High Throughput Expression Screening of Arabinofuranosyltransferases from Mycobacteria

José Rodrigues 1,†, Vanessa T. Almeida 1,†, Ana L. Rosário 1, Yong Zi Tan 2,3,4, Brian Kloss 5, Filippo Mancia 2,* and Margarida Archer 1,*

1 Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa (ITQB NOVA), 2780-157 Oeiras, Portugal; jose.rodrigues@itqb.unl.pt (J.R.); vcalmeida@itqb.unl.pt (V.T.A.); anarosar@gmail.com (A.L.R.)
2 Department of Physiology and Cellular Biophysics, Columbia University Irving Medical Center, New York, NY 10032, USA; yongzi.tan@sickkids.ca
3 National Resource for Automated Molecular Microscopy, Simons Electron Microscopy Center, New York Structural Biology Center, New York, NY 10027, USA
4 Molecular Medicine Program, The Hospital for Sick Children, Toronto, ON M5G 1X8, Canada
5 Center on Membrane Protein Production and Analysis, New York Structural Biology Center, New York, NY 10027, USA; bkloss@nysbc.org
* Correspondence: fm123@cumc.columbia.edu (F.M.); archer@itqb.unl.pt (M.A.)
† These authors contributed equally to this publication.

Abstract: Studies on membrane proteins can help to develop new drug targets and treatments for a variety of diseases. However, membrane proteins continue to be among the most challenging targets in structural biology. This uphill endeavor can be even harder for membrane proteins from Mycobacterium species, which are notoriously difficult to express in heterologous systems. Arabinofuranosyltransferases are involved in mycobacterial cell wall synthesis and thus potential targets for antituberculosis drugs. A set of 96 mycobacterial genes coding for Arabinofuranosyltransferases was selected, of which 17 were successfully expressed in E. coli and purified by metal-affinity chromatography. We herein present an efficient high-throughput strategy to screen in microplates a large number of targets from Mycobacteria and select the best conditions for large-scale protein production to pursue functional and structural studies. This methodology can be applied to other targets, is cost and time effective and can be implemented in common laboratories.

Keywords: membrane proteins; overexpression in E. coli; protein purification; high-throughput protocol; Arabinofuranosyltransferases; Mycobacteria

1. Introduction

Membrane proteins represent 20 to 30% of open-reading frames of all genomes sequenced [1,2] and perform essential functions in cells, such as transportation, signal transduction and energy production [3]. They also play important roles in several diseases and, as a result, are attractive therapeutic targets, estimated to represent more than 30% of all marketed drugs [4–6]. However, biochemical and structural characterization of membrane proteins have several bottlenecks, namely toxicity by excess of mRNA levels of the target protein [7], toxicity caused by heterologous expression [8], membrane lipid composition [9,10], detergent extraction and solubility [11,12], which ultimately results in low amounts of membrane protein produced.

Many efforts have been devoted on the development of protocols to efficiently produce membrane proteins in Escherichia coli. An elegant approach to accelerate this process involves the fusion of green fluorescent protein (GFP) to monitor the expression and purification processes [13,14]. A commonly used strategy consists of varying different parameters simultaneously, such as expression vectors with different tags and promoters, host strains, homologues or solubilizing detergents [15,16]. High-throughput (HTP) protein
production platforms have also been developed by Structural Genomics Consortia, such as JCSG [17], Northeast SGC [18] or the New York Consortium on Membrane Protein Structure (NYCOMPS, New York, NY, USA) [19,20].

Based on the HTP strategy to express prokaryotic membrane proteins previously developed at NYCOMPS, we have devised an accessible protocol to screen mycobacterial membrane proteins, which are difficult to express in heterologous systems [21]. An initial set of 96 target genes was assembled from the genomic sequences coding for Arabinofuranosyltransferases (AraTs) from 14 different Mycobacterium species. Even though Mycobacteria (Mb) proteins display low yield of production in E. coli, it remains one of preferred hosts for heterologous expression of Mb proteins and is compatible with HTP strategies.

The AraT targets are integral membrane enzymes that play a pivotal role in the synthesis of arabinan, an important component of Mycobacteria cell envelope [22,23]. AraTs transfer arabinose in the furanose conformation (Araf) from a single donor decaprenylphosphoryl-β-D-arabinose (DPA) to the arabinan domain of arabinogalactan (AG) of mycobacterial cell wall [24]. The cell envelope is crucial for growth and virulence of pathogenic Mycobacteria [25] and is a major contributor to resistance against common antibiotics [26]. AraT family comprises 7 sub-families: Embs (EmbA, EmbB and EmbC), which are inhibited by the first-line antitubercular drug Ethambutol (EMB), and Afts (AftA, AftB, AftC and AftD), which are potential novel therapeutic targets against tuberculosis.

In this study, we present a simple and cost-effective methodology to screen AraTs from different Mycobacterium species and chose the most promising targets to proceed biochemical and structural studies.

2. Materials and Methods

2.1. High-Throughput Cloning of Arabinofuranosyltransferase Genes

96 genes of AftA, AftB, AftC, AftD, EmbA, EmbB and EmbC were identified from 14 Mycobacterium genomes using a bioinformatics approach [19] (Table A1). Ligation Independent Cloning (LIC) was used to clone all selected targets, using the protocols previously described by Bruni and Kloss [27], with the following modifications: (1) target sequences were amplified by Polymerase Chain Reaction (PCR) using oligonucleotides that were compatible with LIC-adapted expression vectors (pNYCOMPS-N23 and pNYCOMPS-C23) containing appropriate overhangs with the start codon ATG, instead of the endogenous GTG start codon for some of the target genes; (2) XL10 E. coli was used for cloning purposes. All liquid handling was performed by hand, using multichannel pipettes. The resulting constructs were used to transform E. coli C41, C43 and BL21 (DE3) pLysS strains. Transformants were selected on LB [10 g/L Tryptone, 5 g/L Yeast Extract and 10 g/L NaCl] plates containing appropriate antibiotics: 100 µg/mL of ampicillin for E. coli C41 and C43 cells, 100 µg/mL of ampicillin and 34 µg/mL of chloramphenicol for E. coli BL21 (DE3) pLysS cells.

2.2. High-Throughput Expression Screening and Purification

Pre-cultures were grown overnight at 37 °C, 200 rpm in 600 µL of LB medium supplemented with the appropriate antibiotics. For this purpose, 96-well plates (VWR) were used. The overnight cultures were used to inoculate 2.5 mL of 2xYT medium [16 g/L Tryptone, 10 g/L Yeast Extract and 5 g/L NaCl], supplemented with antibiotics, in 24-well plates (UNIPLATE Collection and Analysis Microplate), at an initial optical density (OD600) of 0.03–0.08. Cells were grown at 37 °C, 200 rpm, until the cultures reached OD600 of 0.4–1.2 (2 to 2.5 h) (see Table A2), then cultures were cooled to 22 °C and gene expression was induced overnight (~16 h) with 0.25 mM isopropyl-β-D-thiogalactoside (IPTG). Cells were harvested the next day by centrifugation at 3200 × g, for 20 min at 4 °C. OD600 measurements were done in TECAN Spark 10 M. Cell pellets were re-suspended in 300 µL of lysis buffer (BugBuster supplemented with 0.1 mg/mL Lysozyme, 3 U/mL Benzonase, 2 mM MgCl₂ and 0.5 mM PMSF) through 10 min of vigorous shaking, using a plate shaker, at room temperature. Detergent n-Dodecyl β-D-maltoside (DDM) was added to the lysate at
1% (w/v) final concentration and the plate containing the samples was incubated for 2 h at 4 °C, with gentle agitation. For the separation of the insoluble cell debris, plates were centrifuged at 3200 × g for 20 min, 4 °C and 250 µL of the supernatants were transferred to a 96-well filter plate containing a bed of 50 µL Ni-NTA agarose resin (HisPur™ Ni-NTA Spin plate (Thermo Scientific™, Waltham, MA, USA)), previously washed with double distilled water and equilibrated with buffer (20 mM HEPES pH 7.5, 200 mM NaCl, 10 mM Imidazole, 0.1% DDM). Imidazole at 10 mM final concentration was added to each sample, to avoid unspecific binding of contaminants to the Ni-NTA resin. Plates were incubated for 15 min in a plate-shaker at 4 °C. The plates were then centrifuged, the flowthrough fractions collected and reloaded to the resin bed, repeating the 15 min incubation with the Ni-NTA resin bed. After the second incubation step, plates were centrifuged to remove unbound proteins. The resin was washed three times with 250 µL washing buffer (20 mM HEPES pH 7.5, 200 mM NaCl, 60 mM Imidazole and 0.1% DDM) and finally eluted with 250 µL of elution buffer (20 mM HEPES pH 7.5, 200 mM NaCl, 300 mM Imidazole and 0.05% DDM). Eluted samples were run on SDS-PAGE: 10% polyacrylamide gels were used for targets with molecular weights between 73–149 kDa (AftD, EmbA, EmbB and EmbC) and 12% polyacrylamide gels for 47–75 kDa targets (AftA, AftB and AftC).

