PNAS
www.pnas.org

Supplementary Information for

Crystal structure of human mARC1 reveals its exceptional position among eukaryotic molybdenum enzymes

Christian Kubitza, Florian Bittner, Carsten Ginsel, Antje Havemeyer, Bernd Clement, Axel J. Scheidig

Axel J. Scheidig
Email: axel.scheidig@strubio.uni-kiel.de

Bernd Clement
Email: bclement@pharmazie.uni-kiel.de

This PDF file includes:

www.pnas.org/cgi/doi/10.1073/pnas.1808576115

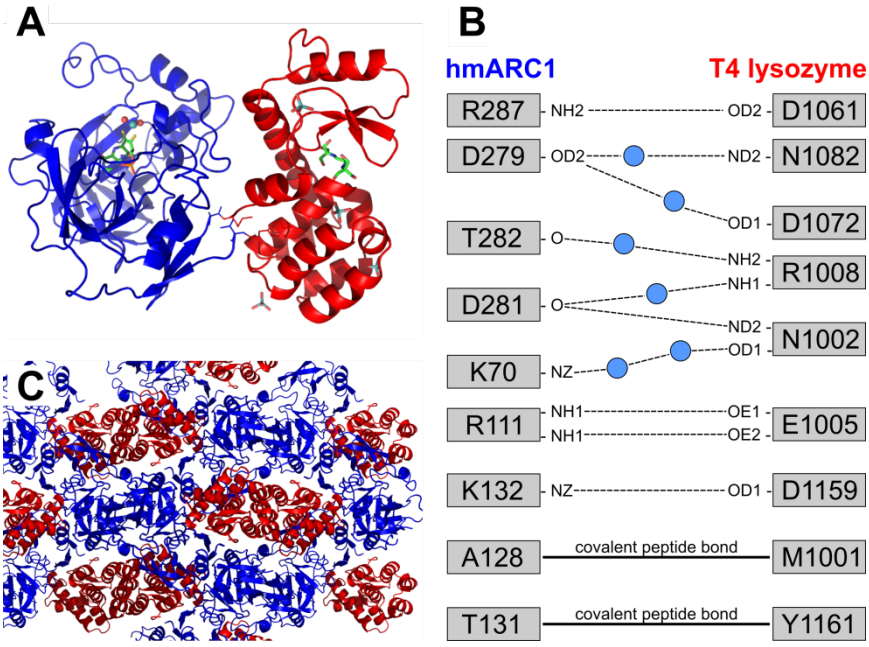

Fig. S1. Interactions between the T4L and *h*mARC1 fusion partners. (*A*) Cartoon representation of the *h*mARC1-T4L fusion protein. The molybdenum cofactor and small molecules are depicted as stick models. Despite the internal integration of the T4L moiety (red) into the *h*mARC1 sequence (blue), both proteins are clearly separated from each other. The two lobes of T4L are rigidly connected with each other by a tightly bound bis-TRIS propane molecule. (*B*) Interaction sites between *h*mARC1 and T4L. In addition to the peptide bonds that covalently tether both moieties to each other, the proteins interact via a few polar direct and indirect contacts. Interacting residues are shown as grey boxes, blue circles represent water molecules and dashed lines indicate hydrogen bonds. (*C*) Crystal packing of the *h*mARC1-T4L fusion protein. Blue, *h*mARC1 molecules; red, T4L molecules. The view is along the b-axis of the unit cell. *h*mARC1 and T4L molecules are packed alternating within the crystal, indicating that the T4L fusion partner is facilitating the crystal contacts.

