BBA - Biomembranes 1866 (2024) 184221

Contents lists available at ScienceDirect

BBA - Biomembranes

journal homepage: www.elsevier.com/locate/bbamem

Liquid-chromatography mass spectrometry describes post-translational modification of *Shewanella* outer membrane proteins

Jessica H. van Wonderen [a,*], Jason C. Crack [a], Marcus J. Edwards [b,c], Thomas A. Clarke [b], Gerhard Saalbach [d], Carlo Martins [d], Julea N. Butt [a,b,*]

[a] School of Chemistry, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK
[b] School of Biological Sciences, University of East Anglia, , Norwich Research Park, Norwich NR4 7TJ, UK
[c] School of Life Sciences, University of Essex, Colchester CO4 3SQ, UK
[d] Proteomics Facility, The John Innes Centre, Norwich Research Park, Norwich NR4 7UH, UK

ARTICLE INFO

Keywords:
Integral membrane protein
Mass spectrometry
Detergent
Intact mass analysis
Electromicrobiology
Lipidation

ABSTRACT

Electrogenic bacteria deliver excess respiratory electrons to externally located metal oxide particles and electrodes. The biochemical basis for this process is arguably best understood for species of *Shewanella* where the integral membrane complex termed MtrCAB is key to electron transfer across the bacterial outer membranes. A crystal structure was recently resolved for MtrCAB from *S. baltica* OS185. However, X-ray diffraction did not resolve the N-terminal residues so that the lipidation status of proteins in the mature complex was poorly described. Here we report liquid chromatography mass spectrometry revealing the intact mass values for all three proteins in the MtrCAB complexes purified from *Shewanella oneidensis* MR-1 and *S. baltica* OS185. The masses of MtrA and MtrB are consistent with both proteins being processed by Signal Peptidase I and covalent attachment of ten *c*-type hemes to MtrA. The mass of MtrC is most reasonably interpreted as arising from protein processed by Signal Peptidase II to produce a diacylated lipoprotein containing ten *c*-type hemes. Our two-step protocol for liquid-chromatography mass spectrometry used a reverse phase column to achieve on-column detergent removal prior to gradient protein resolution and elution. We envisage the method will be capable of simultaneously resolving the intact mass values for multiple proteins in other membrane protein complexes.

1. Introduction

Membrane proteins are responsible for a wide range of biological functions in prokaryote and eukaryote cell types. For example, receptors permit the detection of environmental cues, and respiratory complexes perform electron transfer and proton pumping to maintain the transmembrane electrochemical potential that drives ATP synthesis. Large protein complexes spanning the outer membrane of Gram-negative bacteria have also been recognized to exchange electrons between internal enzymes and external redox partners [1–4]. This electron exchange, sometimes known as extracellular electron transfer, has evolved to enable bacteria to harness energy from the extracellular redox cycling of transition metals including those in iron and manganese rich minerals [5–7]. Extracellular electron transfer also attracts interest for biotechnology [8–10]. Electrons from the bacterial oxidation of organic waste streams can be delivered to anodes producing green electricity in microbial fuel cells. Bacterial electrosynthesis of valuable chemicals can occur when bacteria receive electrons from cathodes.

The biochemical basis for extracellular electron transfer is arguably best understood for species of *Shewanella* [2–4]. In the case of *S. baltica* OS185 a crystal structure was recently resolved for the outer membrane spanning MtrCAB complex [11] that is key to electron transfer across the otherwise electrically insulating lipid bilayer. The structure revealed two membrane spanning proteins, MtrA and MtrB, arranged as a biomolecular wire, Fig. 1. The electrically conductive cytochrome MtrA (~38 kDa) is spanned by a chain of 10 covalently attached *c*-type hemes and insulated from the membrane lipidic environment by its position inside the porin-like beta-barrel configuration of MtrB (~75 kDa). Within the heterotrimeric MtrCAB complex, a terminus of the MtrA heme wire is positioned in close proximity with a terminus of the heme wire in the extracellular cytochrome MtrC (~75 kDa). The latter protein contains 10 covalently bound *c*-type hemes arranged as a staggered cross configuration and presents a large surface area for direct electrical contact with extracellular redox partners.

* Correspondence to: School of Chemistry, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK.
E-mail addresses: j.wonderen@uea.ac.uk (J.H. van Wonderen), j.butt@uea.ac.uk (J.N. Butt).

https://doi.org/10.1016/j.bbamem.2023.184221
Received 8 February 2023; Received in revised form 9 August 2023; Accepted 30 August 2023
Available online 7 September 2023

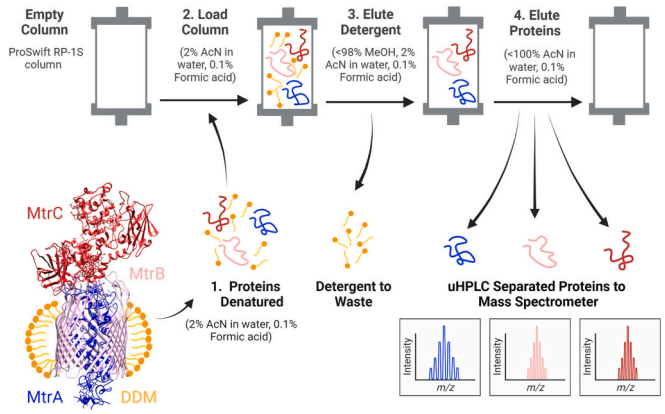


Fig. 1. Overview of LC-MS method to resolve intact masses for proteins of the membrane-spanning MtrCAB complex. The MtrCAB complex of *S. baltica* OS185 MtrCAB (pdb: 6R2Q) is shown: MtrA (blue), MtrB (light pink) and MtrC (red). Dodecyl maltoside (DDM) detergent (yellow) is shown schematically. For LC-MS the complex was first denatured by formic acid with acetonitrile (AcN), then loaded to a ProSwift RP-1S column. DDM was eluted with high methanol (MeOH). Proteins were eluted with a gradient of AcN prior to infusion to the mass spectrometer. Image created with BioRender.com. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

MtrCAB complexes, and paralogs termed MtrDEF, are encoded by the genomes of other *Shewanella* species [12], most notably *S. oneidensis* MR-1 that has become a model for studies of extracellular electron transfer and a chassis for related biotechnology [13]. However, questions remain regarding important details of the MtrCAB structure. For example, MtrC (formerly termed OmcB) of *S. oneidensis* MR-1 was proposed [14] to carry a lipid anchor covalently bound to the N-terminal lipoprotein consensus sequence (LxxC) for signal peptidase II. The same lipid attachment site is encoded by the *mtrC* gene of *S. baltica* OS185, Fig. S1. However, X-ray diffraction did not resolve the N-terminal residues [11] and the lipidation status of MtrC remains poorly described.