2.3. Large Scale Protein Expression and Purification

50 mL cultures of each target in E. coli C41 cells were grown overnight in 250 mL flasks at 37 °C, 200 rpm, in LB medium supplemented with 100 µg/mL of ampicillin. The overnight cultures were used to inoculate 4 × 500 mL of 2xYT medium, supplemented with 100 µg/mL of ampicillin, in 2.5 L Thomson’s Ultra Yield™ Flasks (Oceanside, CA, USA), at an initial OD₆₀₀ around 0.05. Cells were grown at 37 °C, 200 rpm, until the cultures reached OD₆₀₀ of 0.8 (2 to 2.5 h), then cultures were cooled to 22 °C and gene expression was induced overnight (~16 h) with 0.25 mM IPTG. Cells were harvested in the next day by centrifugation at 4472 × g, for 15 min at 4 °C. OD₆₀₀ measurements were done in Ultrospec 10 Cell Density Meter. Cell pellets were re-suspended and homogenized in lysis buffer (20 mM HEPES pH 7.5, 200 mM NaCl, 20 mM MgSO₄, 1 mM TCEP), protease inhibitor EDTA-free cocktail (Thermo Scientific™, Waltham, USA; Catalog number: 88266) and 25 U/mL of Benzonase nuclease (Santa Cruz Biotechnology, Dallas, USA; Catalog number: sc-391121). Cell suspension was passed twice at 15,000 psi on a cell disruptor (Constant Systems Ltd., Daventry, UK). Membranes were collected by ultracentrifugation at 197,215 × g, for 30 min at 4 °C. Membranes were manually homogenized using a Wheaton® glass homogenizer (DWK Life Sciences Limited, Stoke-on-Trent, UK) in 20 mM HEPES pH 7.5 and 200 mM NaCl, to which DDM was added to a final concentration of 1% (w/v). Membranes were solubilized for 2 h, with gentle agitation, at 4 °C. Soluble membrane fraction was collected by ultracentrifugation at 203,756 × g for 30 min, at 4 °C. The supernatants were collected and incubated with 2 mL of equilibrated Ni-NTA agarose resin for 1.5 h at 4 °C, with gentle agitation. Imidazole was added to each sample to a final concentration of 10 mM, to prevent unspecific binding of contaminants. After incubation, the sample was loaded into a column for elution by gravity flow. The resin bed was washed with 10 column volumes (CV) with washing buffer (20 mM HEPES pH 7.5, 200 mM NaCl, 0.1% DDM, 60 mM Imidazole), and the proteins were eluted with 4 CV of elution buffer (20 mM HEPES pH 7.5, 200 mM NaCl, 0.05% DDM and 300 mM Imidazole). Eluted samples were concentrated and injected into a Superdex 200 column (Cytiva Europe GmbH, Freiburg, Germany) to assess protein dispersity. The collected fractions were run on SDS-PAGE: 10% polyacrylamide for 73–149 kDa targets (AftD, EmbA, EmbB and EmbC) and 12% polyacrylamide for those around 47–75 kDa (AftA, AftB and AftC).
3. Results

3.1. Genomic Expansion and High-Throughput Cloning of Arabinofuranosyltransferases

A set of 96 target genes was assembled from the genomic sequences coding for seven AraTs from *Mtb* (AftA, AftB, AftC, AftD, EmbA, EmbB and EmbC). To each “seed” sequence, a “cluster” of homolog sequences, from 14 different *Mycobacterium* genomes was expanded, coding for proteins likely to have similar structure as the seed protein [28] (Table A1). LIC was performed as described previously by Bruni and Kloss [27]. Briefly, all sequences were amplified by PCR, using genomic DNA available from ATCC® (Manassas, VA, USA) [https://www.lgcstandards-atcc.org (accessed on 26 March 2021)] and primer pairs compatible with LIC-adapted expression vectors (pNYCOMPS-N23 and pNYCOMPS-C23) that contained decahistidine affinity tag and Tobacco Etch Virus (TEV) protease cleavage site (ENLYFQS). 56 targets were successfully cloned into pNYCOMPS-N23 and 40 targets into pNYCOMPS-C23. Previous screening experiments had shown no expression for all constructs in pNYCOMPS-C23 vector (data not shown), therefore only the clones in pNYCOMPS-N23 were used for the HTP expression screening approach.

3.2. Small Scale High-Throughput Expression of Arabinofuranosyltransferases

All 56 positive clones in pNYCOMPS-N23 were transformed into C41, C43 and BL21 (DE3) *E. coli* strains. 24 deep-well plates were used to grow the positive clones simultaneously. Growth conditions, 2xYT rich medium, 0.25 mM IPTG and overnight post-induction at 22 °C, were established based on the results obtained in previous experiments. This allowed a fast and reliable comparison between different *E. coli* strains, also leaving room for optimization after target selection.

Cell harvesting by centrifugation and lysis were performed in 24 deep-well plates, maintaining a HTP downstream processing of the samples. Extraction of membrane proteins was achieved by adding detergent directly to each well, after cell lysis, incubating the plate at low temperature. The 24 deep-well plate is centrifuged again to clear the solubilized lysate from the cell debris. The solubilized lysate was transferred to a HisPur™ Ni-NTA Spin 96 well plate (Thermo Scientific™, Waltham, MA, USA) for affinity chromatography purification. In this step, the use of adjustable multichannel pipettes to transfer solutions from 24-well plate to 96-well plate was important for sake of speed and reproducibility/reliability, however, standard multichannel pipettes can also be used although not in an optimal manner. After a single Ni-NTA purification step, the amount of eluted target protein was too low to be detected by SDS-PAGE. Since the sample solution slowly flows from the filter plate by gravity during the incubation period, a second passage was deemed necessary to increase the contact time between the sample and Ni-NTA resin, after which the eluted AraTs could be visualized on the gel. The full pipeline is summarized in Figure 1.

In total, 17 out of 96 distinct proteins were produced and purified, resulting in 18% success rate of protein production (Table 1). All three different *E. coli* host strains were able to produce target proteins: 16 in C41, 6 in C43 and 8 in BL21 (DE3) pLysS (Table A2). AftB and EmbC proteins were not detected in any *E. coli* strain using this HTP method, suggesting that different, perhaps more tailored conditions may be needed to successfully produce these proteins. We found that a single His-tag purification step was not very efficient, considering that persistent contaminants from the host cell are present across all targets (Figure 2; see Figures A1 and A2). Moreover, the production yields for the target proteins herein studied were low in all *E. coli* host strains. Nevertheless, we were still able to successfully identify bands in the SDS-PAGE that could correspond to our target proteins, based on their predicted molecular weight (MW) and considering the gel shifting for membrane proteins in denaturing protein gels [29]. Due to this anomalous migration pattern, bands related to membrane proteins in SDS-PAGE most often appear ~20–30% below their predicted MW.
to successfully identify bands in the SDS-PAGE that could correspond to our target proteins, based on their predicted molecular weight (MW) and considering the gel shifting for membrane proteins in denaturing protein gels [29]. Due to this anomalous migration pattern, bands related to membrane proteins in SDS-PAGE most often appear ~20–30% below their predicted MW.

