

SCIENTIFIC REPORTS

OPEN

Highly selective tungstate transporter protein TupA from *Desulfovibrio alaskensis* G20

Ana Rita Otrelo-Cardoso¹, Rashmi R. Nair¹, Márcia A. S. Correia¹, Raquel S. Correia Cordeiro^{1,4}, Alejandro Panjkovich², Dmitri I. Svergun², Teresa Santos-Silva¹ & Maria G. Rivas³

Molybdenum and tungsten are taken up by bacteria and archaea as their soluble oxyanions through high affinity transport systems belonging to the ATP-binding cassette (ABC) transporters. The component A (ModA/TupA) of these transporters is the first selection gate from which the cell differentiates between MoO_4^{2-} , WO_4^{2-} and other similar oxyanions. We report the biochemical characterization and the crystal structure of the apo-TupA from *Desulfovibrio desulfuricans* G20, at 1.4 Å resolution. Small Angle X-ray Scattering data suggests that the protein adopts a closed and more stable conformation upon ion binding. The role of the arginine 118 in the selectivity of the oxyanion was also investigated and three mutants were constructed: R118K, R118E and R118Q. Isothermal titration calorimetry clearly shows the relevance of this residue for metal discrimination and oxyanion binding. In this sense, the three variants lost the ability to coordinate molybdate and the R118K mutant keeps an extremely high affinity for tungstate. These results contribute to an understanding of the metal-protein interaction, making it a suitable candidate for a recognition element of a biosensor for tungsten detection.

Tungsten is a rare transition metal element from the group VI of the periodic table, together with chromium and molybdenum. It presents several important characteristics that include a high melting point (the highest of all metals, 3422 °C), good conductivity and a large variety of oxidation states (from -2 to +6)^{1,2}. Due to its versatility, W is an important component of a broad range of civil and military industries. The applications include X-ray equipment, implanted medical devices, microwaves, ammunitions, light bulb filaments, jewelry, metal tools, among others^{3,4}. Since the end of the 19th century, Portugal has an important W deposit, being the ninth major producer in the world (by the U.S. Geological Survey in 2016)⁵. In the past, W was viewed as a non-toxic metal but the scientific community became aware of its potential impact on the environment, through the accumulation in air, soils and water (where WO_4^{2-} is the dominant species), mainly in mining sites, but also in some industrial and military sites. Recent data indicate that high concentrations of W are associated with peripheral arterial disease⁶, increased prevalence of stroke⁷, fasting plasma glucose levels and diabetes^{8,9}.

In biological systems, the incorporation of tungsten as a cofactor is limited to bacteria and archaea, with the W-containing enzymes almost restricted to obligate anaerobic prokaryotes^{10,11}. In contrast, the similar element molybdenum can be also found in several enzymes from Eukarya. In most cases, the function of Mo and W-enzymes is to catalyze oxygen atom transfer reactions. One of the challenges of biological systems is to differentiate one metal from the other to avoid its incorrect insertion in the catalytic site of enzymes. Misincorporation of metals into proteins often leads to compromised activity or even inactive enzymes¹²⁻¹⁴.

Both tungsten and molybdenum are taken up by cells in the form of their soluble oxyanions (MoO_4^{2-} or WO_4^{2-}) by high specific ABC transporters^{15,16}. In prokaryotes, there are three types of tungstate/molybdate transporters: ModABC¹⁷, TupABC¹⁸ and WtpABC¹⁹. They are composed of a periplasmic protein (component A), a transmembrane pore forming protein (component B) and a cytoplasmic protein (component C), which hydrolyzes ATP to transport the oxyanion into the cell cytoplasm¹⁶. The genes encoding the three components

¹UCIBIO/REQUIMTE, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516, Caparica, Portugal. ²European Molecular Biology Laboratory-Hamburg Outstation, c/o DESY, Notkestrasse 85, D-22607, Hamburg, Germany. ³Department of Physics, Facultad de Bioquímica y Ciencias Biológicas, Universidad Nacional del Litoral, Santa Fe, 3000, Argentina. ⁴Present address: Ruhr-Universität Bochum, Universitätsstraße, 150/44780, Bochum, Germany. Correspondence and requests for materials should be addressed to T.S.-S. (email: tsss@fct.unl.pt) or M.G.R. (email: mrivas@fcb.unl.edu.ar)

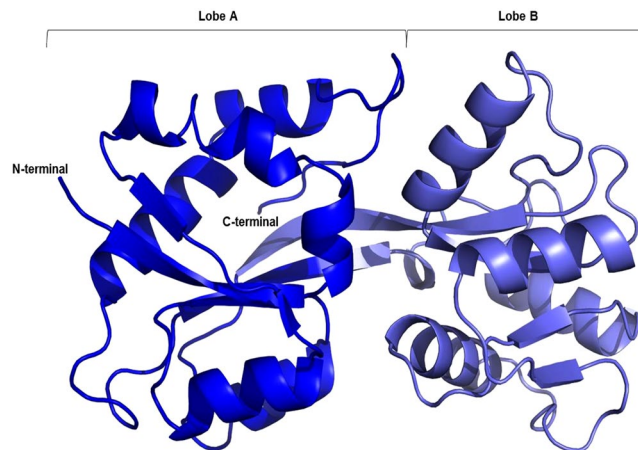


Figure 1. Cartoon representation of the *DaG20* TupA tertiary structure. The different colors represent the two lobes, A in blue and B in light blue. Picture prepared with Pymol⁵⁰.

are organized in an operon (*mod/wtpABC*) or gene cluster and regulated by a transcription factor known as ModE (in the case of the ModABC operon)²⁰. The *modA/tupA/wtpA* gene encodes a periplasmic molybdate/tungstate-binding protein (protein A) which constitutes the first selection gate from which cells should differentiate between Mo and W. The ModA, WtpA, and TupA proteins differ in their binding affinity for molybdate versus tungstate, primary sequence and oxoanion coordination chemistry^{16, 18, 19, 21–23}. The dissociation constants reported to date suggest that TupA and WtpA are strongly selective for tungstate whereas ModA cannot discriminate between molybdate and tungstate¹⁶.

Several crystal structures of ModA/WtpA proteins are available in Protein Data Bank (PDB), namely Mod/WtpA from *Pyrococcus furiosus*²⁴ and ModA from *Escherichia coli*²⁵. In contrast, TupA has been poorly studied. Despite the existence of the TupA crystal structure of *Geobacter sulfurreducens* (*GsTupA*) deposited in the Protein Data Bank (PDB code 3lr1), this system has only been characterized in *Eubacterium acidaminophilum*¹⁸ and *Campylobacter jejuni*²⁶. Atomic detail of tungstate-binding proteins is crucial to understand ligand coordination and the mechanism of metal differentiation, mainly if biotechnological applications are considered.

The present work reports the 1.40 Å crystal structure of the recombinant apo-form of TupA protein from the sulfate reducing bacteria *Desulfovibrio alaskensis* G20 (*DaG20* TupA) together with the biochemical characterization of three variants: R118K, R118E and R118Q. The Arg118 substitution by a lysine yields a variant that is not able to bind molybdate but keeps an extremely high affinity for tungstate. Small-angle X-ray scattering (SAXS) and X-Ray data of the recombinant protein provides valuable information about structural characteristics of TupA and residues involved in ion binding. The last, together with the unique properties of the R118K mutant, contribute for the development of highly selective tungstate sensors.

