Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

The structure of Rph, an exoribonuclease from *Bacillus anthracis*, at 1.7 Å resolution

Andrea E. Rawlings, Elena V. Blagova, Vladimir M. Levdikov, Mark J. Fogg, Keith S. Wilson and Anthony J. Wilkinson*

York Structural Biology Laboratory, Department of Chemistry, The University of York, York YO10 5YW, England

Correspondence e-mail: ajw@ysbl.york.ac.uk

Received 5 June 2008
Accepted 8 December 2008

PDB Reference: Rph, 3dd6, r3dd6sf.

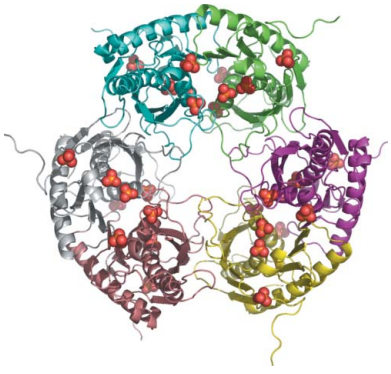
Maturation of tRNA precursors into functional tRNA molecules requires trimming of the primary transcript at both the 5′ and 3′ ends. Cleavage of nucleotides from the 3′ stem of tRNA precursors, releasing nucleotide diphosphates, is accomplished in *Bacillus* by a phosphate-dependent exoribonuclease, Rph. The crystal structure of this enzyme from *B. anthracis* has been solved by molecular replacement to a resolution of 1.7 Å and refined to an *R* factor of 19.3%. There is one molecule in the asymmetric unit; the crystal packing reveals the assembly of the protein into a hexamer arranged as a trimer of dimers. The structure shows two sulfate ions bound in the active-site pocket, probably mimicking the phosphate substrate and the phosphate of the 3′-terminal nucleotide of the tRNA precursor. Three other bound sulfate ions point to likely RNA-binding sites.

1. Introduction

Transfer RNA molecules are transcribed as precursors with additional nucleotides at both their 5′ and 3′ ends. These flanking sequences must be cleaved off to reveal the mature functional tRNA species. In *Bacillus*, as in most organisms, endonucleolytic cleavage of the 5′ extension is performed by RNase P, a heterotetrameric ribonucleoprotein complex (Stams *et al.*, 1998; Fang *et al.*, 2001). However, the manner in which the 3′ nucleotides are removed varies from species to species. In *Escherichia coli*, the most studied system, the 3′ flanking sequence is first trimmed by an endoribonuclease such as RNase E and this is followed by exoribonucleolytic degradation by one or more of a number of enzymes including RNase II, RNase BN, RNase T and RNase PH. In *Bacillus* just two enzymes are known to process the 3′ termini: an endoribonuclease called RNase Z and the exoribonuclease RNase PH (Pellegrini *et al.*, 2003; Wen *et al.*, 2005).

RNase PH is a member of the phosphate-dependent 3′→5′ exoribonuclease (PDX) family (Zuo & Deutscher, 2001; Mian, 1997). This family includes polynucleotide phosphorylase (PNPase) and a shared characteristic is the utilization of inorganic orthophosphate in the cleavage reactions (Deutscher *et al.*, 1988) to generate nucleoside diphosphates rather than the more common nucleoside monophosphates formed by the hydrolytic ribonucleases.

B. anthracis, the causative agent of anthrax, is a soil-dwelling Gram-positive endospore-forming bacterium. Its genome has been sequenced (Read *et al.*, 2003), revealing the presence of an open reading frame that shares a high degree of identity to the well characterized RNases PH from *B. subtilis*, *E. coli* and other bacteria. Here, we report the determination of the structure of Rph from *B. anthracis* Ames, a target in a high-throughput structural genomics study.



