acid in water/5% methanol; 50 min, 100% methanol. The salts were eluted in the first 5 min and the residual 8-hydroxyquinoline was eluted with methanol in the final step.

N.m.r. spectroscopy

N.m.r. tubes were rendered iron-free by washing sequentially with HNO_3 , EDTA and membrane-filtered water. All experiments were performed on the gallium complex, 4–5 mmol in dimethylsulphoxide (DMSO), at 308 K using a Bruker AM500 spectrometer. For one-dimensional experiments, 16 K data points were recorded over 64 transients. For two-dimensional spectra, 2 K data points were collected in f_2 using 512 increments; f_1 was zero-filled to 1 K. The nuclear Overhauser enhancement spectroscopy (NOESY) experiments used a mixing time of 300 ms. Total correlation spectroscopy (TOCSY) experiments were performed with the machine in reverse mode, employing a spin-locking field of 10 kHz. Suppression of the residual water signal was achieved by presaturation. Standard Bruker pulse programs were used. Data was processed using either a Gaussian or shifted squared sinebell window function, and were referenced to the DMSO peak ($\delta = 2.52$ p.p.m.).

Microderivatization reactions and m.s.

Esterification

0.1 M HCl in anhydrous methanol was prepared by blowing a gentle stream of HCl gas over the alcohol until the appropriate gain in mass was achieved. A 50 μl aliquot of this solution was added to about 25 nmol of exochelin in a 0.3 ml vial. The mixture was left to stand for 24 h, after which excess reagents were removed under vacuum and the residue dissolved in acetonitrile/water (1:1, v/v) to give a 50 μM solution for analysis by electrospray m.s.

Acetylation

Acetic anhydride/0.1 M ammonium acetate, pH 8.4, (50 μl ; 1:1, v/v), was added to a small sample (25 nmol) of the exochelin and the mixture subjected to ultrasound for 2 h at room temperature. Excess reagents were removed under vacuum and a solution of the residue prepared as above for analysis by electrospray m.s.

Acetylation II

Pyridine/acetic anhydride (2:1, v/v), 50 μ l, was added to small samples of the exochelin. The mixture was left at room temperature for 4 h after which excess reagents were removed under vacuum and the residue analysed by electrospray m.s.

Electrospray m.s.

All experiments were performed on a VG BioQ instrument using a source temperature of 70 °C and a capillary potential of 4 kV. The mobile phase used was acetonitrile/water (1:1, v/v) at 4 μ l/min. Positive ion mode was used for analysis of the methyl esters but negative ion mode was used for the analysis of acetyl derivatives, as these had no sites for protonation. Typically, 10 μ l of 50 μ M solution was injected, data were collected over 20 scans, smoothed and peak-top mass assignment carried out. Calibration was achieved using polyethylene glycol.

G.c. analysis of amino acid stereochemistry

Preparation of standard derivatives

A 100 μ l aliquot of 3 M HCl in propan-2-ol, was added to about 2 mg of β -alanine and each of the four possible isomers of threonine in separate 0.3 ml vials. Each was heated at 100 °C for 30 min after which excess reagents were removed under vacuum. Trifluoroacetic anhydride/dichloromethane (1:1, v/v), 100 μ l, was then added and the mixture heated at 100 °C for 10 min. Excess reagents were removed with a stream of argon. Dichloromethane, 100 μ l, was added and the resulting *N*-trifluoroacetyl propan-2-yl esters analysed by g.c.

This derivative proved unsuitable for the detection of the enantiomers of ornithine, so the *N*-pentafluoropropionyl methyl esters were used. HCl, 5 M in methanol, was prepared by adding 0.5 ml of acetyl chloride to 1.5 ml of anhydrous methanol with stirring; 100 μ l was then added to *R*-ornithine (2 mg) in a 0.3 ml vial. The mixture was heated at 90 °C for 30 min, after which time excess reagents were removed under vacuum. Pentafluoropropionic anhydride (50 μ l) was added and the mixture left to stand for 30 min. Excess reagents were removed under a stream of argon. Dichloromethane, 100 μ l, was added and the resulting solution analysed by g.c. This was repeated for the *S*-isomer.

Hydrolysis and derivatization of exochelins

Concentrated HI, 100 μ l, was added to exochelin (approx. 100 nmol) and the mixture heated at 100 °C for 24 h. Excess reagents were removed under vacuum and the mixture of amino acids was derivatized as above; two hydrolyses were carried out so both derivatives could be made.