Figure 1. Schematic representation of the high-throughput screening for membrane protein production and purification.

Table 1. Summary of high-throughput screening of membrane protein production.

|                      | Number | Success Rate (%) |
|----------------------|--------|------------------|
| Targets              | 96     | -                |
| Positive Clones      | 56     | 58               |
| (pNYCOMPS-N23)       |        |                  |
| Positive Clones      | 40     | 42               |
| (pNYCOMPS-C23)       |        |                  |
| Proteins purified (total) | 17 | 18               |
| E. coli C41::pNYCOMPS-N23 | 16 | 17               |
| E. coli C43::pNYCOMPS-N23 | 6  | 6                |
| E. coli BL21 (DE3)   | 8      | 8                |
| pLysS::pNYCOMPS-N23  |        |                  |

1 Expression of these constructs was not determined in this study.
3.3. Validation of HTP Target Selection by Large-Scale Protein Production

Based on SDS-PAGE analysis, we selected one target from each cluster for large-scale production: AftA and AftC from *M. neoaurum*, AftD from *M. abscessus* 1948 F5/8, EmbA from *M. marinum* M. and EmbB from *M. vanbaalenii* PYR-1, all produced in *E. coli* C41. Growth conditions were similar to the ones used in the HTP screening, although cell lysis and membrane extraction steps were modified according to the cell mass. Most importantly, the incubation time of solubilized membranes with Ni-NTA resin was increased to improve protein binding and purification yield. All chosen targets were successfully produced in large scale, thus validating the selection made from the HTP screening. Size exclusion chromatography (SEC) was performed after affinity chromatography to further purify the protein and as tool for preliminary biophysical characterization of each protein (Figure 3). Although all targets show some aggregation in the presence of DDM, it was still possible to identify heterogeneous protein populations in most target samples. Upon SDS-PAGE analysis of the SEC elution fractions, we observed that the dominant protein bands correspond to the desired targets, however there were still contaminants present. AftA (Figure 3A), EmbA (Figure 3D) and EmbB (Figure 3E) showed the least amount of contaminant proteins.
Figure 3. Size-exclusion chromatography (SEC) elution profiles of large-scale expression experiments of AraTs and respective SDS-PAGE. Vertical arrows indicate void volume; asterisks (*) and (†) indicate different populations observed in the SEC elution profiles. Ni: Ni-NTA elution samples; HTP: Small-scale HTP results; horizontal arrows indicate the bands corresponding to the target protein of interest. (A) AftA from *M. neoaurum*, (B) AftC from *M. neoaurum*, (C) AftD from *M. abscessus* 1948 F5/8, (D) EmbA from *M. marinum* M. and (E) EmbB from *M. vanbaalenii* PYR-1.

For EmbB, a second SEC step was performed (Figure 4), running each population separately. We could observe that the high molecular weight EmbB population behaves as a stable monodisperse population (Figure 4A,B), while the low molecular weight EmbB population splits into the same two populations observed in the first SEC run (Figure 3E), suggesting that EmbB monomers are prone to form an equilibrium with stable EmbB dimers.
Figure 4. Size-exclusion chromatography elution profiles of EmbB from *M. vanbuilenii* PYR-1. (A) SEC elution profiles of EmbB: after Ni-NTA (full black line), peak (*) after 1st SEC (red dashed lines), peak (†) after 1st SEC (spotted blue line). (B) Normalized SEC elution profiles of EmbB: after Ni-NTA (full black line), peak (*) after 1st SEC (red dashed lines), peak (†) after 1st SEC (spotted blue line). (C) SDS-PAGE analysis of peak (*) and (†) SEC. Vertical arrows indicate void volume. Colored asterisks correspond to the peaks observed in each SEC run, respective to different EmbB populations. Arrow indicates the corresponding band of EmbB (115 kDa).

4. Discussion

The need to screen the expression of a large number of membrane protein targets, as well as the selection of optimal conditions for production and purification of desired targets, led to the development of several HTP strategies. The strategy used in this study is not novel and was intended to set up a protocol to search for the best candidates to pursue functional and structural studies on AraTs from *Mycobacteria*. EmbS are targets of ethambutol, whereas Afts are potential targets to develop new drugs to treat tuberculosis. Nevertheless, the protocol herein described can also be applied to evaluate the expression and purification of other membrane proteins.

The methodology involved the selection of 13–14 orthologue genes of each AraT subfamily (EmbA-C, AftA-D) from a variety of host genomes, gene expression with vectors harboring a poly-histidine affinity tag at either N- or C-terminus, transformation into three different *E. coli* strains, membrane extraction and protein solubilization by DDM detergent, and purification by Ni-NTA chromatography. By using this simple combinatorial approach, we were able to clone 56 genes at pNYCOMPS-N23 and 40 at pNYCOMPS-C23, and produce 17 proteins out of 96 chosen targets, corresponding to a success rate of 18%. Such rate is not surprising, considering that membrane proteins are often difficult to express and purify [7,30]. Heterologous expression of mycobacterial proteins in *E. coli* has previously been reported not to exceed 40% [30–32].

No expression of AraTs cloned into pNYCOMPS-C23 was observed (data not shown). It is well known that type and location of the fused affinity tag has a significant effect
at all stages of protein production [33], however it is not possible to know a priori the impact caused by tag addition. This unpredictability is somehow the foundation of HTP approaches—try as many conditions as is reasonably possible and assess what works to proceed with further studies.

Concerning the host organism, *M. smegmatis* could be a viable alternative for the heterologous expression of *Mb* proteins [31], yet we considered it not appropriate for a HTP approach due to its slower growth rate compared to *E. coli*, and mostly due to its waxy surface [34], which promotes clumping, film formation and cell adhesion to surfaces, especially plastic, preventing an optimal use of 96- or 24-well plates for cell growth. Instead, we used *E. coli* C41 and C43 suited for overexpression of toxic and membrane proteins [35], and BL21 (DE3) pLysS for controlled expression [36]. The different expression levels observed among *E. coli* hosts suggests that the type of strain plays a pivotal role in the number of well-expressed AraTs, as also reported for other target proteins [37]. Indeed, regulation of T7 RNA polymerase expression either by mutations in its promoter (C41 and C43 Walker strains) or by its natural inhibitor T7 lysozyme (T7Lys, pLysS strain) can significantly influence membrane protein overexpression yields [38].

Large scale production (from 2 L of culture) of 5 targets, chosen based on the results of small-scale experiments, yielded purified proteins in milligram amounts, although differences are observed on the intensity of their respective bands (SDS-PAGE) in small and large-scale experiments. The shorter incubation time with the Ni-NTA agarose in the small-scale screening (2 × 15 min vs. 1.5 h) could account for this discrepancy. Moreover, the aeration rate related with the size and shape of the growth vessel (24-well plate vs. 2.5 L flasks) may also affect the overexpression levels. In addition, switching to a cobalt spin plate, instead of nickel, may increase binding specificity of the target protein and thus further improve the results for the small-scale screening. We expect similar results will be obtained on the scale-up production of other targets that showed expression on the HTP screening. Fusions with GFP tag could be advantageous to monitor the various steps of protein production by measuring fluorescence, a very sensitive detection method [13]. However, this methodology is not suitable for membrane proteins with periplasmic C-terminus [13,20,39], which is the case for most of the AraTs herein studied, so it was not considered.

Detergents are required to extract and purify target proteins and their choice is a key parameter on the entire process. We chose to use DDM since it is a mild detergent and one of the most commonly used for this purpose [20,27,33,40]. The aggregation detected in the large-scale production experiments suggests further detergent screens may be needed to select the best detergent formula for each individual target. We were able to separate two different populations of EmbB from *M. vanbaalenii* PYR-1, likely constituted by monomer and dimer, respectively. Despite the aggregation, all targets showed soluble populations in the SEC elution profiles, which represents a good starting point for optimization towards structural studies. Interestingly, 3D structures of both oligomeric states have been already characterized by single particle cryo-electron microscopy (cryo-EM) for EmbB from *M. smegmatis* [41,42]. Noteworthy, different detergents or solubilizing agents may be needed for structural studies. The cryo-EM structures of several AraTs, namely EmbA-EmbB complex, *M. tuberculosis* EmbB [43], *M. smegmatis* EmbB [42] and AftD [44], have been recently characterized and different solubilization agents were used, namely glyco-diosgenin (GDN) detergent, amphipols or nanodiscs. Interestingly, the structure of EmbC solubilized in DDM has been determined by X-ray crystallography [43].

