International Journal of
Molecular Sciences

MDPI

Article

Mitigating the Blurring Effect of CryoEM Averaging on a Flexible and Highly Symmetric Protein Complex through Sub-Particle Reconstruction

Diana S. Suder and Shane Gonen *

Department of Molecular Biology and Biochemistry, University of California Irvine, Irvine, CA 92697, USA
* Correspondence: gonens@uci.edu

Abstract: Many macromolecules are inherently flexible as a feature of their structure and function. During single-particle CryoEM processing, flexible protein regions can be detrimental to high-resolution reconstruction as signals from thousands of particles are averaged together. This "blurring" effect can be difficult to overcome and is possibly more pronounced when averaging highly symmetric complexes. Approaches to mitigating flexibility during CryoEM processing are becoming increasingly critical as the technique advances and is applied to more dynamic proteins and complexes. Here, we detail the use of sub-particle averaging and signal subtraction techniques to precisely target and resolve flexible DARPin protein attachments on a designed tetrahedrally symmetric protein scaffold called DARP14. Particles are first aligned as full complexes, and then the symmetry is reduced by alignment and focused refinement of the constituent subunits. The final reconstructions we obtained were vastly improved over the fully symmetric reconstructions, with observable secondary structure and side-chain placement. Additionally, we were also able to reconstruct the core region of the scaffold to 2.7 Å. The data processing protocol outlined here is applicable to other dynamic and symmetric protein complexes, and our improved maps could allow for new structure-guided variant designs of DARP14.

Keywords: CryoEM; protein design; single-particle; cage; scaffold; sub-particle; flexibility

check for updates

Citation: Suder, D.S.; Gonen, S. Mitigating the Blurring Effect of CryoEM Averaging on a Flexible and Highly Symmetric Protein Complex through Sub-Particle Reconstruction. *Int. J. Mol. Sci.* **2024**, *25*, 5665. https://doi.org/10.3390/ijms25115665

Academic Editor: Slavica Jonić

Received: 31 March 2024
Revised: 13 May 2024
Accepted: 14 May 2024
Published: 23 May 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Over the last decade, single-particle cryogenic electron microscopy (CryoEM) has become a powerful structural biology method routinely used to study proteins at high resolution [1]. As instrumentation and processing software improve, increasingly higher resolutions are achievable [2], less starting material is required [3,4], and the technique can be applied to ever more complex and difficult targets [5]. As images of individual proteins are used, different conformational states can also be observed by separation and classification during data processing [6,7]. However, major challenges deterring high resolution studies still exist, including protein size [8] and dynamics [9]. Small proteins of ~50 kilodaltons (kDa) or smaller are difficult or sometimes impossible to study on their own [8], and flexible protein regions can blur out high-resolution information [10,11].

CryoEM reconstructions rely on aligning and averaging images of thousands of individual proteins. The overall movement of proteins in ice during data collection is typically alleviated by motion correction, where individual movie frames are realigned [12,13]. However, dynamics inherent to the proteins themselves are still a barrier to high resolution due to averaging [14]. This issue can be even more pronounced when symmetry-related subunits are involved [15]. Many different strategies have been described to try to mitigate these issues, including biochemical stabilizers [16,17], binding of ligands aiming to lock the protein into one conformation [18], and fiducial markers for alignment [19]. Additionally, many data processing methods have also been developed to address this issue, including focused refinement, whereby small areas of the map are used as targets while others are

masked or subtracted away from the data [20]; non-uniform refinement [10]; symmetric expansion [21]; multi-body refinement [22]; and sub-particle reconstruction [23], whereby each individual subunit making up the complex is isolated and averaged as a separate particle. Recently, computationally designed proteins called scaffolds, large proteins designed to bind and display small proteins of interest for high-resolution reconstructions, have been developed [24–26]. While these scaffolds aim to bypass the protein size barrier, flexibility of the scaffold architecture and/or the binding of target proteins can still exist, limiting high-resolution detail.

