SUPPLEMENTAL MATERIALS for the paper

"Profound asymmetry in the structure of cAMP-free CRP from *Mycobacterium tuberculosis*" by Gallagher et al.

These Supplemental Materials comprise four parts, as follows:

1. Supplemental Table 1, with information on previously reported, homologous structures.

2. Two pages text with information on Materials and Methods re protein production, etc.

3. Supp. Tables 2, 3, and 4, with information on crystals, diffraction, and structure refinement.

4. Supplemental Figures S1, S2, S3, and S4.

Supp. Table 1. Structures in the Protein Data Bank that are close homologs of MtbCRP.

Gene name/ family	***source(s)***	***PDB ids***	***seq. ident. w/ MtbCRP, %***	***comments***
crp	*M. tuberculosis*	3d0s	100	this study
crp	*E. coli*	1g6n, 1hw5, 1i5z, 2gzw, etc. 1cgp, 2cgp, 1zrf, etc.	30	+cAMP +cAMP+DNA
crp/fnr family	*S. coelicolor*	2pqq	45	N-domain only
crp/fnr family	*P. aeruginosa*	2oz6	28	+cAMP
crp/fnr family	*P. gingivalis*	2gau	28	no ligand
sdrP	*T. thermophilus*	2zcw	27	no ligand
cprK	*D. hafniense*	2h6b, 2h6c	22	off state
cooA	*R. rubrum,* *C. hydrogenoformans*	1ft9. 2fmy, 2hkx	22	off state
prfA	*L. monocytogenes*	1omi, 2beo, 2bgc	17	no ligand

Supplement to Materials and Methods

Reagents: The *E. coli* strain BL21(DE3) and the selenomethionine auxotroph strain B834(DE3) were purchased from EMD Biosciences (Madison, WI). The M9 salts growth medium (Cat. No. MD045003) for the incorporation of selenomethionine into CRP was purchased from Medicilon (Chicago, IL). The pET15b expression vector for the N-terminal histidine tagged fusion with a target protein was obtained from EMD Biosciences. The relevant strains for cloning and protein production are *E. coli* Novablue (K12) and BL21(DE3), respectively. Other reagent purchases were from sources as in (Reddy, 2001).

DNA Procedures: The Rv3676 gene coding sequence for CRP was amplified by PCR using the forward primer:

5'GGAATTC**CATATG**GACGAGATCCTGGCCAGGGCAG (with *NdeI* restriction recognition sequence in bold letters) and the reverse primer: 5'CCG**CTCGAG**TTACCTCGCTCGGCGGGCCAGTCTTTC (with *XhoI* restriction recognition sequence in bold letters). Amplification conditions were: 96 ^{o}C for 60 s for initial melting of DNA, followed by 30 cycles of amplification with each cycle consisting of melting at 96 ^{o}C for 45 s, annealing at 60 ^{o}C for 45 s, and polymerization at 72 ^{o}C for 60 s. Polymerization was continued at the end for 10 min at 72 ^{o}C. Two hundred nanograms of *M. tuberculosis* $H_{37}R_{v}$ chromosomal DNA and 100 ng of primers were used in the amplification. The amplified DNA was digested with *NdeI* and *XhoI* and cloned into the respective sites of pET15b plasmid. A recombinant was isolated from *E. coli* Novablue (K12). *E. coli* BL21 (DE3) was transformed with the recombinant plasmid pET15b:Rv3676 for protein production based on induction with IPTG.

Protein expression and purification: *E. coli* BL21(DE3) harboring pET15b:Rv3676 plasmid was grown at 37 ^{o}C in 1 L LB containing ampicillin (100 μg/mL) to an A_{600} of ~0.5. The culture was briefly cooled in ice, and CRP production was induced overnight with 30 μM IPTG at 25 ^{o}C. Cells were harvested, washed with 25 mM Tris-HCl, pH 7.5, and suspended in 40 mL of lysis buffer consisting of 50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 10 mM

