

# UCLA

## UCLA Previously Published Works

### Title

Protein structural biology using cell-free platform from wheat germ.

### Permalink

<https://escholarship.org/uc/item/6xk390s9>

### Journal

Advanced Structural and Chemical Imaging, 4(1)

### ISSN

2198-0926

### Authors

Novikova, Irina

Sharma, Noopur

Moser, Trevor

et al.

### Publication Date

2018

### DOI

10.1186/s40679-018-0062-9

Peer reviewed

METHODOLOGY

Open Access



# Protein structural biology using cell-free platform from wheat germ

Irina V. Novikova<sup>1</sup>, Noopur Sharma<sup>1</sup>, Trevor Moser<sup>1</sup>, Ryan Sontag<sup>1</sup>, Yan Liu<sup>2</sup>, Michael J. Collazo<sup>3,4</sup>, Duilio Cascio<sup>3,4</sup>, Tolou Shokuhfar<sup>5</sup>, Hanjo Hellmann<sup>2</sup>, Michael Knoblauch<sup>2</sup> and James E. Evans<sup>1,2\*</sup> 

## Abstract

One of the biggest bottlenecks for structural analysis of proteins remains the creation of high-yield and high-purity samples of the target protein. Cell-free protein synthesis technologies are powerful and customizable platforms for obtaining functional proteins of interest in short timeframes, while avoiding potential toxicity issues and permitting high-throughput screening. These methods have benefited many areas of genomic and proteomics research, therapeutics, vaccine development and protein chip constructions. In this work, we demonstrate a versatile and multiscale eukaryotic wheat germ cell-free protein expression pipeline to generate functional proteins of different sizes from multiple host organism and DNA source origins. We also report on a robust purification procedure, which can produce highly pure (> 98%) proteins with no specialized equipment required and minimal time invested. This pipeline successfully produced and analyzed proteins in all three major geometry formats used for structural biology including single particle analysis with electron microscopy, and both two-dimensional and three-dimensional protein crystallography. The flexibility of the wheat germ system in combination with the multiscale pipeline described here provides a new workflow for rapid production and purification of samples that may not be amenable to other recombinant approaches for structural characterization.

**Keywords:** Cell-free protein expression, Wheat germ, Electron microscopy, Cryo-EM, X-ray, Structural biology, Protein purification

## Background

Systems biology seeks to understand genomic and proteomic changes between species or between individual cells to link compositional changes in the genome (or proteome) to the observed phenotype. However, many organisms still have an array of genes and proteins of unknown structure and function while other annotated proteins may have limited homology to proteins in the same class. For example, more than 2000 of the 8000 genes identified in the smallest eukaryote *Ostreococcus tauri* are designated as proteins of unknown function [1, 2]. Understanding the structure, mechanism and function of these unknown proteins is the key to connecting information across scales from the individual protein to

whole cell/organism. Thus, a better link is needed to connect proteomic output with structural biology pipelines.

Today, a variety of cell-free expression options have emerged as powerful alternatives to laborious in vivo methods of protein synthesis to support the growing demand for easy and cost-effective ways for protein production. Depending on the biochemical properties, the origin and the potential use of the proteins to be expressed, several cell-free protein expression systems can be explored and are commercially available today. These are based on *E. coli*, wheat germ, rabbit reticulocyte, *L. tarentolae*, insect and human cell extracts [3–9]. Although not currently commercially available, efficient cell-free lysates can also be prepared from tobacco BY-2 cells, Chinese hamster ovary (CHO) and yeast [10–13]. All cell-free translation systems rely on two major components: (1) a specifically designed vector or a PCR template with your gene of interest, and (2) a cell extract of

\*Correspondence: james.evans@pnnl.gov

<sup>1</sup> Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, 3335 Innovation Blvd, Richland, WA 99354, USA  
Full list of author information is available at the end of the article

choice that matches the experimental needs for yield or post-translational modifications. The reagents are supplemented with additional energy sources, amino acids and various cofactor molecules to aid continuous protein synthesis and folding. Key advantages of these systems are the ease of the setup, fast turnaround times, linear scalability for the reaction volumes and amenability to high-throughput screening. Additionally, these platforms allow significant customization as one can directly provide additives to aid solubilization and folding of difficult targets [14–16]. One can also include unnatural amino acids to facilitate labeling and characterization or introduce additional DNA/RNA templates to generate diverse protein hetero-assemblies by co-expression [17–22]. Since there is no need to sustain a living organism, toxic proteins that failed to be produced by cell-based methods can be readily expressed [23]. These numerous benefits are being widely exploited in applications of NMR and X-ray structure determination, functional genomics and proteome research, protein chip construction, therapeutics and vaccine development [17, 23–26].

While thousands of proteins can be manufactured in a high-throughput fashion using the cell-free translation format [26, 27], structural biology primarily requires a supply of specific protein components. Thus, the most critical parameters are high yield and high purity of the protein sample. High-resolution structural studies require at least 50 µg of highly purified protein for cryo-EM; whereas, X-ray crystallography can require 1 mg or more. As of today, in comparison with other cell-free systems, *E. coli* cell-free protein production is the most frequently used for the needs of structural biology in X-ray and NMR due to its cost and availability [24]. However, a wheat germ lysate has the highest solubility and the highest translation yields for eukaryotic proteins. For instance, in the high-throughput proteome study, 12,996 human clones (out of 13,364 tested) gave rise to proteins; whereas, 97.6% of those were detectable in a soluble fraction [26]. In addition, codon optimization is not required for wheat germ, which makes it compatible with various DNA templates from different organisms of both prokaryotic and eukaryotic origins [28]. Despite these benefits, the use of wheat germ cell-free system for structural biology has been primarily limited to NMR studies [17]. There is only one report on the use of this system for X-ray structure determination [29], yet that study utilized methods for the purification of PabI protein based on its heat-resistance properties, which are not broadly applicable for other protein targets.

The Environmental Molecular Sciences Laboratory is a United States Department of Energy user facility that is primarily focused on organisms relevant to bioenergy and the environment such as bacteria, algae, fungi

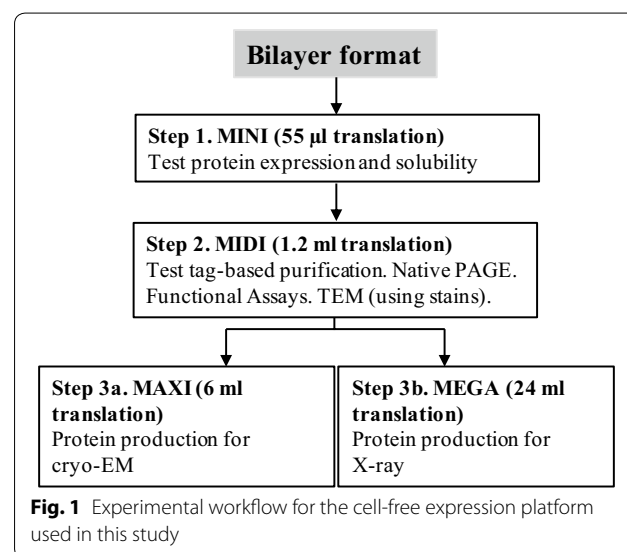
and plants. Using a wheat germ cell-free expression system from CellFree Sciences, invented by the Endo group from Japan [4, 5], we sought to build a cell-free expression pipeline to link the systems biology and structural biology capabilities at this user facility, while maintaining applicability to a wide range of organisms and being compatible with all sample geometries used for high-resolution electron microscopy and X-ray structural analysis. Here, we describe a multiscale pipeline that allows quick screening for protein solubility, quick optimization of expression yield along with purification strategies suitable for high-resolution structural investigations using several target genes from various diverse organisms.

## Results and discussions

Our multiscale protein production workflow included four different scales of protein expression (MINI, MIDI, MAXI and MEGA) to allow screening for solubility and impact of additives (MINI), overall expression levels and purification optimization (MIDI), and scale-up for cryo-EM (MAXI) or X-ray crystallography applications (MEGA) (Fig. 1).

### Quick translation trials to determine the protein expression potential (MINI)

PCR templates for cell-free expression can be prepared in short amount of time (compared to construction of vectors) and can, therefore, be used to quickly test target protein for expression and solubility. Our first stage involves MINI-translation reactions using PCR templates and fluorescently labeled lysine-charged tRNA. These reactions are only 55 µl in volume allowing for small-scale testing of various conditions with minimal

