

Birkbeck ePrints: an open access repository of the research output of Birkbeck College

<http://eprints.bbk.ac.uk>

Cohen-Gonsaud, Martin; Barthe, Philippe; Bagn ris, Clair; Henderson, Brian; Ward, John; Roumestand, Christian and Keep, Nicholas H. (2005) The structure of a resuscitation-promoting factor domain from *Mycobacterium tuberculosis* shows homology to lysozymes. **Nature Structural and Molecular Biology** 12 (3): 270-273

This is an author-produced version of a paper published in *Nature Structural and Molecular Biology* (ISSN 1545-9993). This version has been peer-reviewed but does not include the final publisher proof corrections, published layout or pagination.

All articles available through Birkbeck ePrints are protected by intellectual property law, including copyright law. Any use made of the contents should comply with the relevant law.

Citation for this version:

Cohen-Gonsaud, Martin; Barthe, Philippe; Bagn ris, Clair; Henderson, Brian; Ward, John; Roumestand, Christian and Keep, Nicholas H. (2005) The structure of a resuscitation-promoting factor domain from *Mycobacterium tuberculosis* shows homology to lysozymes. *London: Birkbeck ePrints*. Available at: <http://eprints.bbk.ac.uk/archive/00000343>

Citation for the publisher's version:

Cohen-Gonsaud, Martin; Barthe, Philippe; Bagn ris, Clair; Henderson, Brian; Ward, John; Roumestand, Christian and Keep, Nicholas H. (2005) The structure of a resuscitation-promoting factor domain from *Mycobacterium tuberculosis* shows homology to lysozymes. **Nature Structural and Molecular Biology** 12 (3): 270-273

<http://eprints.bbk.ac.uk>

Contact Birkbeck ePrints at lib-eprints@bbk.ac.uk

RESEARCH ARTICLE:

The structure of a resuscitation promoting factor domain from *Mycobacterium tuberculosis* shows homology to lysozymes

Martin Cohen-Gonsaud¹, Philippe Barthe², Claire Bagn ris¹, Brian Henderson³, John Ward⁴, Christian Roumestand² and Nicholas H. Keep¹.

¹ School of Crystallography and Institute for Structural Molecular Biology, Birkbeck College, University of London, Malet Street, London, WC1E 7HX, UK.

² Centre de Biochimie Structurale, CNRS UMR5048 - INSERM UMR554 - UMI, Facult  de Pharmacie, BP 14491, 15 Avenue Charles Flahault, 34093 Montpellier Cedex 5, France.

³ Division of Microbial Diseases, Eastman Dental Institute, University College London, 256 Gray's Inn Road, London WC1X 8LD, UK.

⁴ Department of Biochemistry and Molecular Biology, University College London, Gower St, London WC1E 6BT, UK.

Correspondence should be addressed to N.H.K. (n.keep@mail.cryst.bbk.ac.uk)

Keywords: RPF, NMR, Resuscitation promoting factors, *Mycobacterium tuberculosis*, dormant culture, bacterial cytokine.

Resuscitation-promoting factor (RPF) proteins reactivate stationary-phase cultures of GC-rich Gram-positive bacteria including the causative agent of tuberculosis, *Mycobacterium tuberculosis*. We report the solution structure of the RPF domain from *M. tuberculosis* Rv1009 (RpfB) solved by heteronuclear multidimensional NMR. Structural homology with various glycoside hydrolases, and biochemical studies including requirement for the conserved active site glutamate, indicate that oligosaccharide cleavage is likely to be the signal for revival from dormancy.

One of the keys to *Mycobacterium tuberculosis*' success as a pathogen is its ability to persist in its host organism in a latent state for years after the first phase of infection¹. The World Health Organization estimates that one third of the world's population harbour a latent tuberculosis infection. The reactivation of the bacteria and, as a consequence, the development of clinical tuberculosis occurs in 5% to 10% of non-HIV infected people. The reactivation rate is much higher in HIV patients (up to 30%), with tuberculosis accounting for 13% of HIV-related deaths (WHO fact sheet N^o.104; <http://www.who.int/mediacentre/factsheets/fs104/en/>). Actively dividing *M. tuberculosis* infection can be cleared by a combination of drug therapy and the patient's immune response, but once the dormant state has been entered the bacteria are no longer cleared. Understanding and controlling the entry and exit from dormancy is therefore very important in the development of novel anti-tubercular therapies.

A secreted protein from *Micrococcus luteus*, that decreases the growth lag time and increases the number of colony forming units when added to dormant cultures of this organism, has been purified and its gene sequenced. This 17 Kda protein, called resuscitation-promoting factor (RPF), is active at very low concentrations (pM)². Homologues of this Rpf are found in several high GC Gram-positive bacteria including *M. tuberculosis*, which has five such genes (*rpfA-E*) that contain a region homologous to the *M. luteus* Rpf. The growth-stimulating activity of the five tuberculosis RPFs has been demonstrated on dormant cultures of *Mycobacterium bovis* BCG³. It was proposed that these proteins were bacterial 'cytokines' that signal the bacteria to exit dormancy by binding to a receptor. By determining the molecular structure of an RPF, we aim to gain insight into the mechanism of action of these important molecules.