Mass spectrometry (MS) offers opportunities to gain information complementary to that afforded by X-ray diffraction. For example, MS of membrane proteins and their complexes in their native states can provide direct insight into oligomeric state, protein-lipid and protein-ligand interactions [15–20]. For detection of post-translational modification MS is more informative when proteins are completely separated from one another and any co-purified non-covalently bound lipids and/or detergents. Such measurements report the intact mass of an individual protein. The corresponding experiments are typically performed by analysis of the denatured protein(s) by Liquid Chromatography-Mass Spectrometry (LC-MS) such that denatured proteins, detergent and other solutes are separated by chromatography prior to injection into the mass spectrometer [21–24].

Reverse-phase chromatography is used to separate solutes from denatured membrane proteins in the main LC-MS approaches reported to date for intact mass determination. The BioResolve RP mAb Polyphenol column with 90 % n-propanol as the mobile phase allows for rapid elution of membrane protein with slower elution of detergent [24]. This is a high throughput method ideal for screening of pure membrane protein constructs for intact mass determination, however the elution is too rapid for extensive separation of individual proteins in multi-protein mixtures. The Zorbax 300SB-C3 column with methanol as the mobile phase can separate proteins from several different detergents under specific conditions [22,23]. HPLC resolution of the protein from the detergent usually requires a longer column length (150 mm) and, as the detergent co-elutes with the protein(s), some non-covalent attachment of detergent to protein can be retained. Although for some sample compositions it is possible to optimize the elution gradients for detergent/protein and multi-protein separation, this method requires high sample purity, an investment of time for optimization, and the resulting protein-detergent spectra can be difficult to deconvolute, particularly for novel protein complexes.

In this work we build on the aforementioned LC-MS methods [21–24] to present a protocol allowing us to reveal the intact mass of each protein in the MtrCAB complexes purified from *S. baltica* OS185 and *S. oneidensis* MR-1. The results are consistent with lipidation of the MtrC proteins, but not MtrA or MtrB proteins. Our LC-MS protocol loads MtrCAB onto a ProSwift RP-1S reversed-phase column to allow complete removal of detergents by application of aqueous methanol while the proteins remain tightly bound to the column. Proteins are subsequently resolved by gradient elution with aqueous acetonitrile and directly infused to a Bruker microQTOF-QIII mass spectrometer for positive mode ESI. This protocol allowed the membrane proteins, initially suspended in the non-ionic detergent dodecyl maltoside (DDM), to be cleanly separated from detergent prior to gradient protein elution. Furthermore, the acetonitrile gradient is readily optimized to assist the separation of different proteins prior to infusion to the mass spectrometer as demonstrated here for proteins of the MtrCAB complexes. Thus, we anticipate this protocol with its two-phase elution profile will be widely applicable and readily amenable to LC-MS analysis of other membrane proteins and their complexes.

2. Materials and methods

2.1. Sample preparation

MtrCAB was purified from *S. baltica* OS185 and *S. oneidensis* MR-1 as described previously [11,25]. Purified MtrCAB was exchanged from 2 % (v/v) Triton X-100, 20 mM HEPES, 50 mM NaCl, pH 7.8 into 5 mM lauryldimethylamine-N-oxide, 20 mM TRIS-HCl, pH 7.5 using anionic exchange chromatography (HiTrap Q-Sepharose, 5 mL) and a NaCl gradient for elution as described previously [25]. A 15 mL volume 100 kDa molecular weight cut off centrifugal concentrator (Merck Life Sciences) was then used to exchange the protein into 0.015 % (v/v) (0.3 mM) DDM, 20 mM TRIS-HCl, pH 7.5, by concentrating to 10 % volume and diluting 10-fold in the new buffer four times. DDM micelles are ~70 kDa so the 100 kDa molecular weight cut-off was used to allow non-protein containing micelles through the filter [26]. Plasticware was avoided and, where necessary, was washed thoroughly with milliQ water to minimize PEG contamination and suppression of MS signals. Protein concentration was estimated by the Beer-Lambert law using an extinction coefficient of 2000 mM^{-1} cm^{-1} at 410 nm for the air equilibrated (oxidized) protein.

Engineered MtrC from *S. oneidensis* MR-1 with the signal peptide replaced by that from MtrB of the same organism and a C-terminal Strep II tag to facilitate purification was prepared as described by Lockwood et al. [25].

2.2. LC-MS

Ultra-pure reagents were used, bought as recommended by Laganowsky et al. [26]. Samples of ~10 μM MtrCAB in 0.3 mM DDM, 20 mM TRIS-HCl, pH 7.5 were denatured by dilution to ~1 μM MtrCAB with an aqueous mixture of 2 % (v/v) acetonitrile, 0.1 % (v/v) formic acid. Samples were loaded onto a ProSwift RP-1S column (4.6 × 50 mm) (Thermo Scientific) using an Ultimate 3000 uHPLC system (Dionex, Leeds, UK). Eluants were continuously infused into a Bruker microQTOF-QIII mass spectrometer, running Hystar (Bruker Daltonics, Coventry, UK), using positive mode electrospray ionization (ESI). During the first part of the protocol that eluted DDM, a 2-minute steep curved step gradient was applied to take the system from 98 % (v/v) water, 2 % (v/v) acetonitrile, 0.1 % (v/v) formic acid to 98 % (v/v) methanol, 2 % (v/v) acetonitrile, 0.1 % (v/v) formic acid, with a flow rate of 0.5 mL/min. At 98 % (v/v) methanol a 3-minute curved gradient was applied returning the system to the initial conditions. The flow rate

was slowed to 0.2 mL/min for 2 min at initial conditions before the eluate composition was changed to 50 % (v/v) acetonitrile, 50 % (v/v) water, 0.1 % (v/v) formic acid over 2 min (4 min for *S. oneidensis* MR-1 at 0.1 mL/min). Bound proteins were then resolved and eluted by a linear gradient to 100 % (v/v) acetonitrile, 0.1 % (v/v) formic acid over 10 min, flow rate 0.2 mL/min. The mass spectrometer was operated in positive ion mode and data acquired from 50 to 3000 *m/z* with the following parameters: Drying gas 240 °C, flow 8 L/min, nebulizer 1.8 Bar, capillary 4500 V, offset of 500 V, ion energy 5.0 eV, collision energy and radio frequency of 7.5 eV and 650 Vpp. The mass spectrometer was calibrated with ESI-L tuning mix (Agilent Technologies, California, USA). Compass Data Analysis, with Maximum Entropy v1.3, (Bruker Daltonics, Coventry) was used for processing of spectra under the LC peak. Exact masses are reported from peak centroids representing the isotope average neutral mass.