On one hand, the production of AftB and EmbC targets was not achieved using the HTP workflow with “standard” conditions. Therefore, other parameters must be explored, such as growth media, temperature, incubation time and type of detergent, host strains or expression vectors, which will likely lead to better success rates. On the other hand, AftA and AftC from *M. neoaurum* were expressed and are attractive targets for drug development [45–47] and structural elucidation, since their structures are not yet known.
The presented protocol stands as a simplified approach based in previous HTP strategies developed at NYCOMPS [19,20] and by others [15], to identify the best candidates for further biochemical and structural studies in a fast and affordable manner for most laboratories. Most importantly, this methodology delivers similar results to conventional medium throughput approaches and, by reducing the variables throughout, it allows target optimization for large-scale protein production.

**Author Contributions:** Conceptualization, M.A., F.M., B.K. and J.R.; methodology, J.R., V.T.A., A.L.R. and Y.Z.T.; data analysis, all authors; original draft preparation, J.R. and V.T.A.; review and editing, all; supervision and funding acquisition, M.A. and F.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Fundação para a Ciência e Tecnologia (FCT), Lisbon, Portugal; grants (PTDC/BIA-BQM/30421/2017 and PTDC/BIA-BQM/4056/2020 to M.A. and PD/BD/128261/2016 to J.R.), European Union’s Horizon 2020 research and innovation programme under grant agreements: No. 857203 (Twinning), No. 823780 (MSCA-RISE) and No. 731005, Instruct-ULTRA, a project to further develop the services of Instruct-ERIC (M.A, J.R. and V.T.A.). This work was also supported by National Institutes of Health (NIH), Bethesda, MD, USA; grants GM132120 (to F.M.) and GM116799 (to Wayne A. Hendrickson).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** We thank Alessio Bortoluzzi, Catarina Paiva and Sandra Santos for advice and help with equipment for high throughput protein expression.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Appendix A.**

**Table A1.** Protein targets used in the high-throughput screening and expression results for each *E. coli* strain tested.

| ID  | GenBank ID      | Organism                                      | Predicted Protein | C41 | C43 | BL21 (DE3) pLysS |
|-----|-----------------|-----------------------------------------------|-------------------|-----|-----|------------------|
| A1  | SIU02450.1      | *Mycobacterium bovis* AF2122/97               | AftA              | −   | −   | −                |
| A8  | EUA63955.1      | *Mycobacterium abscessus* 1948                | AftA              | −   | −   | +                |
| C2  | ABP43658.1      | *Mycobacterium gilvum* PYR-GCK                | AftA              | +   | −   | −                |
| C9  | AFC41461.1      | *Mycobacterium intracellulare* ATCC 13950     | AftA              | −   | −   | −                |
| F1  | ABM16394.1      | *Mycobacterium vanbaalenii* PYR-1             | AftA              | +   | −   | +                |
| F3  | AGP61782.1      | *Mycobacterium yongonense* 05-1390            | AftA              | −   | −   | −                |
| F7  | CDQ43571.1      | *Mycobacterium neoaurum*                     | AftA              | +   | −   | +                |
| F8  | AAS02550.1      | *Mycobacterium aurum* subsp. paratuberculosis K-10 | AftA          | −   | −   | −                |
| A2  | SIU02464.1      | *Mycobacterium bovis* AF2122/97               | AftB              | −   | −   | −                |
| A9  | EUA63936.1      | *Mycobacterium abscessus* 1948                | AftB              | −   | −   | −                |
| C7  | CCP46634.1      | *Mycobacterium tuberculosis* H37Rv            | AftB              | −   | −   | −                |
| D9  | AGP61763.1      | *Mycobacterium yongonense* 05-1390            | AftB              | −   | −   | −                |
| D10 | AFC41442.1      | *Mycobacterium intracellulare* ATCC 13950     | AftB              | −   | −   | −                |
| D11 | ABM16411.1      | *Mycobacterium vanbaalenii* PYR-1             | AftB              | −   | −   | −                |
| E2  | AIR19061.1 *    | *Mycobacterium kansasii* 662                 | AftB              | −   | −   | −                |
| E5  | CDQ43590.1      | *Mycobacterium neoaurum*                     | AftB              | −   | −   | −                |
Table A1. Cont.

| ID   | GenBank ID       | Organism                               | Predicted Protein | C41 | C43 | BL21 (DE3) pLysS |
|------|------------------|----------------------------------------|-------------------|-----|-----|------------------|
| A3   | AMC65006.1       | Mycobacterium bovis AF2122/97          | AftC              | +   | −   | +                |
| A10  | EUA61591.1       | Mycobacterium abscessus 1948           | AftC              | +   | −   | −                |
| C11  | AFC44620.1       | *Mycobacterium intracellular*          | AftC              | +   | +   | −                |
| C12  | AGZ53302.1       | *Mycobacterium kansasi* ATCC 12478     | AftC              | +   | +   | −                |
| D1   | ABP46386.1       | Mycobacterium gilvum PYR-GCK           | AftC              | +   | +   | +                |
| D2   | CDQ43952.1       | Mycobacterium neoaurum                | AftC              | +   | +   | +                |
| E6   | EHB50241.1       | *Mycobacterium rhodesiae* JS60         | AftC              | −   | −   | −                |
| E7   | AFM16967.1       | *Mycobacterium chubunense* NBB4        | AftC              | −   | −   | −                |
| E8   | AAS05110.1       | *Mycobacterium avium* subsp. paratuberculosis K-10 | AftC | +   | +   | −                |
| F9   | ABM13300.1       | Mycobacterium vanbaelenii PYR-1        | AftC              | −   | −   | −                |
| A4   | CAB527947.1      | Mycobacterium bovis AF2122/97          | AftD              | −   | −   | −                |
| D4   | AGZ51741.1       | *Mycobacterium kansasi* ATCC 12478     | AftD              | −   | −   | −                |
| H11  | EUA63217.1       | Mycobacterium abscessus 1948 F5/8      | AftD              | +   | −   | +                |
| A5   | SIU02452.1       | Mycobacterium bovis AF2122/97          | EmbA              | −   | −   | −                |
| A11  | EUA63951.1       | Mycobacterium abscessus 1948           | EmbA              | +   | −   | −                |
| B4   | AAC45280.1       | Mycobacterium tuberculosis H37Rv       | EmbA              | −   | −   | −                |
| B8   | ACC43760.1       | *Mycobacterium marinum* M              | EmbA              | +   | −   | −                |
| C5   | ABP43656.1       | *Mycobacterium gilvum* PYR-GCK         | EmbA              | −   | −   | −                |
| G5   | AAS02546.1       | *Mycobacterium avium* subsp. paratuberculosis K-10 | EmbA | −   | −   | −                |
| G11  | CDQ43576.1       | *Mycobacterium neoaurum*              | EmbA              | −   | −   | −                |
| H1   | ABM16396.1       | Mycobacterium vanbaelenii PYR-1        | EmbA              | −   | −   | −                |
| H2   | AGZ51276.1       | *Mycobacterium kansasi* ATCC 12478     | EmbA              | −   | −   | −                |
| A6   | SIU02453.1       | Mycobacterium bovis AF2122/97          | EmbB              | −   | −   | −                |
| A12  | EUA63949.1       | Mycobacterium abscessus 1948           | EmbB              | +   | −   | −                |
| B3   | AAC45281.1       | Mycobacterium tuberculosis H37Rv       | EmbB              | −   | −   | −                |
| B7   | ACC43761.1       | *Mycobacterium marinum* M              | EmbB              | −   | −   | −                |
| C4   | ABP43655.1       | *Mycobacterium gilvum* PYR-GCK         | EmbB              | −   | −   | −                |
| G3   | AFC41455.1       | *Mycobacterium intracellular*          | EmbB              | −   | −   | −                |
| G6   | AAS02545.1       | *Mycobacterium avium* subsp. paratuberculosis K-10 | EmbB | −   | −   | −                |
| G9   | AGP61776.1       | Mycobacterium yongonense 05-1390       | EmbB              | −   | −   | −                |
| H7   | KEP38884.1       | *Mycobacterium kansasi*               | EmbB              | −   | −   | −                |
| H9   | ABM16397.1       | Mycobacterium vanbaelenii PYR-1        | EmbB              | +   | −   | +                |
| B1   | EUA63954.1       | Mycobacterium abscessus 1948           | EmbC              | −   | −   | −                |
| C3   | ABP43657.1       | *Mycobacterium gilvum* PYR-GCK         | EmbC              | −   | −   | −                |
| G4   | AFC41460.1       | *Mycobacterium intracellular*          | EmbC              | −   | −   | −                |
| G7   | AAS02549.1       | *Mycobacterium avium* subsp. paratuberculosis K-10 | EmbC | −   | −   | −                |
| G10  | AGP61781.1       | Mycobacterium yongonense 05-1390       | EmbC              | −   | −   | −                |
| H4   | CDQ43572.1       | *Mycobacterium neoaurum*              | EmbC              | −   | −   | −                |
| H5   | AGZ51274.1       | *Mycobacterium kansasi* ATCC 12478     | EmbC              | −   | −   | −                |