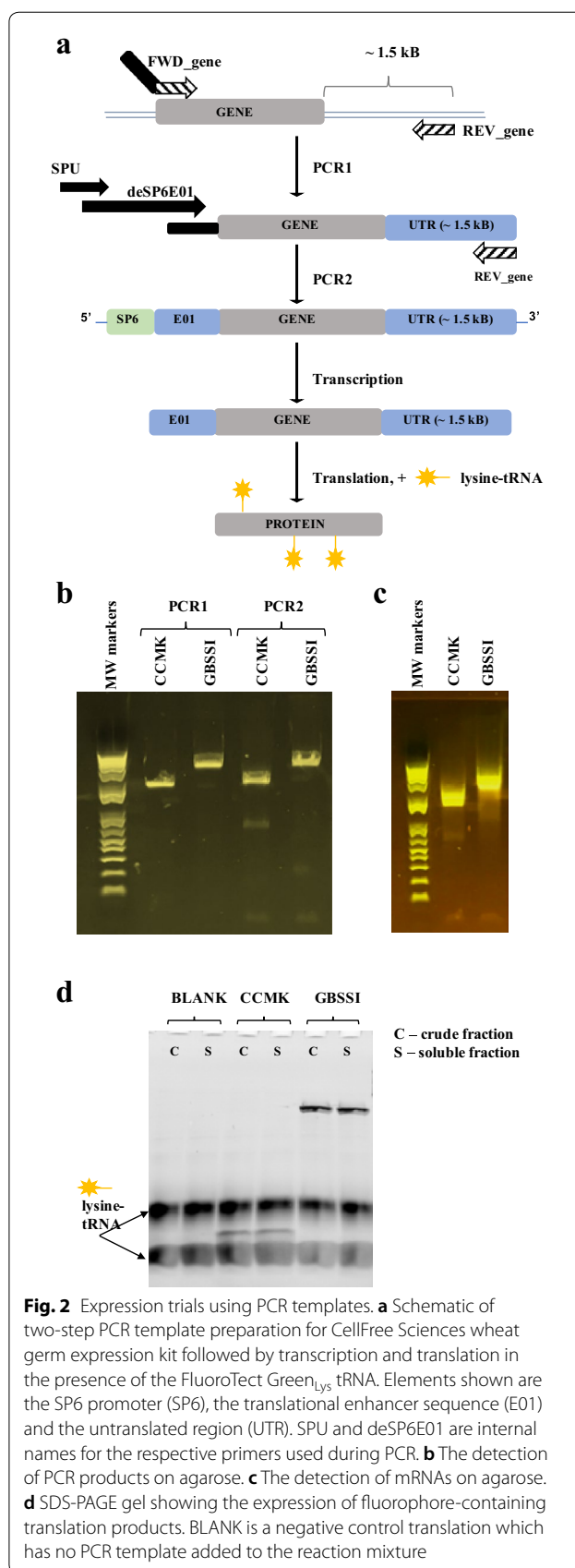


expenditure of reagents. The PCR template requires just three elements: a SP6 transcriptional promoter, an E01 translational enhancer sequence and a 3'-UTR for better protection of mRNA against degradation (Fig. 2a) [30]. To incorporate those elements, a two-step PCR using a split primer design is suggested [30, 31]. In the first round of PCR (PCR1), the gene of interest is amplified using a pair of gene-specific primers, where a reverse primer is designed to bind at ~1.4–1.6 kB away from the termination codon. At the second round of PCR (PCR2), the SP6 promoter and the E01 sequence are introduced through two forward primers SPU and deSP6E01, described previously [31]. The final PCR construct serves as the template for transcription and MINI-scale translation. To create a universal approach in handling DNA from different sources, we carried out all PCR reactions using the Q5 DNA polymerase from NEB, a high-fidelity enzyme with ultra-low error rates and suitable for low, medium and high GC contents of DNA template [32]. In addition, the Q5 enzyme demonstrates better capacity to produce long amplicons.

PCR-based expression trials were conducted on two genes of different origin and size: carbon dioxide-concentrating mechanism protein (CCMK, 10.6 kDa) from *Prochlorococcus marinus* and granule bound starch synthetase (GBSSI, 62 kDa) from *Ostreococcus tauri*. The PCR template for CCMK was successfully amplified from a commercially purchased cDNA clone, while the GBSSI gene was specifically amplified directly from the genome of *O. tauri* (Fig. 2b, see PCR1). The subsequent PCR2 reactions yielded the desired amplicons and were good templates for high-yield production of full-length mRNAs (Fig. 2c), which were further introduced into MINI-translation mix (Fig. 2d). For quick detection of the protein product, all MINI-translation reactions were supplemented with the fluorophore-labeled lysine-charged tRNA (FluoroTect™ Green<sub>Lys</sub>) for incorporation of fluorophore-tagged lysine into the synthesized protein. The use of the fluorophore-tagged lysine allows detection of picomolar levels of protein with SDS-PAGE without any need for purification. Both CCMK and GBSSI proteins were expressed efficiently using this cell-free format and displayed high solubility. Although not seen in the current experiment, it is important to note that this quick screening procedure can also be useful in detection of premature termination, which may indicate needs for additional supplementation or codon optimization.

#### pEU vectors, protein synthesis and purification (MIDI)

If PCR screening trials were successful, the gene of interest was further cloned into our modified pEU vector. For high-yield applications, the optimal expression template is the pEU vector [33], which was modified



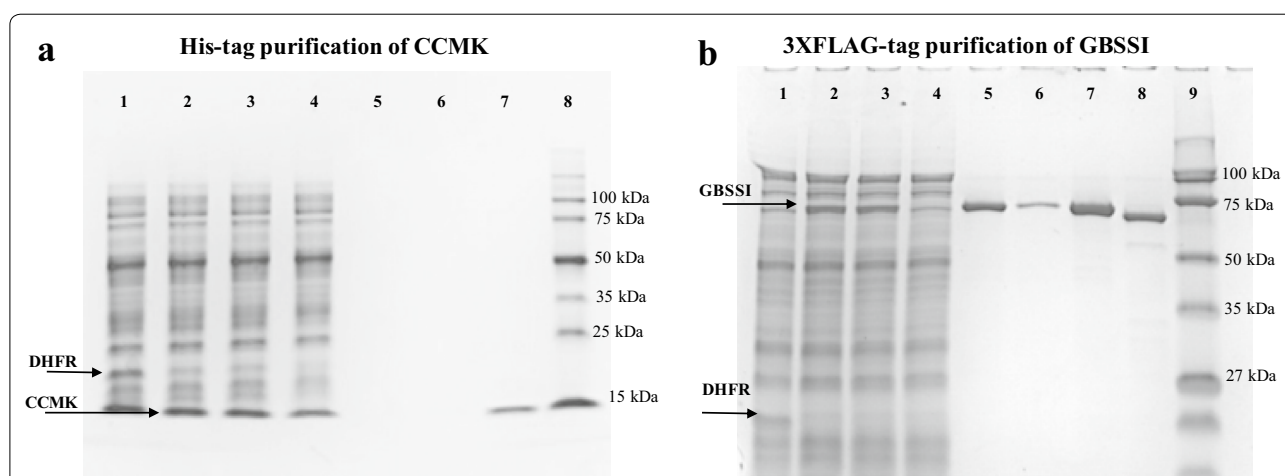
**Fig. 2** Expression trials using PCR templates. **a** Schematic of two-step PCR template preparation for CellFree Sciences wheat germ expression kit followed by transcription and translation in the presence of the FluoroTect Green<sub>Lys</sub> tRNA. Elements shown are the SP6 promoter (SP6), the translational enhancer sequence (E01) and the untranslated region (UTR). SPU and deSP6E01 are internal names for the respective primers used during PCR. **b** The detection of PCR products on agarose. **c** The detection of mRNAs on agarose. **d** SDS-PAGE gel showing the expression of fluorophore-containing translation products. BLANK is a negative control translation which has no PCR template added to the reaction mixture

in house to include a 3XFLAG tag (resulting in vectors pEU-3XFLAG-gene, pEU\_3XFLAG\_gene\_10His and pEU\_gene\_3XFLAG). Every newly produced plasmid was also tested at a MINI scale to assess the quality of produced mRNA and protein. The quality of the plasmid itself was found to be critical for obtaining high yields of mRNA, and thus high yields of protein sequentially. An example of low-quality mRNA is demonstrated in Additional file 1: Figure S1A, lane 4. To avoid the latter, the plasmid should be free of RNases, which can originate from plasmid preparation kit itself. However, we usually find that the main cause of bad-quality RNA is the plasmid mechanical shearing such as nicking during general plasmid purification procedures. This nicking is not critical for many other applications, but, for *in vitro* transcription reactions, premature runs-off are observed. Therefore, during plasmid productions, gentle cell resuspension and minimal vortexing of DNA are warranted.

The main advantage of cell-free synthesis is the linear scalability of the reaction volumes where reactions can be run as MINI-expression trials to determine gene suitability or scaled up for high-resolution structural studies (Fig. 1). If the plasmid passed the MINI-scale stage, we moved to a MIDI-scale production. The goal here is to determine if the tag attached to either the 3' or 5' end of the protein is fully available for efficient binding. As mentioned above, CCMK and GBSSI were found to be good candidates for the cell-free production using PCR screening. Therefore, both genes were cloned into pEU plasmids and tested for protein synthesis and purification using two different tags: 6×His and 3XFLAG. Initially,

the CCMK protein was fused with 6×His tag and purified using HIS-SELECT magnetic beads (Fig. 3a). Naturally, the wheat germ extract has many endogenous proteins which will nonspecifically bind to nickel; therefore, the wheat germ extract optimized for His-tag purification was also employed. A single band of CCMK of desirable purity is clearly seen in the final fraction, which was later used for 2D crystallization trial (discussed later).