RESULTS

NMR Structure of an active RPF domain

Recently, we published a proposal, based on bioinformatic analysis, that the common core domain of all the RPF proteins is structurally related to the c-type lysozyme family (CAZy family: GH22) and the active site glutamate is conserved⁴. We have sub-cloned, over-expressed and purified the C-terminal 108 amino acid core domain of Rv1009 (RpfB), called RpfBc. This domain corresponds to the Rpf homologous region found in all RPFs (residues 280 to 362) and an additional 25 N-terminal residues⁵. In our analysis the conserved sequence region was structurally related to the residues between helices $\alpha 2$ and $\alpha 5$ of the c-type lysozymes but enough residues to form an equivalent to helix $\alpha 1$ were included.

Firstly, the ability of the RpfBc region to resuscitate dormant culture of *M. luteus* was tested and showed that this protein was active at picomolar concentrations

(data not shown). The RPF regions from Rv1884c (RpfC) and Rv2389c (RpfD) were also tested and showed a similar high activity. These experiments formally demonstrate that the conserved domain of the RPF is responsible for the resuscitation of the bacteria, and is fully active as an isolated domain, in agreement with experiments on the conserved region of the *M. luteus* RPF⁶.

Resonance assignments of RpfBc were carried out with a series of standard 2 and 3D experiments using ¹⁵N- and ¹⁵N,¹³C-labelled protein⁵. The first 22 residues from the N-terminal segment (Asn255–Val276) exhibited negative ratio values in the ¹H-¹⁵N heteronuclear NOE experiment corresponding to highly mobile residues⁵. This demonstrates that there is no equivalent in RpfBc of helix α 1 of the c-type lysozyme. For the remaining residues structures were calculated compatible with 1613 distance restraints. The ensemble of 30 best structures shows an RMSD for the main chain of 0.57Å (**Fig. 1a**). Experimental details, results and statistics for the structure determination are described in **Table 1**.

RpfBc resembles several glycosides hydrolases

The experimentally determined structure of RpfBc is most similar to c-type lysozymes with the top hit in FSSP⁷ PDB3LZT having a Z score of 6.4 (RMSD 2.9 Å over 79 residues). The oligosaccharide binding site cleft of c-type lysozymes is well defined (e.g. PDB1LZS⁸). An equivalent cleft is present in RpfBc and we show below that this does indeed bind polysaccharide.

The peptidoglycan binding cleft can be thought of as being made by two sides (**Fig. 2a**). The first side is formed predominantly by the two largest helices of the c-type lysozyme fold (helix α 2 and α 4). The RpfBc helices α 1 and α 4 match these two helices well. Consequently, the catalytic glutamate found at the C-terminal end of the first matching helix is at the same position in those two proteins (**Fig. 2a**). The α 5 helices in the two structures are also equivalent but differ slightly in relative orientation to the larger helices, probably due to the sole disulfide bridge in RpfBc (conserved in all RPFs) being at a different point between helices α 1 and α 5 compared to the disulfide between the equivalent helices α 2 and α 5 in the c-type lysozymes. In RpfBc, there is no equivalent of helices α 1 and α 6 in the c-type lysozyme fold, which lie away from the cleft. Consequently, the RPF fold is the most compact lysozyme-like structure determined.

The second side of the substrate cleft differs between the RPFs and c-type lysozymes. The equivalent region to the section forming the two β -strands in c-type lysozyme is four residues shorter in the RPF fold so although the backbone near the cleft superimposes well, the strands are not present (**Fig. 2a**). The two structures then diverge the most, with helix α 2 in RpfBc not matching the c-type lysozyme fold (**Fig. 2a**). However, this side of the cleft closely resembles the bacterial Soluble Lytic Transglycosidase (SLT) proteins^{9,10} (CAZy family: GH23) particularly Slt70 (PDB1QSA; **Fig. 2b**) as Slt35 (PDB1D0K) has a 45 residue helical insertion between the helices equivalent to α 1 and α 2 of RpfBc. After the GXXQ turn (conserved among RPFs, α -lactalbumins, SLTs, c-type and g-type lysozymes) RpfBc and the SLTs possess an additional helix (α 2 in RpfBc; **Fig. 2b**). At this position in the c-type lysozymes, the O4-hydroxyl of the N-acetylglucosamine (NAG) in position +2 form a hydrogen bond with the indole part of a tryptophan (e.g. Trp316 in PDB1LZS; **Fig. 2a**). In Slt35 and Slt70, this role is played respectively by a serine or a threonine, conserved among the various SLTs. The superposition of RpfBc with the SLTs shows

the presence of a threonine, conserved among the RPFs, at this position (Thr315; **Fig. 2b**).