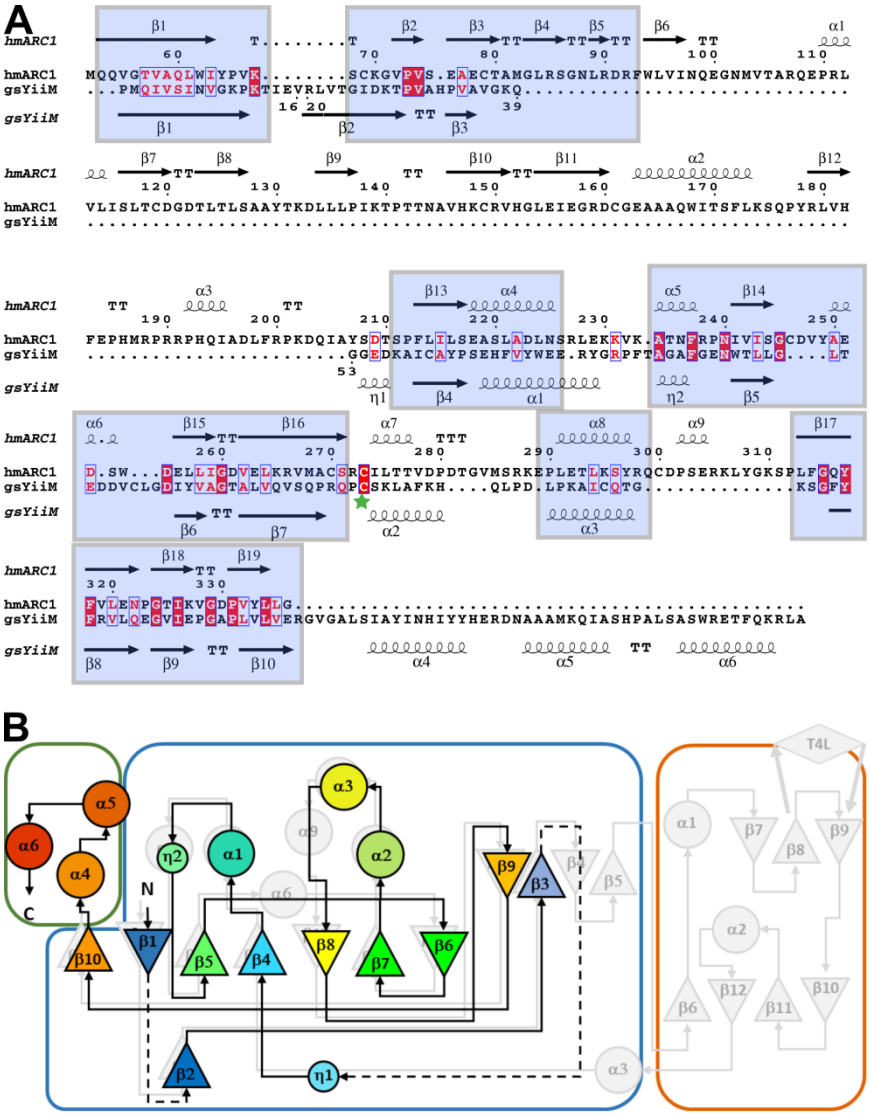

Fig. S2. Structure-based alignment of *h*mARC1 and *gs*YiiM. (*A*) Sequence comparison based on structural alignment with DALI (1). Identical residues are shown as white letters with red background, similar residues are shown as red letters with white background. Secondary structure elements are depicted above and below the sequences. α: α-helix; η: 3_{10}-helix, β: β-sheet; T: turn. Blue shaded areas highlight those parts of the three-dimensional structure, which are superimposed by the structure-based alignment. The green star indicates the cysteine residue coordinating the Moco. (*B*) Topology diagram of *gs*YiiM superimposed onto the *h*mARC1-topology. Triangles represent β-strands, large circles represent α-helices, and small circles 3_{10}-helices. The symbols for *gs*YiiM are rainbow-colored from blue (N-terminus) to red (C-terminus). The symbols for *h*mARC1 are shaded in grey. The superimposed MOSC domain is represented by a blue-colored frame, the mARC-specific MOSC_N domain by an orange-colored frame and the *gs*YiiM-specific C-terminal helix-bundle by a green-colored frame.