2. Materials and methods

2.1. Cloning, expression and purification

The *rph* coding sequence was amplified by PCR from *B. anthracis* Ames genomic DNA using BA4715F (5′-CACCACCACCACATG-CGAGTAGATGGTAGAGAGAAAA-3′) as a forward primer and BA4715R (5′-GAGGAGAAGGCGCGTTACTACTCTATATGAG-

ATACGATGTCACCTAAC-3′) as a reverse primer and cloned using a ligation-independent cloning method (Alzari *et al.*, 2006; Au *et al.*, 2006; Fogg & Wilkinson, 2008). The resulting fragment was treated with T4 DNA polymerase in the presence of dATP to generate single-stranded DNA overhangs complementary to those present on a modified pET28a vector cut with *Bse*RI and then similarly treated with T4 DNA polymerase in the presence of dTTP. The vector and PCR products were incubated together for 20 min to allow annealing of the single-stranded overhang regions (Aslanidis & de Jong, 1990) and the mixture was directly added to *E. coli* Novablue competent cells (Novagen), which were plated and grown overnight on antibiotic-containing media. Kanamycin-resistant colonies were picked the following day and used to inoculate overnight cultures, from which recombinant plasmids were purified. The resulting plasmid pETBaRph contains the coding sequence of *rph* fused to the coding sequence for an N-terminal MGSSHHHHHH tag. pETBaRph was introduced into *E. coli* BL21 (DE3) for overproduction of recombinant Rph.

N-terminally hexahistidine-tagged Rph was expressed from *E. coli* BL21 (DE3)/pETBaRph by overnight growth in kanamycin-supplemented autoinduction media (Studier, 2005) at 310 K. Cells from 0.5 l culture were harvested by centrifugation at 5000 rev min^{-1} for 15 min. Pelletted cells were resuspended in buffer *A* (20 m*M* Na_2HPO_4, 0.5 *M* NaCl, 10 m*M* imidazole pH 7.5) and lysed by sonication. Further centrifugation at 15 000 rev min^{-1} generated a cleared lysate which was loaded onto a 5 ml HiTrap nickel-chelation column, from which Rph was eluted following the application of a linear gradient of 10 to 500 m*M* imidazole in buffer *A*. Peak fractions were automatically directed onto a HiLoad 16/60 Superdex 200 prep-grade gel-filtration column (GE Healthcare) previously equilibrated with buffer *B* (50 m*M* Tris–HCl pH 7.5, 150 m*M* NaCl). The protein eluted with high purity and by comparison with the elution profiles of standard proteins the oligomerization state of Rph was determined to be hexameric. The 37.6 mg of purified protein obtained was concentrated to 20 mg ml^{-1} in buffer *B*, aliquoted and stored at 193 K.

2.2. Crystallization

A Mosquito nanolitre pipetting robot was used to screen Rph against different crystallization conditions using sitting-drop vapour diffusion in a 96-well plate format. 150 nl drops were equilibrated against 80 µl reservoir solution at 293 K. Small crystals were obtained in the Clear Strategy Screen (Brzozowski & Walton, 2001) with a reservoir solution comprising 2.7 *M* ammonium sulfate and 0.1 *M* bis-tris pH 5.5. These conditions were altered systematically, including a change to hanging-drop vapour diffusion in 24-well plates in order to optimize the crystal quality. Large single crystals appeared in 2 µl drops made up from 1 µl 2 *M* ammonium sulfate and 0.2 *M* potassium/sodium tartrate and 1 µl protein solution at a concentration of 20 mg ml^{-1}.

2.3. Data collection and processing

A single crystal was mounted in a cryo-loop (Hampton) and flash-cooled in liquid nitrogen following brief immersion in a cryoprotectant composed of reservoir solution supplemented with 30%(*v*/*v*) glycerol.

Diffraction data were collected on beamline ID14.3 at the European Synchrotron Radiation Facility (ESRF). The crystal was maintained at a constant 100 K throughout the experiment and was exposed to radiation of wavelength 0.931 Å. Images were collected on an ADS Q4R detector. Indexing of the diffraction patterns with *DENZO* from the *HKL*-2000 program suite (Otwinowski & Minor, 1997) showed that the crystal belonged to space group *R*32, with unit-cell parameters $a = b = 86.87$, $c = 179.31$ Å. The diffraction images revealed that the low-resolution spots were surrounded by regions of diffuse scattering arising from a semi-ordered water–ice structure.