Reduction of C-terminal ornithine

Acetic anhydride/0.1 M ammonium acetate, pH 8.4 (1:1, v/v), 50 μ l, was added to 100 nmol of exochelin in a 0.3 ml vial. The mixture was subjected to ultrasound for 2 h. Excess reagents were then removed under vacuum to give the acetylated exochelin. A 50 μ l aliquot of 0.1 M HCl in methanol was added and the mixture left at room temperature for 24 h. Again excess reagents were removed under vacuum to give the acetylated methyl ester. The vial was flushed with argon and 100 μ l of 2 M LiBH₄ in tetrahydrofuran was added. The mixture was heated for 6 h at 60 °C after which time 100 μ l of water, then 30 μ l of 1 M HCl were added to terminate the reaction. Solvent was removed under vacuum and 50 μ l of 6 M HCl was added to ensure all

basic salts were neutralized. Again, the solvent was removed under vacuum and the product hydrolysed and derivatized as the *N*-pentafluoropropionyl methyl ester as above.

Chiral g.c. separation of derivatives

The derivatized amino acids were analysed by g.c. using a Carlo Erba 4130 instrument fitted with a Chirasil-Val capillary column (0.2 mm $\times$ 50 m, Altech Chromatography). Helium was used as carrier gas with both injector and detector maintained at 250 °C; 0.1 μ l of each derivative was injected, with a splitting ratio of 20:1. A temperature program running from 80 °C to 170 °C at 4 °C/min was used for the trifluoroacetyl derivatives, whereas the pentafluoropropionyl derivatives were run isothermally at 175 °C. The components of the mixture of amino acids from the exochelins were identified by comparison of retention times with those of the standards and by spiking with the suspected component.

Molecular modelling

All modelling was performed on a Silicon Graphics Iris Indigo R4000 computer using Macromodel. The AMBER force field was used throughout with SOLVENT set to water. Parameters for iron (III) were created by changing the charge and radius of Ba²⁺ to 3+ and 0.69 Å respectively. Parameters for N-O- stretch were approximated as being equal to those of N-OH. Since electrostatic forces were the only ones between ligand and metal, it was necessary to equalize the default partial charges for the two types of oxygen in the hydroxamate groups. The average of the two was used. This resulted in good octahedral geometry on minimization.

All minimizations were run to convergence (typically less than 1000 iterations) using PRCG. For the isothermal dynamics run, a bath temperature of 600 K was used and samples were taken every 1 ps over the 100 ps run. For simulated annealing, the initial temperature was 1000 K, reducing to 0 K over 100 ps. The SHAKE algorithm was used in both calculations.

RESULTS AND DISCUSSION

Isolation and purification of exochelin MS

The purification procedure described in the Experimental section for the isolation of exochelin MS was developed from an earlier method (Macham et al., 1977). Although this procedure required the use of mildly alkaline conditions (see Figure 1a), we were able to establish that when the exochelin was isolated by an alternative but more laborious process that avoided any pH change and which involved membrane filtration and gel-permeation chromatography, identical material was still produced. The given method was preferred as it was the easier procedure leading to exochelin of high purity. Glucose was used as carbon source for growth of cells instead of glycerol, the previous carbon source, as this led to production of much less viscous solutions of exochelin which were easier to process.

Two cation-exchange chromatography steps were needed to purify the exochelin. The first removed a large amount of neutral and acidic peptides and other materials which interfered with the second, high-resolution ion-exchange stage. The elution profile of the exochelin from this second ion-exchange column is given in Figure 1(a). The elution profile of the material from the subsequent semi-preparative h.p.l.c. column is shown in Figure 1(b).

The final purity of preparations of exochelin isolated in this way was never less than 98% as judged by h.p.l.c. analysis (Figure 1c) or by capillary electrophoresis (Figure 1d). The yield

Figure 1 H.p.l.c. analysis and capillary electrophoresis of ferri-exochelin

(a) Elution profile of partially purified ferri-exochelin from a high-resolving cation-exchange (sulphonic acid) column (Bio-Rad AG 50W-X4, 200–400 mesh, NH_4^+ form, 40 cm $\times$ 2 cm). Fractions (4 ml) were taken and their A_{420} values measured. The eluate between the arrows was collected. (b) H.p.l.c. of ferri-exochelin from previous step after concentration using a semi-preparative, reverse-phase column with a three-solvent elution programme (see Experimental section). The eluate was monitored continuously at 220 nm. The eluate between the arrows was collected. (c) H.p.l.c. of final, purified ferri-exochelin from the previous step using a two-solvent elution programme (see Experimental section). The eluate was monitored continuously at 220 nm. Area under peak was given as 100% of total eluate by the electronic integrator attached to the h.p.l.c. (d) Capillary electrophoresis of final, purified ferri-exochelin. [The procedure was carried out at the Bio-Rad Laboratories, Hemel Hempstead, U.K., using a 24 cm $\times$ 25 μm coated capillary column with 0.1 M phosphate buffer, pH 7, under constant voltage (4 kV) with + to – change in polarity.]

of pure ferri-exochelin was usually about 3 mg/l starting with 2.5 l of culture filtrate.