Results and Discussion

Overall structure. TupA is a globular protein with approximate dimensions of $55.7 \times 34.5 \times 23.8$ Å (Fig. 1). The protein has a butterfly shape with a central beta strand (T77-P89) that supports two lobes (lobe A: A2-R76 and Y201-E251; lobe B: A91-Q200). This architecture was first described for ModA from *E. coli*²⁵ and later for ModA from *Archaeoglobus fulgidus*²⁷. The central beta strand is hydrogen bonded to three beta strands of each of the lobes forming two 4-stranded beta sheets. The TupA metal binding site is located in a central cleft assembled by the two lobes. Despite the very low sequence identity between lobes A and B (5.2%), their secondary structure elements and overall fold is very similar. In fact, superposition of the two lobes gives a RMSD of 3.0 Å (for 57 C α) showing high resemblance between the two parts of the protein that hold the metal ion. This could be associated with the fact that *DaG20* TupA is a monomeric protein suspected to interact with dimeric components of the ABC transporting system. The structural similarities observed between the two lobes could be a result of an evolutionary strategy adopted to improve the interaction between the monomeric and dimeric membrane associated components. This structural characteristic extends to other members of the TupABC family such as *GsTupA* (UniProt Q749P2, PDB code 3lr1) (superposition of the two lobes gives a RMSD of 3.1 Å (for 37 C α), but also to ModABC transporting family (eg ModA from *E. coli*²⁵, where a RMSD of 2.92 Å is obtained for superposition of 63 C α) and WtpABC (eg ModA/WtpA from *Pyrococcus horikoshii*²⁴ with a RMSD of 3.05 Å from 65 C α).

Comparison of *DaG20* TupA with related structures. There are several models available at the PDB of molybdate/tungstate binding proteins. The ModA proteins from eubacteria or archaea share low sequence homology with *DaG20* TupA (less than 22%) and, even though the proteins fold is similar, the RMSD upon superposition is over 3 Å. When the 3D model reported here is used as a search model at the PDB, three bacterial proteins with over 45% sequence identity are retrieved. These are the already mentioned *GsTupA* a protein with unknown function from *Vibrio parahaemolyticus* (Q87PK2, PDB code 3muq) and a LysR substrate binding domain from *Wolinella succinogenes* (Q7M8V9, PDB code 3kn3), all obtained by the Protein Structure Initiative (PSI). In common, the proteins contain the TTTS motif and some other important residues for tungstate binding (see below), indicating that they should be classified as the component A of the TupABC system (Fig. 2).

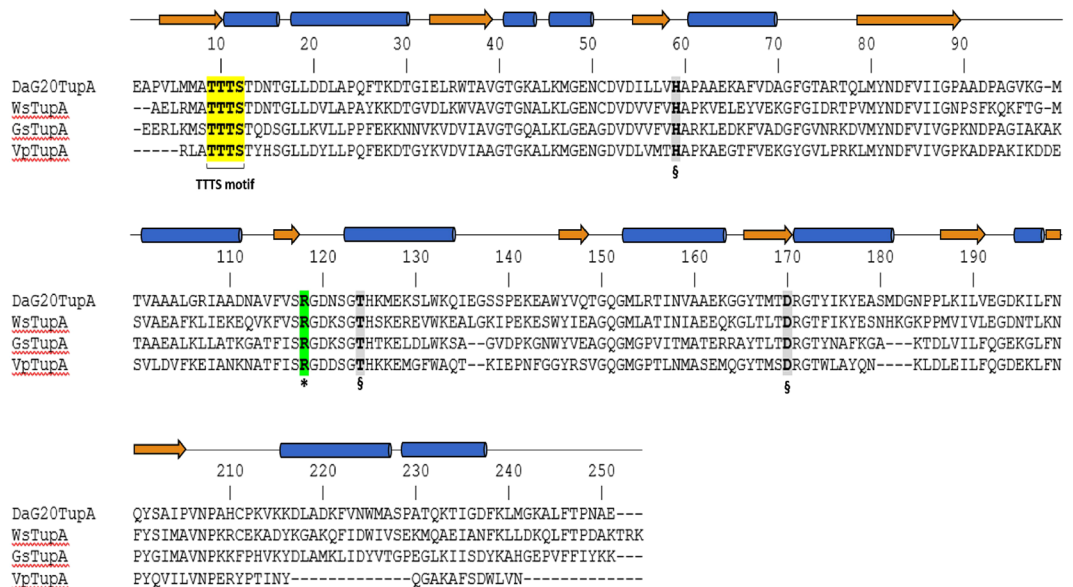


Figure 2. Multiple sequence alignment of mature TupA proteins from different organisms. DaG20TupA - TupA from *Desulfovibrio alaskensis* G20 (locus tag Dde_0234); WsTupA - TupA from *Wolinella succinogenes* (locus tag WS1370); GsTupA - TupA from *Geobacter sulfurreducens* (locus tag GSU2700) and VsTupA - TupA from *Vibrio parahaemolyticus* (locus tag VP1501). Important (*) and putative (§) residues involved in oxyanion binding.

The *DaG20* TupA shares 45.6% sequence identity with *GsTupA*, with conserved secondary structure elements, suggesting that the two proteins are structurally very similar. However, superposition of the entire protein provides an unexpectedly high RMSD value (2.4 Å upon superposition of 208 C α). A more realistic comparison can be achieved when considering the superposition of the independent lobes of the two proteins, yielding much lower deviations (RMSD for the superposition of lobes A and B from the two proteins is 1.24 Å and 1.16 Å for 106 and 96 C α , respectively). These values support a high structural similarity between *DaG20* TupA and *GsTupA* but also indicate some degree of flexibility of the lobes with respect to one another, as expected for the periplasmic A proteins of these transporters. *GsTupA* was likely crystallized in the presence of an ionic form of tungsten, and the deposited model reports a free W⁶⁺ accommodated in the metal binding cleft. In this *GsTupA*-W⁶⁺ holo form, the central cleft volume (363.9 Å³) is four times smaller than the apo form of *DaG20* TupA here reported (1480 Å³). Unexpectedly, the cation is not coordinated to water molecules or protein residues. The water molecules surrounding the W⁶⁺ are at a distance of 3.23–3.82 Å and the closest residues are at 3.9 Å (the OG1 of T9 and NH1 of R118 (*GsTupA* numbering) from lobes A and B, respectively). In *DaG20* TupA these residues are at the same position although not superimposable with *GsTupA*. The data suggest that conformational changes take place upon metal binding, where the protein in the holo form adopts a more compact conformation.

Analysis of the overall structure and stability upon ligand binding. To elucidate the possible conformational changes of *DaG20* TupA upon ligand binding, synchrotron SAXS data were collected both in the presence and absence of tungstate and molybdate. The scattering data of the apo form indicate a monomeric globular TupA protein with a radius of gyration (R_g) of 24.2 Å and a maximum particle size (D_{max}) of about 95 Å (see Table 1). In the presence of tungstate or molybdate (datasets *DaG20* TupA-WO₄²⁻ and *DaG20* TupA-MoO₄²⁻, respectively), the overall shape of the protein remains globular but becomes more spherical and compact, leading to a decreasing R_g (23 Å) and D_{max} (90 Å). Importantly, the datasets obtained for TupA in the presence of tungstate and molybdate are very similar: the CorMap approach²⁸ reveals no statistically significant difference ($C = 12$, p -value = 0.12) between both datasets, indicating that in solution the protein adopts the same conformation upon binding both oxyanions.