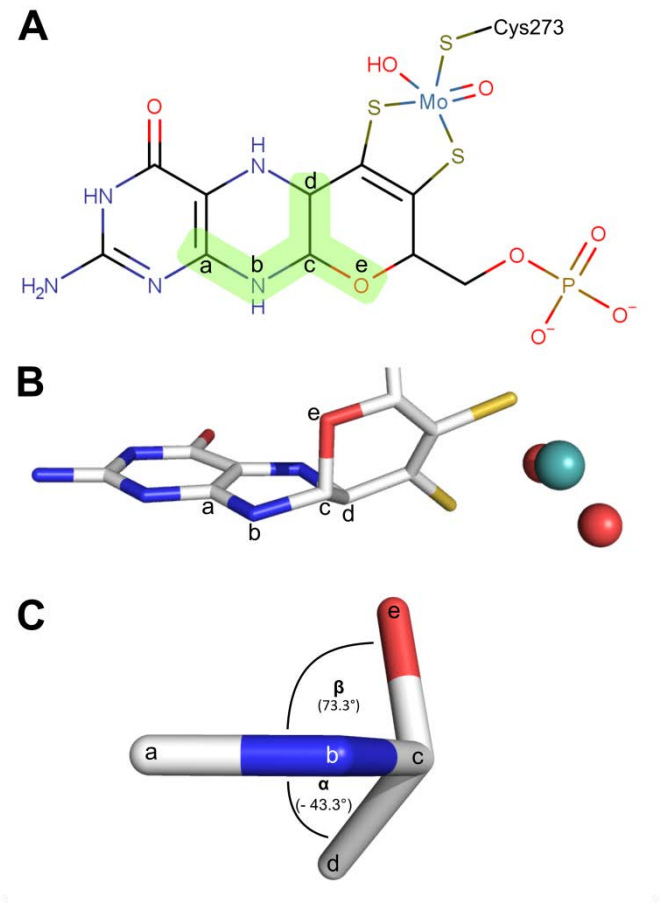

Fig. S3. Molybdopterin conformation observed in *h*mARC1. (*A*) 2D representation of the Moco. Green highlights, atoms used for determination of specific dihedral angles. (*B*) 3D depiction of the Moco conformation as observed in the *h*mARC1 crystal structure. The atoms used for dihedral angle determination are labeled with different letters. (*C*) Representation of the α and β dihedral angles. α is defined by atoms a-b-c-d; β is defined by a-b-c-e.

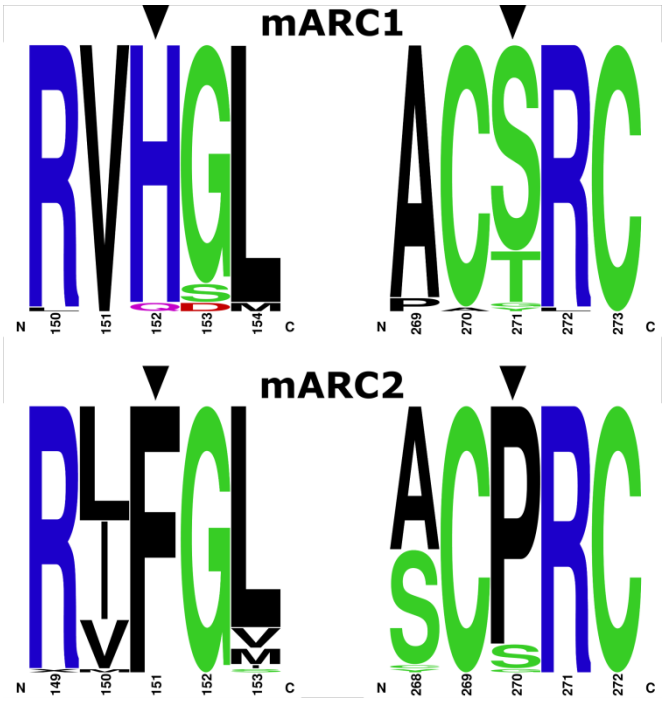

Fig. S4. WebLogo representation of identified mARC paralogue discriminators. Arrows indicate the position of the amino acid residue, which is highly conserved among either mARC1 or mARC2 enzymes, respectively. The height of the letters indicates the percentage with which a residue occurs in this position among all analysed sequences.