Here, we describe and detail the use of sub-particle CryoEM averaging and present updated reconstructions of DARP14, a multi-subunit tetrahedrally symmetric protein complex. Originally described by Liu and Gonen et al. (2018) [25], DARP14 was computationally designed as a scaffold capable of displaying small proteins for CryoEM (Figure 1A). DARPins are small, ~17 kDa helical repeat proteins that feature variable residue loops capable of selective non-covalent binding to other proteins [27], akin to antibodies. Using helical extension, DARPins were genetically fused to one of two trimeric subunits making up a previously designed, self-assembling cage called T33-21 [28]. Since T33-21 assembles with 4 copies of each trimeric subunit (A and B), 12 copies of DARPin are displayed on the final 24-subunit, ~600 kDa chimera termed DARP14 (Figure 1A). Variants of DARP14 and similar scaffolds have since been generated and used to bind and image small proteins, such as GFP [29] and KRAS [30]. In the original 2018 study of DARP14, the core of the protein making up the original T33-21 design was found to be very stable and could be resolved to 3.2 Å resolution when applying tetrahedral symmetry during averaging. In contrast, the DARPin proteins were found to have a marked deviation from the original designed helical extension. Owing to their flexibility away from ideal tetrahedral symmetry and possible terminal unwinding, they could not be resolved to the same resolution ranges as the core and could only be observed at much lower resolution ranges.

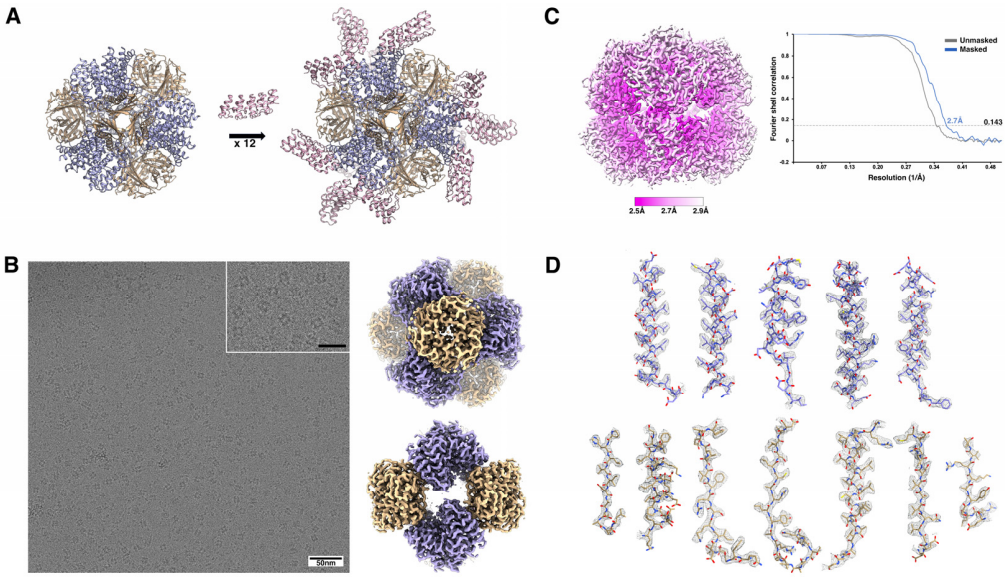

Figure 1. Design overview of DARP14 and CryoEM map of the core scaffold. (**A**) A tetrahedrally symmetric protein scaffold was designed [25] to incorporate small proteins called DARPins via helical extensions in each chain of each subunit A. Subunit A now features three DARPins, totaling 12 DARPins displayed on the protein scaffold, together termed DARP14. (**B**) (**left**) CryoEM micrograph of DARP14 (inset—zoomed-in view, scale bar = 25 nm). (**right**) 2.7 Å map of the core tetrahedral scaffold of DARP14, viewed along the 3- and 2-fold axes, with DARPins masked out. (**C**) (**left**) Local resolution of the map from (**B**) with the corresponding Fourier Shell Correlation (**right**). (**D**) Our refined atomic model was fit into the EM map from (**B**) using UCSF ChimeraX (v1.5) with the "fit in map" command and displayed using the mesh map view. Trimeric A subunits in purple, B subunits in yellow, and DARPin subunits in pink.

The methodology we outline here resulted in new reconstructions of DARP14, improving on the original core T33-21 protein to 2.7 Å resolution (Figure 1B–D), and, crucially, significantly improves the DARPin reconstruction. Using new software and various refinement methods, mainly sub-particle reconstruction, our new maps allowed us to accurately observe the DARPin–cage helical extension as well as areas of close contact between the two proteins. We describe our structural data processing approach, detail the new maps, and discuss their observable features. The approach used here should be applicable to the refinement of most flexible proteins, including highly symmetric proteins and newly designed variants of DARP14.