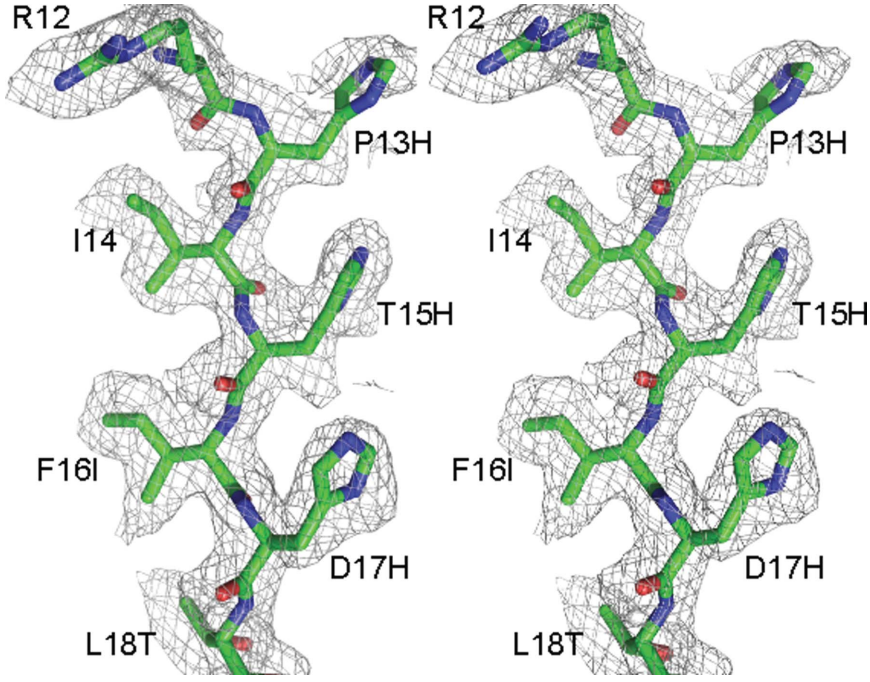


Figure 1
Stereoview of residues 12–18 of strand β1 shown in stick form and coloured by atom type. The electron density, shown as a grey mesh at a contour level of 1.0σ, is from an $F_{obs} - F_{calc}$ OMIT map. The map was generated by deleting residues 12–18 from the molecular-replacement model after carrying out TLS/rigid-body refinement. Ten cycles of *REFMAC*5 refinement were carried out prior to calculation of the map. This map is displayed on the final refined *B. subtilis* Rph model. Residues are labelled with the changes in sequence from the *B. subtilis* search model to the *B. anthracis* structure indicated. This stereo figure was generated with *PyMOL* (DeLano, 2008).

Integration and scaling of the data were performed using *XDS* (Kabsch, 1988) and *SCALA*, respectively (Collaborative Computational Project, Number 4, 1994). The data statistics are presented in Table 1.

2.4. Structure solution and refinement

MOLREP (Vagin & Teplyakov, 1997; Murshudov *et al.*, 1997) was used to solve the structure by molecular replacement using a single subunit of the RNase PH structure from *B. subtilis* (which shares 69% sequence identity with Rph; PDB code 1oyr; Harlow *et al.*, 2004) as the search model. After performing rotation and translation functions, a clear solution was found with an initial *R* factor of 49.1%; the next best solution had an *R* factor of 63.4%. The solution shows a single molecule in the asymmetric unit, indicating a solvent content of 48.8%. Rigid-body and TLS refinement were performed using *REFMAC* (Murshudov *et al.*, 1997). At this stage, changes in sequence between the *B. subtilis* Rph search model and the *B. anthracis* Rph structure were apparent in the electron-density maps (Fig. 1). Structure refinement continued with successive rounds of model building in *Coot* (Emsley & Cowtan, 2004) followed by *REFMAC* refinement. The final *R* factor for the structure consisting of 2090 atoms and 168 water molecules was 19.26% ($R_{\rm free}$ = 24.07%). Protein-structure and ligand analysis was performed using the *PDBSUM* server (Laskowski, 2001).