Oligosaccharide binds to the RPF domain

Using ^1H - ^{15}N -HSQC NMR experiments we tested the binding of RpfBc to NAG, and its trimer, N,N',N''-triacylchitotriose. NAG is one of the two sugars that makes up the polysaccharide of peptidoglycan and polymers of NAG have been bound to lysozyme in a number of crystal structures. With monomeric NAG no chemical shift changes were observed, while in the trimer case, chemical shift changes were observed for residues of the putative binding groove (**Fig. 1b** and **Supplementary Fig. 1**). All the residues we propose above as involved in the binding are affected by the addition of N,N',N''-triacylchitotriose (*e.g.* Glu292, Thr315). The main perturbation is observed for the loop $\alpha 4/\alpha 5$, with the Gln347 residue being the most affected. The equivalent conserved glutamine in the lysozyme (Gln104) is a key residue in binding oligosaccharide and forms an H-bond between the NAG +2 and +3. This loop moves the most between apo and holo crystal structures of c-type lysozyme (**Supplementary Fig. 2**), suggesting that the same binding mechanism is present in the c-type lysozyme and the RPFs.

Active site glutamate is required for *in vivo* activity

Furthermore, to test the importance of the putative catalytic glutamate of the RPF, we produced a mutant protein RpfC-E80A and compared the resuscitation activity with the wild-type recombinant RpfC. CD spectra show that both the wild type and the mutant protein are folded very similarly (**Supplementary Fig. 3**). The experiment shows a complete suppression of the resuscitation activity for the mutant protein, highlighting the critical role of the conserved glutamate residue (**Fig. 3**). However the RPFs we have expressed and described here do not degrade peptidoglycan to the point where the bacteria are lysed in assay conditions where lysozyme completely lysed the cells, nor does it have significant activity in a commercial fluorescence assay for lysozyme (EnzCheck[®] Lysozyme Assay Kit - Molecular Probes; data not shown).

DISCUSSION

The structure of the functional RPF domain is therefore a compact hybrid of the SLT and c-type lysozyme folds, both of which cleave peptidoglycan. Although the overall sequence conservation between the RPFs and either of these protein families is low, a few crucial features such as the catalytic glutamate are conserved. We have demonstrated that this residue is critical for the resuscitation activity. The need for the active site glutamate, combined with a clear binding pocket for a large molecule that undergoes rearrangement in presence of a polysaccharide, strongly suggests that the function of RPFs is to cleave oligosaccharide and this is responsible for the role in revival of dormant bacteria. However the RPFs we have tried do not degrade peptidoglycan in the same way as lysozyme to degrade the cell wall entirely. This does not preclude an oligosaccharide cleavage being the mechanism of action. A second enzyme might be required to release the product of the cleavage from peptidoglycan; a second factor has been shown to act synergistically with RPF in *C. glutamicum*¹¹. Alternatively, the true substrate may be a modified form of

peptidoglycan¹² only found in dormant cells or in a limited region related to cell septation. Our results cannot rule out a model in which RPF may simply bind to an oligosaccharide in a lectin-like manner. However, the data suggests that RPF most likely acts enzymatically on the cell wall of the bacteria. The identification of the physiological substrates for the RPFs coupled with structural and biochemical analysis will improve our understanding of their functional activity and will be a first step in the development of inhibitors to understand and prevent the reactivation of pathogenic mycobacteria such as *Mycobacterium tuberculosis* or *M. leprae*.

METHODS

NMR structure determination. NMR spectra were recorded on samples of 0.5 mM, ¹⁵N or ¹⁵N, ¹³C labelled RpfBc domain dissolved in 25 mM Na-acetate pH 4.6, 2 mM β-mercaptoethanol with 5–10% D₂O for the lock. Some spectra were also collected on sample prepared in D₂O with the same buffer. NMR experiments were carried out at 20°C on Bruker AVANCE 500 and 600 spectrometers equipped with 5 mm z-shielded gradient ¹H-¹³C-¹⁵N triple resonance cryogenic probes. NMR data processing and analysis were carried out using GIFA¹³ and NMRVIEW¹⁴. Nucleus resonances (¹H, ¹⁵N and ¹³C) were assigned using standard three-dimensional (3D) experiments⁵. The backbone dihedral angle restraints were obtained from chemical shift index analysis using the TALOS program¹⁵. The two cysteine residues have up-field-shifted ¹³Cβ resonances (38.33 and 38.16 ppm, respectively) consistent with a disulfide bond formation¹⁶ which was enforced for structure calculations. Hydrogen bond restraints were incorporated if they were protected from H₂O-D₂O exchange as revealed from an HSQC spectrum recorded on a freshly lyophilised sample dissolved in a D₂O buffer, and were supported by NOE analysis. NMR-derived restraints are shown in Table 1. Structure calculations were performed with ARIA 1.2 (Ref. 17) implemented in CNS 1.1(Ref. 18). The 30 structures with the lowest energy were further refined by molecular dynamics calculations in explicit solvent to remove artefacts and analysed with PROCHECK-NMR¹⁹. Analysis of the Ramachandran plot showed that all modelled residues were found in allowed regions with 90.2% in the most favoured region. The r.m.s. deviations and visual display were performed with MOLMOL²⁰.