Comparison of the distance distribution functions ($P(r)$) between holo and apo forms shows a more compact conformation for the holo form, in agreement with the idea that the protein ‘closes’ upon binding of the metal ion. The apo-form $P(r)$ shape supports the notion of the two lobes being more separated than in the holo-state. The slowly decaying long interdistance tail observed in the $P(r)$ function (Fig. 3, insert), suggests an unstructured N-terminal region containing the 23 residues long expression-tag.

The crystal structure reported here was used as a starting point to create models for both the holo-form and apo-form states. The 23 N-terminal residues that are missing in the crystal structure were calculated using BUNCH²⁹. Based on *GsTupA*-W⁶⁺ adduct, a tungstate group was added to the expected binding site of the BUNCH model, generating a theoretical holo-form model. The theoretical scattering curves calculated using this holo-form model of *DaG20* TupA are in very good agreement with the SAXS measurements of the protein in the presence of tungstate or molybdate (discrepancy $\chi^2 = 0.9$) (Fig. 4). The apo-form model was created by refining the initial BUNCH model with the SREFLEX program³⁰. This approach revealed a slight opening of

	TupA	TupA*	TupA W	TupA W*	TupA Mo
Data collection parameters					
Instrument	P12	BM29	P12	BM29	P12
Wavelength, Å	1.24	0.99	1.24	0.99	1.24
s-range, Å ⁻¹	0.0027–0.45	0.003–0.49	0.0027–0.45	0.003–0.49	0.0027–0.45
Exposure time, s	1	10	1	10	1
Concentration range, mg/ml	1.0–8.5	0.8–6.0	1.0–6.0	0.8–12.0	1.0–6.0
Temperature, K	280	278	280	278	280
Structural parameters					
R_g , Å [from $P(r)$]	25.5 ± 0.3	25.4 ± 0.3	24.1 ± 0.2	23.8 ± 0.2	24.4 ± 0.2
R_g , Å (from Guinier)	24.2 ± 0.4	24.0 ± 0.4	23.0 ± 0.3	22.8 ± 0.3	22.9 ± 0.3
D_{max} , Å	96 ± 3	95 ± 3	90 ± 3	89 ± 3	92 ± 3
Porod volume, Å ³ × 10 ³	51.4 ± 3.5	48.5 ± 3.3	45.7 ± 3.1	44.4 ± 3.0	45.8 ± 3.1
Mol. mass determination (kDa)					
From $I(0)$	29 ± 2	32 ± 2	26 ± 2	29 ± 2	26 ± 2
From Porod volume	31 ± 2	29 ± 2	28 ± 2	27 ± 2	28 ± 2
Model vs. data discrepancy (χ^2)					
Holo-form model	4.0	2.4	0.9	2.2	0.9
MX	7.8	5.8	1.8	11.6	1.8
Apo-form model	1.0	0.7	2.0	11.9	1.7

Table 1. Data collection and structural parameters obtained by SAXS. Molecular mass (M) was estimated from forward scattering $I(0)$ and Porod volume respectively, the theoretical molecular mass predicted from sequence is 29.7 kDa. Radius of gyration, R_g (Å) was calculated using the Guinier approximation and also the distance distribution function ($P(r)$) using GNOM, which also estimates maximum particle dimension (D_{max}). χ^2 values correspond to discrepancies between models and experimental data, the lowest χ^2 value per dataset is highlighted. MX: crystal structure. *Data collected at BM29 beamline, ESRF, Grenoble, France. The rest of the datasets were collected at EMBL P12 beamline, DESY, Hamburg, Germany.

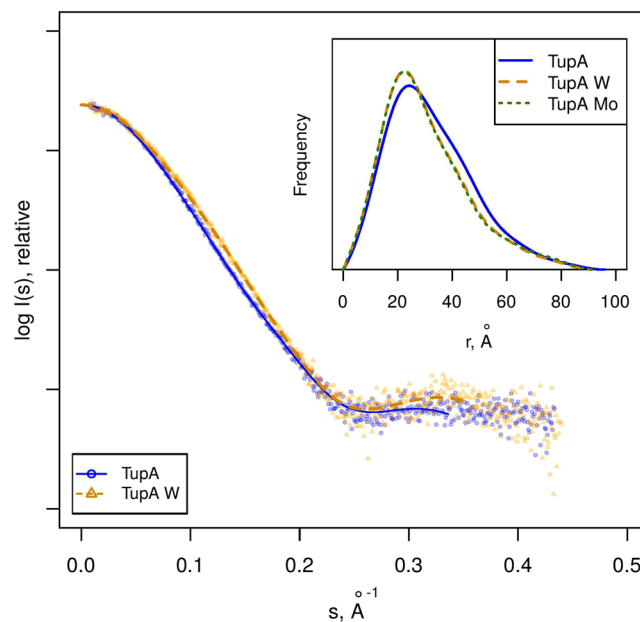


Figure 3. SAXS scattering data (points) and GNOM fits (lines) for TupA in the absence (TupA) and presence of tungstate (TupA W). The data collected in presence of molybdate was omitted from the main plot for clarity as it matches the TupA W data up to noise. Insert: distance distribution functions ($P(r)$) for the different conditions measured.

the lobes (RMSD of 1.5 Å for 274 C α atoms) yielding an excellent agreement to the TupA SAXS data measured in the absence of tungstate or molybdate ($\chi^2 = 1.0$). The optimum holo-form model yielding the best χ^2 to the experimental data and the smallest RMSD to the original structure is presented in Fig. 5 and the conformational

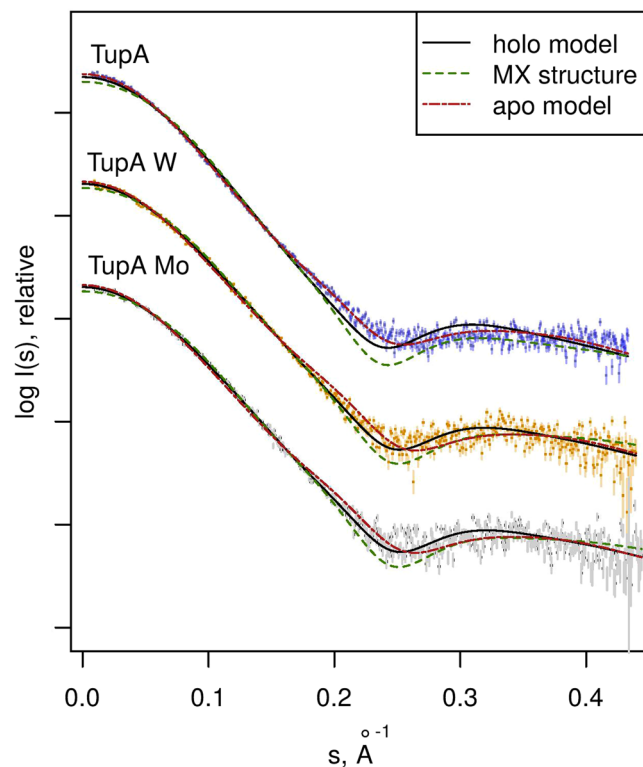


Figure 4. SAXS scattering data (points) for the three experimental conditions, TupA in the absence (TupA) and presence of tungstate (TupA W) or molybdate (TupA Mo). Each dataset has been scaled for display purposes. For each experimental curve, CRY SOL fits are displayed for the crystallographic structure reported in this work (MX) and the holo- and apo-form models generated thereof. The best fit for TupA is the apo-form model ($\chi^2 = 1.0$), while the best fit for both TupA W and TupA Mo is the holo-form model ($\chi^2 = 0.9$).