LC-MS of engineered MtrC from *S. oneidensis* MR-1 was as previously described [27,28]. The equipment was as for LC-MS of the MtrCAB complexes, specifically, a ProSwift RP-1S column (4.6 × 50 mm), Ultimate 3000 uHPLC system and Bruker microQTOF-QIII mass spectrometer, running Hystar, using positive mode electrospray ionization.

2.3. LC-MS/MS

Proteins were precipitated with acetone [29]. Pellets were resuspended in 1.5 % sodium deoxycholate (SDC) with 0.2 M EPPS (Merck), pH 8.0 and digested with sequencing grade trypsin (Promega) according to standard procedures. The SDC was precipitated by adjusting to 0.5 % formic acid, and the peptides were purified from the supernatant using C18 OMIX tips (Agilent).

Aliquots were analyzed by nanoLC-MS/MS on an Orbitrap Eclipse™ Tribrid™ mass spectrometer coupled to an UltiMate® 3000 RSLCnano LC system (Thermo Fisher Scientific, Hemel Hempstead, UK). The samples were loaded onto a pre-column (Acclaim™ PepMap™ NEO C18, 5 μm, 0.3 × 5 mm, Thermo) with 0.1 % trifluoracetic acid at 15 μL min^{-1} for 3 min. The trap column was then switched in-line with the analytical Aurora Frontier™ column (60 cm × 75 μm ID, 1.7 μm C18; IonOpticks, Victoria, Australia) using a 3 h gradient at a flow rate of 0.25 μL min^{-1}, as follows: 0–3 min at 3 % B (parallel to trapping); 3–148 min increase B to 45 %; followed by a ramp to 99 % B and re-equilibration to 3 % B. Data were acquired with the following mass spectrometer settings in positive ion mode: MS1/OT: spray voltage 1600 V, resolution 120 K, profile mode, mass range *m/z* 300–1800, normalized AGC 100 % for long run, fill time 50 ms; MS2/IT: data dependent analysis was performed using HCD fragmentation with the following parameters: 2 s cycle time in IT turbo mode, centroid mode, isolation window 1 Da, charge states 2–7, threshold $1e^4$, CE = 30, AGC target $1e^4$, dynamic inject time, dynamic exclusion 1 count, 15 s exclusion with a mass window of ±10 ppm.

The acquired raw data were converted to .mgf format with msconvert (v.3.0.2, Proteowizard, https://proteowizard.sourceforge.io/index.html) and processed in Mascot 2.8.2 (www.matrixscience.com) using an in-house server, with precursor tolerance of 6 ppm, fragment tolerance of 0.6 Da, semi-trypsin as the endoprotease with 2 missed cleavages allowed, variable modification oxidation (M), acetylation at the protein N-terminal and diacylation (556.4648, 548.4805 and 550.4961 Da) at cysteine residues, no fixed modification. These results were loaded in Scaffold (version 4) for visualization.

Using the Peptide Prophet algorithm with Scaffold delta-mass correction [30], peptide identifications of ≥95.0 % probability and protein identifications of ≥99.0 % probability were accepted with at least one identified peptide. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the Proteomics IDEntifications (PRIDE) [31–33] partner repository with the dataset identifier PXD044260 and https://doi.org/10.6019/PXD044260.

3. Results and discussion

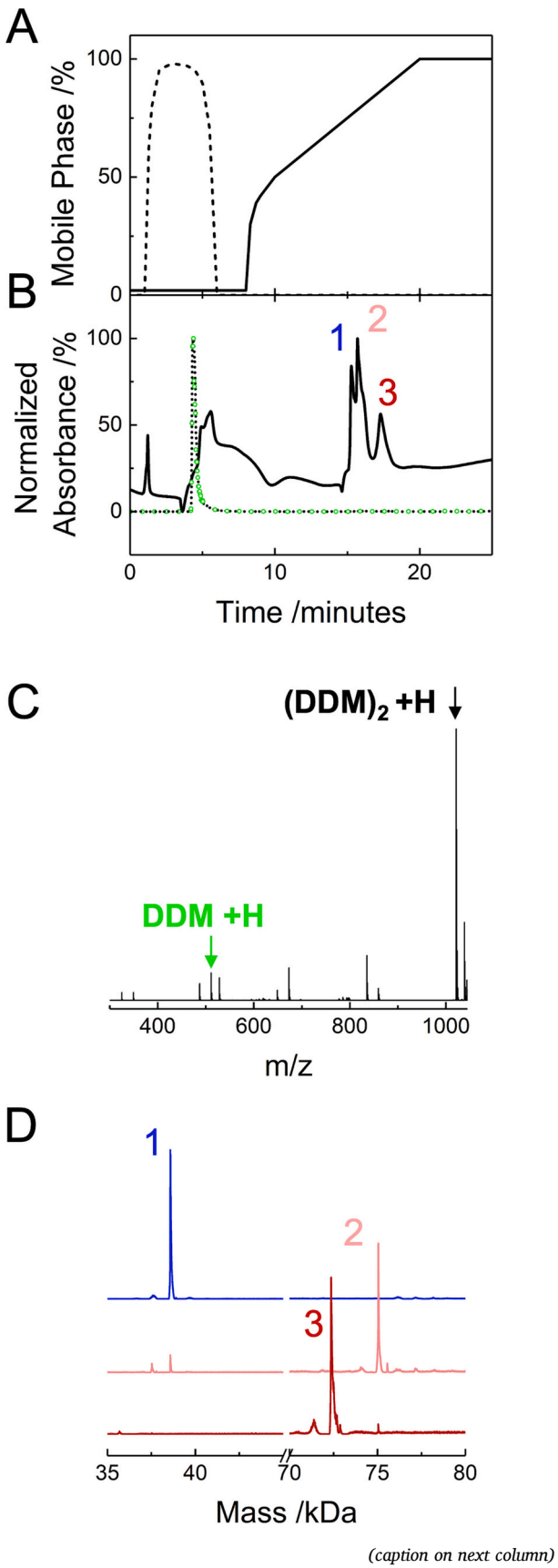
3.1. LC-MS resolves the intact mass of MtrA, MtrB and MtrC

The previously reported protocols to purify MtrCAB complexes from *S. baltica* OS185 [11] and *S. oneidensis* MR-1 [25] provide samples in 2 % Triton X-100, 20 mM Tris-HCl, pH 7.8. We considered that accurate LC-MS with these samples would be difficult because previous studies have found that Triton X-100 greatly suppresses the protein ionization essential for MS [34]. An alternative non-ionic detergent affording less suppression of protein ionization is DDM [22,24,26]. As a consequence, we sought a method for essentially complete removal of Triton X-100 that retained the aqueous solubility of MtrCAB with DDM. Attempts to use ion exchange chromatography for direct replacement of Triton X-100 with DDM were unsuccessful; the MtrCAB proteins remained bound to the column at >1 M NaCl and were likely denatured. Instead, the samples were exchanged into 5 mM lauryldimethylamine-N-oxide, 20 mM Tris-HCl, pH 7.5 using anion exchange chromatography. Then exchange into 0.3 mM DDM, 20 mM Tris-HCl, pH 7.5 was achieved using a centrifugal concentrator as detailed in Materials and Methods. SDS-PAGE, Fig. S2, confirmed the presence of MtrA, MtrB and MtrC in each of the resultant samples.