2. Results and Discussion

While symmetry can be very beneficial for CryoEM analysis, essentially multiplying the number of particles, for some proteins where dynamics are present, symmetry can have a multiplicative negative effect. The processing pipeline we chose leveraged symmetry at the low-resolution level (initial alignment and three-dimensional (3D) refinement) and mitigated the detrimental blurring effects of symmetry when accuracy was required at high resolution (focused refinement of the flexible DARPins). Sub-particle reconstruction and our processing pipelines are summarized in Figures 2 and 3. High-resolution reconstructions of DARP14 can be calculated using standard methods of CryoEM, and medium- to low-resolution maps only require a few thousand particles. These reconstructions typically have strong signal for the core cage, but the DARPins, due to their flexible nature, are not prominent. For the DARPin signal, we turned to sub-particle reconstruction coupled with signal subtraction. These two processing methods rely on accurate alignment at high resolution so that only the correct signal is subtracted, and particle coordinates can be shifted accurately to the corresponding protein sections for focused refinement. We therefore first reconstructed the full DARP14 complex to an overall resolution of 3.1 Å. For this initial reconstruction, particle counts were favored over slight improvements to resolution to keep as much data intact for future rounds of sub-particle analysis and classifications. This reconstruction therefore served as a foundational starting point for all subsequent sub-particle refinements.

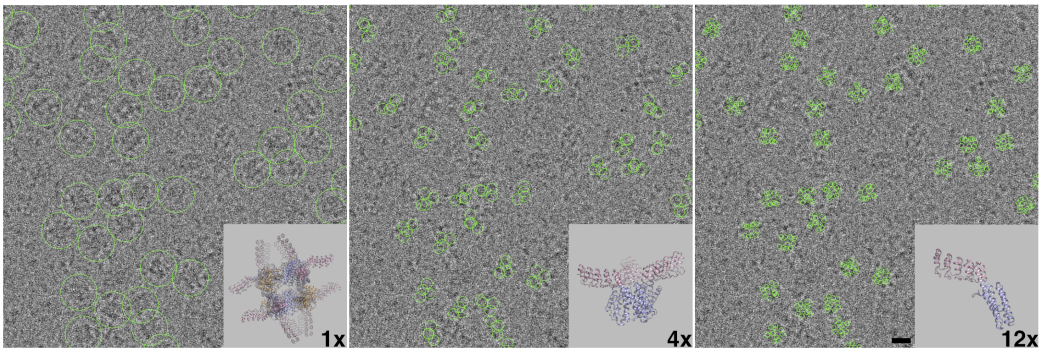

Figure 2. Sub-particle reconstruction overview. (**Left**) A representative section of a micrograph with many copies of the full tetrahedral DARP14 protein, circled. (**Center**) The same representative micrograph with C3 symmetrical trimer coordinates circled; each full DARP14 protein is composed of four trimers. (**Right**) The same representative micrograph with individual DARPin coordinates circled; each full DARP14 protein can result in up to 12 monomeric DARPin coordinates. Scale bar = 10 nm. These micrographs visually represent how one DARP14 cage ultimately gave rise to signal for 12 DARPins that could be combined during data processing to mitigate the effect of symmetry of this small flexible subunit.

In addition to focusing on the DARPin components, we also set out to extract as much high-resolution information as possible from the stable core cage component, this time by masking out the dynamic parts of the complex. Using available and established processing software and scripts combined with careful 3D classifications, signal subtractions, and sub-

particle reconstructions, we determined CryoEM maps including the core cage component and the isolated symmetric pieces of DARP14 containing trimeric and monomeric DARPins.