Chemical shift mapping. RpfBc ¹⁵N and ¹H_N backbone chemical shift perturbations (Δδ) were measured from HSQC experiments upon titration with N,N',N''-triacetylchitotriose (Tri-NAG) (Sigma). In these experiments, the concentration of ¹⁵N-RpfBc was kept constant (50 μM) and the concentration of unlabeled Tri-NAG was varied from 0.5 mM to 20 mM. An additional reference spectrum was performed on a RpfBc sample (50 μM) without Tri-NAG. All ¹H-¹⁵N HSQC spectra were recorded using a time domain data size of 64 *t1* X 512 *t2* complex points, and 32 transients per *t1* increment. For analysis, ¹⁵N and H_N chemical shift changes were combined using the equation:

$$\Delta\delta = \sqrt{((\Delta\delta_H)^2 + (\Delta\delta_N/10)^2)}$$

where values of Δδ above 0.85 p.p.m. have been related to significant perturbations.

Assay of Rpf activity. *M. luteus* NCIMB 13267 was passaged on Nutrient Agar (Oxoid) plates. Overnight cultures of *M. luteus* in lactate minimal media (LMM) were grown to an OD₆₀₀ of 1.8-2.2. The cells were removed by centrifugation and the supernatant was filter sterilised, stored at 4 °C and used within 2 hours of preparation as a source of endogenous *M. luteus* Rpf. For assay of Rpf activity an overnight

culture of *M. luteus* in Nutrient Broth-2 (Oxoid) was diluted 1:10³ in LMM and added to Universals in 4ml aliquots. 40 µl of suitable dilutions of recombinant Rpf proteins were added to the Universals and bacteria were cultured overnight on a shaking incubator at 37 °C. The growth of the cultures were monitored by the absorbance at 600 nm and compared to a negative control of *M. luteus* on its own and a positive control of *M. luteus* with sterile culture supernatant containing the endogenous *M. luteus* Rpf. The data presented in **Figure 3** is the growth after 4 days (average of duplicate cultures) with error bars of twice the Standard Error. The results were confirmed on a second separate growth assay.

CD spectra. Circular dichroism spectra of unmutated and E80A recombinant Rpf were recorded on an AVIV 202 spectrometer at 0.25 mg ml⁻¹ in 20mM sodium phosphate pH 7.0. Protein concentrations were determined using calculated extinction coefficients. All data are normalized to MRE (Mean Residue Ellipticity) using the program CDTTool²¹.

Coordinates. The chemical shifts of the RpfBc domain and the atomic coordinates have been respectively deposited in the BioMagResBank (accession number BMRB-6221) and in the Protein Data Bank (accession number PDB1XSF).

ACKNOWLEDGEMENTS

This work was supported by EU Grant QLK2–2001–02018 (M.C-G. and N.H.K.). NMR experiments were recorded and analysed using the facilities of the Structural Biology platform RIO (C.B.S. – Montpellier – France). We thank Bob Sarra for help with CD.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

REFERENCES

1. Tufariello, J.M., Chan, J. & Flynn, J.L. Latent tuberculosis: mechanisms of host and bacillus that contribute to persistent infection. *Lancet Infect. Dis.* **3**, 578–590 (2003).
2. Mukamolova, G.V., Kaprelyants, A.S., Young, D.I., Young, M. & Kell, D.B. A bacterial cytokine. *Proc. Natl. Acad. Sci. U S A* **95**, 8916–8921 (1998).
3. Mukamolova, G.V. *et al.* A family of autocrine growth factors in Mycobacterium tuberculosis. *Mol. Microbiol.* **46**, 623–635 (2002).
4. Cohen–Gonsaud, M. *et al.* Resuscitation-promoting factors possess a lysozyme-like domain. *Trends in Biochemical Sciences* **29**, 7–10 (2004).
5. Cohen-Gonsaud, M. *et al.* 1H, 15N and 13C chemical shift assignments of the Resuscitation Promoting Factor domain of Rv1009 from Mycobacterium tuberculosis. *J. Biomol. NMR* **30**, 373–374 (2004).
6. Mukamolova, G.V. *et al.* The rpf gene of Micrococcus luteus encodes an essential secreted growth factor. *Mol. Microbiol.* **46**, 611–621 (2002).
7. Holm, L. & Sander, C. Touring protein fold space with Dali/FSSP. *Nucleic Acids Res.* **26**, 316–319 (1998).