Properties of exochelin MS

The u.v.–visible spectrum of ferri-exochelin in water showed a

single peak at 420 nm. Given a molecular mass of ferri-exochelin of 661 Da (see below), a molar absorption coefficient of $1044 \text{ M}^{-1} \cdot \text{cm}^{-1}$ was calculated; ($A^{1\%} = 15.8$). The material is stable in its ferri-form at room temperature, both in solution and when dry. Although it is freely soluble in water and in a number of polar solvents, it cannot be extracted from the original culture

Figure 2 Electrospray mass spectrum of ferri-exochelin MS

Peaks at 662.3 and 331.8 are from MH^+ and MH_2^{2+} respectively and give a mass of 661 Da.

medium using either ethanol, propanol, propan-2-ol, etc. in which K_2CO_3 is used to produce two liquid phases. Ferri-exochelin is strongly basic having a pI of 9.5–9.8 (L. P. Macham, unpublished work). Desferri-exochelin is colourless, although dry material is usually slightly yellow due to the presence of some residual ferri-material. The stability of the pure desferri-material in solution and when dry appears to be satisfactory, though some impure preparations apparently underwent spontaneous hydrolysis.

Although an exact stability constant for iron has not been calculated, desferri-mycobactin, which has a K_s of more than 10^{30} (Snow, 1970), could readily remove iron from the ferri-exochelin.

The biological functions of exochelin MS have been described earlier (Macham et al., 1977). Briefly, the molecule is able to solubilize iron from ferric phosphate and ferritin, and can reverse the bacteriostatic effect of serum (which is due to the presence of transferrin) on the growth of *M. smegmatis*, presumably by removing the iron from the transferrin molecule. Exochelin MS mediates in the active transport of iron into *M. smegmatis* and also into *M. vaccae*, which produces a similar exochelin (Messenger et al., 1986), but not into *M. bovis* BCG, *M. intracellulare* (Stephenson and Ratledge, 1980) or *M. leprae* (Hall and Ratledge, 1987), indicating a specific uptake system is present in *M. smegmatis*. [However, *M. smegmatis* can take up iron from exochelins of other mycobacteria including BCG by a separate, non-active transport process (Stephenson and Ratledge, 1980).]

Structural elucidation of the exochelin

Preliminary work using acid hydrolysis and identification of the products by automated amino acid chromatography indicated the probable presence of δ -N-hydroxyornithine, β -alanine and threonine in the molar ratio of 3:1:1 (L. P. Macham and A. J. Messenger, unpublished work; see Ratledge, 1987). However, an attempted Edman degradation gave no useful results (L. Packman, personal communication).

The component amino acids could be identified by initial 1H -n.m.r. studies, but an unexpected, non-exchanging proton at $\delta = 8.3$ p.p.m. was also present. The molecular masses of desferri- and ferri-form peptides were found to be 608 and 661 Da by electrospray m.s. (Figure 2); the difference of 53 mass units can be accommodated by the binding of iron with associated loss of three protons. When the same samples were run in 2H_2O , masses of 621 and 671 Da were observed, indicating 13 and 10 ex-

changeable protons respectively. A linear pentapeptide of the composition suggested should have 14 exchangeable protons, the iron-bound peptide 11. Thus, one exchangeable proton anticipated for such a structure was missing. Furthermore, the observed mass of 608 Da is 28 more than that calculated for a linear pentapeptide of the above composition. On the basis of this evidence, the presence of a formyl group on the δ -N of one ornithine was postulated (Groves, 1994). This would account for one less exchangeable proton, would explain the downfield proton in the n.m.r. and would add 28 mass units. Furthermore, one of the common iron-binding moieties, *cis*-N-hydroxamide, would be formed.

Functional group identification

Knowledge of the functional groups present in a peptide provides good evidence for an amino acid composition and can indicate if it is linear or if any blocking groups are present (Mortishire-Smith et al., 1991; Nutkins and Williams, 1989; Nutkins, 1991). The functional groups of exochelin MS were investigated by microderivatization reactions and electrospray m.s. Small samples of the exochelin, typically about 50 nmol, were esterified and acetylated in order to determine the numbers of carboxyl, amino and hydroxyl groups present. Esterification with methanol/HCl adds 14 mass units for each carboxylate in the molecule. Acetylation with acetic anhydride/water adds 42 mass units for each amino group, but with acetic anhydride/pyridine both amino and hydroxyl groups react. Thus the number of hydroxyls can be inferred from the difference between the result of the two acetylations. Table 1 gives details of the experiments carried out, together with the change in mass observed.