3. Results and discussion

The structure of Rph from *B. anthracis* was solved to 1.7 Å resolution with a single protein molecule in the asymmetric unit. The protein chain assumes the $\beta\alpha\beta\alpha$ topology (Fig. 2*a*), as observed in other RNase PH structures (Harlow *et al.*, 2004), consisting of nine β-strands and five α-helices. It can be thought of as a double sandwich with strands 1–5 and 6–9 surrounding helices 1–3 and with helices 1–3 and 4–5 surrounding strands 6–9. Analysis of the molecular packing revealed Rph to be a hexamer (Fig. 2*b*) arranged as a trimer of dimers as indicated by the crystal symmetry. The β9 strands from a pair of

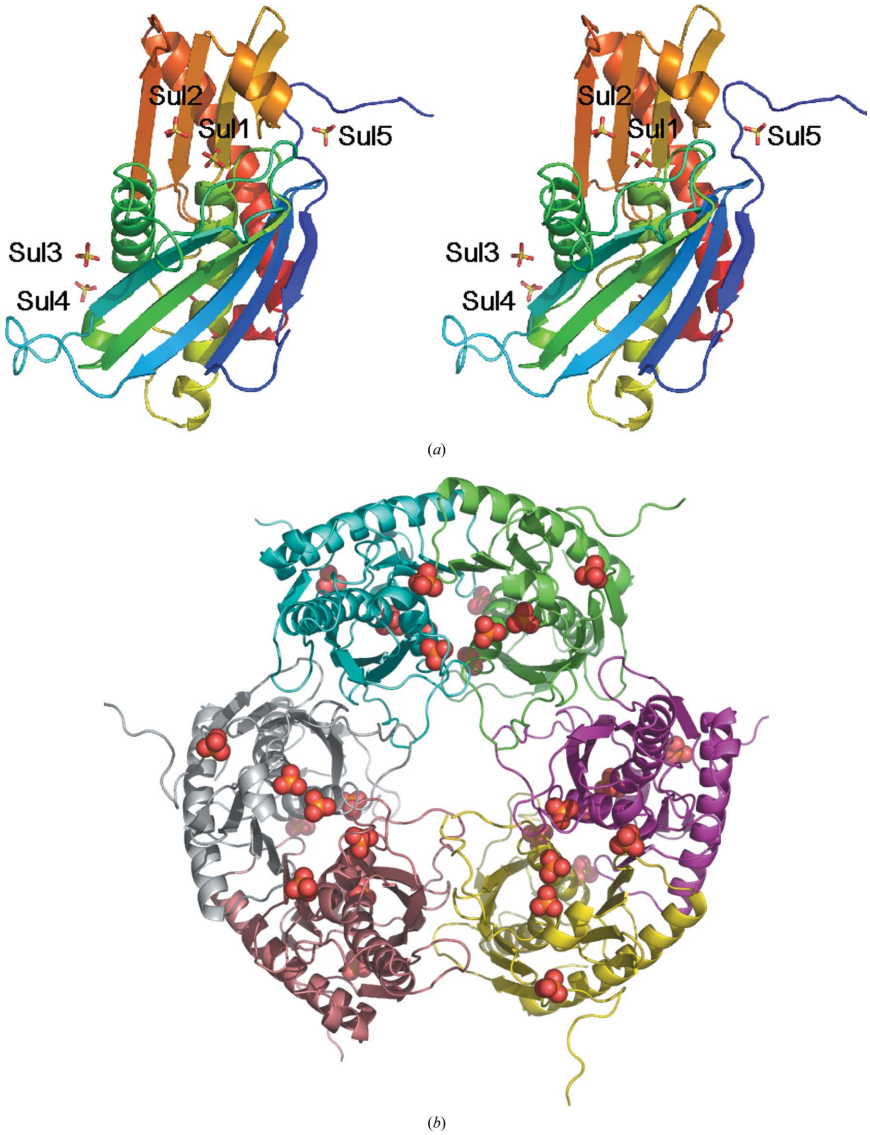


Figure 2
(*a*) Stereo ribbon diagram of Rph from *B. anthracis* in which the chain is colour-ramped from deep blue at the N-terminus to red at the C-terminus. Bound sulfates are shown in stick format and are coloured by atom type: sulfur, orange; oxygen, red. The labels correspond to those in Table 2. (*b*) Ribbon diagram showing the hexamer of Rph viewed along the threefold axis and coloured by subunit, with the bound sulfates shown as spheres. Both figures were generated with *PyMOL* (DeLano, 2008)

Table 1
X-ray data and refinement statistics.