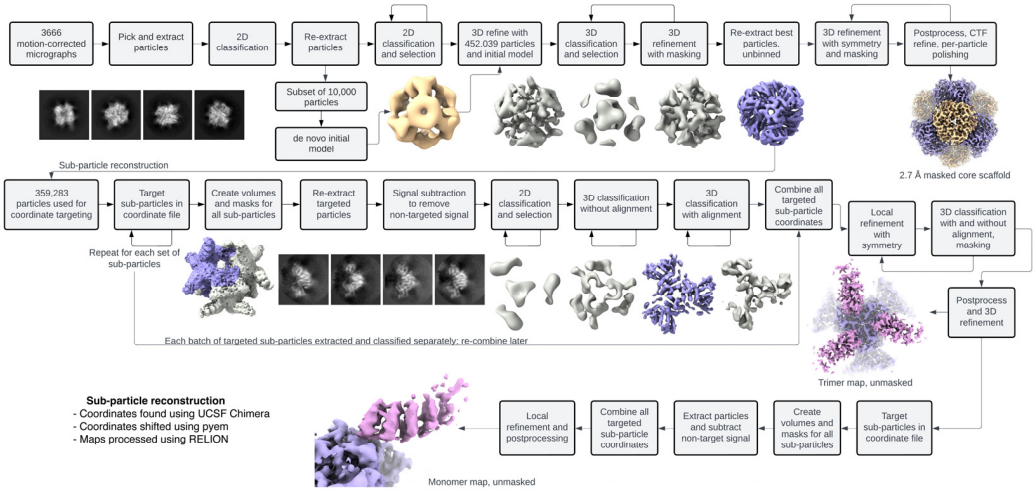

Figure 3. Data processing overview. The CryoEM processing pipeline used in this study is outlined using the sub-particle averaging approach to reach the final trimeric and monomeric maps. The software required for the protocol used here of sub-particle reconstruction is outlined.

2.1. Sub-Particle Reconstruction of Trimeric and Monomeric Subunits of DARP14

For sub-particle reconstruction and signal subtraction, ideally only the components of interest (in this case each DARPin trimer or monomer) would be the sole focus. However, since these components are very small, they are still difficult to align during refinement. Therefore, for both the trimeric and subsequent monomeric reconstructions, extended parts of the core cage were kept intact to allow for sufficient signal to remain for classifications and local refinement. Coordinates were shifted from the main cage to the center of each of the four trimeric A subunits, followed by masking and subtraction (Figures 2 and 3). These sets were then combined, quadrupling the initial particle count. We were then able to proceed with further classifications and refine the particles using applied C3 symmetry. An example 2D classification is shown in Figure S1, whereby high-resolution features for DARPin were chosen for the next steps. Using the newly calculated C3 symmetric map, the DARPins were more defined at a higher overall resolution than the full tetrahedrally symmetric DARP14 map, with observable secondary structure and some side-chain features (Figure 4A).

Stemming from the previous trimeric particle dataset, further subtraction and sub-particle reconstructions were made whereby each monomeric subunit was isolated. Using the same methodology as in the preceding run, the coordinates were shifted from the trimers to the center of each of the three monomers and combined, tripling the dataset. At this point, up to 12 DARPin monomers were retained from each full DARP14 particle found in the micrographs (Figure 2). Since the monomers lack symmetry, our final reconstructions were therefore made using no applied (C1) symmetry.

In contrast to the full complex, the resulting trimeric and monomeric DARPin maps feature vastly improved high-resolution information, particularly for the helical extension linker and residues close to the cage, facilitating model building and demonstrating the effectiveness of reducing the effect of high symmetry on averaging (Figure 4). The most significant improvement to the map came from reducing T-symmetry to C3. The resolution range for both C3 and C1 was comparable; however, C3 DARPin information looked clearer closer to the core cage, while C1 DARPin looked clearer further away from the core cage. The latter residues of the DARPins remained at lower resolutions, likely owing to a significantly higher degree of flexibility in those regions and possible unwinding of the structure [25]. A significantly higher number of particles will need to be collected to attempt further classifications to see if signal can be improved at those locations.

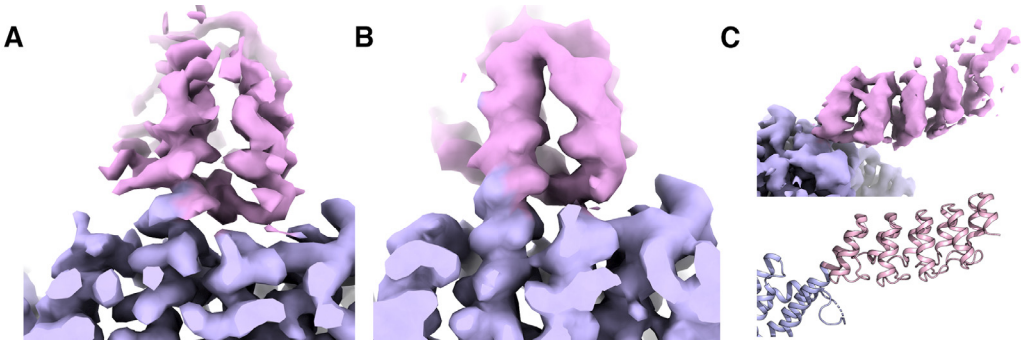

Figure 4. Trimeric and monomeric DARPin CryoEM maps obtained through sub-particle reconstruction and subtraction. (**A**) Trimeric map showing the designed helical extension into the DARPin. Clear secondary and side-chain information can be observed. (**B**,**C**) Monomeric map showing the helical extension view (**B**) and side view, where the helical repeats can be observed and oriented to the model (**C**).