In developing our protocol for LC-MS of MtrCAB proteins we chose conditions compatible with analyses of water-soluble proteins that are routinely analyzed on the same instrument [27,35,36]. To characterize eluants from our protocol, mass spectra from different regions of the chromatogram were collected. This was achieved by continuous infusion into the mass spectrometer for positive mode ESI. The resulting data informed our selection of conditions for optimal separation of DDM from Mtr proteins, and resolution of the latter. Prior to application to the LC column described below, 10 μL of the MtrCAB complex (approximately 10 μM) was denatured by addition of 90 μL of 2 % acetonitrile, 0.1 % formic acid in water, Fig. 1 Step 1. All % values are volume/volume. We note that for native MS it is important to keep the DDM as micelles [26] but for LC-MS the proteins are denatured and the denaturation step used here takes the concentration of DDM below the critical micelle concentration of 0.15 mM.

Our LC-MS protocol employed a ProSwift RP-1S uHPLC column (4.6 × 50 mm) equilibrated with 2 % acetonitrile, 0.1 % formic acid in water. The diluted DDM-containing denatured MtrCAB samples were loaded to this column which bound all proteins and DDM detergent, Fig. 1 Step 2. The mobile phase in all subsequent steps included 0.1 % formic acid and its composition for LC-MS of MtrCAB from *S. baltica* OS185 is presented in Fig. 2A. In summary, to elute DDM from the column, the mobile phase was first taken to high methanol conditions: 98 % methanol, 2 % acetonitrile with a 2-minute gradient followed by a 3-minute gradient back to the starting conditions. Then the denatured proteins were eluted with a 2-minute gradient to 50 % acetonitrile, 50 % water followed by a linear 10-minute gradient that reached 100 % acetonitrile. The corresponding chromatogram for absorbance at 280 nm, Fig. 2B, shows elution in response to both the methanol and acetonitrile gradients.

To identify DDM elution in our protocol, Bruker Data Analysis 4.1 was used to predict the mass spectrum of the singly charged monomeric [DDM: $C_{24}H_{46}O_{11}$ + H, 511.3113 *m/z*] and dimeric [$(DDM)_2$: $(C_{24}H_{46}O_{11})_2$ + H, 1021.6153 *m/z*] forms from their chemical formulae, Fig. S3 (red lines). Inspection of mass spectra from across the chromatogram identified features typical of DDM only for the elution between 4 and 5 min corresponding to high methanol concentrations. A mass spectrum from this region of the chromatogram is shown in Fig. 2C. That spectrum is dominated by a feature at 1021.6144 *m/z* arising from $(DDM)_2$ + H and a feature at 511.3127 *m/z* reports the presence of DDM + H (mass error of <6 ppm ct predicted). Within these features the isotope patterns are in good agreement with those predicted by Bruker Data Analysis 4.1, Fig. S3. Some additional features in Fig. 2C can be attributed to breakdown fragments of DDM and $(DDM)_2$ (analysis not shown). To identify the regions of the chromatogram where DDM



(caption on next column)

Fig. 2. LC-MS of the membrane-associated MtrCAB complex from *S. baltica* OS185. A) Composition of the mobile phase: methanol (dashed black line) and acetonitrile (solid black line). B) Normalized LC chromatogram for absorbance at 280 nm (solid black line) and the base peak chromatogram using charge peaks, *m/z*: 511.31 (green circles) and 1021.62 (dashed black line) for DDM detergent. C) Mass spectrum of DDM (taken from LC elution at 4.3–4.4 min) recorded under optimized conditions for protein ion transmission rather than small molecules. Singly charged DDM + H monomer (green arrow, *m/z*: 511.31) and singly charged $(DDM)_2$ + H dimer (black arrow, *m/z* 1021.61). D) Deconvoluted mass spectra for MtrA (1. blue), MtrB (2. light pink) and MtrC (3. red) where the numbers correspond to the peaks labelled in the chromatogram of panel B. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

eluted, the base peak chromatograms were extracted for the 511.31 and 1021.62 ± 0.01 *m/z* features, Fig. 2B dashed lines. This analysis confirmed that elution of DDM was confined to regions of the chromatogram corresponding to high methanol concentration, illustrated schematically in Fig. 1 Step 3.

Elution of the MtrA, MtrB and MtrC proteins as peaks numbered 1 to 3 in the chromatogram of Fig. 2B, was revealed through a process detailed below and comparable to that which identified DDM elution as described above. For each protein the base peak chromatogram (data not shown) showed elution only in response to the acetonitrile gradient.

Intact masses for the Mtr proteins, Table 1, were predicted from their gene sequences, Fig. S1, and information in the crystal structure [11]. The *mtrA* gene of *S. baltica* OS185 encodes for a protein of 333 residues and the *mtrB* gene encodes for a protein of 695 amino acids. The N-terminal residues of both proteins are predicted to contain a signal peptide for transport to the periplasm by the Sec translocon [37,38], with the mature proteins produced by cleavage with Signal Peptidase I between residues 34 and 35 for MtrA and residues 21 and 22 for MtrB. On this basis the intact mass for MtrA was calculated by including the presence of ten covalently bound *c*-type hemes (615.17 g (mol heme)$^{-1}$ [39]) to give a predicted mass of 38,570 Da. The intact mass for MtrB was predicted as 75,052 Da. The *mtrC* gene encodes for a protein of 650 amino acids and the N-terminal residues are predicted to encode a lipoprotein signal peptide with cleavage by Signal Peptidase II [40] between residues 24 and 25. For this scenario the N-terminal residue of the matured protein would be lipidated Cys^{25} of the conserved lipobox motif (LxxC). On that basis the intact mass of the mature MtrC peptide with 10 *c*-type hemes but no lipidation is predicted as 71,812 Da.