Values in parentheses are for the highest resolution shell (1.79–1.70 Å).

Data collection	
Space group	*R*32
Unit-cell parameters (Å, °)	a = 86.75, b = 86.75, c = 179.08, α = 90.0, β = 90.0, γ = 120.0
Resolution range (Å)	43.36–1.70
No. of unique reflections	28380
Completeness (%)	98.7 (95.6)
Redundancy	7.3 (5.7)
$R_{\rm merge}$†	0.040 (0.38)
$I/\sigma(I)$	25.2 (4.8)
Refinement	
Resolution range (Å)	35.09–1.70
R factor‡	0.193
$R_{\rm free}$§	0.241
No. of subunits in ASU	1
No. of protein non-H atoms	1901
No. of water molecules	168
No. of sulfates	5
No. of TLS groups	6
Average *B* factors (Å^2)	
Solvent	11.2
Main chain	7.9
Side chain	10.3
R.m.s.d bond lengths¶ (Å)	0.022
R.m.s.d bond angles¶ (°)	1.980
Ramachandran plot	
Most favoured (%)	92.6
Additionally allowed (%)	7.4

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity measurement of a reflection with indices hkl and $\langle I(hkl)\rangle$ is the statistically weighted average reflection intensity. ‡ *R* factor = $\sum(F_{\rm o} - F_{\rm c})/\sum F_{\rm o}$, where $F_{\rm o}$ and $F_{\rm c}$ are the observed and calculated structure-factor amplitudes, respectively. § $R_{\rm free}$ is the *R* factor calculated with 5% of the reflections chosen at random and omitted from refinement. ¶ Root-mean-square deviation of bond lengths and bond angles from ideal geometry.

subunits come together so that an eight-stranded intersubunit β-sheet is formed featuring strands 6–9 from the two adjacent subunits. The dimer has a buried surface area of 1523 Å^2, 10% of which is accounted for by Arg212 present on strand β9, which also makes four hydrogen bonds across the interface. The trimer is formed predominantly through contacts between the loops linking strands 1–5, with 1465 Å^2 buried at this surface. His23, Arg68, Arg76, Arg73, Asp115, Asp117 and Gln120, all of which are conserved residues, collectively contribute over 40% of this surface area and 70% of the hydrogen-bond interactions.

The loop region between strand β5 and helix α2, which is found at the centre of the hexamer, was poorly defined in the electron-density maps; as a result, the lysine at position 80 was omitted from the refined model. It has been suggested that this loop, with its arginine-rich character, is prone to proteolytic degradation and that the hexamer may be a storage mechanism to protect the loops (Harlow *et al.*, 2004).

Rph from *B. anthracis* shares a high degree of sequence similarity to its orthologues from *B. subtilis* and *E. coli*, with 69% and 56% identity, respectively (Fig. 3). The five sulfate ions bound to *B. anthracis* Rph (Table 2, Figs. 2*a* and 4) make contacts with residues that are conserved, possibly indicating a role as tRNA phosphate mimics; two of them are located in the active-site pocket (Fig. 5). The first sulfate (Sul1), a likely mimic of the phosphate substrate, makes contacts with the highly conserved residues Arg86, Gly124, Thr125 and Arg126 (situated on helices 1 and 3), which have been shown to play important roles in the phosphorolytic reaction. This sulfate is located in a position equivalent to that of the bound tungstate in PNPase from *Streptomyces antibioticus* (Symmons *et al.*, 2000), the phosphate bound in the structure of Rph from *Aquifex aeolicus* (Ishii