Despite the use of His-tag-optimized wheat germ extract, the desired purity of CCMK was achieved at the expense of the product yield where 50 mM imidazole washes were employed to remove nonspecifically bound proteins. The same purification procedure with 10 mM imidazole washes generated much more material. However, a significant fraction of endogenous proteins was recovered as well unsuitable for our applications (Additional file 1: Figure S1B, lane 7). Ideally, one can try a range of imidazole concentrations to find the optimal value. However, to avoid optimization trials in this regard, we decided to employ a 3XFLAG-based purification scheme. The GBSSI protein was first fused with 3XFLAG sequence on its N-terminus and was tested in the purification trial using ANTI-M2 monoclonal magnetic beads. Expressed GBSSI showed good binding to the beads, as indicated by the full disappearance of the expressed protein band in the flow-through fraction (Fig. 3b, compare lanes 3 and 4). Beads were washed, and then two competitive elution reactions with 3XFLAG peptide were followed to release the protein from the matrix. Both elution fractions had high amounts of released protein (Fig. 3b, lanes 5 and 6). Purified protein appeared to be



**Fig. 3** Protein synthesis and purification. **a** SDS-PAGE gel, showing the expression and purification of CCMK protein using 6×His-tag. Lane 1, control DHFR expression; lane 2, crude mixture; lane 3, soluble fraction; lane 4, flow-through; lane 5, elution fraction 1; lane 6, elution fraction 2; lane 7, elution fractions 1 and 2 combined and concentrated; lane 8, MW markers. **b** SDS-PAGE gel of expression and purification of GBSSI using 3XFLAG-tag. Lane 1, control DHFR expression; lane 2, crude mixture; lane 3, soluble fraction; lane 4, flow-through; lane 5, elution fraction 1; lane 6, elution fraction 2; lane 7, elution fractions 1 and 2 combined and concentrated; lane 8, 3XFLAG tag removed; lane 9, MW markers

(See figure on next page.)

**Fig. 4** The applications of cell-free produced proteins in conventional electron microscopy. **a** STEM image of negatively stained 2D crystals of CCMK protein (inverted contrast). **b** The image from (**a**) has been processed using a temperature factor of  $-100$ , and an IQ cutoff of 2 for the final map. The crystal exhibits P6 symmetry and a unit cell of  $a=b=57$  Å and  $\gamma=120^\circ$ . **c, d** TEM images of negatively stained PDX1.2 protein complexes. Top- and side-view orientations of the complex are depicted by black arrows. **e–h** TEM images of negatively stained SEO1 proteins. The individual fibril unit is indicated by a black arrow

full size, homogenous and free of other contaminating proteins coming from the extract itself (Fig. 3b, lane 7). Using enterokinase, the 3XFLAG tag was then removed with >93% efficiency (Fig. 3b, lane 8). The same purification procedure has been tested on several other proteins including glutamine synthetase (GS, 75 kDa) and pyruvate, phosphate dikinase (PPDK, 100 kDa) proteins from *O. tauri*, sieve element occlusion protein from *M. truncatula* (SEO1, 75 kDa) [34, 35] and pyridoxal 5'-phosphate synthase-like subunit (PDX1.2, 34 kDa) from *A. thaliana* [36]. All SDS-PAGE gels for these proteins are included (Additional file 1: Figure S2A). High purity of each final sample was achieved along with high protein recovery. Only PDX1.2 has a somewhat reduced binding to the beads due to its tight structural fold (demonstrated in the next section). It is worth noting that the major advantage of using the 3XFLAG tag is that we observed a complete lack of nonspecific binding. As the result, no optimization is required, and a general one-step purification “bind–wash–elute” protocol can be used for all samples.

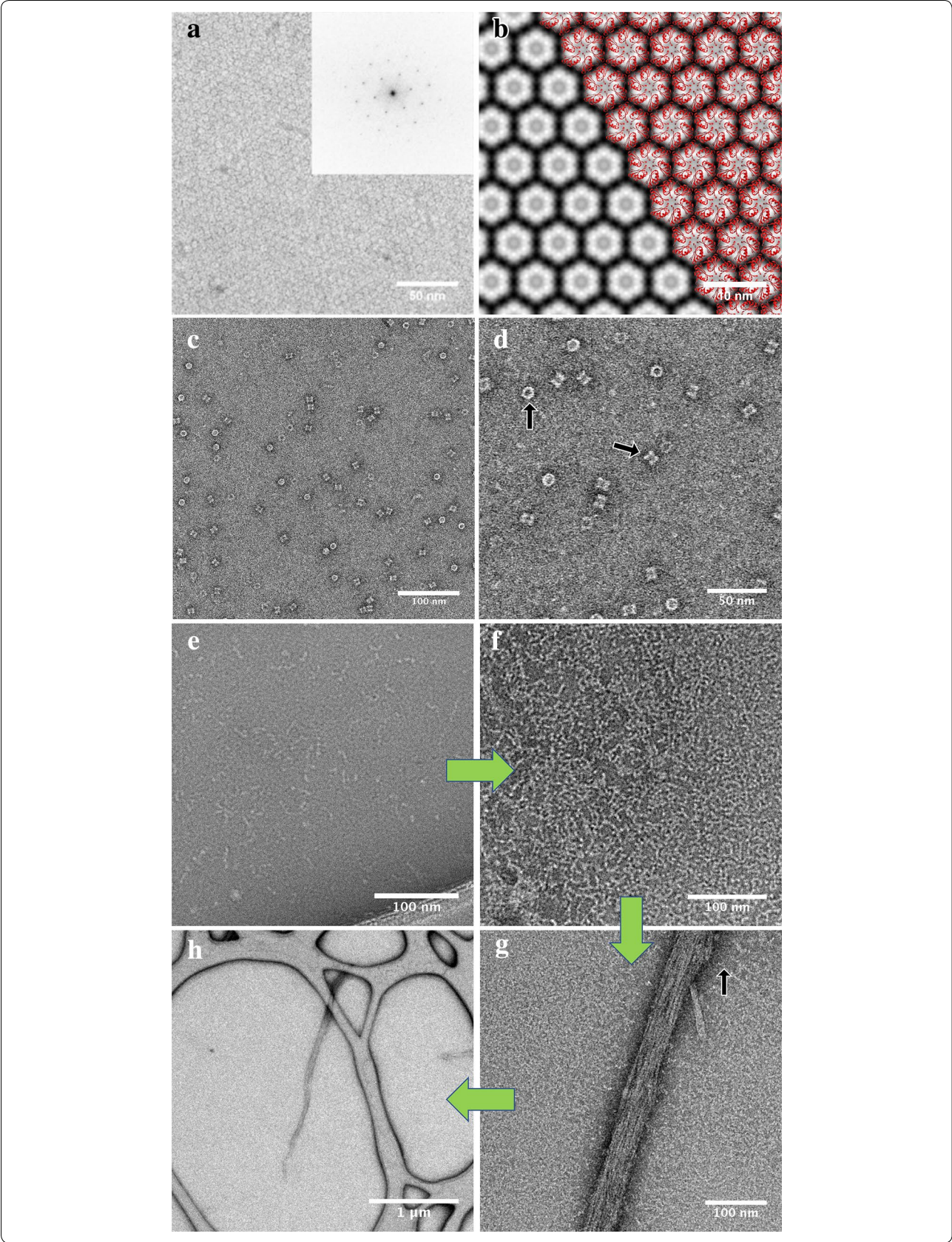
#### Diverse structural biology applications of cell-free produced proteins

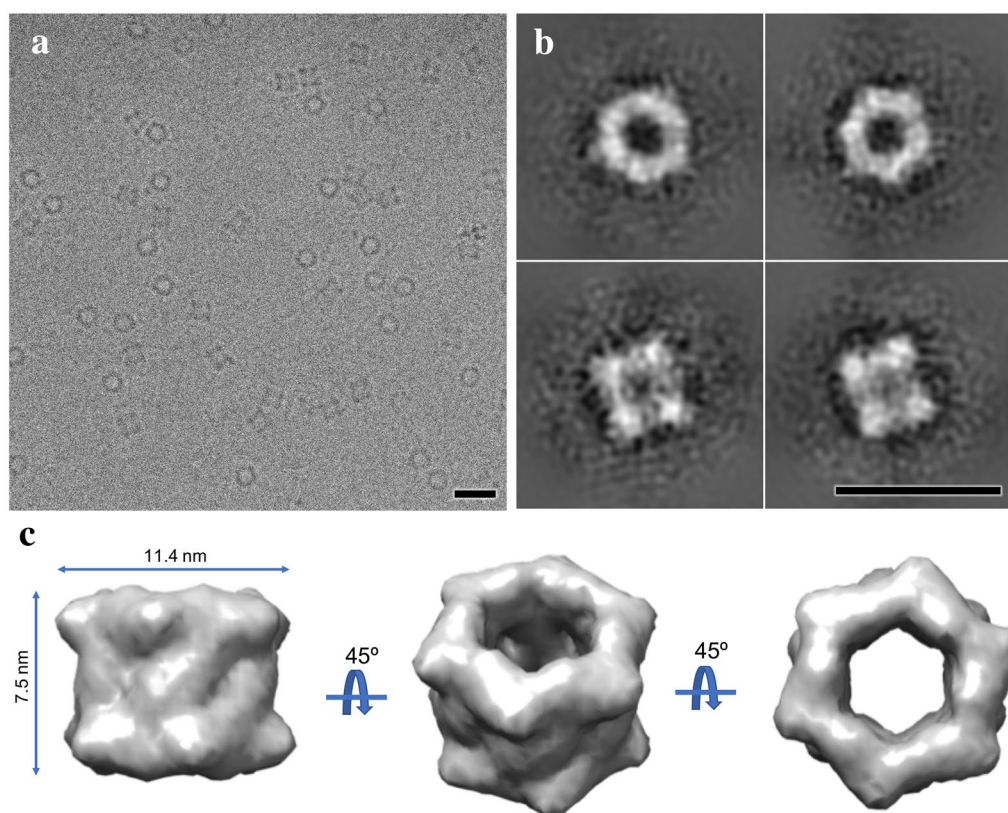
Once we had an isolated protein in hand, the next step was to verify its activity, assembly state and identify whether it is a suitable candidate for high-resolution structural studies. At MIDI scale, the general yields were determined to be around 20–30 µg at 0.2–0.3 mg/ml concentrations. These protein amounts are sufficient for various functional assays, Native PAGE electrophoresis, and preliminary TEM imaging characterization. The structure of CCMK from *P. marinus* is known from X-ray crystallography (PDB ID: 4OX8), and the native protein has been shown to pack functionally as hexamers in the 3D crystals. Similarly, 2D crystallization for a CCMK homolog from *Synechocystis* PCC6803 [37] using a nickelated lipid monolayer also resulted in hexameric assemblies. To validate our pipeline, cell-free expressed and purified CCMK (6×His tagged) in this study was subjected to a similar 2D crystallization trial. From the first attempt, initial 2D crystal patches were formed and analyzed by TEM imaging (Fig. 4a). Negatively stained CCMK proteins were clearly visualized as hexameric structures, mimicking the packing in the previous study [37]. Image processing of the collected EM micrographs revealed a lattice with reflections extending to 13 Å,

limited by the negative stain itself. This 2D crystal lattice was able to be directly overlaid by the known atomic structure of *P. marinus* CCMK (RCSB: 4OX8) [38] verifying that the expected structural fold has been adopted by the cell-free-synthesized CCMK protein (Fig. 4b).