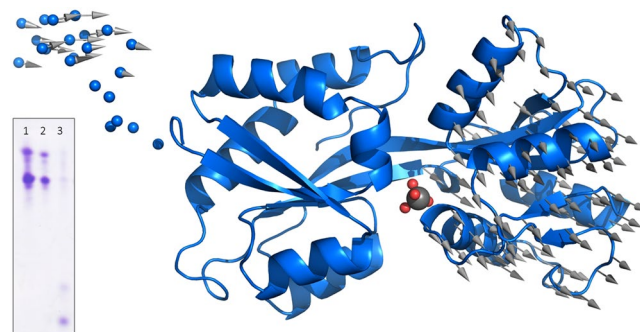


Figure 5. Three-dimensional coordinates for the holo-form hybrid model are displayed using a cartoon representation. The N-terminal section modeled with BUNCH is shown as small spheres. The large gray sphere in the center corresponds to the tungstate group (small red spheres represent O atoms) modeled by homology with PDB entry 3cfz and 3lr1. Vectors have been drawn connecting C α s from the holo-form model to the apo-form model generated by SREFLEX, after superposition of lobe A, to display the ‘opening’ conformational transition. Upon optimal superposition including all C α (274), the RMSD is 1.5 Å. Insert: Urea-polyacrylamide gel electrophoresis of (1) TupA, (2) TupA + MoO $_4$, (3) TupA + WO $_4$. The samples in presence of ligand were first passed through a size exclusion PD-10 minitrapp G-25 columns to eliminate the excess.

transition holo-apo is depicted by arrows. The SAXS data corroborate the hypothesis that *DaG20* TupA is a flexible protein that is able to adopt a loose conformation in the free form but, when binding to molybdate or tungstate, switches to a more compact fold. This feature can be explored to design new detection systems, as discussed later.

To study the impact of the metal binding in the protein stability urea-polyacrylamide gel electrophoresis was carried out for the recombinant *DaG20* TupA in the presence/absence of tungstate and molybdate. TupA-ligand samples were prepared using a 10-fold excess of metal and the excess was removed by size exclusion chromatography. The gel shows that, upon tungstate binding, the wild-type protein migrates in the gel in a larger extend

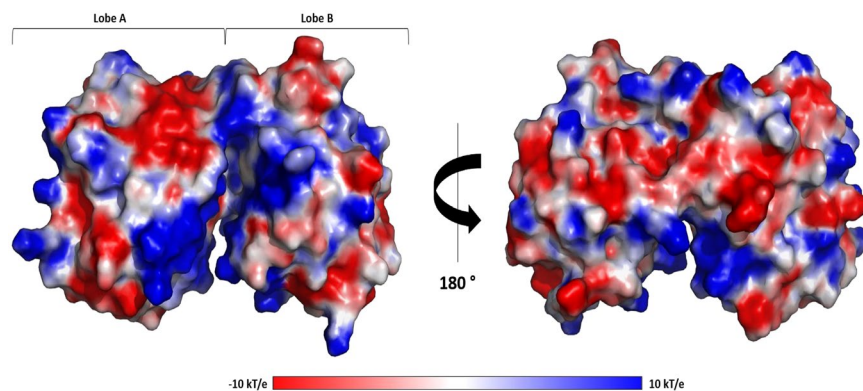


Figure 6. Electrostatic potentials of TupA surface. Electrostatic surface potentials were calculated using Pymol⁵⁰. Surface potentials varies from -10.0 kT/e (red) to 10 kT/e (blue).

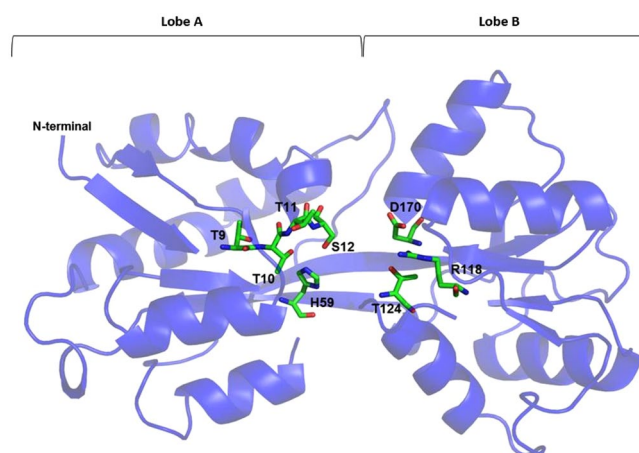


Figure 7. Cartoon representation of the *DaG20* TupA tertiary structure with the conserved residues involved in the metal binding site highlighted. The His59, Thr9-11, Ser12, Arg118, Thr124 and Asp170 are represented as sticks and color by element. Picture prepared with Pymol⁵⁰.

than in the absence of metal or in the presence of molybdate. This indicates that TupA adopts a more compact conformation that is likely to increase stability under a urea gradient, in agreement with our SAXS analysis (Fig. 5, insert).

Oxyanion binding site. Most of the residues that form the binding cleft of ModA are essentially polar but poorly conserved among proteins from different organisms, as discussed for *EcModA*²⁵. In *DaG20* TupA, the cleft is decorated with positively charged residues as seen by the electrostatic surface potential calculations (Fig. 6). The pronounced positive environment of the pocket must be an advantage to enable capture of the oxyanion, even when the extracellular concentration is low. Several residues are likely to be involved in ligand binding, attracting, accommodating and delivering the oxyanion to the membrane component of the transporter system, TupB. The TTTS motif is in lobe A, with the serine pointing towards the metal binding site. Opposite to these are R118, T124 and D170, also likely involved in oxyanion interaction (Fig. 7). These residues are highly conserved among other TupA proteins from *Desulfovibrio* species but also extending to proteobacteria, Green Non Sulfur bacteria and even to Gram-positive bacteria such as Firmicutes. When searching for similar sequences with the TTTS motif and excluding the *Desulfovibrio* genus, over 500 sequences were found with more than 44% identity.

Moreover, one chloride anion, arising from the crystallization conditions, was found at 3.8 Å from R118, indicating the propensity of the pocket to attract negatively charged ions. Although phosphate buffer was used during protein purification this ion is not occupying the WO_4^{2-} binding site, in agreement with what is already known for this type of transporters and with our previous experimental data where we showed that *DaG20* TupA is highly specific for WO_4^{2-} / MoO_4^{2-} and not for other oxyanions as SO_4^{2-} , PO_4^{3-} and ClO_4^- ³¹. To examine the importance of R118 in ligand binding, site-directed mutagenesis has been carried out, as described in the following section.

The relevance of arginine 118 in the oxyanion binding. Using ITC, we reported the binding affinity of *DaG20* TupA for tungstate and molybdate (K_d of 6.30 ± 0.02 pM and 6.1 ± 0.9 nM, respectively)³¹. The described dissociation constants agree with what has been observed for the putative TupA Cj1540 from *C. jejuni*, with this protein binding more tightly tungstate (K_D 1.0 ± 0.2 pM) than molybdate (K_D 50 ± 10 nM)²⁶.