2.2. Stabilizing the DARPin–Cage Interaction

All DARPin proteins displayed on the scaffold contain a single anchoring point, the extended helical linker. While further interactions between neighboring DARPins have proved to help stabilize a variant of the complex [30], additional interactions could be made using structure-guided design between the core cage and DARPin. Several residues can be found in close proximity to the core cage, potentially allowing for non-covalent interactions and/or covalent disulfide bridges. A secondary anchor point could stabilize the movement of the entire DARPin in relation to the core cage, allowing for symmetrical reconstruction and refinement and eliminating or reducing the need for focused refinements. Additionally, results from the original 2018 manuscript [25] discovered a deviation of the helical extension away from the design. It may be possible to further strengthen the helical extension without modifying the cage by finding residues amenable to mutations compatible with cystine cross-linkers for intra-residue interactions [31].

2.3. Reconstruction of the Core Cage of DARP14

To push for the highest resolution possible given the dataset, we created a mask that excluded the flexible DARPin components while keeping the core cage of DARP14 intact. Multiple rounds of 3D classifications and refinements using tetrahedral symmetry were made until no improvements to the map were observed. An example 2D classification is shown in Figure S2, whereby high-resolution features were chosen for the next steps. A final set of 277,091 particles were postprocessed using multiple rounds of Bayesian polishing and CTF refinement to obtain a final map of the core cage at 2.7 Å (Figures 1B–D and 3). Local resolution is shown in (Figure 1C). This is a marked improvement in overall resolution compared to the original 3.2 Å study. At this resolution, we could model nearly every side chain of the complex. The map also shows possible sites of ions (mainly in the 3-fold trimeric interfaces) and water molecules (not modeled). The previous DARPin core model (PDB ID 6C9I) was initially fit into our new map using UCSF Chimera (v1.15) with the "fit in map" command and refined using Phenix v1.20.1 and COOT v0.9.5 (see Materials and Methods) and displayed using UCSF ChimeraX (v1.5) (Figure 1D). At this point the resolution of the core cage is likely reaching the limits inherent to the sample and data collection parameters.

2.4. Processing Highly Symmetric Protein Complexes by CryoEM

Sub-particle reconstruction can be combined with new advances and methods of CryoEM processing to extract information from other highly symmetric and non-symmetric particles, even those with minimal flexibility. The tools for sub-particle reconstruction are readily available; however, studies and protocols for using it are currently sparse. Here, we described the processing pipeline when applied to a tetrahedrally symmetric scaffold

protein DARP14. Our resulting maps could allow for structure-guided redesign of the original DARPin construct.

3. Materials and Methods

3.1. CryoEM Data Processing

The previous dataset of 3666 movies was imported into RELION 3.0.8 [32]. Raw movies were corrected for beam-induced motion using the RELION implementation algorithm and all 24 movie frames. CTF estimation was carried out using the GCTF [33] executable in RELION. Particle coordinates were picked automatically with optimized parameters using Gautomatch v0.53 [34] and imported into RELION, then extracted and binned to speed up processing. All subsequent steps were performed in RELION unless otherwise specified. Reference-free 2D classification was carried out initially using 200 classes, and those displaying recognizable features were selected. Selected particles were re-extracted and re-centered. To generate a C1 (non-symmetric) de novo initial model, a subset of these particles having already undergone one round of selection were subjected to another 2D classification, this time using 100 classes. The classes with the strongest and clearest signal were selected and limited to 10,000 particles as the input for initial model creation. An additional initial model was made from the C1 model by applying T-symmetry with relion_image_handler. The full set of selected classified particles was re-extracted and split into three groups of approximately 500,000 particles each to speed up processing and enable finer sampling, then subjected to two more rounds of 2D classification. Classes displaying recognizable particles were selected with increasingly stringent criteria each time; selected particles were eventually re-extracted with a box size of 192 pixels, binned by two (1.31 Å/pixel), and joined into one set of 452,039 particles to continue to 3D processing.