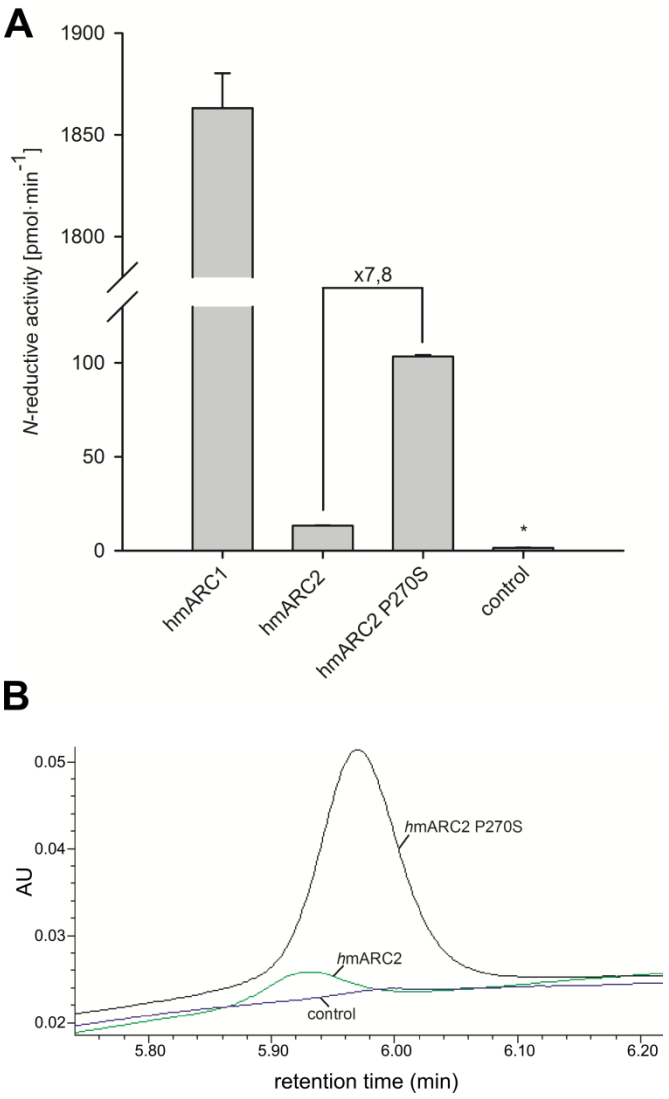

Fig. S5. Reductase activity towards CP544439. (*A*) NADH-dependent activity of the reconstituted enzyme system containing mARC (either *h*mARC1, *h*mARC2 or *h*mARC2_P270S), cytochrome b_5 and cytochrome b_5 reductase. Mixtures without mARC protein were used as negative control. Data are means + SD of biological duplicates. *: below limit of quantification (7.5 pmol/min). (*B*) HPLC chromatogramms showing the peak of the amide-metabolite of CP544439.

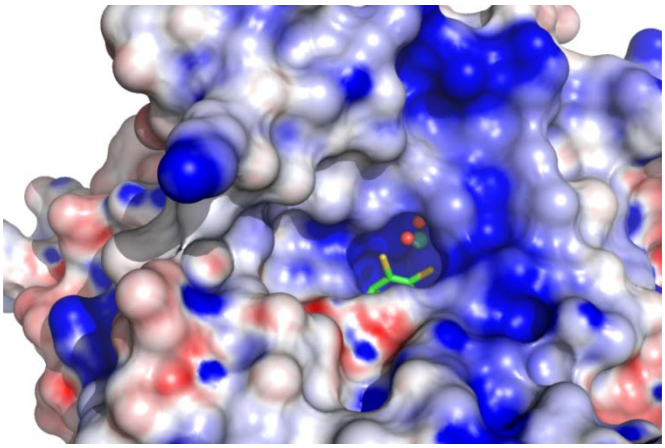

Fig. S6. Electrostatic potential of the surface surrounding the active site. The electrostatic potential is represented as a colour gradient from red (-30 k_BT/e_c) over white (0 k_BT/e_c) to blue (+30 k_BT/e_c).