mercaptoethanol. The cell suspension was passed through a French pressure cell twice at 10,000 psi and the cell-free extract was centrifuged at 48,000 x g for 1 hr. The supernatant was mixed with 1 mL of Ni-NTA agarose in a 250 mL bottle kept on a rocker for 1 hr at 4 ^{o}C. The resin-supernatant was centrifuged at 2,500 rpm for 10 min and most of the supernatant was poured out. The resin was transferred to a 10 mL column and washed with 40 mL of the lysis buffer. Washing was continued successively with 20 mL of the lysis buffer containing 10 mM and 20 mM imidazole. Some of the CRP was eluted in these two washings along with a lot of contaminating proteins. Next, CRP was eluted with four 5 mL successive washings with the lysis buffer containing 100 mM imdazole. CRP in the 100 mM imidazole eluate was nearly homogeneous as judged by SDS-PAGE. The CRP eluate was dialyzed overnight against 25 mM Tris-HCl, pH 8.0 containing 150 mM NaCl. We observed that removal of NaCl by dialysis resulted in the precipitation of CRP and therefore kept NaCl at 150 mM to keep the protein soluble. The vector encoded histidine tag was removed by digestion with thrombin (biotinylated) according to the manufacturer's recommendation. The histidine tag and the biotinylated thrombin were removed by successively passing through Ni-NTA agarose and streptavidine agarose, respectively. The CRP solution was then dialyzed against 25 mM Tris-HCl, pH 8.0, and 150 mM NaCl.

Production and purification of selenomethionine CRP: *E. coli* B834 (DE3), a methionine auxotroph, was transformed with the pET15b:Rv3676 recombinant plasmid. Incorporation of selenomethionine into CRP was performed using the M9 salts selenomethionine growth medium according to the manufacturer's recommendation. Briefly, cells were grown in 1L LB medium containing ampicillin (100 μg/mL) overnight at 37 ^{o}C. Cells were harvested, washed twice with sterile water, and suspended in 100 mL of M9 salts medium. Four 1L M9 salts media containing ampicillin were inoculated with 25 mL of the culture per 1 L. Cells were grown at 37 ^{o}C to A_{600}= 0.4. At this stage, selenomethionine was added and expression was induced with 30 μM IPTG at 37 ^{o}C overnight. Purification of selenomethionine CRP was essentially as described for native CRP.

Crystallization and diffraction: Despite extensive screening in hanging drops by vapor

diffusion using various commercial screens at both 4º and 24º C, the only crystals grew from Crystal Screen (Hampton Research) reagent #41. Crystallization drops produced heavy precipitate immediately on mixing, but became clear again overnight. Milder conditions that obviated the precipitation failed to produce crystals. Severe, systematic growth defects, probably linked to rotational disorder arising from the dimer's pseudosymmetric structure, required extensive optimization. It was found that crystal growth results were highly sensitive to the batch of PEG, and that the best crystals grew from conditions in which the isopropyl alcohol concentration was initially high (15%) and then lowered to 3% after 2 hours, suggesting separate processes of nucleation and growth with different alcohol dependencies. The conditions were optimized to 10% PEG 8000, 100 mM Na Hepes, pH 7.5 with the abovementioned timecourse of isopropyl alcohol from 15% to 3%. Crystals belong to space group $P2_12_12_1$ with lattice parameters 52.8 Å, 83.8 Å, 98.4 Å. This cell is similar to that reported in (Akif et al., Acta Cryst. F62, 873-875, 2006). The unit cell volume leads to the prediction of two subunits (one dimer) per asymmetric unit with $V_m = 3.0$ Å^3 Da^{-1}.

Supplemental Table 2. Crystal information.

Macromolecule details	
Database code(*s*)	UNP code: O69644 MYCTU
Mass (Da)	24791
Predicted pI	10.2
Source organism	*Mycobacterium tuberculosis* H_{37}Rv
Crystallization	
Crystallization method	Vapor diffusion, hanging drop
Temperature (K)	298

Macromolecule solution	20 mg ml^{-1}, 150 m*M* NaCl, 25 mM Na Tris, pH 8.0
Precipitant	l0% PEG 8K, 15% IPA, 0.1M Na Hepes, pH 7.5
Reservoir	l4% PEG 4K, 3% IPA, 0.1 M Na Hepes, pH 7.5
Cryo treatment	
Final cryoprotection solution	16% glycerol in reservoir
Soak time	3 s
Cooling	immersion in liquid N
Annealing	none
Crystal data	
Crystal shape	rectangular bars
Crystal size (mm)	0.03 × 0.08 × 0.2 mm
Matthews coefficient, *V*M ($Å^3$ Da^{-1})	2.17
Solvent content (%)	43.24
Unit-cell data	
Cell setting, space group	$P2_12_12_1$
a, *b*, *c* (Å)	52.80, 83.80, 98.39
α, β, γ (°)	90, 90, 90
No. of molecules in unit cell, *Z*	4 dimers

Supplemental Table 3. Data collection and structure determination statistics

Values for the outer shell are given in parentheses.