These results are compatible with a linear peptide containing three δ -N-hydroxyornithines, one threonine, a β -alanine and a formyl group. It was not clear at this stage whether the amino groups were NH_2 or $NHOH$ as the reactivity of the latter under the above conditions was not known. Interestingly, the tri-acetylated product and the ester still bound iron but the hepta-acetyl derivative did not. This suggested neither the amino groups nor the carboxylate group were involved in chelation. Also, when the desferri-material was acetylated with acetic anhydride/water, six acetyl groups were added. This was readily explained by the acetylation of the three amino groups plus the three N-hydroxyls, the reactivity of which must be greater than that of a normal hydroxyl (e.g. that of threonine). This strongly suggested that the N-hydroxyl groups were involved in binding the iron, as expected.

N.m.r. spectroscopy

Ferri-exochelins cannot be studied by n.m.r. because the paramagnetic iron causes fast relaxation and severe broadening of lines. Although desferri-exochelin is suitable for n.m.r. work, and was used in the initial work described above, its conformation is unlikely to be the same as the ferri-form and so any conformational information from, for example, NOESY experiments might not be relevant for the ferri-form. Accordingly, the gallium complex was used. Ga^{3+} is of similar size to Fe^{3+} but it is not paramagnetic so a sharp spectrum can be obtained while retaining the conformation of the iron-bound material.

The solvent of choice for n.m.r. work would have been water because of its biological relevance but it proved impossible to obtain any nuclear Overhauser enhancements (NOEs) in this solvent. This was a problem as NOEs were likely to contain useful information on the sequence and conformation. One- and two-dimensional double-quantum-filtered correlation spectro-

Table 1 Results of micro-derivatization reactions on ferri-exochelin MS

Results are for ferri-form unless stated otherwise.

Reagents	Mass (Da)	Change (Da)	Conclusion
None	661	—	—
(desferri-form)	608	—	—
0.1 M HCl in methanol	675	+14	1 carboxylic acid
Acetic anhydride/water (1:1, v/v)	787	+126	3 amino groups
(desferri-form)	(860)	(+252)	(3 more basic groups)
Acetic anhydride/pyridine (1:2, v/v)	902	+241	Loss of iron + 7 acetyls
(desferri-form)	(902)	(+294)	(7-3) = 4 hydroxyls

scopy (DQF-COSY) (Piantini et al., 1987), TOCSY (Braunschweiler and Ernst, 1983) and NOESY (Bodenhausen et al., 1984) spectra of the gallium complex were therefore recorded in DMSO at 308 K. Table 2 gives the full ^1H assignment.

One-dimensional ^1H n.m.r. of the gallium complex (Figure 3a) revealed very little about the exochelin. The formyl proton and the threonine methyl group were the only discernible features. The DQF-COSY spectrum identified the spin systems of β -alanine and threonine (Figure 3b). Notably, all of the protons of β -alanine were at different chemical shifts, and one of the β protons was not coupled to the NH, suggesting a specific conformation. By using COSY and TOCSY spectra, the spin

systems of the three ornithines were also assigned. However, the most notable feature was that only β -alanine and threonine had amide protons; the spin systems of the three hydroxyornithines all traced through to three broad $\alpha\text{-NH}_3^+$ signals and there were no signals attributable to NHOH. This showed that the three ornithines were not incorporated as standard amino acids. They all had to be linked through their $\delta\text{-N}$ atoms so removing the $\delta\text{-N}$ protons and leaving the α -amino groups free. One of these links would have to be to the formyl group; the others to the α -carbonyls of other amino acids. Not only does this interpretation fit the n.m.r. data, it also provides three hydroxamate chelating groups of the kind so common in siderophores. The fact that such a structure would have three amino groups also fitted with the results of the functional group identification and the strong basicity of the molecule. The structure therefore had to be a linear peptide (a free carboxylate had been detected, Table 1), with the hydroxyornithines linked through their $\delta\text{-N}$ atoms. Attention now turned to possible sequences.

Sequence determination

Fast-atom-bombardment m.s. failed to provide fragment ions for ferri- or desferri-exochelins. Pulse hydrolysis also proved unhelpful. The sequence was eventually determined by a careful analysis of the n.m.r. spectra. It proved possible to 'walk' from one end of the molecule to the other, using COSY and TOCSY to go along spin systems and NOESY to go between them.