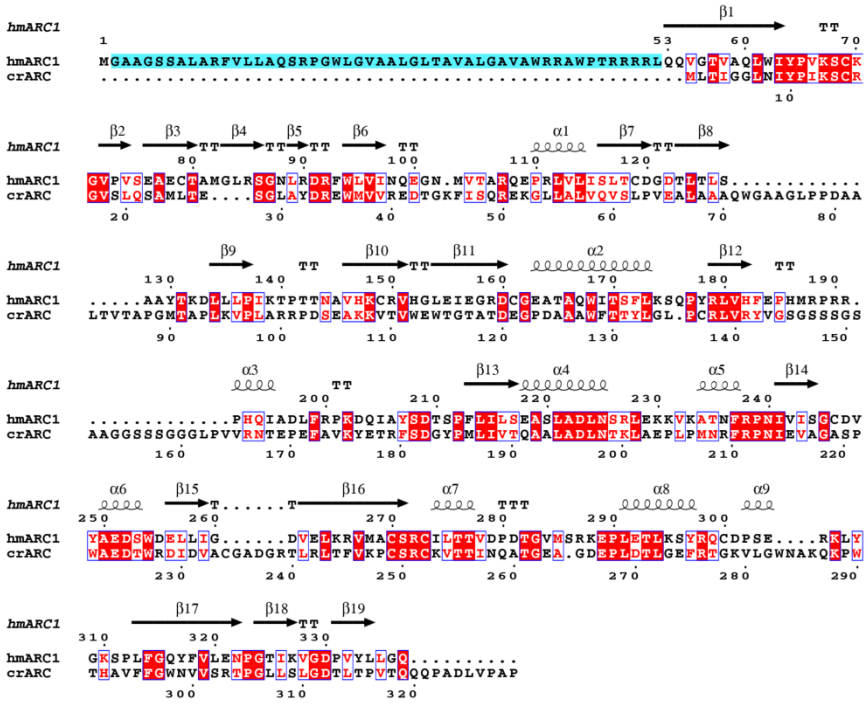

Fig. S7. Sequence comparison of *h*mARC1 and *cr*ARC. Identical residues are shown as white letters with red background, and similar residues are shown as red letters with white background. The residues 2-52 of *h*mARC1 (cyan background) were not part of the crystallized construct. The secondary structure of *h*mARC1 is depicted above the sequences. α: α-helix; β: β-sheet; T: turn.