* Record removed.
Table A2. List of genes coding for AraTs from several *Mycobacterium* species selected for cloning and expression screening.

| ID | GenBank ID | Organism | Predicted Protein | % Identity \((M. \text{tuberculosis})\) | Predicted Molecular Weight (kDa) | Predicted Transmembrane Helixes |
|----|------------|----------|-------------------|---------------------------------|--------------------------------|-------------------------------|
| A1 | SIU02450.1 | *Mycobacterium bovis* AF2122/97 | AftA | 100 | 70 | 13 |
| A2 | SIU02464.1 | *Mycobacterium bovis* AF2122/97 | AftB | 99 | 69 | 9 |
| A3 | AMC65006.1 | *Mycobacterium bovis* AF2122/97 | AftC | 99 | 49 | 8 |
| A4 | CAB5247947.1 | *Mycobacterium bovis* AF2122/97 | AftD | 99 | 146 | 9 |
| A5 | SIU02452.1 | *Mycobacterium bovis* AF2122/97 | EmbA | 99 | 116 | 13 |
| A6 | SIU02453.1 | *Mycobacterium bovis* AF2122/97 | EmbB | 99 | 118 | 13 |
| A7 | SIU02451.1 | *Mycobacterium bovis* AF2122/97 | EmbC | 99 | 118 | 13 |
| A8 | EUA63955.1 | *Mycobacterium abscessus* 1948 | AftA | 65 | 68 | 13 |
| A9 | EUA63936.1 | *Mycobacterium abscessus* 1948 | AftB | 67 | 71 | 10 |
| A10 | EUA61591.1 | *Mycobacterium abscessus* 1948 | AftC | 64 | 47 | 8 |
| A11 | EUA63951.1 | *Mycobacterium abscessus* 1948 | EmbA | 65 | 114 | 12 |
| A12 | EUA63949.1 | *Mycobacterium abscessus* 1948 | EmbB | 68 | 73 | 8 |
| B1 | EUA63954.1 | *Mycobacterium abscessus* 1948 | EmbC | 68 | 117 | 11 |
| B2 | AAC45279.1 | *Mycobacterium tuberculosis* H37Rv | EmbC | 100 | 117 | 13 |
| B3 | AAC45281.1 | *Mycobacterium tuberculosis* H37Rv | EmbB | 100 | 118 | 12 |
| B4 | AAC45280.1 | *Mycobacterium tuberculosis* H37Rv | EmbA | 100 | 116 | 13 |
| B5 | CCP49261.1 | *Mycobacterium tuberculosis* H37Rv | AftD | 100 | 146 | 9 |
| B6 | ACC43759.1 | *Mycobacterium marinum* M | EmbC | 86 | 117 | 14 |
| B7 | ACC43761.1 | *Mycobacterium marinum* M | EmbB | 89 | 116 | 12 |
| B8 | ACC43760.1 | *Mycobacterium marinum* M | EmbA | 87 | 118 | 13 |
| B9 | AFM19671.1 | *Mycobacterium chibunese* NBB4 | EmbB | 72 | 115 | 13 |
| B10 | AFM19669.1 | *Mycobacterium chibunese* NBB4 | EmbC | 75 | 116 | 12 |
| B11 | AFM19670.1 | *Mycobacterium chibunese* NBB4 | EmbA | 69 | 115 | 13 |
| B12 | AFM19668.1 | *Mycobacterium chibunese* NBB4 | AftA | 67 | 71 | 10 |
| C1 | ACC43758.1 | *Mycobacterium marinum* M | AftA | 83 | 70 | 13 |
| C2 | ABP43658.1 | *Mycobacterium guilium* PYR-GCK | AftA | 68 | 67 | 13 |
| C3 | ABP43657.1 | *Mycobacterium guilium* PYR-GCK | EmbC | 74 | 115 | 14 |
| C4 | ABP43655.1 | *Mycobacterium guilium* PYR-GCK | EmbB | 70 | 115 | 13 |
| C5 | ABP43656.1 | *Mycobacterium guilium* PYR-GCK | EmbA | 69 | 114 | 12 |
| C6 | AIU11367.1 | *Mycobacterium smegmatis* str. MC2 155 | AftA | 68 | 67 | 12 |
| C7 | CCP46634.1 | *Mycobacterium tuberculosi* H37Rv | AftB | 100 | 69 | 9 |
| C8 | AAS02532.1 | *Mycobacterium avium* subsp. Paratuberculosis | AftB | 82 | 70 | 9 |
| C9 | AFC41461.1 | *Mycobacterium intracellulare* ATCC 13950 | AftA | 77 | 68 | 13 |
| C10 | ABK72123.1 | *Mycobacterium smegmatis* str. MC2 155 | AftC | 70 | 49 | 8 |
| C11 | AFC44620.1 | *Mycobacterium intracellulare* ATCC 13950 | AftC | 84 | 50 | 8 |
| C12 | AGZ53302.1 | *Mycobacterium kansasii* ATCC 12478 | AftC | 89 | 49 | 8 |
Table A2. Cont.