Table 2 Complete ^1H -n.m.r. assignment of exochelin MS

Residue	Assignment	δ (p.p.m.)	NOESY	COSY	TOCSY
Formyl	H	8.23	Orn-1 ($\delta 1$, $\delta 2$, $\beta 1$, $\beta 2$)	—	—
Orn-1	$\delta 1$	3.60	Formyl, $\gamma 1$, $\gamma 2$	$\delta 2$	$\gamma 1$, $\gamma 2$, $\beta 1$, $\beta 2$
	$\delta 2$	3.50	Formyl, $\beta 1$, $\gamma 2$	$\delta 1$	$\gamma 1$, $\gamma 2$, $\beta 1$, $\beta 2$
	$\gamma 1$	1.64	$\gamma 2$, $\delta 1$	—	$\delta 1$, $\delta 2$
	$\gamma 2$	1.45	$\gamma 1$, $\delta 1$, $\delta 2$	—	$\delta 1$, $\delta 2$, α
	$\beta 1$	1.76	α , $\beta 2$, $\gamma 2$, $\delta 2$	α , $\beta 2$	$\delta 1$, $\delta 2$, α
	$\beta 2$	1.62	α , $\beta 1$, β ala (NH)	$\beta 1$	$\delta 1$, $\delta 2$, α
	α	3.90	$\beta 1$, $\beta 2$, β ala (NH)	$\beta 1$, NH_2	$\gamma 2$, $\beta 1$, $\beta 2$
	NH_2	8.10	—	α	α , $\beta 1$, $\beta 2$, $\gamma 2$
	NH	8.35	$\beta 2$, $\alpha 2$, Orn-1 (α , $\beta 2$)	$\beta 1$	$\beta 1$, $\beta 2$, $\alpha 1$, $\alpha 2$
	$\beta 1$	3.67	$\beta 2$, $\alpha 1$	NH, $\beta 2$, $\alpha 2$	NH, $\beta 2$, $\alpha 1$, $\alpha 2$
β -Alanine	$\beta 2$	3.05	NH, $\beta 1$, $\alpha 1$, $\alpha 2$	$\beta 1$	NH, $\beta 1$, $\alpha 1$, $\alpha 2$
	$\alpha 1$	2.95	$\beta 1$, $\beta 2$, $\alpha 2$, Orn-2 ($\delta 1$, $\beta 1$, $\beta 2$)	$\alpha 2$	NH, $\beta 1$, $\beta 2$, $\alpha 2$
	$\alpha 2$	2.33	NH, $\alpha 1$, $\beta 2$	$\beta 1$, $\alpha 1$	NH, $\beta 1$, $\beta 2$, $\alpha 1$
	$\delta 1$	3.80	$\delta 2$, $\gamma 2$, $\beta 2$, β ala ($\alpha 1$)	$\delta 2$, $\gamma 1$	$\beta 1$, $\beta 2$, $\gamma 1$, $\gamma 2$
	$\delta 2$	3.50	$\delta 1$, $\gamma 1$, $\gamma 2$	$\delta 1$, $\gamma 1$	$\beta 1$, $\beta 2$, $\gamma 1$, $\gamma 2$
	$\gamma 1$	1.65	$\delta 2$	$\beta 1$, $\gamma 2$, $\delta 1$, $\delta 2$	$\delta 1$, $\delta 2$, α
	$\gamma 2$	1.42	$\delta 1$, $\delta 2$, $\beta 1$	$\gamma 1$, $\beta 2$	$\delta 1$, $\delta 2$, α
	$\beta 1$	1.87	α , $\beta 2$, $\gamma 2$	α , $\gamma 1$	$\delta 1$, $\delta 2$, α
	$\beta 2$	1.50	α , $\beta 1$, Thr (NH)	$\gamma 2$	$\delta 1$, $\delta 2$, α
	α	3.98	$\beta 1$, $\beta 2$	$\beta 1$, NH_2	$\beta 1$, $\beta 2$, $\gamma 1$, $\gamma 2$
Threonine	NH_2	8.14	—	α	—
	NH	8.80	β , methyl, Orn-2 (α , $\beta 2$)	α	α , β , methyl
	α	4.60	Methyl, Orn-3 ($\delta 1$, $\delta 2$)	NH, β	NH, β , methyl
	β	3.88	NH, methyl	α , methyl	NH, α , methyl
	Methyl	1.20	NH, α , β	β	α , β , NH
	OH	5.20	—	—	—
	Orn-3				
	$\delta 1$	3.84	Thr (α)	$\delta 2$, $\gamma 1$	$\beta 1$, $\beta 2$, $\gamma 1$, $\gamma 2$
	$\delta 2$	3.79	Thr (α)	$\delta 1$, $\gamma 2$	$\beta 1$, $\beta 2$, $\gamma 1$, $\gamma 2$
	$\gamma 1$	1.94	—	$\delta 1$	$\delta 1$, $\delta 2$, α
	$\gamma 2$	1.91	—	$\delta 2$	$\delta 1$, $\delta 2$, α
	$\beta 1$	1.82	—	$\beta 2$	$\delta 1$, $\delta 2$, α
	$\beta 2$	1.99	—	$\beta 1$	$\delta 1$, $\delta 2$, α
	α	3.85	—	—	$\beta 1$, $\beta 2$, $\gamma 1$, $\gamma 2$
	NH_2	8.20	—	—	—