The rest of the purified proteins (GBSSI, GS, PPDK, PDX1.2, SEO1) have been initially tested by Native PAGE electrophoresis to gain insight about their potential structural organization and assembly state (Additional file 1: Figure S2B). GS and PDX1.2 proteins were found to form defined higher-molecular weight complexes; whereas, the other proteins showed multiple assembly states. TEM images collected for PDX1.2 showed a two-ring stacked architecture (Fig. 4c, d). No aggregation events or partially disordered complexes were seen, which suggests that the population is quite homogenous where the average diameter of formed complexes was 12 nm. The protein structure closely resembles the structural arrangements of similar PDX1.1 and 1.3 proteins, which also form two-ring dodecamers [39]. As a result, PDX1.2 protein seems to form proper structural arrangements and would be a good candidate system for single particle analysis studies by cryo-EM after scale-up expression.

We also explored expression of SEO proteins from *Medicago truncatula*, which are known to form large protein bodies, called forisomes [40]. These spindle-shaped bodies achieve tens of micrometers in length and look like elongated needles, which swell and deform upon  $\text{Ca}^{2+}$  introduction. This swelling/deformation is reversible, and it is a critical mechanism in sieve tube flow control [41]. Natural forisomes consist of several members of SEO protein family [34, 35]. However, SEO proteins have not been analyzed at the individual protein level previously. As seen in Native PAGE, cell-free produced SEO1 (with the monomer molecular weight of 75 kDa) runs higher than the 480 kDa NativeMark molecular weight standard and upon 3XFLAG removal, SEO1 proteins form even higher-order molecular weight structures (Additional file 1: Figure S2B). TEM images of SEO1 proteins show that, at the initial stages of assembly, they form individual fibrils, which serve as building blocks for higher-order congregation (Fig. 4e–h). Additional incubation of SEO1 (in  $\text{Ca}^{2+}$ -free environment, 10 mM EGTA) results in the formation of micrometer-long strands, composed of multiple individual fibrils aligned





**Fig. 5** Cryo-EM applications of cell-free-synthesized protein PDX1.2. **a** Raw cryo-EM image of PDX1.2 particles distributed in vitreous ice and adopting random orientations. **b** Single-particle class averages of cryo-EM dataset from **(a)** showing class averages for the top, side and two intermediate views. A double stacked hexameric ring geometry is evident. **c** Preliminary 3D SPA cryo-EM reconstruction of PDX1.2 complex shown in three different views at 15 Å resolution. All scale bars correspond to 20 nm

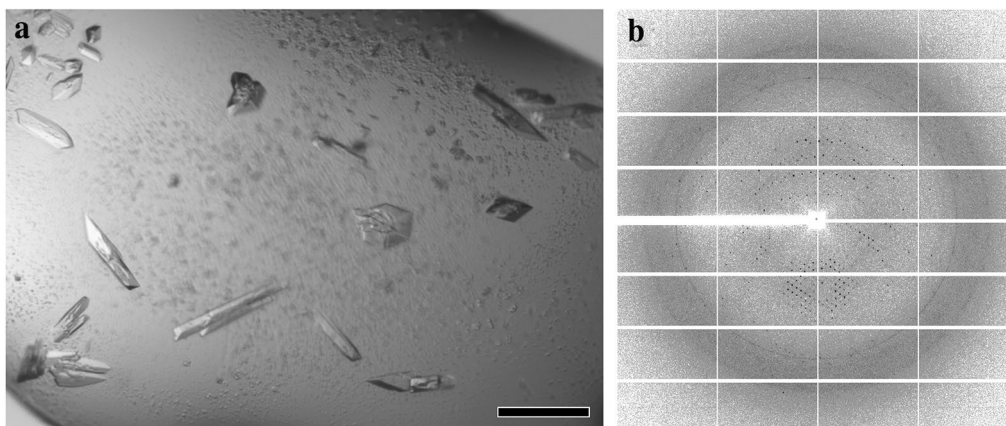
and interacting cooperatively. The obtained data are consistent with the functional roles of these proteins *in vivo*, where they participate in the molecular self-assembly of large 3D bodies. This is the first demonstration that such protein bodies can be generated *in vitro* from individual monomers.

#### Scaling-up for cryo-EM (MAXI) and X-ray (MEGA) structural approaches

Due to significant instrumental advances in the electron microscopy such as the development of direct electron cameras and phase plates [42, 43], cryo-electron microscopy (cryo-EM) holds a promise to be major imaging technology to study the structures of biological molecules in its native environment. However, the success of the cryo-EM experiment is also defined by the quality of protein sample such as its purity and homogeneity and typically requires about 50 µg of sample for screening and data collection.

Depending on the outcome of the above MIDI-scale tests, the protein synthesis can be further scaled up for high-resolution cryo-EM and X-ray crystallography characterization. Since the negative stain imaging of PDX1.2 showed a fairly homogenous sample, a MAXI-scale cell-free expression was performed to generate suitable protein for cryo-EM. While the current 2D class averages and initial 3D volume are limited in resolution due to the conventional microscope used and available for this work (Fig. 5), preliminary 3D single-particle reconstruction of PDX1.2 complex clearly shows that PDX1.2 adopts a 2-ring stacked dodecameric architecture of sixfold symmetry with the diameter of 11.4 nm and height of 7.5 nm. Based upon the Fourier Shell Correlation of two data set half-maps, the resolution of this preliminary map is ~15 Å, and the files have been uploaded to the wwPDB repository under ascension EMD-9190 (<https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-9190>). Experiments utilizing new state-of-the-art cryo-EM instrumentation are ongoing and will be the focus of a separate paper. Nevertheless, this work demonstrates the ability of this pipeline





**Fig. 6** X-ray applications of cell-free-synthesized protein GS. **a** Optical image of 3D microcrystals for GS. The scale bar corresponds to 50  $\mu\text{m}$ . **b** X-ray diffraction pattern collected from the crystals seen in **(a)**. Data was collected at APS-NECAT beamline 24-ID-E with an EIGER 16 M pixel detector. The detector distance was 500 mm, and the wavelength was 0.9792  $\text{\AA}$

to generate ample yields for cryo-EM studies especially since the PDX1.2 protein used for cryo-EM validation had the lowest yields of any of the proteins tested.

Finally, we sought to test if the same pipeline could generate enough protein material of good quality for 3D crystallization screening. As other GS classes from cell-based recombinant expression have been successfully crystallized previously [44, 45], we, therefore, performed a MEGA-scale synthesis of the GS complex and generated enough material ( $\sim 10$  mg/ml, 100  $\mu\text{l}$ ) for 3D crystallization trial screening. Several conditions were found suitable for the crystal growth and, without any additional optimization, diffractions up to 5.5  $\text{\AA}$  were obtained (Fig. 6). The preliminary X-ray characterization and refinement statistics of this dataset can be found in the Additional file 1: Table S1. Additional experiments to improve the resolution are ongoing and will be the subject for a separate mechanistic study on the GS enzyme.

## Conclusions

The preparation of a wheat germ cell lysate is a tedious and a quite complex procedure [4]; therefore, the use of established commercial products is often recommended for obtaining reproducible results [28]. As of 2018, wheat germ cell-free expression (WGCF) kits are available from two manufacturers: CellFree Sciences and Promega. To minimize sample handling, the Promega WGCF kit offers a coupled setup where transcription and translation are run simultaneously. The CellFree Sciences WGCF kits have transcription and translation processes rather decoupled where each step is optimized individually to maximize the yield. If one needs to use an additive for the translation or change the reaction temperature, the latter setup offers a benefit of not interfering with the

transcription stage. Due to the above-described reasons and a wider selection of the reagents for a MINI/MIDI or MAXI/MEGA-protein synthesis, we opted to use the commercial system from CellFree Sciences.

While the use of the specialized pEU vector is preferred to obtain higher yields of the given protein in this cell-free format, the most time-consuming steps are still the cloning of the respective gene, purification of the plasmid product, its subsequent sequencing and scale-up preparation. While this is the final route to get high quantities of the desired protein product, it is always beneficial to know in advance if the protein is a good candidate to be expressed in this system and is worth the time invested. Fortunately, cell-free expression with somewhat reduced protein yields can be performed using PCR templates. The latter offers an ideal rapid screening to explore the protein potential to be expressed in a soluble and active form. Based on our experiences in expressing a variety of proteins for different applications, we had a very high probability of producing good quality proteins from different organisms using this system.