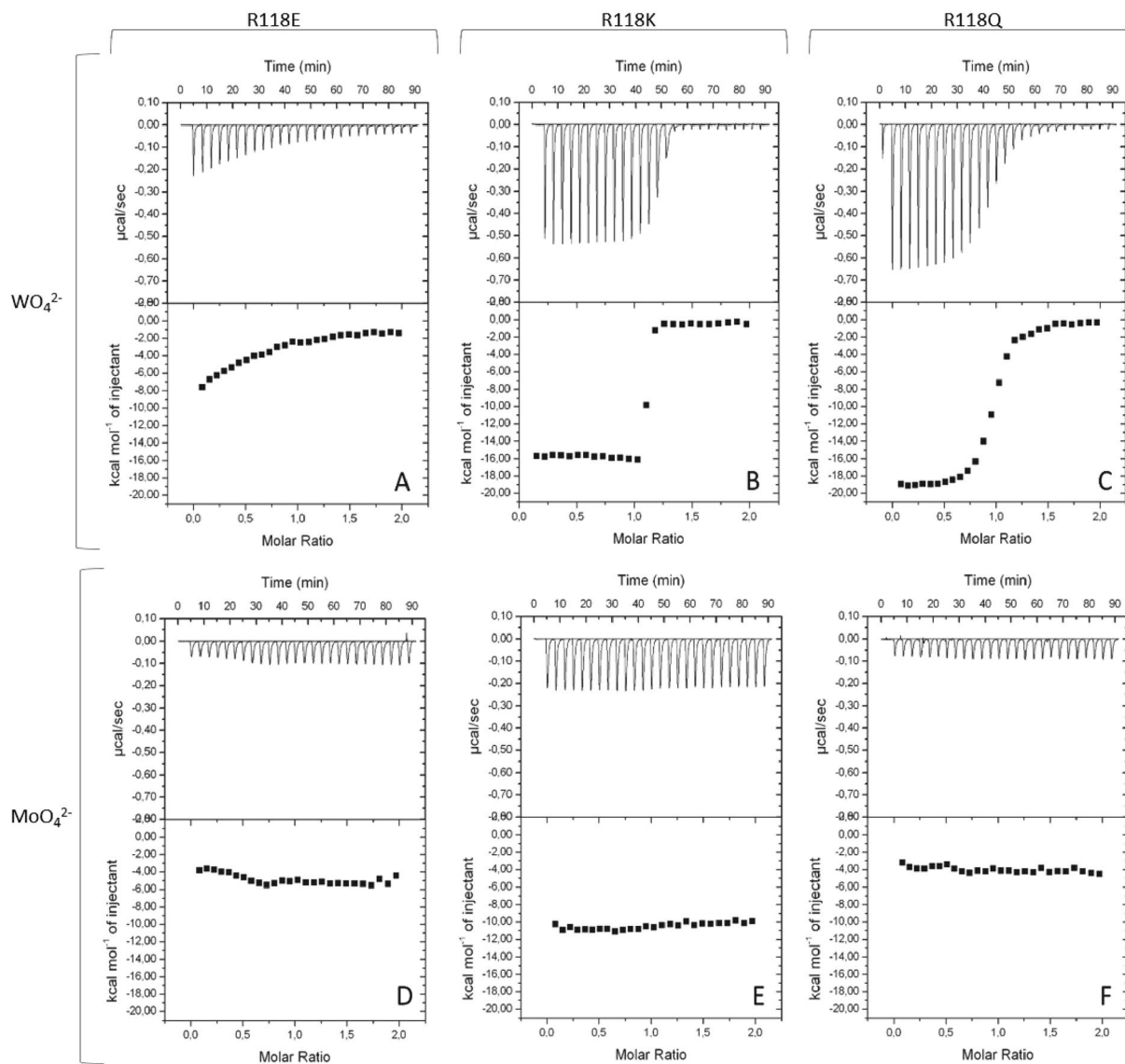


Figure 8. Isothermal titration calorimetry of ligand binding to TupA variants. 10 μM of the variant R118E (A,D), R118K (B,E) or R118Q (C,F) was titrated with injections of 100 μM of WO_4^{2-} (A,B,C) and 100 μM of MoO_4^{2-} (D,E,F). Data were fitted with ORIGIN software. The raw ITC data are shown in the top graphs.

Variant	n	K_A (M^{-1})	K_D (nM)	ΔH (kcal mol^{-1})
TupA ³¹	0.845 ± 0.003	$(1600 \pm 6) \times 10^8$	$(6.30 \pm 0.02) \times 10^{-3}$	-14.60 ± 0.04
R118K	1.080 ± 0.003	$(6 \pm 2) \times 10^8$	1.8 ± 0.8	-15.8 ± 0.1
R118Q	0.950 ± 0.003	$(0.111 \pm 0.006) \times 10^8$	90 ± 50	-19.40 ± 0.08
R118E	No binding			

Table 2. ITC analysis of tungstate binding to TupA variants at 30 °C. In each case 10 μM protein was used for the titrations. n = measured stoichiometry of binding.

The same methodology was used to attend the ability of the R118E/Q/K variants for oxyanion binding (see Fig. 8 and Table 2). The results show that substitution of the positive side chain by a carboxylic acid (R118E variant) prevents the binding to both molybdate and tungstate, confirming the relevance of this conserved residue. When R118 is replaced by a glutamine (R118Q variant), the protein binds tungstate with much less affinity than the wild-type (K_D of 90 ± 50 nM), also losing the ability to bind molybdate. However, if the arginine is replaced by a lysine, the results are remarkably different. Analysis of the ITC thermograms of R118K titrated with tungstate shows a very steep binding curve that indicates a strong interaction. The obtained curve hampers the determination of K_D , extremely smaller as the determined for the wild-type. This problem was also observed before for

Primer	Sense	Sequence	Variant
fTupA_R118K	forward	CCGTATTTGTAAGCAAGGGCGACAACCTCGG	R118K
rTupA_R118K	reverse	CCGAGTTGTGCGCCCTTGCTTACAAATACGG	
fTupA_R118E	forward	CCGTATTTGTAAGCGAGGGCGACAACCTCGG	R118E
rTupA_R118E	reverse	CCGAGTTGTGCGCCCTCGCTTACAAATACGG	
fTupA_R118Q	forward	CCGTATTTGTAAGCCAGGGCGACAACCTCGG	R118Q
rTupA_R118Q	reverse	CCGAGTTGTGCGCCCTGGCTTACAAATACGG	

Table 3. Primers used to mutate the R118 residue.

the *DaG20* TupA-tungstate binding but was overcome by competition studies using molybdate. However, R118K totally lost the affinity to bind molybdate and such binding competition strategy cannot be adopted. Nevertheless, the results clearly show that R118K variant has an extremely high affinity and selectivity for tungstate.

This contrasts with what has been observed for *E. acidaminophilum* where the authors reported that the R/K mutant strongly diminishes the binding of tungstate^{16,32}.

Conclusions

Tungstate is a widely used heavy metal which is raising environmental concerns due to bioaccumulation in soils and water. Currently, tungsten detection requires expensive equipment (such as ICP-mass spectrometry) with only very few reports on new detection methodologies that can be applied directly on the field. A promising research field is the use of enzyme biosensors for determining toxic compounds since the associated analytical systems are simple, rapid and selective. Insertion of tungsten (and molybdenum) in bacterial enzymes is crucial for its activity and requires an uptake system, usually in the form of oxyanion. The organisms developed an efficient biological system, an ABC-type transporter that allows the metal uptake as well as differentiation between the two similar elements, W and Mo. Three classes of ABC transporting systems are described in the literature, in particular ModABC, WtpABC and TupABC.