| ID | GenBank ID | Organism | Predicted Protein | % Identity (M. tuberculosis) | Predicted Molecular Weight (kDa) | Predicted Transmembrane Helixes |
|----|------------|----------|-------------------|-----------------------------|---------------------------------|--------------------------------|
| D1 | ABP46386.1 | *Mycobacterium gilvum* PYR-GCK | AftC | 70 | 48 | 9 |
| D2 | CDQ43952.1 | *Mycobacterium neoaurum* | AftC | 70 | 48 | 8 |
| D3 | ABK71542.1 | *Mycobacterium smegmatis* str. MC2 155 | AftD | 71 | 148 | 12 |
| D4 | AGZ51741.1 | *Mycobacterium kansasii* ATCC 12478 | AftD | 82 | 148 | 13 |
| D5 | ACC38960.1 | *Mycobacterium marinum* | AftD | 80 | 146 | 7 |
| D6 | AFM15049.1 | *Mycobacterium chubuense* NBB4 | AftD | 70 | 145 | 12 |
| D7 | ADT97050.1 | *Mycobacterium gilvum* Spyr1 | AftD | 70 | 148 | 13 |
| D8 | ABP43645.1 | *Mycobacterium gilvum* PYR-GCK | AftB | 71 | 70 | 9 |
| D9 | AGP61763.1 | *Mycobacterium yongonense* 05-1390 | AftB | 82 | 72 | 10 |
| D10 | AFC41442.1 | *Mycobacterium intracellulare* ATCC 13950 | AftB | 82 | 72 | 10 |
| D11 | ABM16411.1 | *Mycobacterium vanbaalenii* PYR-1 | AftB | 70 | 72 | 9 |
| D12 | EHB54870.1 | *Mycobacterium rhodesiae* JS60 | AftB | 68 | 73 | 10 |
| E1 | CDM79377.1 | *Mycobacterium marinum* E11 | AftB | 85 | 73 | 10 |
| E2 | AIR19061.1 | *Mycobacterium kansasii* 662 | AftB | 86 | 73 | 10 |
| E3 | ABK75671.1 | *Mycobacterium smegmatis* str. MC2 155 | AftB | 73 | 70 | 9 |
| E4 | AFM19681.1 | *Mycobacterium chubuense* NBB4 | AftB | 68 | 72 | 11 |
| E5 | CDQ43590.1 | *Mycobacterium neoaurum* | AftC | 67 | 67 | 11 |
| E6 | EHB50241.1 | *Mycobacterium rhodesiae* JS60 | AftC | 71 | 49 | 7 |
| E7 | AFM16967.1 | *Mycobacterium chubuense* NBB4 | AftC | 71 | 49 | 9 |
| E8 | AAS05110.1 | *Mycobacterium avium* subsp. paratuberculosis K-10 | AftC | 83 | 49 | 8 |
| E9 | AGP64972.1 | *Mycobacterium yongonense* 05-1390 | AftC | 83 | 50 | 8 |
| E10 | ACC40492.1 | *Mycobacterium marinum* | AftC | 87 | 50 | 8 |
| E11 | AAS06236.1 | *Mycobacterium avium* subsp. paratuberculosis K-10 | AftD | 79 | 145 | 9 |
| E12 | EHB55421.1 | *Mycobacterium rhodesiae* JS60 | AftD | 71 | 149 | 13 |
| F1 | ABM16394.1 | *Mycobacterium vanbaalenii* PYR-1 | AftA | 67 | 68 | 13 |
| F2 | CDQ42439.1 | *Mycobacterium neoaurum* F5/8 | AftD | 67 | 148 | 13 |
| F3 | AGP61782.1 | *Mycobacterium yongonense* 05-1390 | AftA | 77 | 68 | 13 |
| F4 | EHB54850.1 | *Mycobacterium rhodesiae* JS60 | AftA | 71 | 66 | 13 |
| F5 | AGZ51271.1 | *Mycobacterium kansasii* ATCC 12478 | AftA | 84 | 70 | 13 |
| F6 | AGP66385.1 | *Mycobacterium yongonense* 05-1390 | AftD | 80 | 146 | 13 |
| F7 | CDQ43571.1 | *Mycobacterium neoaurum* | AftA | 67 | 67 | 11 |
| F8 | AAS02550.1 | *Mycobacterium avium* subsp. paratuberculosis K-10 | AftA | 79 | 75 | 11 |
| F9 | ABM13300.1 | *Mycobacterium vanbaalenii* PYR-1 | AftC | 69 | 48 | 9 |
| ID  | GenBank ID    | Organism                                      | Predicted Protein | % Identity (M. tuberculosis) | Predicted Molecular Weight (kDa) | Predicted Transmembrane Helixes |
|-----|---------------|-----------------------------------------------|-------------------|------------------------------|----------------------------------|---------------------------------|
| F10 | ABM11102.1    | Mycobacterium vanbaalenii PYR-1               | AftD              | 70                           | 147                              | 10                              |
| F11 | AFP42646.1    | Mycobacterium smegmatis str. MC2 155         | EmbA              | 69                           | 117                              | 13                              |
| F12 | ABK72840.1    | Mycobacterium smegmatis str. MC2 155         | EmbB              | 69                           | 117                              | 13                              |
| G1  | ABK72375.1    | Mycobacterium smegmatis str. MC2 155         | EmbC              | 75                           | 115                              | 10                              |
| G2  | AFC41456.1    | Mycobacterium intracellulare ATCC 13950      | EmbA              | 83                           | 115                              | 14                              |
| G3  | AFC41455.1    | Mycobacterium intracellulare ATCC 13950      | EmbB              | 85                           | 115                              | 12                              |
| G4  | AFC41460.1    | Mycobacterium intracellulare ATCC 13950      | EmbC              | 85                           | 114                              | 13                              |
| G5  | AAS02546.1    | Mycobacterium avium subsp. paratuberculosis K-10 | EmbA            | 83                           | 117                              | 14                              |
| G6  | AAS02545.1    | Mycobacterium avium subsp. paratuberculosis K-10 | EmbB            | 84                           | 115                              | 12                              |
| G7  | AAS02549.1    | Mycobacterium avium subsp. paratuberculosis K-10 | EmbC            | 85                           | 117                              | 13                              |
| G8  | AGP61777.1    | Mycobacterium yongonense 05-1390             | EmbA              | 83                           | 116                              | 14                              |
| G9  | AGP61776.1    | Mycobacterium yongonense 05-1390             | EmbB              | 85                           | 115                              | 12                              |
| G10 | AGP61781.1    | Mycobacterium yongonense 05-1390             | EmbC              | 85                           | 114                              | 13                              |
| G11 | CDQ43576.1    | Mycobacterium neoaurum                      | EmbA              | 69                           | 116                              | 13                              |
| G12 | EHB54852.1    | Mycobacterium rhodesiae JS60                 | EmbA              | 72                           | 116                              | 13                              |
| H1  | ABM16396.1    | Mycobacterium vanbaalenii PYR-1              | EmbA              | 70                           | 116                              | 13                              |
| H2  | AGZ51276.1    | Mycobacterium kansasii ATCC 12478            | EmbA              | 88                           | 115                              | 13                              |
| H3  | AEV72559.1    | Mycobacterium rhodesiae NBB3                 | EmbC              | 76                           | 114                              | 14                              |
| H4  | CDQ43572.1    | Mycobacterium neoaurum                      | EmbC              | 72                           | 115                              | 13                              |
| H5  | AGZ51274.1    | Mycobacterium kansasii ATCC 12478            | EmbC              | 88                           | 117                              | 14                              |
| H6  | ABM16395.1    | Mycobacterium vanbaalenii PYR-1              | EmbC              | 73                           | 116                              | 10                              |
| H7  | KEP38884.1    | Mycobacterium kansasii                      | EmbB              | 90                           | 117                              | 11                              |
| H8  | AEV72557.1    | Mycobacterium rhodesiae NBB3                 | EmbB              | 76                           | 114                              | 13                              |
| H9  | ABM16397.1    | Mycobacterium vanbaalenii PYR-1              | EmbB              | 70                           | 115                              | 13                              |
| H10 | AHC23140.2    | Mycobacterium neoaurum VKM Ac-1815D         | EmbB              | 69                           | 115                              | 13                              |
| H11 | EUA63217.1    | Mycobacterium abscessus 1948 F5/8            | AftD              | 62                           | 149                              | 12                              |
| H12 | AFC46048.1    | Mycobacterium intracellulare ATCC 13950      | AftD              | 79                           | 146                              | 13                              |

* Record removed.
Table A3. Optical Density at 600 nm (OD$_{600}$) of small-scale cultures. In-plate target coordinates.

|     | AftA | A1  | A8  | C2   | C9   | F1  | F3  | F7  | F8   |
|-----|------|-----|-----|------|------|-----|-----|-----|------|
| AftB| 0.92 | 0.99| 1.04| 0.85 | 1.03 | 1.05| 0.80| 1.10|      |
| AftC| 0.93 | 1.00| 0.72 | 0.80 | 1.05 | 0.84 | 0.88| 1.07|      |
| AftD| 0.93 | 0.72 | 0.79 | 0.74 | 0.93 | 0.93 | 1.16| 1.11|      |

Table A4. OD$_{600}$ upon induction of expression. Color scheme: green color—OD$_{600}$ interval (0.5 – 1.0), red color—OD$_{600}$ below 0.5 or above 1.0. The black contoured squares show the values obtained for the targets selected for large scale.