Figure 3 One-dimensional ¹H-n.m.r spectrum (a) and DQF-COSY spectrum (b) of the gallium complex of exochelin MS taken at 308 K in DMSO

(a) The three peaks at $\delta = 7$ p.p.m. are an impurity from an ammonium used in the isolation (these were of variable intensity in different samples). (b) The spin systems of β -alanine (full line) and threonine (broken line) are identified.

Figure 4 Covalent structure of exochelin MS

Broken lines indicate observed NOEs.

Figure 4 shows the resulting structure; broken lines show the NOEs described below. Starting from the formyl proton, NOEs were seen to the diastereotopic δ -protons of an ornithine. The

TOCSY spectrum unambiguously identified the α -proton of this residue, which showed NOEs to the β -alanine NH. Therefore β -alanine was the next residue. One of the β -alanine α -protons gave an NOE to a δ -proton of the next ornithine and the α -proton of this residue gave an NOE to the threonine-NH, showing these were the next two amino acids in the sequence. Finally, an NOE was found from the threonine α -proton to the δ -protons of the third ornithine. The assignment of this last residue left no protons unaccounted for and, notably, no NOEs were seen from its α -proton. It also showed much smaller shift differences between the diastereotopic β -, γ - and δ -protons, a reflection of the much greater motion present in this residue, only one end being bound to the metal.

Determination of stereochemistry of residues

The chirality of the amino acids was determined by chiral g.c. Using standards derivatized as the *N*-trifluoroacetyl isopropyl

Figure 5 Observed NOEs and structural conformations of exochelin MS

(a) Around the ornithine-1 region; (b) around the β -alanine region; and (c) around the threonine region.

esters, a single assay for the analysis of β -alanine and all four isomers of threonine was established. This derivative was unsuitable for the resolution of ornithine enantiomers, but the *N*-pentafluoropropionyl methyl esters (Demange et al., 1988) gave a good assay. Ornithine, and not the δ -*N*-hydroxyl derivative, was used because the latter was unavailable in optically pure form. The exochelin was therefore hydrolysed with HI instead of HCl. As well as effecting hydrolysis, HI reduces hydroxylamines to amines (Stephan et al., 1993b). The exochelin was thus cleaved into its constituent amino acids and the δ -*N*-hydroxyornithines were reduced to ornithines. The *N*-trifluoroacetyl propan-2-yl esters and *N*-pentafluoropropionyl methyl esters were made from the hydrolysate and the two g.c. assays carried out (results not shown); peaks were identified by comparison with the standards. The results clearly demonstrated the presence of

Figure 6 Three-dimensional Macromodel structure of ferri-exochelin MS

The ferric atom forms a hexadentate structure with the neighbouring six O atoms.

the achiral β -alanine, *R*-allo-threonine, two *R*- δ -*N*-hydroxyornithines and one *S*- δ -*N*-hydroxyornithine.

Location of the *S*-ornithine residue

From the preceding section, there were three possible locations for the single *S*-ornithine residue. Space-filling CPK models were made of all three possibilities, from which it was clear that only *R*-*R*-*S* and *S*-*R*-*R* (going from N- to C-terminus) could satisfy the complete set of NOEs (see below). These were distinguished by acetylating the exochelin, esterifying it and reducing it with LiBH₄. [LiBH₄ reduces esters in the presence of amides which do not react (Nystrom et al., 1949).] Thus the C-terminal residue was effectively removed. After hydrolysis and derivatization as described above, only *R*-ornithine remained, indicating that the C-terminal residue was *S*. The chirality of the residues in Figure 4 is therefore: *R*- δ -*N*-hydroxyornithine¹- β -alanine-*R*- δ -*N*-hydroxyornithine²-*R*-allo-threonine-*S*- δ -*N*-hydroxyornithine³.

Determination of conformation

The exochelin gave many other NOEs in addition to those used in sequence determination (see Figure 5). From the primary structure of the peptide, and the stereochemistry of its residues, it seemed that the spatial constraints provided by the NOEs might be enough to define its conformation. CPK models were made, from which it became clear that, although there were potentially many ways in which the peptide could fold, only one could satisfy the complete set of constraints.