In combination with the wheat germ cell-free extract, this work presents a useful methodology for a multi-scale pipeline covering all stages from cloning, screening potential candidates of protein synthesis to final production of micrograms quantities of a highly pure and tag-free protein. We show that the developed 3XFLAG purification protocol allowed high protein recoveries with purities reaching 98%, and it was found to be our preferred method for preparation of all protein samples. The only exception was to use a His-tag-based purification for the 2D protein crystallography method which relies on a Ni-NTA functionalized lipid monolayer. Nevertheless, His-tag purification can be considered a

good alternative strategy. In terms of yield, 0.6 mg/ml wheat germ extract has been generally obtained in the bilayer format used here. Not demonstrated here, but for large-scale productions, impressive protein yields up to 20 mg per ml of wheat germ extract can be achieved by a two-chamber dialysis or a “filter-feed” method using dedicated robotic synthesizers [28]. While we show that such robotic platforms are not necessary, if available at a nearby institution, they could be used to replace the MEGA-scale expression using lower reaction volumes at potentially lower cost. In terms of structural function/fold, cell-free-produced proteins in this study exhibited expected structure and assembly state in comparison with similar proteins produced and characterized by other means.

The generated protein products are compatible with three major protein sample formats for high-resolution structural studies: (1) single-particle analysis, (2) 2D crystallography and (3) 3D macromolecular crystallography. Obviously, the developed pipeline is applicable to any science area where interests lie in expressing of proteins in high yields and purities without the use of any specialized equipment and minimal time invested. Due to the fact that the cell-free expression platforms are gaining significant popularity, we believe that these comprehensive experiments will be useful for a wide range of scientific audience interested in exploring these systems especially as this platform can be replicated without major investments in instrumentation.

## Methods

### DNA template construction

PCR reactions were carried out using the Q5 Hot Start High-Fidelity 2× Master Mix from New England Biolabs. NEB Tm Calculator (<http://tmcalculator.neb.com/#/>) has been used to assess annealing temperatures. DNA primers were acquired from either IDTNA or Invitrogen. The pEU vector (pEU-MCS-TEV-HIS-C1) was from CellFree Sciences. Gene sequences were either amplified from genomic DNA (*Ostreococcus tauri*—cultured in-house, *Medicago truncatula*—gift from M. Knoblauch, or *Arabidopsis thaliana*—gift from H. Hellmann), or purchased as synthetic genes from GENSCRIPT. To obtain genomic DNA from *O. tauri*, the DNeasy Plant kit from QIAGEN has been used for extraction but tissue homogenization procedures were omitted. Genomic DNA was further ethanol precipitated and re-suspended in water. To amplify GS and PPK genes from the genomic DNA, we used nested PCR strategy, described in detail in the Additional file 1: Discussion.

The quality of genomic DNA and PCR products was assessed using the E-gel Electrophoresis System (E-gel pre-cast 1% EX Agarose gels with SYBR safe along with

E-gel 1kb Plus DNA Ladder) from Invitrogen. Gel purification of DNA was carried out with the help of Zymo-clean Gel DNA recovery kit from Zymo Research as instructed. Zymo-clean-purified DNA was additionally desalted using the Bio-Rad mini spin columns, pre-filled with the Bio-Gel P-6DG soaked in water. For cloning in the pEU vector, the Gibson Assembly Master Mix from New England Biolabs was used. The linearized pEU vector and the insertion fragment containing the desired gene were prepared by PCR. These fragments were further gel purified and quantified with NanoDrop 2000c from ThermoFisher. As an example of Gibson reaction, linear pEU plasmid (2 µl, 20 ng/µl) was mixed with linear gene PCR template (3.8 µl, 15 ng/µl), water (4.2 µl) and 2× Gibson Assembly Mix (10 µl) and incubated at 50 °C for 1 h in the thermocycler. The assembled product (2 µl) has been used to transform an NEB10 *E. coli* supplied with the Gibson Assembly kit. Transformed *E. coli* was later plated on a carbenicillin-containing agar plate and grown at 37 °C overnight. Several individual colonies were then selected and grown overnight in several milliliters of LB broth supplemented with 100 µg/ml carbenicillin. The QIAGEN Plasmid Mini Kit was then employed for plasmid extractions, which were then sent for sequencing to Eurofins Genomics or MCLAB. Successful clones were later grown in larger quantities of LB broth (with carbenicillin) and harvested using the QIAGEN Plasmid Maxi kit. To note, a 3XFLAG tag was initially introduced through PCR primers in the first construct to generate pEU\_3XFLAG\_LA\_gene. For all subsequent cloning experiments for other genes, the obtained pEU\_3XFLAG\_LA\_gene vector was used as a template for other constructs. The general sequence design of pEU vectors is provided in Additional file 1: Figure S3.

### Cell-free expression

Protein synthesis was carried out using Wheat Germ Protein Research Kits (WEPRO7240 and WEPRO7240H) from CellFree Sciences as suggested by the manufacturer. All translation reactions were carried out in the bilayer format where total translation volume accounts for the SUB-AMIX layer. For MIDI-scale plasmid-based protein expression (1.2 ml total translation volume), 5 µl of pEU plasmid (1 µg/µl) was mixed with 10 µl of 5× Transcription Buffer, 5 µl of 25 mM NTP mix, 0.625 µl of RNase inhibitor and 0.625 µl of SP6 polymerase and incubated at 37 °C for 3–4 h with gentle shaking. Upon completion, 1 µl of the transcription mixture was loaded on the E-gel precast agarose gel (Invitrogen) to assess the mRNA integrity. Then, 50 µl of transcription reaction was mixed with 50 µl of wheat germ extract (WEPRO7240) and 4.25 µl of 1 mg/ml creatine kinase by a pipette, and the entire mixture was then transferred to the bottom of

the individual well of a standard 24-well plate, pre-filled with 1.1 ml of translation buffer (1× SUB-AMIX SGC). To prevent evaporation, wells were sealed with Parafilm. These reactions were carried out for at least 20 h at 15 °C with no mixing at Thermomixer C from Eppendorf. At the end, the contents were gently mixed by a pipette and centrifuged at 20,000g for 15 min. The supernatant fraction was then buffer exchanged to TBS (50 mM Tris-HCl, pH 7.5, 150 mM NaCl) using pre-equilibrated in TBS Zeba Spin Desalting Columns (5 ml) from Thermo Fisher (this step is critical to remove DTT) and further subjected to the protein purification procedure of choice. For MAXI- and MEGA-scale plasmid-based protein expression (6- and 24 ml translation volumes), all reagents were scaled linearly, and a 6-well plate was used instead.

For PCR-based protein expression (MINI scale, 55 µl translation), the following changes have been implemented in the transcription and translation protocols: (1) for transcription, 0.5 µl of the PCR product (2 µg/µl) was mixed with 4.5 µl of the Transcription premix and left incubated at 37 °C for 4 h, (2) for translation, 2 µl of the resulting mRNA was mixed with 0.5 µl of FluoroTect Green<sub>Lys</sub> and 2.5 µl of the wheat germ extract (WEPRO7240) and transferred under 50 µl of the translation buffer (1× SUB-AMIX SGC), placed in the 96-well half-area plate. The FluoroTect Green<sub>Lys</sub> in vitro Labeling System was purchased from Promega. The translation was carried out overnight and the protein products were analyzed by SDS-PAGE.

#### Protein purification using a His-tag

Purification using a His-tag has been carried out using HIS-Select Nickel Magnetic Agarose beads from Sigma-Aldrich in combination with the MagneSphere technology Magnetic Separation Stand from Promega. Per 1.2 ml translation (MIDI-scale), 200 µl of gel suspension (100 µl of packed gel) was used. The needed volume of gel suspension was first equilibrated in 50 mM Tris, pH 8.0, 300 mM NaCl, 50 mM imidazole buffer and then combined with the DTT-free translation mixture. The binding was carried out for 1 h at room temperature with shaking on the ThermoMixer C (~1000 rpm). Using the magnetic stand, the supernatant was removed and discarded, and the beads were then washed 3 times with 1 ml of 50 mM Tris, pH 8.0, 300 mM NaCl, 50 mM imidazole wash buffer. To elute the bound protein, 0.5 ml of 50 mM Tris, pH 8.0, 300 mM NaCl, 500 mM imidazole elution buffer was applied to the beads, and the mixture was left shaking for 30 min. Then, the supernatant fraction was collected, and the elution procedure was repeated one more time. Then, the elution fractions were combined, concentrated using 3 kDa Amicons from EMD Millipore and buffer exchanged to 20 mM Tris, pH 7.5, 600 mM NaCl

for crystallization needs. For MAXI scale, all reagents were scaled linearly.

#### Protein purification using a 3XFLAG tag

Purification using a 3XFLAG tag has been carried out using ANTI-FLAG M2 magnetic beads from Sigma-Aldrich. All magnetic separations were done using the MagneSphere technology Magnetic Separation Stand from Promega. Per 1.2 ml translation (MIDI scale), 170 µl of gel suspension (85 µl of packed gel) was used. The needed volume of gel suspension was first equilibrated in TBS buffer and then combined with the DTT-free translation mixture. The binding was carried out for 1 h at room temperature with shaking on the ThermoMixer C (~1000 rpm). Using the magnetic stand, the supernatant was removed and discarded, and the beads were then washed 3 times with TBS buffer (0.5 ml each time). To elute the bound protein, TBS buffer supplemented with 150 µg/ml 3XFLAG peptide (purchased from Sigma) was applied to the beads, and the mixture was left shaking for 20 min. Then, the supernatant fraction was collected, and the elution procedure was repeated one more time. Two elution fractions were combined and concentrated to ~100 µl using 10 kDa or 30 kDa amicons from EMD Millipore. For MAXI- and MEGA-scale reactions, reagents were scaled linearly but final concentrated volume remained around 100 µl.