|     | BL21 (DE3) pLysS |      |      |      |      |      |      |      |      |
|-----|-----------------|------|------|------|------|------|------|------|------|
| AftA| 0.92            | 0.99 | 1.04 | 0.85 | 1.03 | 1.05 | 0.80 | 1.10 |      |
| AftB| 0.93            | 1.00 | 0.82 | 0.87 | 1.00 | 0.75 | 0.78 | 0.82 | 0.58 |
| AftC| 0.93            | 1.00 | 0.72 | 0.79 | 0.72 | 0.80 | 1.05 | 0.84 | 0.88 |
| AftD| 0.93            | 1.00 | 0.72 | 0.79 | 0.72 | 0.80 | 1.05 | 0.84 | 0.88 |

Table A5. Variation of OD$_{600}$ between the induction time and cell harvesting. A graduated scale of three colors was applied to these values, indicating average values in white, the lower values in red and higher values in green. The black contoured squares show the values obtained for the targets selected for large scale.

|     | BL21 (DE3) pLysS |      |      |      |      |      |      |      |      |
|-----|-----------------|------|------|------|------|------|------|------|------|
| AftA| 0.92            | 0.99 | 1.04 | 0.85 | 1.03 | 1.05 | 0.80 | 1.10 |      |
| AftB| 0.93            | 1.00 | 0.82 | 0.87 | 1.00 | 0.75 | 0.78 | 0.82 | 0.58 |
| AftC| 0.93            | 1.00 | 0.72 | 0.79 | 0.72 | 0.80 | 1.05 | 0.84 | 0.88 |
| AftD| 0.93            | 1.00 | 0.72 | 0.79 | 0.72 | 0.80 | 1.05 | 0.84 | 0.88 |

Table A6. Variation of OD$_{600}$ between the induction time and cell harvesting. A graduated scale of three colors was applied to these values, indicating average values in white, the lower values in red and higher values in green. The black contoured squares show the values obtained for the targets selected for large scale.
Figure A1. Protein expression results from the HTP screening of 56 AraTs from Mycobacteria, overexpressed in E. coli C43 cells. A—AftA proteins: 1—A1; 2—A8; 3—C2; 4—C9; 5—F1; 6—F3; 7—F7; 8—F8. B—AftB proteins 1—A2; 2—A9; 3—C7; 4—D9; 5—D10; 6—D11; 7—E2; 8—E5. C—AftC proteins: 1—A3; 2—A10; 3—C11; 4—C12; 5—D1; 6—D2; 7—E6; 8—E7; 9—E8; 10—E9; 11—F9. D—AftD and EmbA proteins: 1—A4; 2—D4; 3—H11; 4—A5; 5—A11; 6—B4; 7—B8; 8—C5; 9—G5; 10—G11; 11—H1; 12—H2. E—EmbB proteins: 1—A6; 2—A12; 3—B3; 4—B7; 5—C4; 6—G3; 7—G6; 8—G9; 9—H7; 10—H9. F—EmbC proteins: 1—B1; 2—C3; 3—G4; 4—G7; 5—G10; 6—H4; 7—H5. Asterisks (*) indicate protein bands corresponding to the predicted MW of the corresponding target.

Figure A2. Protein expression results from the HTP screening of 56 AraTs from Mycobacteria, overexpressed in E. coli BL21 (DE3) pLysS cells. A—AftA proteins: 1—A1; 2—A8; 3—C2; 4—C9; 5—F1; 6—F3; 7—F7; 8—F8. B—AftB proteins 1—A2; 2—A9; 3—C7; 4—D9; 5—D10; 6—D11; 7—E2; 8—E5. C—AftC proteins: 1—A3; 2—A10; 3—C11; 4—C12; 5—D1; 6—D2; 7—E6; 8—E7; 9—E8; 10—E9; 11—F9. D—AftD and EmbA proteins: 1—A4; 2—D4; 3—H11; 4—A5; 5—A11; 6—B4; 7—B8; 8—C5; 9—G5; 10—G11; 11—H1; 12—H2. E—EmbB proteins: 1—A6; 2—A12; 3—B3; 4—B7; 5—C4; 6—G3; 7—G6; 8—G9; 9—H7; 10—H9. F—EmbC proteins: 1—B1; 2—C3; 3—G4; 4—G7; 5—G10; 6—H4; 7—H5. Asterisks (*) indicate protein bands corresponding to the predicted MW of the corresponding target.
References

1. Wallin, E.; Von Heijne, G. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci.* 1998, 7, 1029–1038. [CrossRef] [PubMed]

2. Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E.L. Predicting transmembrane protein topology with a hidden markov model: Application to complete genomes. Edited by F. Cohen. *J. Mol. Biol.* 2001, 305, 567–580. [CrossRef] [PubMed]

3. Von Heijne, G. The membrane protein universe: What’s out there and why bother? *J. Intern. Med.* 2007, 261, 543–557. [CrossRef]

4. Yasi, E.A.; Kruyer, N.S.; Peralta-Yahya, P. Advances in G protein-coupled receptor high-throughput screening. *Curr. Opin. Biotechnol.* 2020, 64, 210–217. [CrossRef] [PubMed]

5. Arinaminpathy, Y.; Khurana, E.; Engelman, D.M.; Gerstein, M.B. Computational analysis of membrane proteins: The largest class of drug targets. *Drug Discov. Today* 2009, 14, 1130–1135. [CrossRef] [PubMed]

6. Davey, J. G-Protein-Coupled Receptors: New Approaches to Maximise the Impact of GPCRs in Drug Discovery. *Expert Opin. Ther. Targets* 2004, 8, 165–170. [CrossRef] [PubMed]

7. Schlegel, S.; Löblöm, J.; Lee, C.; Hjelms, K.; Strous, M.; Drew, D.; Slotboom, D.J.; de Gier, J.-W. Optimizing Membrane Protein Overexpression in the Escherichia coli strain Lemo21 (DE3). *J. Mol. Biol.* 2012, 423, 648–659. [CrossRef]

8. Gubellini, F.; Verdon, G.; Karpowich, N.K.; Luff, J.D.; Boël, G.; Gauthier, N.; Handelman, S.K.; Ades, S.E.; Hunt, J.F. Physiological Response to Membrane Protein Overexpression in *E. coli*. *Mol. Cell. Proteom.* 2011, 10. [CrossRef]

9. Hunte, C. Specific protein–lipid interactions in membrane proteins. *Biochem. Soc. Trans.* 2005, 33, 938–942. [CrossRef]

10. Lee, A.G. How lipids affect the activities of integral membrane proteins. *Biochim. Biophys. Acta Biomembr.* 2004, 1666, 62–87. [CrossRef]

11. Hardy, D.; Mandon, E.D.; Rothnie, A.J.; Jawhari, A. The yin and yang of solubilization and stabilization for wild-type and full-length membrane protein. *Methods* 2018, 147, 118–125. [CrossRef]

12. Smith, S.M. Strategies for the Purification of Membrane Proteins. *Methods Mol. Biol.* 2016, 1485, 389–400. [CrossRef]

13. Drew, D.; Lerch, M.; Kunji, E.; Slotboom, D.J.; de Gier, J.-W. Optimization of membrane protein overexpression and purification using GFP fusions. *Nat. Chem. Biol.* 2006, 3, 303–313. [CrossRef]

14. Bond, L.E.; Rada, H.; Verma, A.; Gasper, R.; Birch, J.; Jennions, M.; Löwe, J.; Moraes, I.; Owens, R.J. Green fluorescent protein-based expression screening of membrane proteins in Escherichia coli. *J. Vis. Exp.* 2015, 95, e52357. [CrossRef]

15. Eschagh, S.; Hedrén, M.; Nasser, M.I.A.; Hammarberg, T.; Thornell, A.; Norgren, P. An efficient strategy for high-throughput expression screening of recombinant integral membrane proteins. *Protein Sci.* 2005, 14, 676–683. [CrossRef] [PubMed]

16. Ma, P.; Varela, F.; Magoch, M.; Silva, A.R.; Rosário, A.L.; Brito, J.; Oliveira, T.F.; Nogly, P.; Pessanha, M.; Stelter, M.; et al. An Efficient Strategy for Small-Scale Screening and Production of Archaeal Membrane Transport Proteins in Escherichia coli. *PLoS ONE* 2013, 8, e76913. [CrossRef]