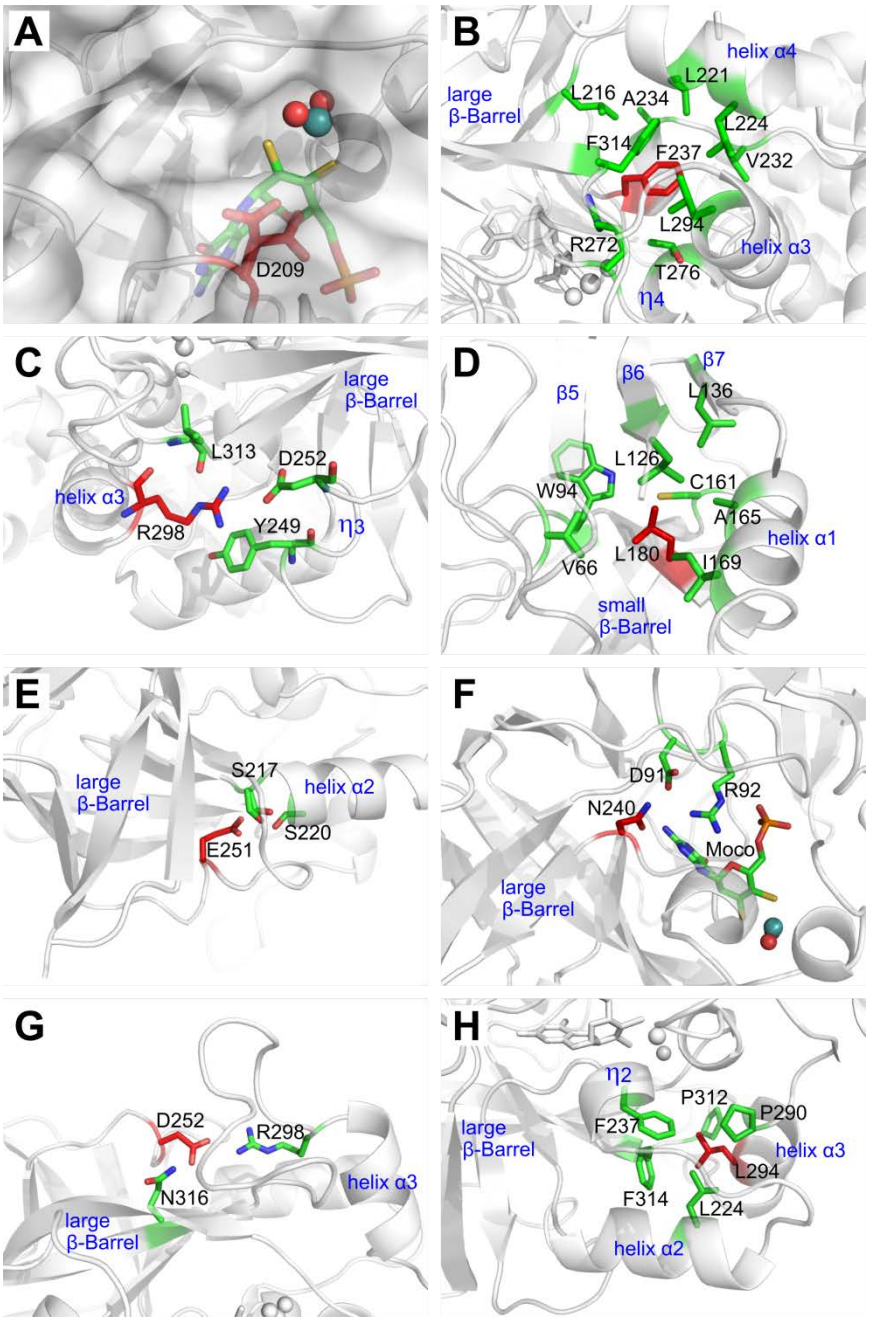

Fig. S8. Position and function of residues that are highly conserved among mARC enzymes from different organisms. The *h*mARC1 backbone is shown in cartoon representation and coloured in white. Highly conserved mARC-residues (red) and neighbouring interacting residues (green) are depicted in stick representation and coloured according to different atom types (red, oxygen; blue, nitrogen; yellow, sulphur; orange, phosphorus). Secondary structure elements of interest are highlighted by blue labels. (*A*) The protein is additionally shown in a semi-transparent surface representation and the Moco is depicted in stick (molbydopterin backbone) and sphere representation (central molybdenum ion and oxygen ligands). The catalytically crucial residue D209 is located in direct proximity to the reactive centre and supposedly recruits the nitrogen atom of *N*-hydroxylated substrates to the active site. (*B*) F237 is the central residue within a large hydrophobic core, which connects the large β-barrel and helices α4, α7 andα8. (*C*) R298 interacts with residues Y249, D252 and L313, thereby connecting helices α8 and α6 with the large β-barrel and 3_{10}-helix η3. (*D*) L180 is is the central residue within a hydrophobic core, which connects helix α2 with the small β-barrel and the

β-sheet comprised of strands β7 to β9. (*E*) Residue E251 interacts with S217 and S220, thereby connecting the large β-barrel with helix α4. (*F*) N240 interacts with the neighbouring residues D91 and R92 as well as the Moco by polar contacts and stabilizes the connection between the large β-barrel and the Moco binding site. (*G*) D252 is the central residue between R298 and N316 and thereby fixates the relative position of helix α8 towards the large β-barrel. (*H*) L294 is the central residue within a hydrophobic core, which connects the helices α4, α5 and α8.

Table S1. Data collection and refinement statistics.

Data collection	
Space group*	$P2_12_12_1$
Cell dimensions*	
a, *b*, *c* (Å)	61.13, 74.82, 110.73
α, β, γ (°)	90.00, 90.00, 90.00
Resolution (Å) *	43.57 – 1.78 (1.81 – 1.78)
R_{merge}*	0.08 (0.40)
$I / \sigma I$ *	10.9 (2.6)
Mean <I> half-set correlation	99.5 (82.9)
Completeness (%) *	91.9 (23.9)
Redundancy *	6.8 (4.2)
Refinement	
Resolution (Å)†	43.57 – 1.78 (1.82 – 1.78)
No. reflections†	311122
R_{work} / R_{free}†	16.9 / 20.8 (27.0 / 34.6)
No. atoms	
Protein	3562
Molybdenum cofactor	27
Bis-TRIS propane	19
Molybdate	20
Phosphate	5
Water	439
B-factors‡	
Protein	29.1
Molybdenum cofactor	23.5
Bis-TRIS propane	26.1
Molybdate	41.1
Phosphate	57.8
Water	34.6
R.m.s. deviations†	
Bond lengths (Å)	0.0241
Bond angles (°)	2.7570

Values in parentheses are for highest-resolution shell. * Values as provided by AIMLESS after data processing, merging and scaling. † Calculated by REFMAC5 (2). ‡ Calculated using the program BAVERAGE. All programs used here are implemented within the CCP4 program package (3).

References

1. Holm L & Laakso LM (2016) Dali server update. *Nucleic Acids Res* 44(W1):W351-355.
2. Murshudov GN, *et al.* (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):355-367.
3. Collaborative Computational Project N (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50(Pt 5):760-763.