8. Song, H., Inaka, K., Maenaka, K. & Matsushima, M. Structural changes of active site cleft and different saccharide binding modes in human lysozyme co-crystallized with hexa-N-acetyl-chitohexaose at pH 4.0. *J. Mol. Biol.* **244**, 522–540 (1994).
9. van Asselt, E.J. *et al.* Crystal structure of Escherichia coli lytic transglycosylase Slt35 reveals a lysozyme-like catalytic domain with an EF-hand. *Structure Fold. Des.* **7**, 1167–1180 (1999).
10. van Asselt, E.J., Thunnissen, A.M. & Dijkstra, B.W. High resolution crystal structures of the Escherichia coli lytic transglycosylase Slt70 and its complex with a peptidoglycan fragment. *J. Mol. Biol.* **291**, 877–898 (1999).
11. Hartmann, M. *et al.* The glycosylated cell surface protein Rpf2, containing a resuscitation-promoting factor motif, is involved in intercellular communication of Corynebacterium glutamicum. *Arch. Microbiol.* **182**, 299–312. (2004).
12. Raymond J.B., Mahapatra S., Crick D.C. & Pavelka M.S. Jr. Identification of the *namH* gene, encoding the hydroxylase responsible for the N-glycolylation of the mycobacterial peptidoglycan. *J. Biol. Chem.* **280**, 326–333 (2005).
13. Pons, J.L., Malliavin, T.E. & Delsuc, M.A. Gifa V 4: A complete package for NMR data set processing. *J. Biomol. NMR* **8**, 445–452 (1996).
14. Johnson, B.A. & R.A., B. NMRView: A computer program for the visualization and analysis of NMR data. *J. Biomol. NMR* **4**, 603–614 (1994).
15. Cornilescu, G., Delaglio, F. & Bax, A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR* **13**, 289–302 (1999).
16. Sharma, D. & Rajarathnam, K. ¹³C NMR chemical shifts can predict disulfide bond formation. *J. Biomol. NMR* **18**, 165–171 (2000).
17. Linge, J.P., O'Donoghue, S.I. & Nilges, M. Automated assignment of ambiguous nuclear Overhauser effects with ARIA. *Methods Enzymol.* **339**, 71–90 (2001).
18. Brunger, A.T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905–921 (1998).
19. Laskowski, R.A., MacArthur, M.W., Hutchinson, E.G. & Thornton, J.M. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst* **26**, 283–291 (1993).
20. Koradi, R., Billeter, M. & Wuthrich, K. MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51–55 (1996).
21. Lees, J.G., Smith, B., Wien, F., Miles, A., and Wallace, B.A. CDTool – an integrated software package for circular dichroism spectroscopic data processing, analysis, and archiving. *Anal. Biochem.* **332**, 285–289 (2004).

FIGURE LEGENDS

Figure 1 Solution structure and oligosaccharide binding to RpfBc. **(a)** Sausage representation of 30 NMR structures of RpfBc. (residues: Thr274–Arg362). The backbone thickness is directly proportional to the r.m.s. deviation of the ensemble. The sausage is coloured by secondary structure, with helices in red, coil and turn in grey and disulfide bond in orange. **(b)** Chemical shift mapping of RpfBc after addition of (NAG)₃. Ribbon representation of the RpfBc structure with (NAG)₄ superimposed based on a structural superposition with c-type lysozyme-(NAG)₄ complex (PDB1LMQ). Residues whose combined weighted N and NH chemical shift change on addition of (NAG)₃ exceeds 1.0 p.p.m. are coloured in red (*e.g.* Gln347) and those whose chemical shift are between 0.85 and 1.0 p.p.m. are coloured in orange (*e.g.* Glu315 and Thr317). Residues which disappear from the HSQC due to chemical exchange line broadening are in purple. This figure was generated with the program MOLMOL²⁰.

Figure 2 Comparison of RpfBc with other glycoside hydrolases. **(a)** Structure superimposition in the cleft area of the RpfBc (in red) and a c-type lysozyme (PDB1LMQ; in blue). The catalytic glutamate at the C-terminal end of the first helix is at the same position in RpfBc and lysozyme, Glu292 and Glu35 respectively. **(b)** Structure superimposition in the cleft area of the RpfBc (in red), Slt35 (PDB1D0K; grey) and Slt70 (PDB1QSA; green). The catalytic glutamate is at the same position for RpfBc, Slt35 and Slt70, Glu292, Glu162 and Glu478 respectively. At the equivalent position of the threonine Thr315 in RpfBc, a serine is present in Slt35 (Ser230) and a threonine in Slt70 (Thr501). These residues bind the O4-hydroxyl of the NAG in position +2 and played the same role as one of the conserved tryptophans Trp316 in PDB1LMQ. This figure was generated with the program PyMol (<http://pymol.sourceforge.net>).

Figure 3 Resuscitation activity assays of recombinant *M. tuberculosis* RpfC and RpfC-E80A. Only 100 pM unmutated RpfC and the endogenous *M. luteus* Rpf control show enhanced resuscitation. Error bars represent 2*standard error of duplicate samples.