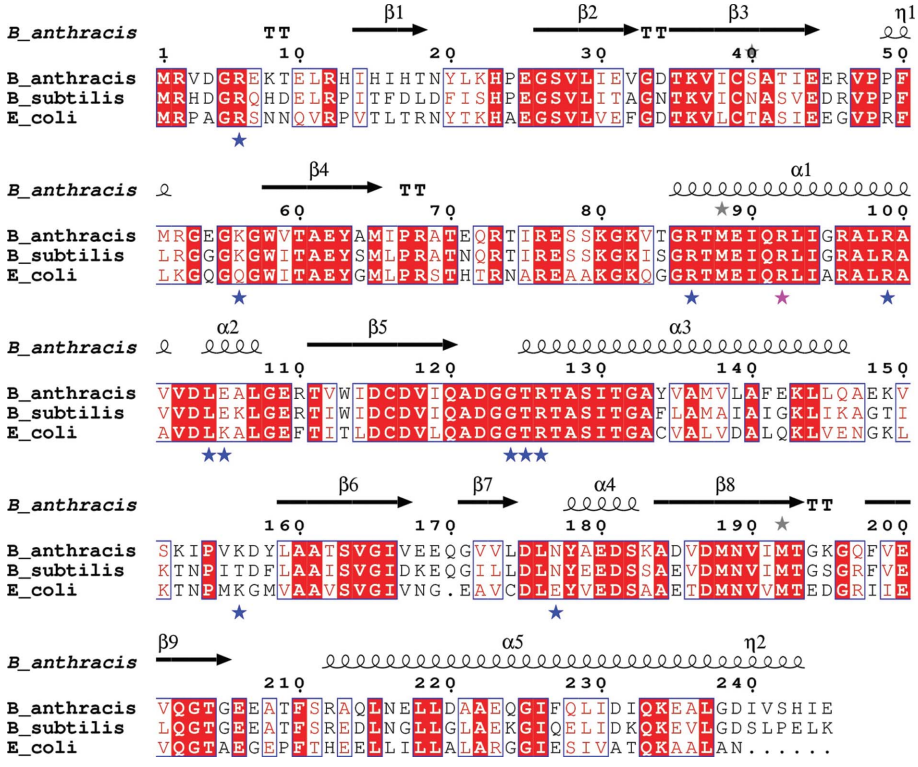


Figure 3
Sequence alignment of Rph from *B. anthracis* (Q81LA9), *B. subtilis* (P28619) and *E. coli* (P03842) made using *ClustalW* (Thompson *et al.*, 1994) and formatted using *ESPript* (Gouet *et al.*, 1999). The UniProt accession numbers of the proteins are given in parentheses. Red boxes represent conserved residues. Blue stars show residues that interact with the bound sulfate ions. Arg92 is highlighted with a magenta star to show that this residue makes contacts with Sul2 of the neighbouring subunit. The secondary-structure elements of Rph from *B. anthracis* are depicted above the alignment. Structural elements are labelled as follows: α, α-helices; β, β-strands; η, 3_{10}-helices; TT, turns.

Table 2
A list of the bound sulfate ions and interacting residues present in the *B. anthracis* Rph structure.

Sulfate identifier	Sulfate interaction with Rph
Sul1	Arg86 NH2, Gly124 N, Thr125 OG1 and N, Arg126 N, NE and NH2
Sul2	Arg86 NE, Arg126 NH2, Arg92† NH1 and NH2
Sul3	Arg99 NH1
Sul4	Lys56 NZ, Leu104 N, Glu105 N, Lys156 NZ
Sul5	Arg6 NE and NH2, Asn177 ND

† Arg92 is from an adjacent dimer-forming subunit.

et al., 2003) and the sulfate found in the active site of RNase PH from *B. subtilis* (Harlow *et al.*, 2004). The second sulfate ion (Sul2) is bound to residues Arg86 and Arg126, as well as Arg92 from the neighbouring dimer-forming subunit. It is probable that this sulfate mimics the phosphate of the 3′-terminal phosphodiester bond found on the tRNA substrate. The presence of Sul2 in the active site may help to order the Arg86 side chain, which was poorly defined in the *B. subtilis* structure that featured just a single sulfate ion. The S atoms of Sul1 and Sul2 are 6.3 Å apart (Fig. 5). It appears that Arg86 and Arg126, which make contacts with both sulfates, ensure that the substrate tRNA and the phosphate nucleophile assume the appropriate juxtaposition within the active site. Attack of the nucleophilic oxygen of the phosphate at the phosphorus centre of the 3′-terminal nucleotide, thought to progress by an S_N2-type mechanism, would result in the transient formation of a pentacoordinate phosphorus species (Mildvan, 1997). The conserved arginine residues in the active site may help to stabilize this negatively charged transition state.