To remove the 3XFLAG tag, a light chain enterokinase from New England Biolabs, NEB, was employed. Due to inhibition of this enzyme at high concentrations of salts and its requirement for the Ca<sup>2+</sup> ions, ~100 µl of purified and concentrated protein was further buffer exchanged using the Bio-Rad spin mini columns, filled with Bio-gel P-6DG (pre-soaked in 20 mM Tris, pH 7.5, 50 mM NaCl and 2 mM CaCl<sub>2</sub>). Because the enterokinase cleavage efficiency is time- and template concentration dependent, small-scale cleavage titration was carried out first to determine the right conditions: 10 µl of concentrated protein was mixed with 2 µl of 1.6 U/µl enterokinase. Aliquots (2 µl) were taken over time up to 3 h and the cleavage efficiency was analyzed by running the samples on the SDS-PAGE. Once the appropriate time was determined, the rest of the protein was cleaved in a similar fashion. For single-particle analysis, the removal of the 3XFLAG by enterokinase was not required and thus, the sample was buffer exchanged back to TBS using Bio-gel P-6DG (pre-soaked in TBS), split in aliquots, frozen in liquid nitrogen and then transferred to -80 °C for a long-term storage. In cases where prolonged protein incubation was later required at ambient temperatures either for crystallization or assembly trials, the enterokinase was removed by stopping the reaction with 1/10th volume of 5 M NaCl and mixing with 15 µl of 100%

pre-equilibrated agarose-trypsin slurry from Sigma. The mixture was further incubated at room temperature for 30 min at 300 rpm on the ThermoMixer C. Upon completion, the mixture was directly applied on spin mini column, containing Bio-Gel P6-DG (presoaked in TBS buffer). The flow-through was collected, quantified, frozen in liquid nitrogen and transferred to  $-80^{\circ}\text{C}$  for a long-term storage.

#### PAGE electrophoresis

Protein expressions were evaluated on 10% Mini-PROTEAN TGX Precast Protein gels from Bio-Rad. For SDS-PAGE, 3  $\mu\text{l}$  of a protein solution was mixed with 4  $\mu\text{l}$  of 500 mM DTT and 7  $\mu\text{l}$  of the 2 $\times$  SDS Sample Buffer. Then, the sample was denatured at  $90^{\circ}\text{C}$  for 5 min, snap-cooled on ice and loaded on the gel. The general running conditions for SDS were 150 V in the 1 $\times$  Tris-Glycine-SDS Running Buffer from Bio-Rad. The gels were further stained with Bio-Safe Coomassie Stain from Bio-Rad. Molecular ladders used were Broad Range Protein Molecular Weight Markers from Promega and Kaleidoscope<sup>TM</sup> Prestained Standards (#161-0324 and #161-0375) from Bio-Rad. Native PAGE was performed using the same 10% Mini-PROTEAN TGX Precast Protein gels where 12  $\mu\text{l}$  of protein solution was mixed with 12  $\mu\text{l}$  of 2 $\times$  Native Sample Buffer from Bio-Rad prior to loading (no heat denaturation employed). For Native PAGE, the molecular ladder used was NativeMark<sup>TM</sup> Unstained Protein Standard from Thermo Fisher (#LC0725).

To analyze the proteins comprising FluoroTect Green<sub>Lys</sub>, the denaturation protocol for SDS-PAGE was changed to  $70^{\circ}\text{C}$  for 2 min, and a protein loading per well was increased to 6  $\mu\text{l}$ . After the run was complete, the fluorescent images of gels were collected immediately by FluorChem Q from Alpha Innotech or a laser-based scanner Typhoon FLA 9500 from GE Healthcare, set for Alexa488 excitation and emission.

#### 2D protein crystallization using a Ni<sup>2+</sup>-chelating lipid monolayer

Purified CCMK protein with 6xHis tag on C-terminal (15  $\mu\text{l}$ , 0.3 mg/ml) was adjusted with imidazole to 10 mM and glycerol to 10% and loaded in the concave well of PELCO Immunostaining Pad from Ted Pella. Lipid mixture was prepared by combining four parts of 10 mg/ml of DOPC (in chloroform) with one part of 2.5 mg/ml of 18:1 DGS-NTA(Ni) (in chloroform), achieving the final weight ratio of 16:1. Both lipids were purchased from Avanti Polar Lipids. To create a monolayer, 2  $\mu\text{l}$  of lipid mixture was deposited on the air–water interface of the protein solution. The PELCO pad was then enclosed inside the sealed Petri dish, containing wet filter papers

to create a humid environment and prevent evaporation. The 2D crystallization was carried out at room temperature for 13 h.

#### TEM imaging and data analysis

Negative staining of PDX1.2 and SEO1 proteins was carried out in the following manner: (1) an ultrathin carbon film on lacey carbon (01824, Ted Pella) was glow discharged at 10 mA for 30 s using PELCO easiGlow from Ted Pella, (2) 3  $\mu\text{l}$  of protein solution was deposited on the carbon side and let to absorb for 1 min, (3) the excess of solution was blotted away, and the grid was floated on a 20  $\mu\text{l}$  drop of the NanoW stain from Nanoprobes for 1 min, (4) then the excess of the stain was wicked away, and the grid was let to air-dry. For 2D protein crystals (CCMK protein), an ultrathin carbon film on lacey carbon grid was not glow discharged. The carbon side was directly dropped down on top of lipid layer at the air–water interface of protein solution. The grid was then lifted and transferred to a drop of Nano-W stain, incubated for 1 min, blotted and air-dried. The EM micrographs were acquired on 300 kV FEI Environmental TEM, 300 kV FEI Scanning TEM and 300 kV JEOL JEM-3000SFF. For CCMK crystals, the Focus suite has been used for image processing, FFT determination, lattice unbending, and mapping [46]. For cryo-EM experiments, C-flat holey grids (CF213, EMS) were used instead. Prior to sample deposition, they were first glow discharged at the same conditions as above and then transferred to the plunge-freezer Leica EM GP, brought to 80% humidity at  $25^{\circ}\text{C}$ . Then, 3  $\mu\text{l}$  of PDX1.2 protein sample (50  $\mu\text{g}/\text{ml}$ , in TBS buffer) was pipetted on the carbon side and blotted for 3 s before plunge-freezing in the liquid ethane. The frozen samples were then transferred to liquid nitrogen and then to 300 kV JEOL JEM-3000SFF, pre-cooled with liquid helium to 4 K. The data were collected at low-dose conditions with defocus values of 2–4  $\mu\text{m}$  using DE20 camera from Direct Electron in movie mode. Each movie consisted of 16 frames encompassing total exposure of 0.5 s with electron flux per frame of 2 electrons per square angstrom and a pixel size of 0.117 nm. All standard image processing from movie alignment to 3D map refinement was performed using CisTEM software [47] with default parameters (apart from inputting the microscope Cs, image pixel size and estimated particle radii and symmetry).

#### X-ray crystallization trials

Cell-free-produced GS protein (tagged with 3XFLAG, 10 mg/ml, 100  $\mu\text{l}$ ) was crystallized in hanging drop format by the addition of 0.2 M sodium chloride, 0.1 M BIS-Tris pH 5.5 and 25% w/v PEG 3350. The TTP Labtech Mosquito nanodispenser was used to set up 210 nl drops of protein and reservoir solution mixed in a 1:1, 1:2, and

2:1 ratio. Drops equilibrated over 100  $\mu$ l of reservoir solution in a 96-well tray. Crystals appeared over several days. Crystals were soaked for 10 s in a 2:1 solution of mother liquor and glycerol.

## Additional file

**Additional file 1.** Additional discussion, figures, table.

## Abbreviations

PCR: polymerase chain reaction; DNA: deoxyribonucleic acid; RNA: ribonucleic acid; NMR: nuclear magnetic resonance spectroscopy; cryo-EM: cryo-electron microscopy; tRNA: transfer RNA; UTR: untranslated region; mRNA: messenger RNA; CCMK: carbon dioxide-concentrating mechanism protein; GBSSI: granule bound starch synthetase; cDNA: complementary DNA; SDS-PAGE: denaturing gel electrophoresis; GS: glutamine synthetase; PPK: pyruvate, phosphate dikinase; pEU: cell-free expression vector; 6 $\times$ His: a polyhistidine tag, consisting of 6 histidines; 3XFLAG: a 22 amino acid tag; 2D: two-dimensional; ANTI-M2: affinity resin for 3XFLAG tags; SEO1: sieve element occlusion protein; PDX: pyridoxal 5'-phosphate synthase-like subunit; Native PAGE: native gel electrophoresis; TEM: transmission electron microscopy; EGTA: ethylene glycol-bis(2-aminoethylether)-*N,N,N',N'*-tetraacetic acid; 3D: three-dimensional; WGCF: wheat germ cell-free expression; Ni-NTA: nickel-charged affinity group for histidines.