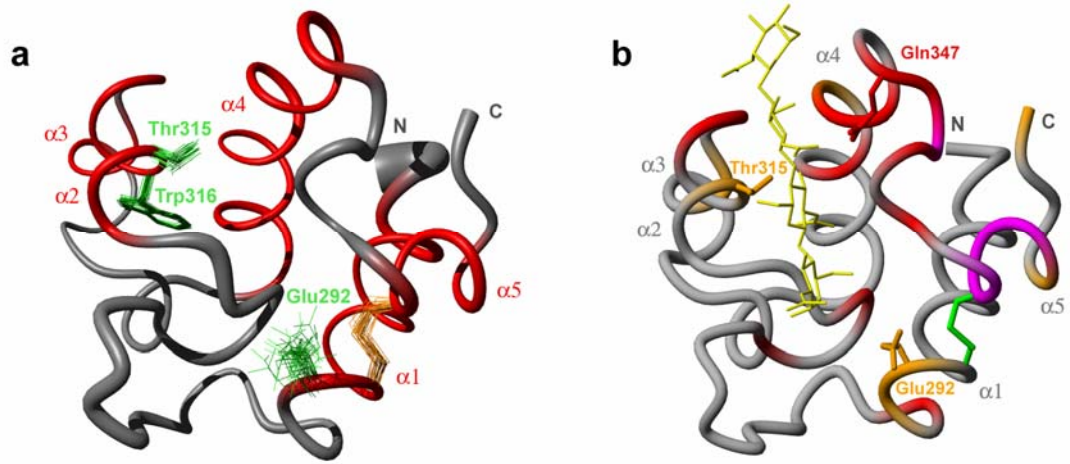


Figure 1

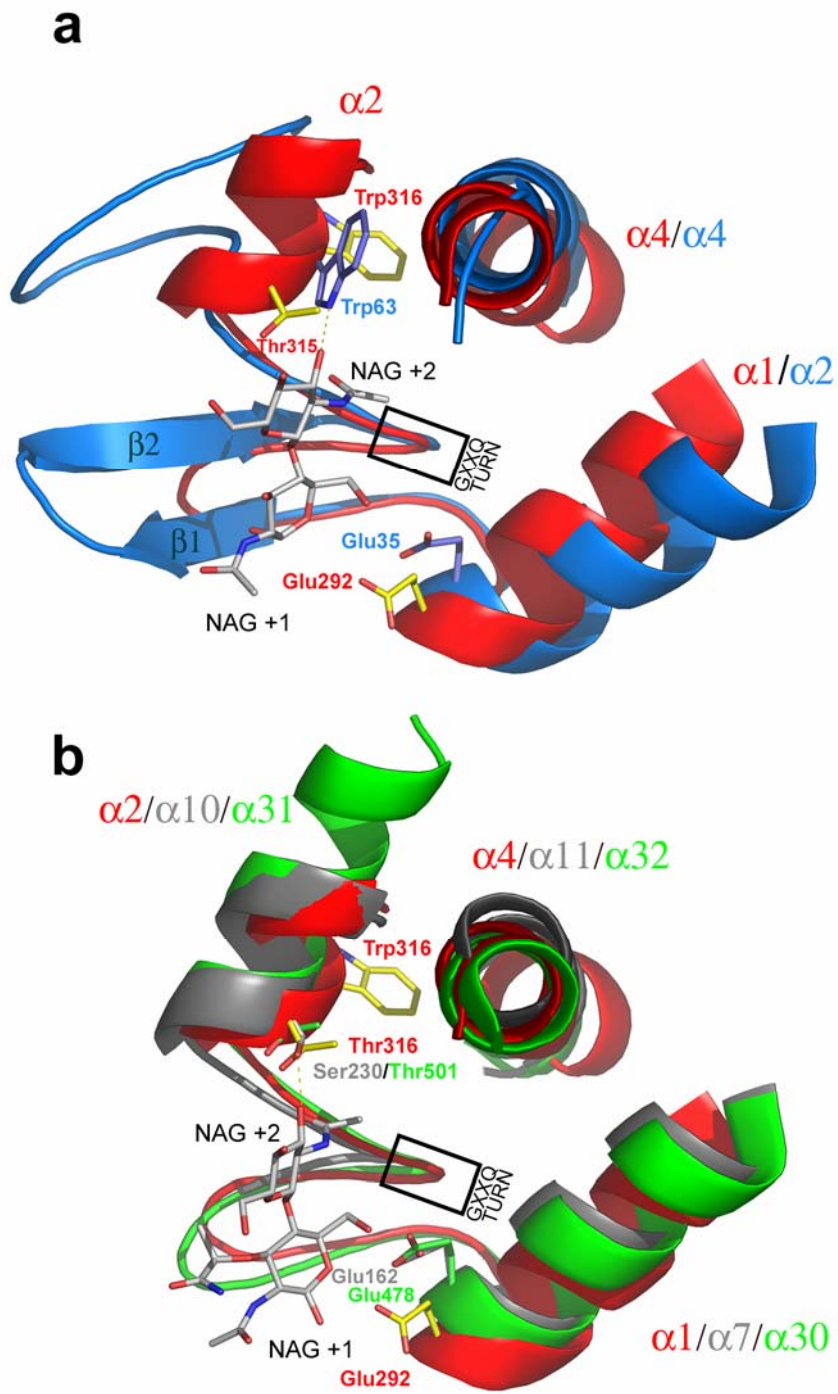