Starting at the formyl proton (Figure 5a), NOEs were seen to the δ - and β -protons of the first ornithine, but not to the γ . Further NOEs were seen from the α - to β -protons and from the NH of the β -alanine to the α - and one of the β -protons. The β -alanine amide proton (Figure 5b) showed an NOE to its β 2-proton but was not coupled to it. The β 1-proton, however, was

coupled to the amide proton, but showed no NOE to it, strongly suggesting the amide and the β 1-proton were anti-periplanar. This proton was also coupled to the α 2-proton, but no NOE was observed between them, suggesting these were also anti. Only one conformation satisfies all of these constraints and it brings the second bidentate hydroxamate back to the iron atom in exactly the right orientation. Moving along the chain, NOEs were seen from the α 1-proton of the β -alanine to the δ 1- and the two β -protons of the next ornithine. NOEs were also seen between Orn α and β protons and between the Thr NH and Orn α and β protons, exactly as for the first ornithine. The threonine NH showed NOEs to its methyl group (Figure 5c) and to the β -proton, but not to the α . It did, however, show a coupling to the α proton, suggesting an anti-conformation. Finally, the α -proton also showed NOEs to the δ -protons of ornithine-3. The conformation of the C-terminal ornithine was not well defined but this was expected as only one end is involved in chelation, the other being free to move around in solution.

A complete three-dimensional structure had now been generated. This was entered into the MacroModel (Mohamadi et al., 1990) molecular modelling package and minimized using a slightly modified version of the AMBER (Weiner et al., 1984, 1986) force-field. Simulated annealing was then carried out in order to find the global minimum, followed by a final round of energy minimization. The resulting structure is shown in Figure 6. The interatomic distances of the atoms with NOE relationships were measured and, in all cases, they were less than 4 Å; those for strong NOEs were all less than 3 Å.

The structure was also subjected to an isothermal dynamics run, with structures being sampled every 1 ps for 100 ps. A moving film was created from the samples from which it was clear that the complex held together tightly and showed no tendency to change conformation or dissociate in a way that would not fit the observed NOEs. Several of these samples were minimized and when superimposed (data not shown) gave a very high degree of overlap apart from the C-terminal ornithine, which, as expected, showed a greater degree of flexibility as it was not involved in the binding of the exochelin to the iron.

It should be remembered that the NOEs used to define the original structure were obtained in DMSO, not water, so the relevance of this structure to that in nature might be questioned. However, a similar conformation is supported in this particular case by the fact that the one-dimensional ^1H -n.m.r spectrum is almost identical in the two solvents.

Comments on the structure

The structure has several noteworthy points. It co-ordinates the iron octahedrally with three *cis*-hydroxamate ligands. Additionally, the arrangement of these ligands is *fac*, such that the three N-bound O atoms are next to each other in the structure (see Figure 6). The octahedral iron therefore has a three-fold axis of symmetry. All other hydroxamate-type siderophores for which a three-dimensional structure is known have been found to co-ordinate in this way (Raymond, 1977; Winkelmann, 1991). The structure also has a high proportion of *R*-(D)-amino acids. *R*-Amino acids are not usual in siderophores but some examples are known: in pseudobactins, pyoverdins and azotobactin found in various *Pseudomonas* species (Abdallah, 1991). However, what is unique about exochelin MS is that all of the conventional peptide linkages involve *R*-amino acids. It seems likely that this is to give the peptide resistance to hydrolytic enzymes. The only *S*-ornithine is at the C-terminus and has no α -linkage through its α -N or C atoms. Perhaps the reason for this retention is that there would be no enhancement of hydrolytic resistance by

inverting it and thus some metabolic energy is saved by retaining it as the natural *S*-isomer. If other exochelins in the pathogenic mycobacteria are found to involve similar *R*-amino acids, this could explain how these siderophores are able to persist when produced inside host cells by invading mycobacteria such as *M. tuberculosis* or *M. leprae*.

This is the first structure to be elucidated for what is clearly a family of siderophores belonging to the mycobacteria. It is already known that exochelin MS is different to those found in other mycobacteria (Ratledge, 1987) but, just as there is a common structure for the various mycobactins of this genus, so we might expect to find a common structural theme running through the exochelins. Work is already underway on the structural elucidation of some of these.

Joint application of modern n.m.r. techniques, derivatization, m.s. and g.c. has allowed complete determination of the three-dimensional structure of this natural product. Such is the sensitivity of these techniques that the whole structural determination was carried out on only a few milligrams of material, though over 120 l of cultures had to be grown in order to make this possible.

The work carried out at the University of Hull was supported by the Science and Engineering Research Council Grant No. GR/F 45387. The SERC is also thanked for financial support for the work carried out at the University of Cambridge. We are most appreciative of the technical assistance given by Mrs. Maureen Ewing in the purification of the exochelins. We are also indebted to Mr. Barry Keenan, Bio-Rad Research Laboratories, Hemel Hempstead, Herts., HP2 7TD, U.K. for carrying out the capillary electrophoresis work.