All 452,039 particles were refined without applied symmetry and without a mask. This refinement provided an EM map displaying at least partial signal for most of the 12 DARPins on the full scaffold, albeit at a low resolution.

3.2. Refinement of the Core Scaffold to 2.7 Å Resolution

To refine the core scaffold to the highest possible resolution, we first created a solvent mask that excluded the flexible DARPins using UCSF Chimera v1.15 [35]. Starting from the first refinement, we ran a 3D classification with 4 classes in C1. The best class with 359,283 particles was selected as the input into a second 3D refinement. This process of 3D classification followed by class selection and 3D refinement was repeated twice, keeping the best particles. An example 3D classification is highlighted in Figure S3, whereby the highest resolution class was chosen for the next steps of processing. A selected class of 309,973 particles was autorefined in T symmetry, then subjected to a round of postprocessing, CTF refinement, Bayesian polishing, and 3D refinement. Then, the particles were 3D classified again without angular sampling, with the intent of keeping the most well-aligned particles. Selected particles were again refined, postprocessed, CTF refined, Bayesian polished, refined, and again subjected to 3D classification where the best class of 277,091 particles was selected. These remaining 277,091 particles were subjected to autorefinement, postprocessing, CTF refinement, and Bayesian polishing. They were re-extracted to a pixel size of 0.9825 Å/pixel and a box size of 296 pixels. The final refined and postprocessed reconstruction containing only the core cage component of DARP14 reached an overall resolution of 2.69 Å using the gold-standard FSC 0.143 threshold.

3.3. Sub-Particle Reconstruction of the Trimeric Subunits of DARP14

As a starting point for breaking up the components of DARP14 into their trimeric A subunits using sub-particle reconstruction, an early refinement of the full DARP14 particle was used (the same as for the previous section), composed of 359,283 particles. Using UCSF Chimera v1.15, coordinates in 3D space that corresponded to the center of each of the four A subunit trimers were found. Using the subparticles.py script available in the

pyem package v0.5 [36], the coordinates of the refinement star file were shifted to the new trimer coordinates four times, once for each trimer.

Using UCSF Chimera v1.15, two maps and corresponding masks were made for each of the four trimers—one containing the trimer only (volume to keep) and one containing everything else (volume to subtract). Some extended parts of the neighboring subunits were kept intact with each trimer to aid alignment. Using the same parameters as for subparticles.py, the map.py script (pyem package [36]) was used to re-center the .mrc files. All 359,283 particles were then re-extracted four separate times using each shifted .star file with a box size of 240 pixels; the box size was chosen to appropriately frame the trimer. Signal subtraction was performed four separate times using these particles and their corresponding .mrc files and masks, resulting in particles containing a trimer of DARPins and most of the main scaffold removed. Successful subtraction was confirmed using a subset of particles and relion_reconstruct. Similarly, a cut down version of the initial model was made, representing only the relevant signal.

Following signal subtraction, multiple rounds of 2D classifications were performed where strong DARPin signal was favored between rounds. Iterative rounds of maskless 3D classifications with and without alignment followed until the best class converged. An example 3D classification is highlighted in Figure S3, whereby the highest resolution class was chosen for the next steps of processing. Finally, a mask containing the three DARPins was made to isolate the flexible structure of interest and used for new rounds of 3D classifications using C1 (no symmetry). All four trimeric particle groups were joined into a single curated set of 366,069 particles. This curated set was then used to push to the highest resolution and map quality possible with new rounds of 2D and 3D classifications using C1 and C3 symmetry, CTF refinement, particle polishing, and postprocessing. The final map used 95,889 particles refined with C3 symmetry. Local resolution estimates ranged from ~3.3–5 Å across the first two helical repeats and ~5–8 Å towards the tip of the DARPin component.

3.4. Sub-Particle Reconstruction of the Monomeric Subunits of DARP14

Using the final trimeric set containing 95,889 particles, a similar protocol was followed as in the previous section to target each individual DARPin, tripling the number of particles in the dataset. Further rounds of 3D classification did not reveal classes that were of sufficiently better resolution or map quality; therefore, the final refinement used all 287,667 particles without symmetry.