Mass spectra corresponding to Peaks 1 to 3, Fig. 2B, were averaged to improve signal to noise, Fig. S4A. This revealed each peak to arise from an essentially homogeneous sample, Fig. 2D, with observed masses reported in Table 1. Peak 1 arose from protein with an intact mass of 38,572 Da such that it was assigned to MtrA. Peak 2 arose from protein with an intact mass of 75,050 Da and was assigned to MtrB. Peak 3 arose from protein with an intact mass of 72,364 Da such that it was not immediately reconciled with the predicted intact mass for MtrA, MtrB or MtrC, Table 1. We were unable to identify elution of protein with a mass of 71,812 Da, as predicted for MtrC with post-translational modification only through the addition of ten *c*-type hemes, Table 1. Nor could we identify elution of any additional species with mass in the range 30–90 kDa. We consider the most likely origin of Peak 3 is MtrC having 10 *c*-types hemes and an additional post-translational modification of 552 Da, noting that the our mass measurement are reproducible ±1 Da.

Having established an LC-MS method for separation of DDM and Mtr proteins of *S. baltica* OS185, an equivalent method was applied to MtrCAB purified from *S. oneidensis* MR-1, Fig. S5. Elution of DDM was again by a methanol gradient prior to resolved elution of the three Mtr proteins by acetonitrile gradient. Comparison of the predicted and observed intact masses, Table 1, allowed assignment of Peak 1 to MtrA and Peak 2 to MtrB. The protein eluted in Peak 3 is proposed to be MtrC with 10 *c*-types hemes and an additional post-translational modification of 560 Da. Inspection of the LC-MS data failed to detect the presence of

Table 1
Intact mass values for the Mtr proteins.

Protein		Mass/Da	
		S. baltica OS185	*S. oneidensis* MR-1
MtrA	Observed[a]	38,572	38,576[c]
	Predicted[b]	38,570	38,574
	Difference	+2	+2
MtrB	Observed[a]	75,050	75,478
	Predicted[b]	75,052	75,479
	Difference	−2	−1
MtrC	Observed[a]	72,364	75,531
	Predicted[b]	71,812	74,971
	Difference	+552	+560

[a] MtrA, MtrB and MtrC correspond to Peaks 1 to 3 respectively in Fig. 2B for *S. baltica* OS185 and in Fig. S5B for *S. oneidensis* MR-1. Values ± 1 Da from 3 measurements.

[b] The predicted masses for: MtrA proteins take residue 35 as the N-terminus and include 10 covalently bound *c*-type hemes; MtrB proteins take residue 22 as the N-terminus; MtrC proteins take residue 25 as the N-terminus and include 10 covalently bound *c*-type hemes.

[c] The main peak for MtrA is at 38,592 Da but is an oxygen adduct, the first peak without oxygen is at 38,576 Da, see Fig. S5D.

proteins other than those described here.

In closing this section we note that our initial attempts to provide intact mass values for the Mtr proteins of *S. oneidensis* MR-1 employed LC-MS with the ProSwift RP-1S column to study samples that had been pre-treated through acetone precipitation. The hope was to remove detergent and other solutes in the liquid phase prior to resolubilizing the denatured proteins for LC-MS. Elution used an acetonitrile gradient, Fig. S6A, typical of that used for LC-MS of soluble proteins. The resulting chromatogram, Fig. S6B, showed a single peak that could be attributed to an Mtr protein. That protein had an intact mass of 38,592 Da, Fig. S6C–D, such that it was attributed to MtrA. We concluded that the MtrB and MtrC proteins were being lost during the pre-treatment and developed the LC-MS method presented here in response to those findings.

3.2. LC-MS/MS of trypsin-digested MtrA, MtrB and MtrC

Evidence for post-translational modification of the Mtr proteins was sought from LC-MS/MS of the trypsin digested MtrCAB complexes. The mass spectrometry proteomics results are presented in the ProteomeXchange Consortium via the Proteomics IDEntifications (PRIDE) partner repository with dataset identifier PXD044260 and https://doi.org/10.6019/PXD044260 for both *S. baltica* OS185 and *S. oneidensis* MR-1.

For *S. baltica* OS185, peptide fragments corresponding to MtrA (64 % coverage), MtrB (94 % coverage) and MtrC (74 % coverage) were observed. For *S. oneidensis* MR-1, peptide fragments corresponding to MtrA (63 % coverage), MtrB (94 % coverage) and MtrC (87 % coverage) were observed. The lower coverage noted for the MtrA and MtrC proteins when compared to MtrB was largely due to the absence of fragments associated with peptides containing the CxxCH *c*-type heme binding motifs. None of the MtrC peptides that were detected provided evidence for post-translational modification(s) revealed by the intact mass analysis. In this context it was significant that the N-terminal cysteine (either with or without modification) was not detected in the N-terminal MtrC peptides from either *Shewanella* species. Thus, mass analysis of fragments from trypsin digestion is consistent with post-translational modification(s) of the N-terminal region of both *Shewanella* MtrC proteins but unable to provide greater insight into the nature or location of the modification(s).

3.3. LC-MS of N-terminally engineered S. oneidensis MR-1 MtrC

Previous work found the N-terminal residue of *S. oneidensis* MR-1 MtrC could not be identified by N-terminal sequencing methods [41,42]. In addition, *S. oneidensis* MR-1 grown anaerobically in the presence of the radiolabelled lipidic fatty acid ^{3}H-palmitoleic acid (16:1ω7) incorporated that radiolabel into MtrC [14]. Both observations are consistent with MtrC being lipidated at the N-terminal cysteine residue following processing by Signal Peptidase II. As a test of this hypothesis, we produced an engineered form of MtrC from *S. oneidensis* MR-1. In the gene for that protein [25], the lipoprotein signal peptide of MtrC was replaced with the secretory signal peptide of MtrB from the same organism and a C-terminal Strep II tag was introduced to facilitate purification. The DNA, and consequently the amino acid sequences Fig. S7A, of the engineered mature and genomic mature MtrC proteins are identical aside from the terminal modifications.

The engineered *mtrC* gene when included in a pBAD202D/TOPO vector and expressed in *S. oneidensis* MR-1 produced a soluble form of MtrC that was released to the spent media [25,28]. LC-MS, Fig. S7B, C, revealed the intact mass of the engineered MtrC protein to be 76,253 Da, in excellent agreement with the value of 76,252 Da predicted for the mature peptide having ten-covalently bound hemes and no further post-translational modification. This result is consistent with the +560 Da modification of genomically encoded MtrC of *S. oneidensis* MR-1 identified in this work arising from post-translational modification of the lipoprotein signal peptide.