REFERENCES

- Abdallah, M. A. (1991) in Handbook of Microbial Chelates (Winkelmann, G., ed.), pp. 139–153, CRC Press, Boca Raton, FL.
- Barclay, R. and Ratledge, C. (1983) J. Bacteriol. **153**, 1138–1146
- Barclay, R. and Ratledge, C. (1988) J. Gen. Microbiol. **134**, 771–776
- Bodenhausen, G., Ernst, R. R. and Kogler, H. (1984) J. Mag. Res. **58**, 370–388
- Braunschweiler, L. and Ernst, R. R. (1983) J. Mag. Res. **53**, 521–528
- Demange, P., Abdallah, M. A. and Frank, H. (1988) J. Chromatogr. **438**, 291–297
- Groves, P. (1994) Ph.D. Thesis, Cambridge, U.K.
- Hall, R. M. and Ratledge, C. (1987) J. Gen. Microbiol. **133**, 193–199
- Macham, L. P. and Ratledge, C. (1975) J. Gen. Microbiol. **89**, 379–382
- Macham, L. P., Ratledge, C. and Nocton, J. C. (1975) Infect. Immun. **12**, 1242–1251
- Macham, L. P., Stephenson, M. C. and Ratledge, C. (1977) J. Gen. Microbiol. **101**, 41–49
- Messenger, A. J. M., Hall, R. M. and Ratledge, C. (1986) J. Gen. Microbiol. **132**, 845–852
- Mohamadi, F., Richards, N. G. J., Guida, W. C., Liskamp, R., Lipton, M., Caufield, C., Chang, C., Henrickson, T. and Still, W. C. (1990) J. Compt. Chem. **11**, 440–467
- Mortishire-Smith, R. J., Nutkins, J. C., Packman, L. C., Bradley, C. L., Raney, P. B., Johnstone, K. and Williams, D. H. (1991) Tetrahedron **47**, 3645–3654
- Nutkins, J. C. (1991) J. Am. Chem. Soc. **113**, 2621–2627
- Nutkins, J. C. and Williams, D. H. (1989) Eur. J. Biochem. **181**, 97–102
- Nystrom, R. F., Chaiken, S. W. and Brown, W. G. (1949) J. Am. Chem. Soc. **71**, 3245–3246
- Piantini, U., Sorensen, O. W. and Ernst, R. R. (1987) J. Am. Chem. Soc. **109**, 6800–6801
- Ratledge, C. (1987) in Iron Transport in Microbes, Plants and Animals (Winkelmann, G., van der Helm, D. and Neilands, J. B., eds.), pp. 207–221, VCH Publishers, Weinheim, Germany
- Ratledge, C. and Hall, M. J. (1971) J. Bacteriol. **108**, 312–317
- Ratledge, C. and Marshall, B. J. (1972) Biochim. Biophys. Acta **279**, 58–74
- Ratledge, C., Macham, L. P., Brown, K. A. and Marshall, B. J. (1974) Biochim. Biophys. Acta **372**, 39–51
- Raymond, K. N. (1977) in Bioinorganic Chemistry 2 (Raymond, K. N., ed.), pp. 33–54, American Chemical Society, Washington, DC
- Snow, G. A. (1970) Bacteriol. Rev. **34**, 99–125
- Sritharan, M. and Ratledge, C. (1990) BioMetals **2**, 203–208
- Stephan, H., Freund, S., Beck, W., Jung, G., Meyer, J.-M. and Winkelmann, G. (1993a) BioMetals **6**, 93–100
- Stephan, H., Freund, S., Meyer, J., Winkelmann, G. and Jung, G. (1993b) Liebig's Ann. Chem. **43**–48
- Stephenson, M. C. and Ratledge, C. (1979) J. Gen. Microbiol. **110**, 193–202
- Stephenson, M. C. and Ratledge, C. (1980) J. Gen. Microbiol. **116**, 521–523

Weiner, S. J., Kollman, P. A., Case, D., Singh, U. C., Alagona, G., Profeta, S. and Weiner, P. (1984) *J. Am. Chem. Soc.* **106**, 765–784
Weiner, S. J., Kollman, P. A., Nguyen, P. T. and Case, D. A. (1986) *J. Compt. Chem.* **7**, 230–252

Wheeler, P. R. and Ratledge, C. (1994) in *Tuberculosis: Pathogenesis, Protection and Control* (Bloom, B. R., ed.), pp. 353–385, Am. Soc. Microbiol., Washington, DC
Winkelman, G. (ed.) (1991) *Handbook of Microbial Iron Chelates*, CRC Press, Boca Raton, FL

Received 24 June 1994/3 August 1994; accepted 8 August 1994