## Authors' contributions

JEE managed all aspects of this work. JEE and IVN devised experiments for the study. IVN performed genetics, molecular biology, cell-free expression and electron microscopy experiments. NS performed cryo-EM screening, data collection and analysis of PDX sample. TM performed all work with CCMK sample. RS assisted IVN with material handling. MJC and DC performed x-ray crystallization screen and diffraction data collection for GS sample. YL, HH, and MK provided clones of PDX and SEO genes and assisted with analysis of the resulting proteins. IVN and JEE wrote initial manuscript draft but all authors contributed to writing the manuscript. All authors read and approved the final manuscript.

## Author details

<sup>1</sup> Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, 3335 Innovation Blvd, Richland, WA 99354, USA. <sup>2</sup> School of Biological Sciences, Washington State University, Pullman, WA 99164, USA. <sup>3</sup> Department of Biological Chemistry, University of California Los Angeles, Howard Hughes Medical Institute, UCLA-DOE Institute for Genomics and Proteomics, Los Angeles, CA 90095, USA. <sup>4</sup> Department of Chemistry and Biochemistry, University of California Los Angeles, Howard Hughes Medical Institute, UCLA-DOE Institute for Genomics and Proteomics, Los Angeles, CA 90095, USA. <sup>5</sup> Department of Bioengineering, University of Illinois at Chicago, Chicago, IL 60607, USA.

## Acknowledgements

A portion of the research was performed using the Environmental Molecular Sciences Laboratory (EMSL), a national scientific user facility sponsored by the Department of Energy's Office of Biological and Environmental Research and located at PNNL. This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357.

## Competing interests

The authors declare that they have no competing interests.

## Availability of data and materials

The plasmid constructs and all datasets, generated in this study, can be provided by the corresponding author upon reasonable request.

## Funding

This work was supported by DOE-BER Mesoscale to Molecules Bioimaging Project FWP# 66382. X-ray crystallization trials were carried out at X-ray Crystallography Core facility at UCLA, supported by DOE Grant DE-FC02-02ER63421. Diffraction data was obtained at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P41 GM103403). The Eiger 16M detector on 24-ID-E beam line is funded by a NIH-ORIP HEI Grant (S100D021527).

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 8 August 2018 Accepted: 31 October 2018

Published online: 10 November 2018

## References

- Blanc-Mathieu, R., Verhelst, B., Derelle, E., Rombauts, S., Bouget, F.Y., Carre, I., Chateau, A., Eyre-Walker, A., Grimsley, N., Moreau, H., Piegu, B., Rivals, E., Schackwitz, W., Van de Peer, Y., Piganeau, G.: An improved genome of the model marine alga *Ostreococcus tauri* unfolds by assessing Illumina de novo assemblies. *BMC Genom.* **15**, 1103 (2014). <https://doi.org/10.1186/1471-2164-15-1103>
- Derelle, E., Ferraz, C., Rombauts, S., Rouze, P., Worden, A.Z., Robbens, S., Partensky, F., Degroev, S., Echeynie, S., Cooke, R., Saeys, Y., Wuyts, J., Jabbari, K., Bowler, C., Panaud, O., Piegu, B., Ball, S.G., Ral, J.P., Bouget, F.Y., Piganeau, G., De Baets, B., Picard, A., Delseny, M., Demaille, J., Van de Peer, Y., Moreau, H.: Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proc. Natl. Acad. Sci. U.S.A.* **103**(31), 11647–11652 (2006). <https://doi.org/10.1073/pnas.0604795103>
- Kigawa, T., Yabuki, T., Matsuda, N., Matsuda, T., Nakajima, R., Tanaka, A., Yokoyama, S.: Preparation of *Escherichia coli* cell extract for highly productive cell-free protein expression. *J. Struct. Funct. Genomics* **5**(1–2), 63–68 (2004). <https://doi.org/10.1023/B:JSGF.0000029204.57846.7d>
- Madin, K., Sawasaki, T., Ogasawara, T., Endo, Y.: A highly efficient and robust cell-free protein synthesis system prepared from wheat embryos: plants apparently contain a suicide system directed at ribosomes. *Proc. Natl. Acad. Sci. U.S.A.* **97**(2), 559–564 (2000). <https://doi.org/10.1073/pnas.97.2.559>
- Sawasaki, T., Hasegawa, Y., Tsuchimochi, M., Kamura, N., Ogasawara, T., Kuroita, T., Endo, Y.: A bilayer cell-free protein synthesis system for high-throughput screening of gene products. *FEBS Lett.* **514**(1), 102–105 (2002)
- Mureev, S., Kovtun, O., Nguyen, U.T., Alexandrov, K.: Species-independent translational leaders facilitate cell-free expression. *Nat. Biotechnol.* **27**(8), 747–752 (2009). <https://doi.org/10.1038/nbt.1556>
- Kovtun, O., Mureev, S., Jung, W., Kubala, M.H., Johnston, W., Alexandrov, K.: Leishmania cell-free protein expression system. *Methods* **55**(1), 58–64 (2011). <https://doi.org/10.1016/j.jymeth.2011.06.006>
- Ezure, T., Suzuki, T., Shikata, M., Ito, M., Ando, E.: A cell-free protein synthesis system from insect cells. *Methods Mol. Biol.* **607**, 31–42 (2010). [https://doi.org/10.1007/978-1-60327-331-2\\_4](https://doi.org/10.1007/978-1-60327-331-2_4)
- Mikami, S., Kobayashi, T., Imataka, H.: Cell-free protein synthesis systems with extracts from cultured human cells. *Methods Mol. Biol.* **607**, 43–52 (2010). [https://doi.org/10.1007/978-1-60327-331-2\\_5](https://doi.org/10.1007/978-1-60327-331-2_5)
- Buntru, M., Vogel, S., Spiegel, H., Schillberg, S.: Tobacco BY-2 cell-free lysate: an alternative and highly-productive plant-based in vitro translation system. *BMC Biotechnol.* **14**, 37 (2014). <https://doi.org/10.1186/1472-6750-14-37>
- Brodel, A.K., Sonnabend, A., Kubick, S.: Cell-free protein expression based on extracts from CHO cells. *Biotechnol. Bioeng.* **111**(1), 25–36 (2014). <https://doi.org/10.1002/bit.25013>
- Hodgman, C.E., Jewett, M.C.: Optimized extract preparation methods and reaction conditions for improved yeast cell-free protein synthesis. *Biotechnol. Bioeng.* **110**(10), 2643–2654 (2013). <https://doi.org/10.1002/bit.24942>