3.4. Proposal for post-translational modification of the genomically encoded MtrC proteins of S. oneidensis MR-1 and S. baltica OS185

The post-translational modification of the MtrC proteins identified in this work is most reasonably assigned to lipidation of the N-terminal cysteine residue in accord with the proposals of previous researchers [14,41,42]. The biosynthetic pathway for lipoproteins in Gram-negative bacteria is well-described [40] and summarised in Fig. S8. For *S. oneidensis* MR-1 and *S. baltica* OS185 both diacyl and triacyl lipoproteins are feasible as the relevant enzymes are encoded in the corresponding genomes: Lgt as the product of SO_1334 and Shew185_1231; LspA as the product of SO_3531 and Shew185_1122; Lnt as the product of SO1177 and Shew185_3307 (for MR-1 and OS185 respectively). The acyl chains of glycophospholipids typically range from 14 to 24 carbons in length. Assuming lipidation by C14 chains, the minimum mass increase associated with diacylation is 495 Da and for triacylation is 706 Da. On this basis we conclude that the MtrC proteins are diacylated lipoproteins. For context, a mass increase of 547 Da would result from diacylation by palmitoleic acid, a C16 chain with a single unsaturated bond, previously shown to lipidate MtrC of *S. oneidensis* MR-1 [14].

AlphaFold [43] was used to investigate the possible arrangement of N-terminal MtrC residues in the *S. baltica* OS185 MtrCAB complex. For the residues resolved previously [11] by X-ray crystallography, Fig. 1, there was excellent agreement with the structure predicted by AlphaFold, Fig. S9. For the 19 N-terminal MtrC residues not seen in the crystal structure, AlphaFold does not predict secondary structure or interaction with the crystallographically defined proteins. Those predictions are consistent with an N-terminal peptide that is disordered in the protein crystals and consequently not detected by X-ray diffraction. It is of interest that AlphaFold predicts the 19 N-terminal residues of MtrC are directed towards the outer membrane such that the sulfur of the lipidated Cys could sit within approximately 6 Å of that bilayer, Fig. S9. That distance would allow the corresponding lipid to embed in the outer membrane and serve as a membrane anchor.

MtrC proteins of several *Shewanella* species are predicted to be lipoproteins based on their signal sequence, see for example Fig. S1C. Such lipidation may facilitate formation of the MtrCAB complex by ensuring the secreted MtrC proteins are retained at the cell surface to bind with the outer membrane embedded MtrAB heterodimer. However, lipidation can also impact the stability and activity of proteins by modulating the protein-membrane interaction in response to change of membrane composition and/or rigidity. Further experiments, beyond

the scope of this contribution, will be needed to distinguish which of these possibilities provides the primary driver for lipidation of MtrC.

4. Conclusions

An LC-MS protocol is presented that allows complete separation of DDM detergent from proteins of the MtrCAB integral membrane complex from *Shewanella* bacteria. The protocol allows for subsequent chromatographic resolution of MtrA, MtrB and MtrC and measurement of their intact mass by positive mode ESI. The MtrC, but not the MtrA and MtrB, proteins are revealed to be diacyl lipoproteins. We envisage the LC-MS protocol described herein could be readily optimized to allow accurate mass resolution of many other membrane proteins solubilized in DDM detergent. In this way information on the post-translational modification(s) of membrane proteins may be revealed more readily by LC-MS than other techniques.

CRediT authorship contribution statement

Jessica H. van Wonderen: Conceptualization, Methodology, Formal analysis, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Funding acquisition. **Jason C. Crack:** Methodology, Formal analysis, Writing – original draft, Writing – review & editing. **Marcus J. Edwards:** Resources, Writing – review & editing. **Thomas A. Clarke:** Writing – review & editing, Funding acquisition. **Gerhard Saalbach:** Methodology, Formal Analysis, Data Curation, Writing - review & editing. **Carlo Martins:** Methodology, Formal Analysis, Data Curation, Writing - review & editing. **Julea N. Butt:** Conceptualization, Data curation, Writing – original draft, Writing – review & editing, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Jessica van Wonderen reports financial support was provided by Biotechnology and Biological Sciences Research Council.

Data availability

Mass spectrometry proteomics results are presented in the ProteomeXchange Consortium via the Proteomics IDEntifications (PRIDE) partner repository with dataset identifier PXD044260 and https://doi.org/10.6019/PXD044260. Datasets used to make figures are deposited at Figshare (DOI: https://doi.org/10.6084/m9.figshare.21725531). For the purposes of open access, the author has applied a Creative Commons Attribution (CC BY) license to any author accepted manuscript version arising.

Acknowledgments

We thank Dr. Simone Payne (University of East Anglia) for purification of MtrCAB from *S. oneidensis* MR-1, Joshua Burton (University of East Anglia) for AlphaFold modelling of *S. baltica* OS185 MtrCAB, and Dr. Rod Chalk (Oxford) for helpful discussion. This research was funded by the U.K. Biotechnology and Biological Sciences Research Council (BBSRC), grant numbers BB/S002499/1 and BB/P01819X/1.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbamem.2023.184221.