In this work, we characterized structurally and biochemically the oxyanion coordination and selectivity of the wild-type *DaG20*TupA and three R118 variants. X-ray crystallography and SAXS, reveals that TupA architecture has a common feature with other substrate binding proteins, which is the existence of two separate lobes. This characteristic may improve the interaction of the periplasmic component with the dimeric transmembrane component (TupB), but also provide the structural flexibility that allows TupA to switch between a loose or compact conformation in the absence or presence of the oxyanion, respectively.

The ITC results clearly indicate the relevance of R118 in the oxyanion coordination where three mutants lost the ability to bind molybdate. Curiously, for the R118K variant, the residue substitution increases the selectivity of the protein towards tungstate with high affinity, depleting its ability to bind molybdate.

Due to the conformational changes of TupA upon ligand binding and the high affinity and selectivity of the genetically modified variant R118K, this protein can be considered for biotechnology applications. An alkaline phosphatase based biosensor was recently described to detect W in aqueous media by Alvarado-Gómez and co-workers. The detection relies on the fact that tungstate affects the activity of the immobilized enzyme. However, this method is highly compromised by the presence of other elements like selenium, iron, calcium or aluminum, that even at low concentration interfere with the heavy metal detection³³. The remarkable advantage of using TupA_R118K as the bioreceptor component of a sensor is that its activity is not affected by the presence of other similar ions like molybdate, sulfate, phosphate or perchlorate, allowing selective detection. Combining different techniques, we could better understand the mode of binding of tungstate in TupABC system, paving the way to the development of a new trend in W detection.

Methods

DNA Cloning and Site-directed Mutagenesis. The pet46-tupA expression vector³¹ containing the *tupA* gene (locus tag Dde_0234) was used as a template to perform site-directed mutagenesis of the R118 residue. Primers listed in Table 3 were designed to construct the pET46-tupA_R118K, pET46-tupA_R118E and pET46-tupA_R118Q variants using the Site-Directed Mutagenesis Kit (QuikChange®, Stratagene) following the manual kit instructions. XL1-Blue supercompetent cells were transformed with each expression vector and plasmids were isolated from a unique colony using the NZY-Tech Miniprep kit (NZY-Tech, Lisboa, Portugal). The variants constructs were confirmed by DNA sequencing using an ABI3700 DNA analyzer (Perkin/Elmer/Applied Biosystems, STABvida, Caparica, Portugal). The sequences were aligned and analyzed using the online tool BLASTp³⁴ and ClustalW³⁵.

Expression of TupA variants. The heterologous expression conditions of TupA previously optimized³¹ were tested for the three variants. These conditions yield a high percentage of protein in the insoluble fraction (data not shown). To overcome this problem, several parameters were tested, such as IPTG concentration (0.1, 0.3, 0.5, 0.8 and 1 mM) and different strains of expression hosts (Rosetta 2(DE3)pLysS, Origami (DE3) and Tuner (DE3), from Merck Millipore). The best results were obtained using *E. coli* Rosetta 2 and Origami cells for pET46-tupA_R118K and pET46-tupA_R118E, respectively. In both cases, the cells were transformed and cultured in sterile Luria-Bertani medium supplemented with ampicillin (100 µg/mL) at 180 rpm and 37 °C. When OD₆₀₀ reached 0.6, the protein expression was induced using 1 mM IPTG and cells were grown for 16 hours at 19 °C. For the pET46-tupA_R118Q plasmid, *E. coli* Tuner cells were transformed and the expression performed as previously described but using 0.3 mM of IPTG during the induction period.

Protein isolation protocol. After protein expression, the cells were harvested at $5000 \times g$ for 20 min and the pellet was resuspended in a ratio of 2 g cells/mL in 50 mM sodium phosphate buffer (pH 8.0) containing 500 mM NaCl, 5 mM DNase and 1 tablet/L of Protease Inhibitor Cocktail - EDTA (Sigma-Aldrich). The cells were disrupted by sonication and the solution clarified by centrifugation ($13000 \times g$ for 30 min). The soluble fraction was filtered through a 0.45 μm membrane.

Protein purification was performed in a one-step protocol using an immobilized-metal affinity chromatography (IMAC), His GraviTrap (GE Healthcare) following the manufacturer's instructions. Target proteins were eluted using 50 mM sodium phosphate buffer (pH 8.0) containing 500 mM NaCl and 250 mM imidazole. The fractions were analyzed by 10% tris-tricine/polyacrylamide gel electrophoresis stained with Coomassie blue. Fractions containing TupA_R118K, TupA_R118E and TupA_R118Q were dialyzed against 5 mM Tris-HCl pH 7.6. All of the steps were performed at 4 °C.

Isothermal Titration Calorimetry. ITC experiments were performed as described previously³¹ using a VP-ITC calorimeter (MicroCal GE Healthcare). Prior to experiments, the protein was extensively dialyzed against the reaction buffer (5 mM Tris-HCl, pH 7.5, prepared with Milli-Q dH₂O). The reaction cell containing 10 μM of protein was equilibrated at 30 °C, titrated with sodium tungstate or molybdate (20 or 23 injections of 10 μl of a 100 μM oxyanion solution) and the heat response recorded. After subtraction of the baseline, the integrated heats were fitted to the single binding site model using the ORIGIN software package supplied with the calorimeter to derive, n , K_a and ΔH values.

Urea-polyacrylamide gel electrophoresis. The stability of TupA after ligand binding (tungstate and molybdate) was analyzed by urea gel electrophoresis using the Novex 6% tris(hydroxymethyl) aminomethane (Tris)-borate (TBE)-urea minigels and a XCell SureLock™ Mini-Cell (Invitrogen), as previously described by Mehtab *et al.*³⁶. The protein (5 μL , 50 μM) was mixed with 5 μL of $2 \times$ Novex sample buffer (without EDTA). The electrophoresis was carried out for 150 min at 180 V and 40 mA. The protein bands were examined using Coomassie blue staining. In order to avoid the metal chelation, EDTA was removed from the electrophoresis solutions.

Crystallization of TupA. Crystallization experiments were performed using the hanging drop vapor-diffusion method, adding 2 μL of TupA (7.5 mg/mL) to 1 μL of precipitating solution containing 0.2 M magnesium chloride, 0.1 M HEPES (pH 7.5) and 30% (w/v) polyethylene glycol 3350. Colorless crystals with maximum dimensions of $0.3 \times 0.15 \times 0.06 \text{ mm}^3$ appeared within four days. For further details, see Otrelo-Cardoso *et al.*³¹.

To obtain a crystal structure of the holo form of TupA ($\text{MoO}_4^{2-}/\text{WO}_4^{2-}$ - TupA) soaking and co-crystallization experiments were performed. For the soaking experiments, crystals obtained in the described condition were stabilized by adding a harvesting buffer solution containing 32% (w/v) PEG 3350. The crystals were then incubated with a 5, 10 or 20-fold excess of ligand (prepared in 0.2 M magnesium chloride, 0.1 M HEPES pH 7.5 and 32% (w/v) PEG 3350) for 10 min to 24 hours. Macroscopically, the crystals were not damaged during the soaking and were afterward flash frozen using the mentioned cryo-protectant. Although more than 120 crystals were tested, almost all had poor to non-existing diffraction.

Data Collection, structure determination and refinement. Crystals were flash-cooled in liquid nitrogen using Paratone oil as a cryoprotectant and maintained at 100 K under a stream of gaseous nitrogen during data collection. A complete dataset was collected at beamline ID23-1 at the European Synchrotron Radiation Facility (ESRF, Grenoble, France). The crystals diffract beyond 1.40 Å resolution at a wavelength of 0.954 Å and belong to space group P1₂1. Data was processed with the XDS³⁷ package and AIMLESS³⁸ from the CCP4 program package v. 6.3.0 (Collaborative Computational Project, Number 4, 1994)³⁹. The data collection and processing statistics are presented in Table 4.