	Diffraction set 1	Diffraction set 2
Temperature (K)	100	100
Diffraction source	NSLS beamline X26C	APS beamline 24-ID-C
Diffraction protocol	Single wavelength	Single wavelength
Monochromator	Coated mirror	Coated mirror
Wavelength (Å)	0.98	1.0
Detector	ADSC CCD	ADSC CCD
No. of frames	360	240
Software	HKL	HKL
Reflection data		
Resolution range (Å)	30.00–2.30 (2.38–2.30)	30.00–2.00 (2.07–2.00)
No. of unique reflections	18644 (779)	29026 (2796)
Completeness (%)	87.6(39.7)	95.200 (93.7)
Redundancy	11.6 (2.2)	6.200 (6.30)

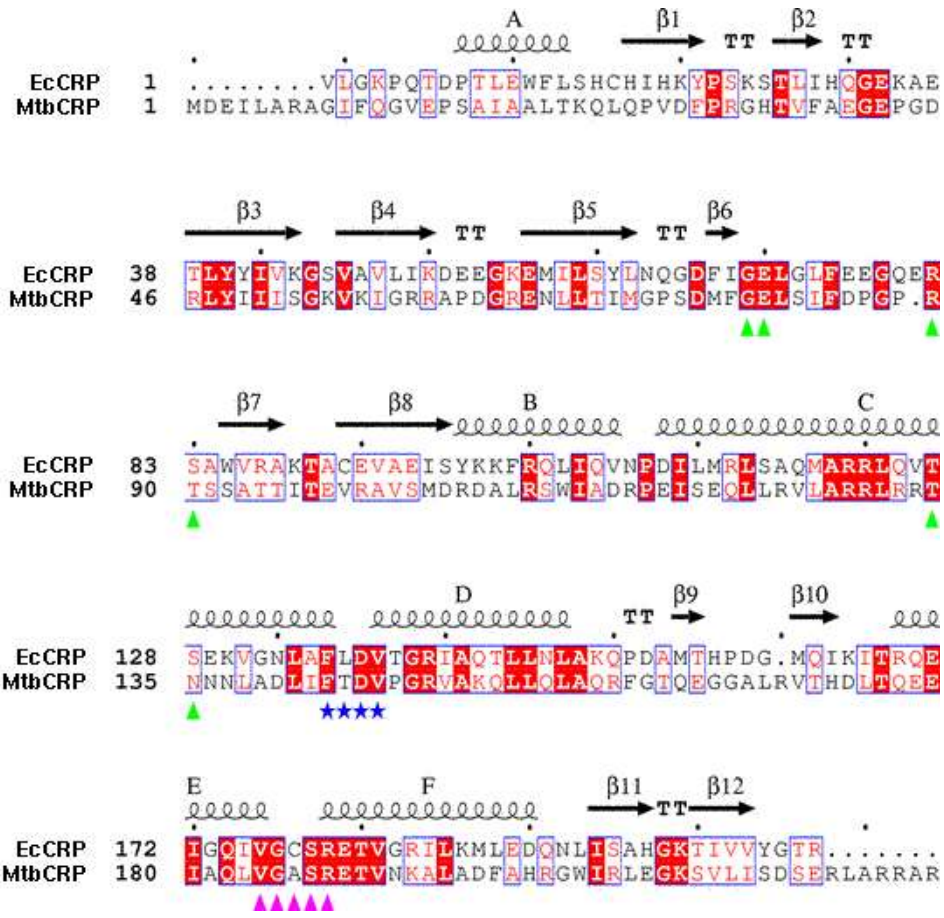
$I/\sigma(I)$	16.1(0.8)	13.9 (4.1)
*d*spacing at $I/\sigma(I)$= 2 (Å)	2.53	1.9
*R*merge	0.056(0.476)	0.078 (0.333)
Phasing method	SeMet SAD	-
Number of scatterers	8 Se/A.U. = 1 Se per 56 residues	
No. of sets used in phasing	1	
Phasing resolution range (Å)	28 - 2.3	
No. of refls used in phasing	18312	
Figure of merit	0.39	
Solution software	*PHENIX*	