3.5. Model Building, Refinement, and Validation

For the core scaffold, a previously published set of coordinates, PDB 6C9I, was aligned into the final refined EM map using UCSF Chimera v1.15 [37] with the "fit in map" command, and residues were allowed to be refined individually against the map using Phenix.real_space_refine [38]. The new atomic coordinates from Phenix (v1.20.1) were imported into COOT [39] (v0.9.5) to be manually inspected for steric clashes and other potential issues. For the DARPin monomer, a subset of previously published coordinates (PDB 6C9K) containing only chain A and the attached DARPin were used, along with the final refined EM map for the DARPin component, and were aligned in UCSF Chimera v1.15. Residues were allowed to be refined individually against the map using Phenix.real_space_refine, and the EM map and coordinates were imported into COOT to be manually inspected. MolProbity [40] was used inside of the Phenix software package to evaluate global and local model quality. Model building and validation statistics are available in Table 1; all the values that appear in this table under Model Validation were generated in Phenix v1.20.1 and COOT v0.9.5 during the described procedures.

Figures were prepared using UCSF Chimera v1.15 and v1.16 [35], UCSF ChimeraX v1.3 and v1.5 [37], PyMOL v2.3.0 [41], Microsoft PowerPoint v16.78.3 (Microsoft, Redmond WA, USA) and Adobe Photoshop 2022 (Adobe, San Jose CA, USA).

Table 1. CryoEM data collection and processing parameters.

Experimental Parameters	Experimental Values	
Imaging		
Grid type	C-Flat 1.2/1.3 200-mesh (EMS)	
Microscope	Titan Krios	
Camera	Gatan K2 direct electron detector	
Voltage	300 kV	
Magnification	22,500×	
Dose per frame	1.2 $e^-/Å^2$	
Frames per movie	24	
Defocus range	−0.7 to −2.5 µm	
Pixel size	0.655 Å/px	
Number of movies	3666	
CryoEM Maps	Core Scaffold	DARPin Trimer/Monomer
Number of particles	277,091	95,889 (Trimer), 287,667 (Monomer)
Pixel size	0.9825 Å/px	1.31 Å/px
Box size (px)	296	240
Overall map resolution	2.7 Å	N/A
FSC for resolution calculation	0.143	0.143
Applied Symmetry	T	C3 (Trimer), C1 (Monomer)
Model Validation		
Ramachandran:	Core Scaffold	DARPin (Fit in Monomer)
Favored %	99.83	99.82
Allowed %	0.17	0.18
Outliers %	0	0
MolProbity score	1.21	1.51

4. Conclusions

Protein flexibility is often detrimental for calculating high-resolution CryoEM maps due to averaging. Here, we detail the use of sub-particle reconstruction combined with signal subtraction when applied to a highly symmetric protein scaffold. Improvement of the map with respect to the flexible DARPin components was observed, with both secondary structure and side-chain information, particularly close to the core cage. Future redesigns of DARP14 informed by our new maps could be made. This methodology can be combined with other processing steps as well as biochemical or construct optimizations to help mitigate issues of flexibility. While we outlined a strategy for reducing the effects of symmetry by realigning the individual building-block subunits, other datasets containing non-symmetric proteins could benefit from this strategy with realignment to focus on areas of flexibility for classifications and local refinements with or without the subtraction steps. Large particle numbers and sufficient signal are still needed for this strategy; however, in the future, new improvements to equipment and data collection strategies should reduce these requirements. As the complexity of proteins studied through CryoEM increases, methods to extract as much information as possible from the data through processing can become a crucial factor in observing important details.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms25115665/s1.

Author Contributions: Conceptualization, S.G.; methodology, S.G. and D.S.S.; analysis and investigation, D.S.S. and S.G.; manuscript writing, S.G. and D.S.S.; supervision, S.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Institute of General Medical Sciences, grant number R35-GM142797. D.S.S. was supported with a graduate fellowship from Graduate Assistance in Areas of National Need (GAANN) (P200A210024) provided by the U.S. Department of Education.

Data Availability Statement: Atomic coordinates and CryoEM maps have been deposited in the Protein Data Bank (www.wwpdb.org) and EM Data Bank (www.ebi.ac.uk/emdb), respectively. PDB ID codes: 8VDZ, core scaffold; 8VE7, DARPin monomer. EMDB codes: EMD-43157, core scaffold; EMD-43158, full DARP14; EMD-43159, trimer; EMD-43167, DARPin monomer; EMD-43265 early C1 full DARP14.