Structure determination was carried out by molecular replacement using PHASER⁴⁰ and several molecular models were selected according to sequence alignment homologies, namely: a conserved functionally unknown protein from *Vibrio parahaemolyticus* serotype O3:K6 (PBD code 3muq) and the GsTupA (PDB code 3lr1), after omitting all the cofactors and solvent molecules. The MR solution could only be obtained when the two models were superposed and small domains of the protein were used separately: Domain I, including the first 81 residues; Domain II comprising residues 82 to 188; and finally, Domain III with residues 189 to 236. After structure solution, Buccaneer was used for the automated model building⁴¹ and REFMAC 5 for restrained refinement⁴². The water molecules were automatically added by REFMAC 5 and manually inspected in Coot⁴³. Geometrical validation and model improvement was carried out using PDB_REDO⁴⁴ and the final values of 0.176 and 0.217 for R and R-free factors were obtained, respectively. The Ramachandran plot has 97.15% of the residues in the preferred regions, without any outliers. Mean bond angle and bond length deviations from ideal values and other refinement statistics are presented in Table 4. The deposited model contains 250 protein residues, 332 water molecules, two chlorides and one sodium ion. During the electron density inspection, two mutated residues were found: R107K and S138P. These mutations are likely to come from the cloning strategy but, since are located far from the tungstate binding pocket were considered irrelevant for structure analysis.

Small-angle X-ray scattering, data collection and analysis. SAXS data was collected at EMBL P12 beamline, DESY, Hamburg, Germany and at EMBL BM29 beamline, ESRF, Grenoble, France with protein concentration ranges of 12–0.5 mg/ml. Data was cropped from the first point of the Guinier region until the end of the useful range defined by SHANUM⁴⁵. High and low concentration curves were merged to counter concentration effects such as interparticle interference using the program PRIMUS from the ATSAS package⁴⁶. GNOM⁴⁷ was used to obtain the $p(r)$ and determine the corresponding and values. BUNCH²⁹ and SREFLEX³⁰ were used

Data collection, processing and structure refinement statistics	
X-ray source	ID23-1 (ESRF, Grenoble)
Detector	PILATUS 6M-F
Wavelength (Å)	0.954
Unit-cell parameters (Å, °)	a = 52.32; b = 42.53; c = 54.75; β = 95.45
Space group	P12 ₁ 1
Molecules per AU [#]	1
Matthews coefficient (Å ³ , Da)	2.09
Mosaicity (°)	0.22
Resolution range (Å)	42.53–1.40 (1.42–1.40)
<I/σI>	12.8 (2.4)
R _{merge} (%)	3.4 (29.7)
R _{pin} (%)	3.3 (28.8)
R _{meas} (%)	4.8 (41.4)
Multiplicity	2.9 (2.9)
No. of observed reflections	135886 (6650)
No. of unique reflections	46519 (2277)
Completeness (%)	98.3 (98.5)
R _{free} [*] (%)	21.7%
R _{factor} (%)	17.6%
Number of water molecules	332
Other heteroatoms	1 sodium, and 2 chloride
Average B factor for all atoms (Å ²)	23.67
RMSD from ideal geometry	
Bond lengths (Å)	0.022
Bond angles (°)	2.017

Table 4. Data collection, processing and structure refinement statistics for TupA crystal. Values in parentheses correspond to the highest resolution shell. [#]AU: asymmetric unit. ^{*}R_{free} was calculated for 5.1% of the reflections randomly chosen from the data set.

to generate and refine high-resolution hybrid models using the crystallographic structure reported here as the starting point. The scattering curves from the high-resolution models were calculated using CRY SOL⁴⁸.

Data Availability. Coordinates and observed structure factor amplitudes of DaG20 TupA have been deposited in the Protein Data Bank under the accession code 5my5. The collected SAXS data and the generated high-resolution hybrid models have been deposited and are available at SASBDB (entries: SASDBD9, SASDBE9, SASDBF9, SASDBG9 and SASDBH9)⁴⁹.

References

- Bolt, A. M. & Mann, K. K. Tungsten: an emerging toxicant, alone or in combination. *Curr. Environ. Heal. Reports* **3**, 405–415 (2016).
- Dermatas, D. *et al.* Solubility, sorption, and soil respiration effects of tungsten and tungsten alloys. *Environ. Forensics* **5**, 5–13 (2004).
- Strigul, N. *et al.* Effects of tungsten on environmental systems. *Chemosphere* **61**, 248–258 (2005).
- Kennedy, A. J. *et al.* Tungsten toxicity, bioaccumulation, and compartmentalization into organisms representing two trophic levels. *Environ. Sci. Technol.* **46**, 9646–9652 (2012).
- Faria, M. S., Lopes, R. J., Malcato, J., Nogueira, A. J. A. & Soares, A. M. V. M. *In situ* bioassays with Chironomus riparius larvae to biomonitor metal pollution in rivers and to evaluate the efficiency of restoration measures in mine areas. *Environ. Pollut.* **151**, 213–221 (2008).
- Navas-Acien, A. *et al.* Metals in urine and peripheral arterial disease. *Environ. Health Perspect.* **113**, 164–169 (2005).
- Tyrrell, J. *et al.* High urinary tungsten concentration is associated with stroke in the National Health and Nutrition Examination Survey 1999–2010. *PLoS One* **8**, e77546 (2013).
- Menke, A., Guallar, E. & Cowie, C. C. Metals in urine and diabetes in US adults. *Diabetes* **65**, 164–171 (2016).
- Feng, W. *et al.* Association of urinary metal profiles with altered glucose levels and diabetes risk: A population-based study in China. *PLoS One* **10**, e0123742 (2015).
- Hille, R. Molybdenum enzymes containing the pyranopterin cofactor: An overview. *Met. Ions Biol. Syst* **39**, 187–226 (2002).
- Schwarz, G., Hagedoorn, P. L. & Fischer, K. Molybdate and tungstate: uptake, homeostasis, cofactors, and enzymes in *Microbiology of Heavy Metals* (eds. Nies, D. H. & Silver, S.) 421–452 (Springer, 2007).
- Gates, A. J. *et al.* Properties of the periplasmic nitrate reductases from *Paracoccus pantotrophus* and *Escherichia coli* after growth in tungsten-supplemented media. *FEMS Microbiol. Lett* **220**, 261–269 (2003).
- May, H. D., Patel, P. S. & Ferry, J. G. Effect of molybdenum and tungsten on synthesis and composition of formate dehydrogenase in *Methanobacterium formicicum*. *J. Bacteriol* **170**, 3384–3389 (1988).
- Pollock, V. V., Conover, R. C., Johnson, M. K. & Barber, M. J. Bacterial expression of the molybdenum domain of assimilatory nitrate reductase: production of both the functional molybdenum-containing domain and the nonfunctional tungsten analog. *Arch. Biochem. Biophys.* **403**, 237–248 (2002).
- Schwarz, G., Mendel, R. R. & Ribbe, M. W. Molybdenum cofactors, enzymes and pathways. *Nature* **460**, 839–47 (2009).
- Hagen, W. R. Cellular uptake of molybdenum and tungsten. *Coord. Chem. Rev.* **255**, 1117–1128 (2011).