Figure2

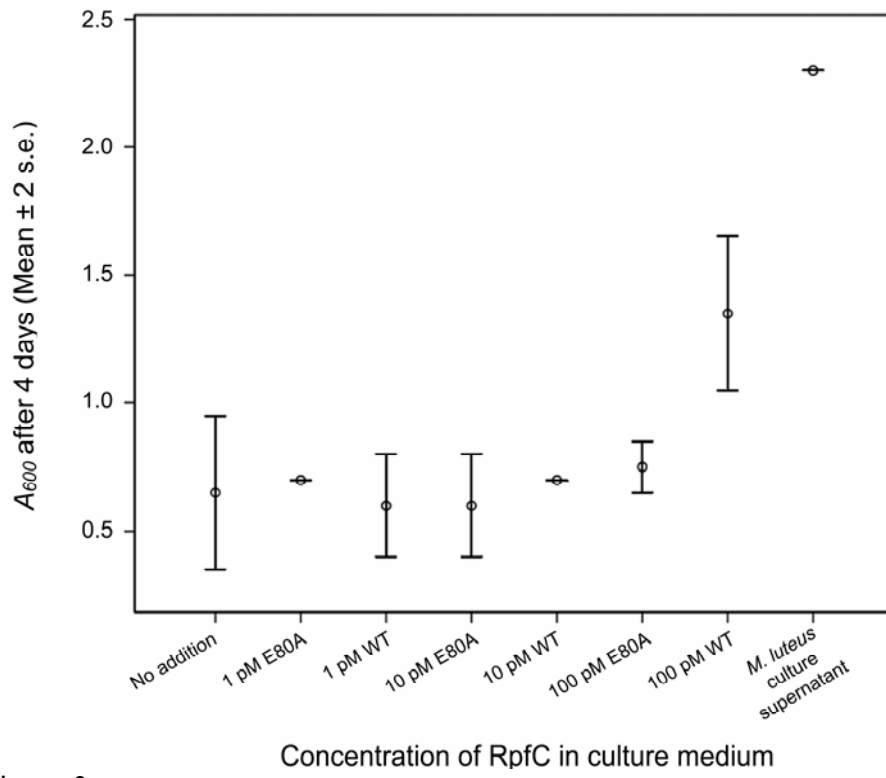
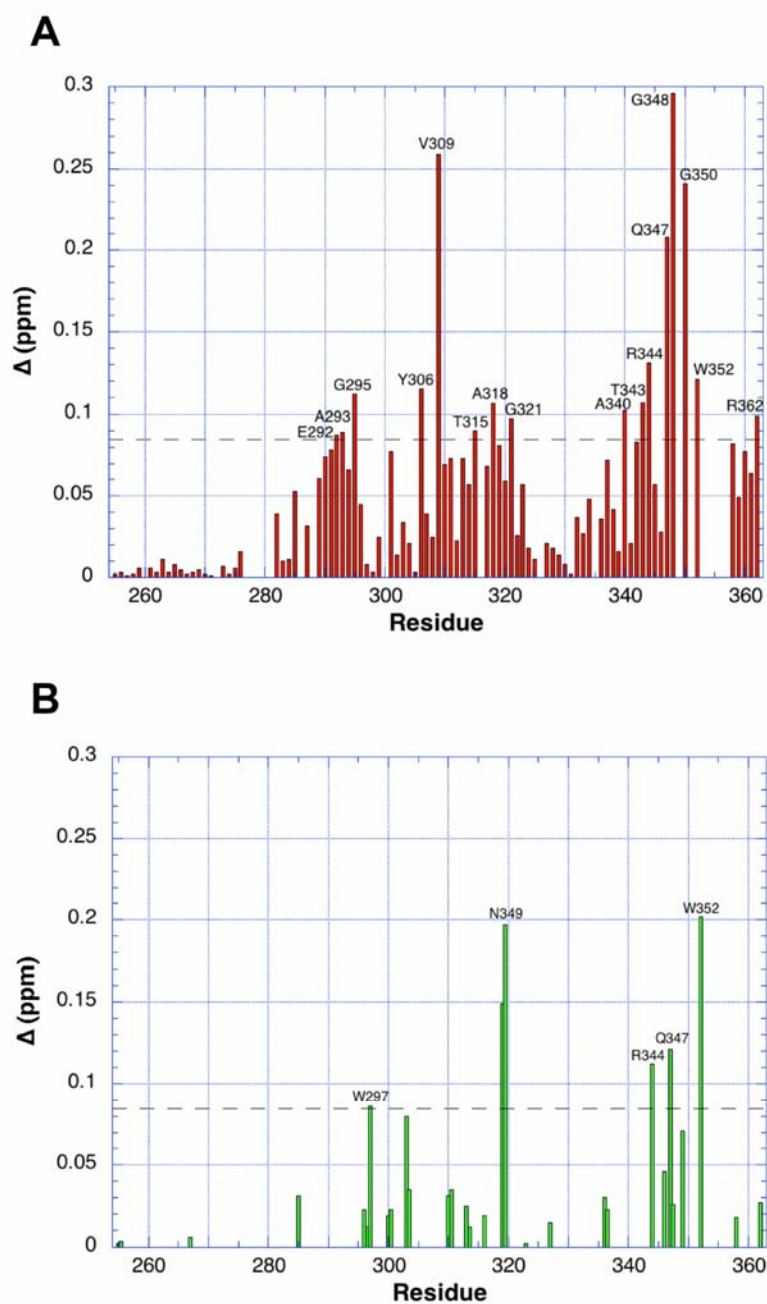