References

[1] G.F. White, M.J. Edwards, L. Gomez-Perez, D.J. Richardson, J.N. Butt, T.A. Clarke, Mechanisms of bacterial extracellular electron exchange, Adv. Microb. Physiol. 68 (2016) 87–138, https://doi.org/10.1016/bs.ampbs.2016.02.002.
[2] T.A. Clarke, Plugging into bacterial nanowires: a comparison of model electrogenic organisms, Curr. Opin. Microbiol. 66 (2022) 56–62, https://doi.org/10.1016/j.mib.2021.12.003.
[3] J.T. Zhao, F. Li, Y.X. Cao, X.B. Zhang, T. Chen, H. Song, Z.W. Wang, Microbial extracellular electron transfer and strategies for engineering electroactive microorganisms, Biotechnol. Adv. 53 (2021), 107682, https://doi.org/10.1016/j.biotechadv.2020.107682.
[4] A. Kouzuma, Molecular mechanisms regulating the catabolic and electrochemical activities of Shewanella oneidensis MR-1, Biosci. Biotech. Bioch. 85 (7) (2021) 1572–1581, https://doi.org/10.1093/bbb/zbab088.
[5] L. Shi, H.L. Dong, G. Reguera, H. Beyenal, A.H. Lu, J. Liu, H.Q. Yu, J. K. Fredrickson, Extracellular electron transfer mechanisms between microorganisms and minerals, Nat. Rev. Microbiol. 14 (10) (2016) 651–662, https://doi.org/10.1038/nrmicro.2016.93.
[6] D.R. Lovley, D.E. Holmes, Electromicrobiology: the ecophysiology of phylogenetically diverse electroactive microorganisms, Nat. Rev. Microbiol. 20 (1) (2022) 5–19, https://doi.org/10.1038/s41579-021-00597-6.
[7] O.N. Lemaire, V. Mejean, C. Iobbi-Nivol, The *Shewanella* genus: ubiquitous organisms sustaining and preserving aquatic ecosystems, FEMS Microbiol. Rev. 44 (2) (2020) 155–170, https://doi.org/10.1093/femsre/fuz031.
[8] Y.G. Yang, M.Y. Xu, J. Guo, G.P. Sun, Bacterial extracellular electron transfer in bioelectrochemical systems, Process Biochem. 47 (12) (2012) 1707–1714, https://doi.org/10.1016/j.procbio.2012.07.032.
[9] J. Finkelstein, J. Swartz, M. Koffas, Bioelectrosynthesis systems, Curr. Opin. Biotech. 74 (2022) 211–219, https://doi.org/10.1016/j.copbio.2021.11.011.
[10] Y.D. Hu, Y.H. Wang, X. Han, Y.W. Shan, F. Li, L. Shi, Biofilm biology and engineering of *Geobacter* and *Shewanella* spp. for energy applications, Front. Bioeng. Biotech. 9 (2021), 786416, https://doi.org/10.3389/fbioe.2021.786416.
[11] M.J. Edwards, G.F. White, J.N. Butt, D.J. Richardson, T.A. Clarke, The crystal structure of a biological insulated transmembrane molecular wire, Cell 181 (3) (2020) 665–673, https://doi.org/10.1016/j.cell.2020.03.032.
[12] K. Thorell, J.P. Meier-Kolthoff, A. Sjoling, A.J. Martin-Rodriguez, Whole-genome sequencing redefines *Shewanella* taxonomy, Front. Microbiol. 10 (2019), https://doi.org/10.3389/fmicb.2019.01861.
[13] S. Ikeda, Y. Takamatsu, M. Tsuchiya, K. Suga, Y. Tanaka, A. Kouzuma, K. Watanabe, *Shewanella oneidensis* MR-1 as a bacterial platform for electro-biotechnology, Essays Biochem. 65 (2) (2021) 355–364, https://doi.org/10.1042/Ebc20200178.
[14] C.R. Myers, J.M. Myers, The outer membrane cytochromes of *Shewanella oneidensis* MR-1 are lipoproteins, Lett. Appl. Microbiol. 39 (5) (2004) 466–470, https://doi.org/10.1111/j.1472-765X.2004.01611.x.
[15] S. Mehmood, T.M. Allison, C.V. Robinson, Mass spectrometry of protein complexes: from origins to applications, Annu. Rev. Phys. Chem. 66 (2015) 453–474, https://doi.org/10.1146/annurev-physchem-040214-121732.
[16] D.P. Donnelly, C.M. Rawlins, C.J. DeHart, L. Fornelli, L.F. Schachner, Z.Q. Lin, J. L. Lippens, K.C. Aluri, R. Sarin, B.F. Chen, C. Lantz, W. Jung, K.R. Johnson, A. Koller, J.J. Wolff, I.D.G. Campuzano, J.R. Auclair, A.R. Ivanov, J.P. Whitelegge, L. Pasa-Tolic, J. Chamot-Rooke, P.O. Danis, L.M. Smith, Y.O. Tsybin, J.A. Loo, Y. Ge, N.L. Kelleher, J.N. Agar, Best practices and benchmarks for intact protein analysis for top-down mass spectrometry, Nat. Methods 16 (7) (2019) 587–594, https://doi.org/10.1038/s41592-019-0457-0.
[17] M.T. Agasid, C.V. Robinson, Probing membrane protein-lipid interactions, Curr. Opin. Struc. Biol. 69 (2021) 78–85, https://doi.org/10.1016/j.sbi.2021.03.010.
[18] A. Konijnenberg, J.F. van Dyck, L.L. Kailing, F. Sobott, Extending native mass spectrometry approaches to integral membrane proteins, Biol. Chem. 396 (9–10) (2015) 991–1002, https://doi.org/10.1515/hsz-2015-0136.
[19] A. Konijnenberg, L. Bannwarth, D. Yilmaz, A. Kocer, C. Venien-Bryan, F. Sobott, Top-down mass spectrometry of intact membrane protein complexes reveals oligomeric state and sequence information in a single experiment, Protein Sci. 24 (8) (2015) 1292–1300, https://doi.org/10.1002/pro.2703.
[20] S. Mehmood, J. Marcoux, J.T.S. Hopper, T.M. Allison, I. Liko, A.J. Borysik, C. V. Robinson, Charge reduction stabilizes intact membrane protein complexes for mass spectrometry, J. Am. Chem. Soc. 136 (49) (2014) 17010–17012, https://doi.org/10.1021/ja510283g.
[21] J.P. Whitelegge, C.B. Gundersen, K.F. Faull, Electrospray-ionization mass spectrometry of intact intrinsic membrane proteins, Protein Sci. 7 (6) (1998) 1423–1430, https://doi.org/10.1002/pro.5560070619.
[22] G. Berridge, R. Chalk, N. D'Avanzo, L.A. Dong, D. Doyle, J.I. Kim, X.B. Xia, N. Burgess-Brown, A. deRiso, E.P. Carpenter, O. Gileadi, High-performance liquid chromatography separation and intact mass analysis of detergent-solubilized integral membrane proteins, Anal. Biochem. 410 (2) (2011) 272–280, https://doi.org/10.1016/j.ab.2010.11.008.
[23] Y.Y. Dong, H. Wang, A.C.W. Pike, S.A. Cochrane, S. Hamedzadeh, F.J. Wyszynski, S.R. Bushell, S.F. Royer, D.A. Widdick, A. Sajid, H.I. Boshoff, Y. Park, R. Lucas, W. M. Liu, S.S. Lee, T. Machida, L. Minall, S. Mehmood, K. Belaya, W.W. Liu, A. Chu, L. Shrestha, S.M.M. Mukhopadhyay, C. Strain-Damerell, R. Chalk, N.A. Burgess-Brown, M.J. Bibb, C.E. Barry, C.V. Robinson, D. Beeson, B.G. Davis, E.P. Carpenter, Structures of DPAGT1 explain glycosylation disease mechanisms and advance TB antibiotic design, Cell 175 (4) (2018) 1045–1058, https://doi.org/10.1016/j.cell.2018.10.037.