17. Eilsliger, M.A.; Deacon, A.M.; Godzik, A.; Lesley, S.A.; Wooley, J.; Wüthrich, K.; Wilson, I.A. The JCSG high-throughput structural biology pipeline. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 2010, 66, 1137–1142. [CrossRef] [PubMed]

18. Xiao, R.; Anderson, S.; Aramini, J.; Belote, R.; Buchwald, W.A.; Ciccossanti, C.; Conover, K.; Everett, J.K.; Hamilton, K.; Huang, Y.J.; et al. The high-throughput protein sample production platform of the Northeast Structural Genomics Consortium. *J. Struct. Biol.* 2010, 172, 21–33. [CrossRef] [PubMed]

19. Love, J.; Mancia, F.; Shapiro, L.; Punta, M.; Rost, B.; Girvin, M.; Wang, D.-N.; Zhou, M.; Hunt, J.F.; Sypererski, T.; et al. The New York Consortium on Membrane Protein Structure (NYCOMPS): A high-throughput platform for structural genomics of integral membrane proteins. *J. Struct. Funct. Genom.* 2010, 11, 191–199. [CrossRef] [PubMed]

20. Mancia, F.; Love, J. High-throughput expression and purification of membrane proteins. *J. Struct. Biol.* 2010, 172, 85–93. [CrossRef]

21. Bellinzoni, M.; Riccardi, G. Techniques and Applications: The heterologous expression of Mycobacterium tuberculosis genes is an uphill road. *Trends Microbiol.* 2003, 11, 351–358. [CrossRef]

22. Jankute, M.; Grover, S.; Rana, A.K.; Besra, G.S. Arabinogalactan and lipoarabinomannan biosynthesis: Structure, biogenesis and their potential as drug targets. *Future Microbiol.* 2012, 7, 129–147. [CrossRef] [PubMed]

23. Umesiri, F.E.; Sanki, A.K.; Boucak, J.; Ronning, D.R.; Sucheck, S.J. Recent advances toward the inhibition of mAG and LAM synthesis in Mycobacterium tuberculosis. *Med. Res. Rev.* 2010, 30, 290–326. [CrossRef] [PubMed]

24. Wolucka, B.; McNeil, M.; De Hoffmann, E.; Chojnacki, T.; Brennan, P. Recognition of the lipid intermediate for arabinogalactan/arabinomannan biosynthesis and its relation to the mode of action of ethambutol on mycobacteria. *J. Biol. Chem.* 1994, 269, 23328–23335. [CrossRef]

25. Jankute, M.; Cox, J.A.; Harrison, J.; Besra, G.S. Assembly of the Mycobacterial Cell Wall. *Annu. Rev. Microbiol.* 2015, 69, 405–423. [CrossRef]

26. Abrahams, K.A.; Besra, G.S. Mycobacterial cell wall biosynthesis: A multifaceted antibiotic target. *Parasitology* 2018, 145, 116–133. [CrossRef]

27. Bruni, R.; Kloss, B. High-Throughput Cloning and Expression of Integral Membrane Proteins in Escherichia coli. *Curr. Protoc. Protein Sci.* 2013, 74, 29.6.1–29.6.34. [CrossRef]

28. Punta, M.; Love, J.; Handelman, S.; Hunt, J.F.; Shapiro, L.; Hendrickson, W.A.; Rost, B. Structural genomics target selection for the New York consortium on membrane protein structure. *J. Struct. Funct. Genom.* 2009, 10, 255–268. [CrossRef]

29. Rath, A.; Glöbowicka, M.; Nadeau, V.G.; Chen, G.; Deber, C.M. Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proc. Natl. Acad. Sci. USA* 2009, 106, 1760–1765. [CrossRef]
30. Moreland, N.; Ashton, R.; Baker, H.M.; Ivanović, I.; Patterson, S.; Arcus, V.L.; Baker, E.N.; Lott, J.S. A flexible and economical medium-throughput strategy for protein production and crystallization. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2005**, *61*, 1378–1385. [CrossRef]

31. Bashiri, G.; Baker, E.N. Production of recombinant proteins in *Mycobacterium smegmatis* for structural and functional studies. *Protein Sci.* **2014**, *24*, 1–10. [PubMed]

32. Goldstone, R.M.; Moreland, N.J.; Bashiri, G.; Baker, E.N.; Lott, J.S. A new Gateway® vector and expression protocol for fast and efficient recombinant protein expression in *Mycobacterium smegmatis*. *Protein Expr. Purif.* **2008**, *57*, 81–87. [CrossRef] [PubMed]

33. Lewinson, O.; Lee, A.T.; Rees, D.C. The Funnel Approach to the Precrystallization Production of Membrane Proteins. *J. Mol. Biol.* **2008**, *377*, 62–73. [PubMed]

34. Kieser, K.J.; Rubin, E.J. How sisters grow apart: Mycobacterial growth and division. *Nat. Rev. Genet.* **2014**, *12*, 550–562. [CrossRef]

35. Miroux, B.; Walker, J.E. Over-production of Proteins in *Escherichia coli*: Mutant Hosts that Allow Synthesis of some Membrane Proteins and Globular Proteins at High Levels. *Protein Expr. Purif.* **1996**, *260*, 289–298. [CrossRef]

36. Mathieu, K.; Javed, W.; Vallet, S.; Lesterlin, C.; Candusso, M.-P.; Ding, F.; Xu, X.N.; Ebel, C.; Jault, J.-M.; Orelle, C. Functionality of membrane proteins overexpressed and purified from *E. coli* is highly dependent upon the strain. *Sci. Rep.* **2019**, *9*, 2654. [CrossRef]

37. Zhang, L.; Zhao, Y.; Gao, Y.; Wu, L.; Gurcha, S.S.; Wu, C.; Wu, F.; Gurcha, S.S.; et al. Structures of cell wall arabinosyltransferases with the anti-tuberculosis drug ethambutol. *Science* **2020**, *368*, 1211–1219. [CrossRef]

38. Alderwick, L.J.; Seidel, M.; Sahm, H.; Besra, G.S.; Eggeling, L. Identification of a Novel Arabinofuranosyltransferase (AftA) Involved in Cell Wall Arabinan Biosynthesis in *Mycobacterium tuberculosis*. *J. Biol. Chem.* **2006**, *281*, 15653–15661. [CrossRef]

39. Tang, Y.Z.; Rodríguez, J.; Keener, J.E.; Zheng, R.B.; Brunton, R.; Kloss, B.; Giacometti, S.I.; Rosário, A.L.; Zhang, L.; Niederweis, M.; et al. Cryo-EM structure of arabinofuranosyltransferase AftD from *Mycobacteria*. *Mol. Cell* **2020**, *78*, 683–699.e11. [CrossRef] [PubMed]

40. Tan, Y.Z.; Rodríguez, J.; Zheng, R.B.; Giacometti, S.I.; Rosário, A.L.; Kloss, B.; Danady, V.P.; Wei, H.; Brunton, R.; et al. Cryo-EM Structures and Regulation of Arabinofuranosyltransferase AftD from *Mycobacteria*. *Mol. Cell* **2020**, *78*, 683–699.e11. [CrossRef] [PubMed]

41. Zhang, L.; Zhao, Y.; Gao, Y.; Wu, L.; Gao, R.; Zhang, Q.; Wang, Y.; Wu, C.; Wu, F.; Gurcha, S.S.; et al. Structures of cell wall arabinosyltransferases with the anti-tuberculosis drug ethambutol. *Science* **2020**, *368*, 1211–1219. [CrossRef]

42. Tan, Y.Z.; Zhang, L.; Rodríguez, J.; Zheng, R.B.; Giacometti, S.I.; Rosário, A.L.; Kloss, B.; Danady, V.P.; Wei, H.; Brunton, R.; et al. Cryo-EM Structures and Regulation of Arabinofuranosyltransferase AftD from *Mycobacteria*. *Mol. Cell* **2020**, *78*, 683–699.e11. [CrossRef] [PubMed]