Figure 3

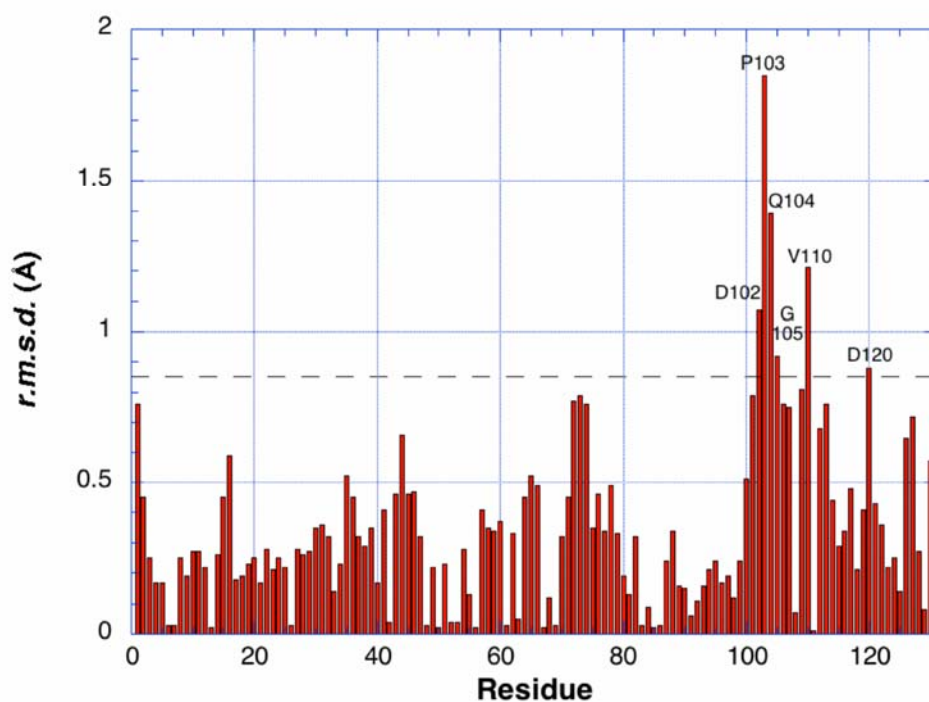
Table 1 NMR and refinement statistics for RpfBc

	Protein
NMR distance & dihedral constraints	
Distance constraints	
Total NOE	1613
Intra-residue	629
Inter-residue	
Sequential ($ i-j = 1$)	432
Medium-range ($ i-j < 4$)	276
Long-range ($ i-j > 5$)	276
Intermolecular	-
Hydrogen bonds	10
Total dihedral angle restraints	
phi	62
psi	62
Structure Statistics	
Violations (mean and s.d.)	
Distance constraints (Å)	0.11 ± 0.09
Dihedral angle constraints (°)	1.21 ± 0.17
Max. dihedral angle violation (°)	3.75 ± 0.44
Max. distance constraint violation (Å)	0.36 ± 0.09
Deviations from idealized geometry	
Bond lengths (Å)	0.01 ± 0.00
Bond angles (°)	0.68 ± 0.02
Impropers (°)	1.78 ± 0.10
Average pairwise r.m.s.d.** (Å)	
Heavy	1.07 ± 0.14
Backbone	0.57 ± 0.11

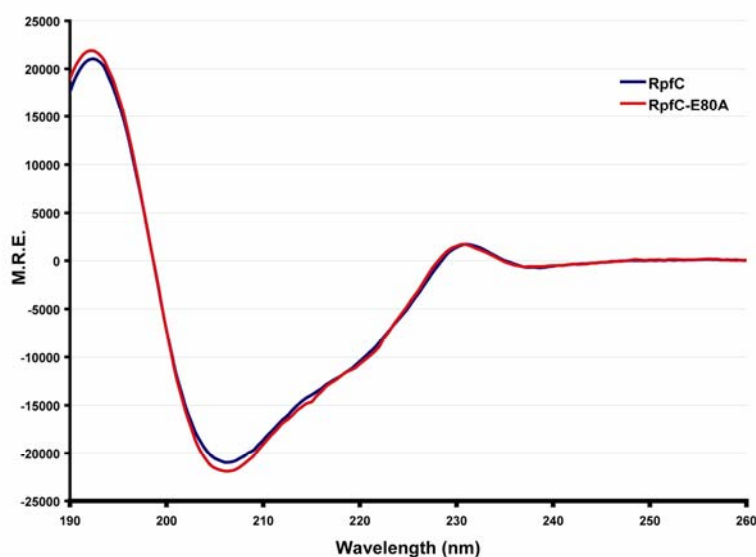
**Pairwise r.m.s.d. was calculated among 30 refined structures for residues 25-105.



Supplementary Figure 1: Weighted chemical shift change plotted against residue number. **A:** Graph of the NH backbone chemical shift. **B:** Graph of the NH side-chain chemical shift.



Supplementary Figure 2: Graph of residue root mean square deviation comparing crystal structures of human c-type lysozyme with and without (NAG)₆ (PDB1LZS and PDB1JSF respectively). The region showing the most movement is structurally equivalent to the region of RpfC in the neighbourhood of Gln347 that shows the biggest chemical shift change on addition of (NAG)₃.



Supplementary Figure 3: CD spectra of recombinant RpfC and RpfC-E80A. Spectra show that both proteins are folded and very similar in structure.