[24] J.L. Lippens, P.F. Egea, C. Spahr, A. Vaish, J.E. Keener, M.T. Marty, J.A. Loo, L.D. G. Campuzano, Rapid LC-MS method for accurate molecular weight determination of membrane and hydrophobic proteins, Anal. Chem. 90 (22) (2018) 13616–13623, https://doi.org/10.1021/acs.analchem.8b03843.
[25] C.W.J. Lockwood, J.H. van Wonderen, M.J. Edwards, S.E.H. Piper, G.F. White, S. Newton-Payne, D.J. Richardson, T.A. Clarke, J.N. Butt, Membrane-spanning electron transfer proteins from electrogenic bacteria: production and investigation, Method Enzymol. 613 (2018) 257–275, https://doi.org/10.1016/bs.mie.2018.10.011.
[26] A. Laganowsky, E. Reading, J.T.S. Hopper, C.V. Robinson, Mass spectrometry of intact membrane protein complexes, Nat. Protoc. 8 (4) (2013) 639–651, https://doi.org/10.1038/nprot.2013.024.
[27] J.H. van Wonderen, D.B. Li, S.E.H. Piper, C.Y. Lau, L.P. Jenner, C.R. Hall, T. A. Clarke, N.J. Watmough, J.N. Butt, Photosensitised multiheme cytochromes as light-driven molecular wires and resistors, ChemBioChem 19 (20) (2018) 2206–2215, https://doi.org/10.1002/cbic.201800313.
[28] S.E.H. Piper, M.J. Edwards, J.H. van Wonderen, C. Casadevall, A. Martel, L.J. C. Jeuken, E. Reisner, T.A. Clarke, J.N. Butt, Bespoke biomolecular wires for transmembrane electron transfer: spontaneous assembly of a functionalized multiheme electron conduit, Front. Microbiol. 12 (2021), 714508, https://doi.org/10.3389/fmicb.2021.714508.
[29] J.L. Nickerson, A.A. Doucette, Rapid and quantitative protein precipitation for proteome analysis by mass spectrometry, J. Proteome Res. 19 (5) (2020) 2035–2042, https://doi.org/10.1021/acs.jproteome.9b00867.
[30] A. Keller, A.I. Nesvizhskii, E. Kolker, R. Aebersold, Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search, Anal. Chem. 74 (20) (2002) 5383–5392, https://doi.org/10.1021/ac025747h.
[31] E.W. Deutsch, N. Bandeira, Y. Perez-Riverol, V. Sharma, J.J. Carver, L. Mendoza, D. J. Kundu, S.B. Wang, C. Bandla, S. Kamatchinathan, S. Hewapathirana, B. S. Pullman, J. Wertz, Z. Sun, S. Kawano, S. Okuda, Y. Watanabe, B. MacLean, M. J. MacCoss, Y.P. Zhu, Y. Ishihama, J.A. Vizcaino, The ProteomeXchange consortium at 10 years: 2023 update, Nucleic Acids Res. 51 (2022) D1539–D1548, https://doi.org/10.1093/nar/gkac1040.
[32] Y. Perez-Riverol, J.W. Bai, C. Bandla, D. Garcia-Seisdedos, S. Hewapathirana, S. Kamatchinathan, D.J. Kundu, A. Prakash, A. Frericks-Zipper, M. Eisenacher, M. Walzer, S.B. Wang, A. Brazma, J.A. Vizcaino, The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences, Nucleic Acids Res. 50 (D1) (2022) D543–D552, https://doi.org/10.1093/nar/gkab1038.
[33] Y. Perez-Riverol, Q.W. Xu, R. Wang, J. Uszkoreit, J. Griss, A. Sanchez, F. Reisinger, A. Csordas, T. Ternent, N. Del-Toro, J.A. Dianes, M. Eisenacher, H. Hermjakob, J. A. Vizcaino, PRIDE Inspector Toolsuite: moving toward a universal visualization tool for proteomics data standard formats and quality assessment of proteomeXchange datasets, Mol. Cell. Proteomics 15 (1) (2016) 305–317, https://doi.org/10.1074/mcp.O115.050229.
[34] S. Opitz, F. Hannika, T. Kruger, H. Rhode, The removal of Triton X-100 by dialysis is feasible!, Anal. Bioanal. Chem. 407 (4) (2015) 1107–1118, https://doi.org/10.1007/s00216-014-8333-3.
[35] L.P. Jenner, J.C. Crack, J.M. Kurth, Z. Soldanova, L. Brandt, K.P. Sokol, E. Reisner, J.M. Bradley, C. Dahl, M.R. Cheesman, J.N. Butt, Reaction of thiosulfate dehydrogenase with a substrate mimic induces dissociation of the cysteine heme ligand giving insights into the mechanism of oxidative catalysis, J. Am. Chem. Soc. 144 (40) (2022) 18296–18304, https://doi.org/10.1021/jacs.2c06062.
[36] J.C. Crack, A.J. Thomson, N.E. Le Brun, Mass spectrometric identification of intermediates in the O_2-driven [4Fe-4S] to [2Fe-2S] cluster conversion in FNR, Proc. Natl. Acad. Sci. U. S. A. 114 (16) (2017) E3215–E3223, https://doi.org/10.1073/pnas.1620987114.
[37] A. Tsirigotaki, J. De Geyter, N. Sostaric, A. Economou, S. Karamanou, Protein export through the bacterial Sec pathway, Nat. Rev. Microbiol. 15 (1) (2017) 21–36, https://doi.org/10.1038/nrmicro.2016.161.
[38] K. Denks, A. Vogt, I. Sachelaru, N.A. Petriman, R. Kudva, H.G. Koch, The Sec translocon mediated protein transport in prokaryotes and eukaryotes, Mol. Membr. Biol. 31 (2–3) (2014) 58–84, https://doi.org/10.3109/09687688.2014.907455.
[39] E.D. Merkley, B.J. Anderson, J. Park, S.M. Belchik, L. Shi, M.E. Monroe, R.D. Smith, M.S. Lipton, Detection and identification of heme *c*-modified peptides by histidine affinity chromatography, high-performance liquid chromatography-mass spectrometry, and database searching, J. Proteome Res. 11 (12) (2012) 6147–6158, https://doi.org/10.1021/pr3007914.
[40] H. Nakayama, K. Kurokawa, B.L. Lee, Lipoproteins in bacteria: structures and biosynthetic pathways, FEBS J. 279 (23) (2012) 4247–4268, https://doi.org/10.1111/febs.12041.
[41] J.M. Myers, C.R. Myers, Role for outer membrane cytochromes OmcA and OmcB of *Shewanella putrefaciens* MR-1 in reduction of manganese dioxide, Appl. Environ. Microb. 67 (1) (2001) 260–269, https://doi.org/10.1128/Aem.67.1.260-269.2001.
[42] J.M. Myers, C.R. Myers, Isolation and sequence of *omcA*, a gene encoding a decaheme outer membrane cytochrome *c* of *Shewanella putrefaciens* MR-1, and detection of *omcA* homologs in other strains of *S. putrefaciens*, BBA-Biomembranes 1373 (1) (1998) 237–251, https://doi.org/10.1016/S0005-2736(98)00111-4.
[43] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Zidek, A. Potapenko, A. Bridgland, C. Meyer, S.A. A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, Nature 596 (7873) (2021) 583–589, https://doi.org/10.1038/s41586-021-03819-2.