The third sulfate ion (Sul3), which is in contact with Arg99, has also been observed in the structures of RNase PH from *B. subtilis* and *A. aeolicus*, where additional interactions with Trp58 and Thr60 are made (Harlow *et al.*, 2004; Ishii *et al.*, 2003). It is possible that the dimerization of two Rph subunits allows the insertion of the 3′ terminus of the tRNA precursor into the active-site pocket of molecule 1, with phosphates present elsewhere in the substrate interacting with residues 99, 58 and 60 of molecule 2. Sulfate 4 is hydrogen bonded to Arg73, a conserved residue from the motif RX_4RX_2R beginning at residue 68 and located at the trimer interface. A fifth sulfate (Sul5) forms hydrogen bonds to Lys156, Leu104, Val102 and Lys56 and is also observed in the structure of RNase PH from *A. aeolicus* (Ishii *et al.*, 2003).

The conserved N-terminal motif RX_3RX_5R beginning at Arg2 (Zuo & Deutscher, 2001), which in this structure is associated with the binding of a fifth sulfate (Sul5) at Arg6, is unusually followed by three regularly spaced histidines at positions 13, 15 and 17 (Fig. 1. If the RX_3RX_5R sequence is involved in interactions with the RNA, as has been proposed for the corresponding motif in PNPase (Symmons *et al.*, 2002), then the histidines may form a base-stacking interaction with RNA bases, although these histidine residues are not strongly conserved in other members of the PDX family.

In conclusion, among the RNase PH structures, the structure of Rph from *B. anthracis* is distinct in exhibiting two sulfate ligands

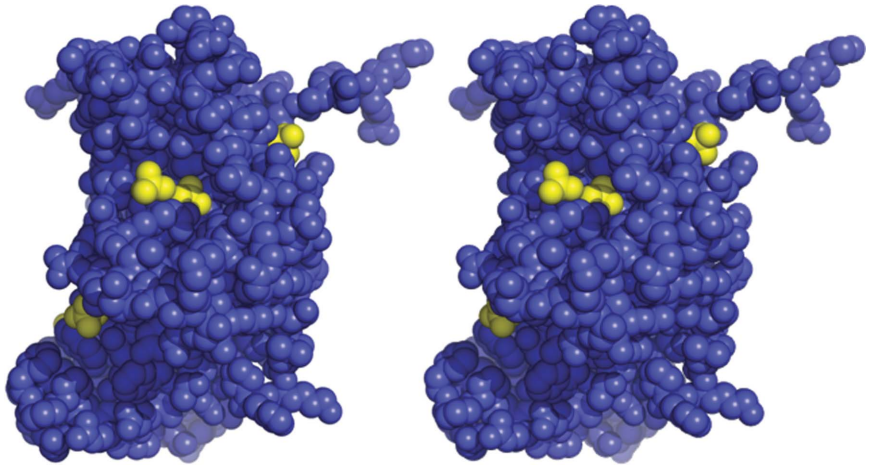


Figure 4
Stereo space-filling diagram of a single subunit of Rph in blue, with bound sulfates in yellow. The Rph molecule is oriented as in Fig. 2(*a*). This figure was generated with *PyMOL* (DeLano, 2008).

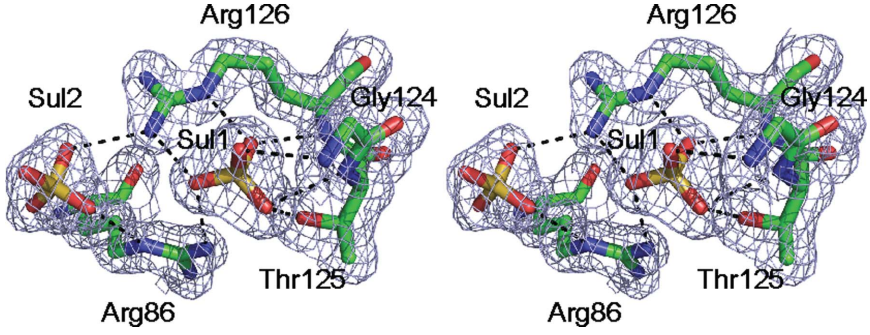


Figure 5
Stereoview of the electron density in the vicinity of the active site. Sulfates (Sul) 1 and 2 and the interacting residues Arg86, Gly124, Thr125 and Arg126 are shown in stick form and coloured according to element. The electron-density map is shown in grey at a contour level of 1.0σ. Hydrogen bonds between the sulfates and the protein are denoted by dashed black lines. This figure was generated with *PyMOL* (DeLano, 2008).