Supplemental Table 4. Structure refinement and model validation

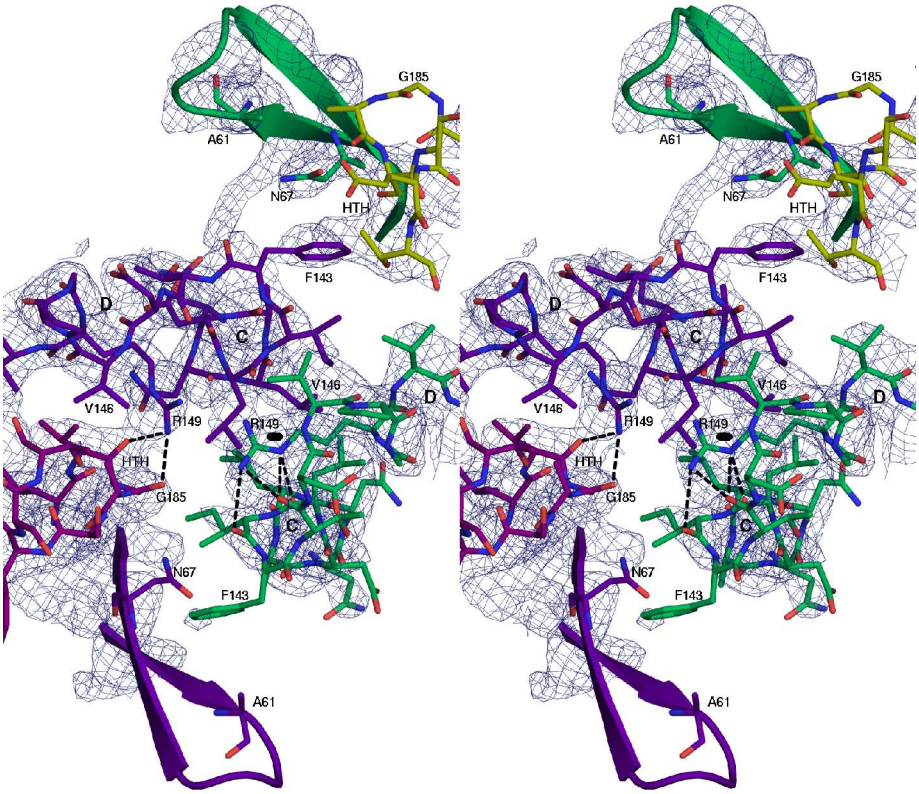
Values for the outer shell are given in parentheses.

Refinement software	*REFMAC*5.2.0019
Refinement on	F
σ cutoff	$F > 0.00\sigma(F)$
Resolution range (Å)	16.000–2.000 (2.051–2.000)
No. of reflections used in refinement	27234 (1913)
Overall average *B* factor ($Å^2$)	37.0
No. of protein atoms	3431
No. of ligand atoms	3
No. of solvent atoms	184
Total No. of atoms	3618
Final *R*work	0.210 (0.212)
No. of reflections for *R*free	1448 (107)
Final *R*free	0.284 (0.288)
Rmsd bond lengths (Å)	0.018
Rmsd bond angles (°)	1.67
Ramachandran data	0 residues in disallowed regions
Missing residues	B subunit 216-224
PDB code	3d0s