13. Thoring, L., Dondapati, S.K., Stech, M., Wustenhagen, D.A., Kubick, S.: High-yield production of "difficult-to-express" proteins in a continuous exchange cell-free system based on CHO cell lysates. *Sci. Rep.* **7**, 11710 (2017). <https://doi.org/10.1038/s41598-017-12188-8>
14. Carraher, C., Nazmi, A.R., Newcomb, R.D., Kralicek, A.: Recombinant expression, detergent solubilisation and purification of insect odorant receptor subunits. *Protein Expr. Purif.* **90**(2), 160–169 (2013). <https://doi.org/10.1016/j.pep.2013.06.002>
15. Genji, T., Nozawa, A., Tozawa, Y.: Efficient production and purification of functional bacteriorhodopsin with a wheat-germ cell-free system and a combination of Fos-choline and CHAPS detergents. *Biochem. Biophys. Res. Commun.* **400**(4), 638–642 (2010). <https://doi.org/10.1016/j.bbrc.2010.08.119>
16. Katzen, F., Peterson, T.C., Kudlicki, W.: Membrane protein expression: no cells required. *Trends Biotechnol.* **27**(8), 455–460 (2009). <https://doi.org/10.1016/j.tibtech.2009.05.005>
17. Vinarov, D.A., Lytle, B.L., Peterson, F.C., Tyler, E.M., Volkman, B.F., Markley, J.L.: Cell-free protein production and labeling protocol for NMR-based structural proteomics. *Nat. Methods* **1**(2), 149–153 (2004). <https://doi.org/10.1038/nmeth716>
18. Kang, S.H., Jun, S.Y., Kim, D.M.: Fluorescent labeling of cell-free synthesized proteins by incorporation of fluorophore-conjugated nonnatural amino acids. *Anal. Biochem.* **360**(1), 1–6 (2007). <https://doi.org/10.1016/j.ab.2006.10.029>
19. Quast, R.B., Mrusek, D., Hoffmeister, C., Sonnabend, A., Kubick, S.: Cotranslational incorporation of non-standard amino acids using cell-free protein synthesis. *FEBS Lett.* **589**(15), 1703–1712 (2015). <https://doi.org/10.1016/j.febslet.2015.04.041>
20. Cappuccio, J.A., Blanchette, C.D., Sulchek, T.A., Arroyo, E.S., Kralj, J.M., Hinz, A.K., Kuhn, E.A., Chromy, B.A., Segelke, B.W., Rothschild, K.J., Fletcher, J.E., Katzen, F., Peterson, T.C., Kudlicki, W.A., Bench, G., Hoeprich, P.D., Coleman, M.A.: Cell-free co-expression of functional membrane proteins and apolipoprotein, forming soluble nanolipoprotein particles. *Mol. Cell. Proteomics* **7**(11), 2246–2253 (2008). <https://doi.org/10.1074/mcp.M800191-MCP200>
21. Shinoda, T., Shinya, N., Ito, K., Ishizuka-Katsura, Y., Ohsawa, N., Terada, T., Hirata, K., Kawano, Y., Yamamoto, M., Tomita, T., Ishibashi, Y., Hirabayashi, Y., Kimura-Someya, T., Shirouzu, M., Yokoyama, S.: Cell-free methods to produce structurally intact mammalian membrane proteins. *Sci. Rep.* **6**, 30442 (2016). <https://doi.org/10.1038/srep30442>
22. Hanawa-Suetsugu, K., Kukimoto-Niino, M., Mishima-Tsumagari, C., Akasaka, R., Ohsawa, N., Sekine, S., Ito, T., Tochio, N., Koshiba, S., Kigawa, T., Terada, T., Shirouzu, M., Nishikimi, A., Uruno, T., Katakai, T., Kinashi, T., Kohda, D., Fukui, Y., Yokoyama, S.: Structural basis for mutual relief of the Rac guanine nucleotide exchange factor DOCK2 and its partner ELMO1 from their autoinhibited forms. *Proc. Natl. Acad. Sci. U.S.A.* **109**(9), 3305–3310 (2012). <https://doi.org/10.1073/pnas.1113512109>
23. Mudeppa, D.G., Rathod, P.K.: Expression of functional *Plasmodium falciparum* enzymes using a wheat germ cell-free system. *Eukaryot. Cell* **12**(12), 1653–1663 (2013). <https://doi.org/10.1128/Ec.00222-13>
24. Terada, T., Kusano, S., Matsuda, T., Shirouzu, M., Yokoyama, S.: Cell-free protein production for structural biology. In: Senda, T., Maenaka, K. (eds.) *Advanced methods in structural biology*. Springer protocols handbooks. Springer, Tokyo (2016)
25. Carlson, E.D., Gan, R., Hodgman, C.E., Jewett, M.C.: Cell-free protein synthesis: applications come of age. *Biotechnol. Adv.* **30**(5), 1185–1194 (2012). <https://doi.org/10.1016/j.biotechadv.2011.09.016>
26. Goshima, N., Kawamura, Y., Fukumoto, A., Miura, A., Honma, R., Satoh, R., Wakamatsu, A., Yamamoto, J., Kimura, K., Nishikawa, T., Andoh, T., Iida, Y., Ishikawa, K., Ito, E., Kagawa, N., Kaminaga, C., Kanehori, K., Kawakami, B., Kenmochi, K., Kimura, R., Kobayashi, M., Kuroita, T., Kuwayama, H., Maruyama, Y., Matsuo, K., Minami, K., Mitsubori, M., Mori, M., Morishita, R., Murase, A., Nishikawa, A., Nishikawa, S., Okamoto, T., Sakagami, N., Sakamoto, Y., Sasaki, Y., Seki, T., Sono, S., Sugiyama, A., Sumiya, T., Takayama, T., Takayama, Y., Takeda, H., Togashi, T., Yahata, K., Yamada, H., Yanagisawa, Y., Endo, Y., Imamoto, F., Kisu, Y., Tanaka, S., Isogai, T., Imai, J., Watanabe, S., Nomura, N.: Human protein factory for converting the transcriptome into an in vitro-expressed proteome. *Nat. Methods* **5**(12), 1011–1017 (2008)
27. Vinarov, D.A., Newman, C.L.L., Markley, J.L.: Wheat germ cell-free platform for eukaryotic protein production. *FEBS J.* **273**(18), 4160–4169 (2006). <https://doi.org/10.1111/j.1742-4658.2006.05434.x>
28. Harbers, M.: Wheat germ systems for cell-free protein expression. *FEBS Lett.* **588**(17), 2762–2773 (2014). <https://doi.org/10.1016/j.febslet.2014.05.061>
29. Watanabe, M., Miyazono, K., Tanokura, M., Sawasaki, T., Endo, Y., Kobayashi, I.: Cell-free protein synthesis for structure determination by X-ray crystallography. *Methods Mol. Biol.* **607**, 149–160 (2010). [https://doi.org/10.1007/978-1-60327-331-2\\_13](https://doi.org/10.1007/978-1-60327-331-2_13)
30. Sawasaki, T., Ogasawara, T., Morishita, R., Endo, Y.: A cell-free protein synthesis system for high-throughput proteomics. *Proc. Natl. Acad. Sci. U.S.A.* **99**(23), 14652–14657 (2002). <https://doi.org/10.1073/pnas.232580399>
31. Takai, K., Sawasaki, T., Endo, Y.: Practical cell-free protein synthesis system using purified wheat embryos. *Nat. Protoc.* **5**(2), 227–238 (2010). <https://doi.org/10.1038/nprot.2009.207>
32. Marx, V.: PCR: the price of infidelity. *Nat. Methods* **13**(6), 475–479 (2016). <https://doi.org/10.1038/nmeth.3868>
33. Sawasaki, T., Hasegawa, Y., Tsuchimochi, M., Kasahara, Y., Endo, Y.: Construction of an efficient expression vector for coupled transcription/translation in a wheat germ cell-free system. In: *Nucleic acids symposium series*, vol. 44, pp. 9–10 (2000)
34. Zielonka, S., Ernst, A.M., Hawat, S., Twyman, R.M., Pruffer, D., Noll, G.A.: Characterization of five subgroups of the sieve element occlusion gene family in *Glycine max* reveals genes encoding non-forisome P-proteins, forisomes and forisome tails. *Plant Mol. Biol.* **86**(1–2), 51–67 (2014). <https://doi.org/10.1007/s11103-014-0211-z>
35. Pellissier, H.C., Peters, W.S., Collier, R., van Bel, A.J., Knoblauch, M.: GFP tagging of sieve element occlusion (SEO) proteins results in green fluorescent forisomes. *Plant Cell Physiol.* **49**(11), 1699–1710 (2008). <https://doi.org/10.1093/pcp/pcn141>
36. Leuendorf, J.E., Mooney, S.L., Chen, L.Y., Hellmann, H.A.: *Arabidopsis thaliana* PDX1.2 is critical for embryo development and heat shock tolerance. *Planta* **240**(1), 137–146 (2014). <https://doi.org/10.1007/s00425-014-2069-3>
37. Dryden, K.A., Crowley, C.S., Tanaka, S., Yeates, T.O., Yeager, M.: Two-dimensional crystals of carboxysome shell proteins recapitulate the hexagonal packing of three-dimensional crystals. *Protein Sci.* **18**(12), 2629–2635 (2009). <https://doi.org/10.1002/pro.272>
38. Cai, F., Sutter, M., Bernstein, S.L., Kinney, J.N., Kerfeld, C.A.: Engineering bacterial microcompartment shells: chimeric shell proteins and chimeric carboxysome shells. *ACS Synth. Biol.* **4**(4), 444–453 (2015). <https://doi.org/10.1021/sb500226j>
39. Rodrigues, M.J., Windeisen, V., Zhang, Y., Guedez, G., Weber, S., Strohmaier, M., Hanes, J.W., Royant, A., Evans, G., Sinning, I., Ealick, S.E., Begley, T.P., Tews, I.: Lysine relay mechanism coordinates intermediate transfer in vitamin B6 biosynthesis. *Nat. Chem. Biol.* **13**(3), 290–294 (2017). <https://doi.org/10.1038/nchembio.2273>
40. Knoblauch, M., Noll, G.A., Muller, T., Pruffer, D., Schneider-Huther, I., Scharner, D., Van Bel, A.J., Peters, W.S.: ATP-independent contractile proteins from plants. *Nat. Mater.* **2**(9), 600–603 (2003). <https://doi.org/10.1038/nmat960>
41. Knoblauch, M., Stubenrauch, M., van Bel, A.J., Peters, W.S.: Forisome performance in artificial sieve tubes. *Plant Cell Environ.* **35**(8), 1419–1427 (2012). <https://doi.org/10.1111/j.1365-3040.2012.02499.x>
42. Khoshouei, M., Radjainia, M., Baumeister, W., Danev, R.: Cryo-EM structure of haemoglobin at 3.2 Å determined with the Volta phase plate. *Nat. Commun.* **8**, 16099 (2017). <https://doi.org/10.1038/ncomms16099>
43. Khoshouei, M., Radjainia, M., Phillips, A.J., Gerrard, J.A., Mitra, A.K., Plietzko, J.M., Baumeister, W., Danev, R.: Volta phase plate cryo-EM of the small protein complex Prx3. *Nat. Commun.* **7**, 10534 (2016). <https://doi.org/10.1038/ncomms10534>
44. van Rooyen, J.M., Abratt, V.R., Belrhali, H., Sewell, T.: Crystal structure of type III glutamine synthetase: surprising reversal of the inter-ring interface. *Structure* **19**(4), 471–483 (2011). <https://doi.org/10.1016/j.str.2011.02.001>
45. Unno, H., Uchida, T., Sugawara, H., Kurisu, G., Sugiyama, T., Yamaya, T., Sakakibara, H., Hase, T., Kusunoki, M.: Atomic structure of plant glutamine synthetase: a key enzyme for plant productivity. *J. Biol. Chem.* **281**(39), 29287–29296 (2006). <https://doi.org/10.1074/jbc.M601497200>
46. Biyani, N., Righetto, R.D., McLeod, R., Caujolle-Bert, D., Castano-Diez, D., Goldie, K.N., Stahlberg, H.: Focus: the interface between data collection and data processing in cryo-EM. *J. Struct. Biol.* **198**(2), 124–133 (2017). <https://doi.org/10.1016/j.jsb.2017.03.007>
47. Grant, T., Rohou, A., Grigorieff, N.: cisTEM, user-friendly software for single-particle image processing. *Elife* **7**, e35383 (2018). <https://doi.org/10.7554/elife.35383>