within the active site, possibly mimicking the 3′-terminal phosphate of both the tRNA substrate and the phosphate nucleophile. Other bound sulfates indicate likely tRNA-binding sites. The conserved motif $RX_5RX_3RX_2R$ beginning at Arg86 and found in helix $\alpha1$ makes contacts with Sul1, Sul2 and Sul3 as well as Sul2 of the adjacent subunit through interaction with residue Arg92. This suggests that the dimer may be required for catalysis.

The work described here was funded by the European Commission as SPINE contract No. QLG2-CT-2002-00988 under the RTD programme 'Quality of Life and Management of Living Resources'. AER was funded by the BBSRC and VL by the Wellcome Trust. The authors would like to thank Sam Hart and Dr Tracey Gloster of the York Structural Biology Laboratory for their help with data collection at the ESRF and Dr Garib Murshudov for useful advice and assistance with model refinement.

References

Alzari, P. M. *et al.* (2006). *Acta Cryst.* D**62**, 1103–1113.
Aslanidis, C. & de Jong, P. J. (1990). *Nucleic Acids Res.* **18**, 6069–6074.
Au, K. *et al.* (2006). *Acta Cryst.* D**62**, 1267–1275.
Brzozowski, A. M. & Walton, J. (2001). *J. Appl. Cryst.* **34**, 97–101.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
DeLano, W. L. (2008). *The PyMOL Molecular Graphics System.* http://www.pymol.org.
Deutscher, M. P., Marshall, G. T. & Cudny, H. (1988). *Proc. Natl Acad. Sci. USA*, **85**, 4710–4714.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Fang, X. W., Yang, X. J., Littrell, K., Niranjanakumari, S., Thiyagarajan, P., Fierke, C. A., Sosnick, T. R. & Pan, T. (2001). *RNA*, **7**, 233–241.
Fogg, M. J. & Wilkinson, A. J. (2008). *Biochem. Soc. Trans.* **36**, 771–775.
Gouet, P., Courcelle, E., Stuart, D. I. & Métoz, F. (1999). *Bioinformatics*, **15**, 305–308.
Harlow, L. S., Kadziola, A., Jensen, K. F. & Larsen, S. (2004). *Protein Sci.* **13**, 668–677.
Ishii, R., Nureki, O. & Yokoyama, S. (2003). *J. Biol. Chem.* **278**, 32397–32404.
Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 916–924.
Laskowski, R. A. (2001). *Nucleic Acids Res.* **29**, 221–222.
Mian, I. S. (1997). *Nucleic Acids Res.* **25**, 3187–3195.
Mildvan, A. S. (1997). *Proteins*, **29**, 401–416.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pellegrini, O., Nezzar, J., Marchfelder, A., Putzer, H. & Condon, C. (2003). *EMBO J.* **22**, 4534–4543.
Read, T. D. *et al.* (2003). *Nature (London)*, **423**, 81–86.
Stams, T., Niranjanakumari, S., Fierke, C. A. & Christianson, D. W. (1998). *Science*, **280**, 752–755.
Studier, F. W. (2005). *Protein Expr. Purif.* **41**, 207–234.
Symmons, M. F., Jones, G. H. & Luisi, B. F. (2000). *Structure*, **8**, 1215–1226.
Symmons, M. F., Williams, M. G., Luisi, B. F., Jones, G. H. & Carpousis, A. J. (2002). *Trends Biochem. Sci.* **27**, 11–18.
Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Wen, T., Oussenko, I. A., Pellegrini, O., Bechhofer, D. H. & Condon, C. (2005). *Nucleic Acids Res.* **33**, 3636–3643.
Zuo, Y. & Deutscher, M. P. (2001). *Nucleic Acids Res.* **29**, 1017–1026.