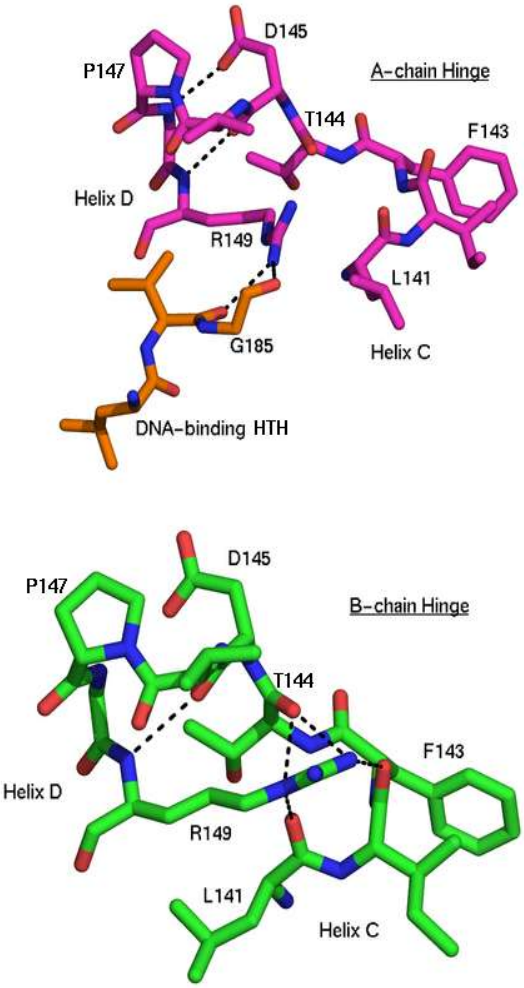
Supplemental Figure 1. Sequence alignment between EcCRP and MtbCRP. Secondary structure labels and cartoons (above sequences) are for EcCRP but correspond closely with MtbCRP, expecially after helix B. The dots mark every tenth residue for the Mtb sequence. Green triangles indicate the six residues that bind cAMP in the EcCRP structure 1ZRF.PDB. Blue stars mark the hinge, and magenta triangles mark the central turn of the HTH DNA-binding region. Note that the alignment is one-to-one, except that EcCRP has a 1-residue insertion E81 and MtbCRP has a 1-residue insertion R170.



Supplemental Figure 2. Stereo image of electron density showing asymmetry of the hinges. The view is down the molecular pseudodyad (black oval near center), from the DNA-binding region. The map shown is the original 2.3-Å Se-phased (thus unbiased by any model) SAD map, contoured at 1.4 sigma. Subunit A colors are purple for the hairpin (bottom) and hinge (upper left), and violet for the HTH (left). Subunit B colors are green for the hairpin (top) and hinge (lower right), and yellow-green for the HTH (upper right). The hairpins are shown as ribbons with most sidechains omitted for clarity and only residues A61 and N67 shown, as sticks. Note that the two hinges have fundamentally different geometries, giving different environments to many dyad-related residues; the two sidechains V146 and R149 are labeled in both hinges as examples. The two N67 sidechains are on opposite sides of the adjacent F143 ring, the two V146 sidesubunits are at very different distances from the dyad, and the two R149 guanidino groups are both to the left of the dyad. The different sets of H-bonds formed by the two R149s are shown as dashed lines. Also note that the two HTH elements are in different positions with respect to the dyad; the Cα - Cα distance between R149 and G185 of the HTH is 8.7 Å in the A subunit and is 18.9 Å in the B subunit. Image made using Pymol (27).



Supplemental Figure 3. Close up views of the MtbCRP hinges in each subunit, emphasizing the different interactions and conformations of R149, which differ at the chi-3 rotamer. It should be noted that the two hinges are actually adjacent in the structure (see also Figures 1,3,4 and S2). H-bonds under 3.2 Å are drawn as dashed lines. (In both subunits, D145 H-bonds with the amide of R149, capping helix D, but the H-bond distance is 3.3 Å in the B subunit). A, in subunit A, R149 has little contact with L141 and H-bonds to the carbonyl oxygens of V184 and G185 in the turn of the A subunit's HTH DNA-binding region (orange). This leaves its helix C capped only by waters (not shown). B) In the B subunit, R149 covers L141 and caps helix C with two H-bonds, along with H-bonds to the carbonyl of T144. Figure made using PyMol (27).



Supplemental Figure 4. Stereo C-alpha trace of the B-chain (green) superimposed on the A-chain (purple). The N-domains were superposed (least-squares superposition based on positions 27-57 and 70-137 had rmsd=2.0 Å for 99 atoms). Helix F is at top. The asterisk indicates the turn of the HTH motif. The hinge is visible just below helix F, at the top of the long vertical helix C. The N-domain ribbon projects outward at upper right. The hinge and the HTH regions are the most discrepant between the two chains.