17. Rech, S., Wolin, C. & Gunsalus, R. P. Properties of the periplasmic ModA molybdate-binding protein of *Escherichia coli*. *J. Biol. Chem.* **271**, 2557–62 (1996).
18. Makdessi, K., Andreesen, J. R. & Pich, A. Tungstate Uptake by a highly specific ABC transporter in *Eubacterium acidaminophilum*. *J. Biol. Chem.* **276**, 24557–24564 (2001).
19. Bevers, L. E., Hagedoorn, P. L., Krijger, G. C. & Hagen, W. R. Tungsten transport protein A (WtpA) in *Pyrococcus furiosus*: the first member of a new class of tungstate and molybdate transporters. *J. Bacteriol.* **188**, 6498–505 (2006).
20. Anderson, L. A. *et al.* Characterisation of the molybdenum-responsive ModE regulatory protein and its binding to the promoter region of the *modABCD* (molybdenum transport) operon of *Escherichia coli*. *Eur. J. Biochem.* **246**, 119–126 (1997).
21. Rech, S., Deppenmeier, U. & Gunsalus, R. P. Regulation of the molybdate transport operon, *modABCD*, of *Escherichia coli* in response to molybdate availability. *J. Bacteriol.* **177**, 1023–1029 (1995).
22. Imperial, J., Hadi, M. & Amy, N. K. Molybdate binding by ModA, the periplasmic component of the *Escherichia coli mod* molybdate transport system. *Biochim. Biophys. Acta* **1370**, 337–346 (1998).
23. Bevers, L. E., Schwarz, G. & Hagen, W. R. A molecular basis for tungstate selectivity in prokaryotic ABC transport systems. *J. Bacteriol.* **193**, 4999–5001 (2011).
24. Hollenstein, K. *et al.* Distorted octahedral coordination of tungstate in a subfamily of specific binding proteins. *J. Biol. Inorg. Chem.* **14**, 663–672 (2009).
25. Hu, Y., Rech, S., Gunsalus, R. P. & Rees, D. C. Crystal structure of the molybdate binding protein ModA. *Nature Struct. Mol. Biol.* **4**, 703–707 (1997).
26. Smart, J. P., Cliff, M. J. & Kelly, D. J. A role for tungsten in the biology of *Campylobacter jejuni*: tungstate stimulates formate dehydrogenase activity and is transported via an ultra-high affinity ABC system distinct from the molybdate transporter. *Mol. Microbiol.* **74**, 742–57 (2009).
27. Hollenstein, K., Frei, D. C. & Locher, K. P. Structure of an ABC transporter in complex with its binding protein. *Nature* **446**, 213–216 (2007).
28. Franke, D., Jeffries, C. M. & Svergun, D. I. Correlation Map, a goodness-of-fit test for one-dimensional X-ray scattering spectra. *Nat. Methods* **12**, 419–422 (2015).
29. Petoukhov, M. V. & Svergun, D. I. Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophys. J.* **89**, 1237–1250 (2005).
30. Panjkovich, A. & Svergun, D. I. Deciphering conformational transitions of proteins by small angle X-ray scattering and normal mode analysis. *Phys. Chem. Chem. Phys.* **18**, 5707–5719 (2015).
31. Otrelo-Cardoso, A. R., Nair, R., Correia, M. A. S., Rivas, M. G. & Santos-Silva, T. TupA: a tungstate binding protein in the periplasm of *Desulfovibrio alaskensis* G20. *Int. J. Mol. Sci.* **15**, 11783 (2014).
32. Andreesen, J. R. & Makdessi, K. Tungsten, the surprisingly positively acting heavy metal element for prokaryotes. *Ann. N. Y. Acad. Sci.* **1125**, 215–229 (2008).
33. Alvarado-Gómez, A. L., Alonso-Lomillo, M. A., Domínguez-Renedo, O. & Arcos-Martínez, M. J. A chronoamperometric screen printed carbon biosensor based on alkaline phosphatase inhibition for W(VI) determination in water, using 2-phospho-l-ascorbic acid trisodium salt as a substrate. *Sensors (Basel)* **15**, 2232–2243 (2015).
34. Altschul, S. F. & Lipman, D. J. Protein database searches for multiple alignments. *Proc. Natl. Acad. Sci. USA* **87**, 5509–5513 (1990).
35. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680 (1994).
36. Mehtab, S. *et al.* Interaction of vanadium(IV) with human serum apo-transferrin. *J. Inorg. Biochem.* **121**, 187–95 (2013).
37. Kabsch, W. XDS. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 125–32 (2010).
38. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. D. Biol. Crystallogr.* **69**, 1204–1214 (2013).
39. Number 4 CCP. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* **50**, 760–63 (1994).
40. Read, R. J. Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **57**, 1373–1382 (2001).
41. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D. Biol. Crystallogr.* **62**, 1002–11 (2006).
42. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D. Biol. Crystallogr.* **53**, 240–255 (1997).
43. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2126–2132 (2004).
44. Joosten, R. P., Long, F., Murshudov, G. N. & Perrakis, A. The PDB_REDO server for macromolecular structure model optimization. *IUCr* **1**, 213–220 (2014).
45. Konarev, P. V. & Svergun, D. I. A posteriori determination of the useful data range for small-angle scattering experiments on dilute monodisperse systems. *IUCr* **2**, 352–60 (2015).
46. Petoukhov, M. V. *et al.* New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **45**, 342–350 (2012).
47. Svergun, D. I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25**, 495–503 (1992).
48. Svergun, D., Barberato, C. & Koch, M. H. J. CRY SOL – a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **28**, 768–773 (1995).
49. Valentini, E., Kikhney, A. G., Previtali, G., Jeffries, C. M. & Svergun, D. I. SASBDB, a repository for biological small-angle scattering data. *Nucleic Acids Res.* **43**, D357–D363 (2015).
50. DeLano, W. The PyMOL Molecular Graphics System. *San Carlos, CA DeLano Sci.* (2002).

Acknowledgements

This work was supported by the Unidade de Ciências Biomoleculares Aplicadas-UCIBIO which is financed by national funds from FCT/MEC (UID/Multi/04378/2013) and co-financed by the ERDF under the PT2020 Partnership Agreement (POCI-01-0145-FEDER-007728) and the following FCT-MEC projects: EXPL/BBB-BEP/0274/2012; PTDC/REQ-MED/1902/2014; RECI/BBB-BEP/0124/2012. ARO-C, RRN and MASC thank FCT-MEC for grants (Reference SFRH/BD/85806/2012, SFRH/BPD/63061/2009 and SFRH/BPD/64917/2009). MGR is a member of the CONICET (Argentina). The authors thank the EMBL P12 (DESY, Hamburg, Germany) and BM29 (ESRF, Grenoble, France) beamline staff for assistance during data collection. AP was supported by the EMBL Interdisciplinary Postdoc Programme under Marie Curie COFUND Actions. Data collection at EMBL P12 beamline was supported by the European Commission, BioStruct-X grant 283570. We acknowledge Faculdade de Medicina Veterinária da Universidade Técnica de Lisboa for the Isothermal titration calorimetry facility.

Author Contributions

A.R.O.-C., M.A.S.C., T.S.S.-S. and M.G.R. designed experiments, analyzed data and wrote the manuscript. A.R.O.-C., R.R.N. and R.C. conducted the experiments. A.P. analyzed SAXS data with assistance from D.I.S. T.S.S.-S. and M.G.R. supervised the project.

Additional Information

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017