Acknowledgments: The authors thank all the members of the Gonen Lab for support and critical discussions and Stephen McCarthy for help with software administration and model validation and analysis.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Danev, R.; Yanagisawa, H.; Kikkawa, M. Cryo-Electron Microscopy Methodology: Current Aspects and Future Directions. *Trends Biochem. Sci.* **2019**, *44*, 837–848. [CrossRef] [PubMed]
2. Cheng, Y. Single-Particle Cryo-EM—How Did It Get Here and Where Will It Go. *Science* **2018**, *361*, 876–880. [CrossRef] [PubMed]
3. Passmore, L.A.; Russo, C.J. Specimen Preparation for High-Resolution Cryo-EM. *Methods Enzymol.* **2016**, *579*, 51–86. [CrossRef] [PubMed]
4. Arnold, S.A.; Albiez, S.; Bieri, A.; Syntychaki, A.; Adaixo, R.; McLeod, R.A.; Goldie, K.N.; Stahlberg, H.; Braun, T. Blotting-Free and Lossless Cryo-Electron Microscopy Grid Preparation from Nanoliter-Sized Protein Samples and Single-Cell Extracts. *J. Struct. Biol.* **2017**, *197*, 220–226. [CrossRef] [PubMed]
5. Ruiz-Arroyo, V.M.; Raj, R.; Babu, K.; Onolbaatar, O.; Roberts, P.H.; Nam, Y. Structures and Mechanisms of tRNA Methylation by METTL1–WDR4. *Nature* **2023**, *613*, 383–390. [CrossRef] [PubMed]
6. Zhao, J.; Benlekbir, S.; Rubinstein, J.L. Electron Cryomicroscopy Observation of Rotational States in a Eukaryotic V-ATPase. *Nature* **2015**, *521*, 241–245. [CrossRef] [PubMed]
7. Zhou, Y.; Moscovich, A.; Bartesaghi, A. Data-Driven Determination of Number of Discrete Conformations in Single-Particle Cryo-EM. *Comput. Methods Programs Biomed.* **2022**, *221*, 106892. [CrossRef] [PubMed]
8. Wu, M.; Lander, G.C. How Low Can We Go? Structure Determination of Small Biological Complexes Using Single-Particle Cryo-EM. *Curr. Opin. Struct. Biol.* **2020**, *64*, 9–16. [CrossRef] [PubMed]
9. Lyumkis, D. Challenges and Opportunities in Cryo-EM Single-Particle Analysis. *J. Biol. Chem.* **2019**, *294*, 5181–5197. [CrossRef] [PubMed]
10. Punjani, A.; Zhang, H.; Fleet, D.J. Non-Uniform Refinement: Adaptive Regularization Improves Single-Particle Cryo-EM Reconstruction. *Nat. Methods* **2020**, *17*, 1214–1221. [CrossRef] [PubMed]
11. Punjani, A.; Fleet, D.J. 3DFlex: Determining Structure and Motion of Flexible Proteins from Cryo-EM. *Nat. Methods* **2023**, *20*, 860–870. [CrossRef] [PubMed]
12. Zheng, S.Q.; Palovcak, E.; Armache, J.-P.; Verba, K.A.; Cheng, Y.; Agard, D.A. MotionCor2: Anisotropic Correction of Beam-Induced Motion for Improved Cryo-Electron Microscopy. *Nat. Methods* **2017**, *14*, 331–332. [CrossRef] [PubMed]
13. Scheres, S.H. Beam-Induced Motion Correction for Sub-Megadalton Cryo-EM Particles. *eLife* **2014**, *3*, e03665. [CrossRef] [PubMed]
14. Rawson, S.; Iadanza, M.G.; Ranson, N.A.; Muench, S.P. Methods to Account for Movement and Flexibility in Cryo-EM Data Processing. *Methods* **2016**, *100*, 35–41. [CrossRef] [PubMed]
15. McCarthy, S.; Gonen, S. Improved Interface Packing and Design Opportunities Revealed by CryoEM Analysis of a Designed Protein Nanocage. *Heliyon* **2022**, *8*, e12280. [CrossRef] [PubMed]
16. Kastner, B.; Fischer, N.; Golas, M.M.; Sander, B.; Dube, P.; Boehringer, D.; Hartmuth, K.; Deckert, J.; Hauer, F.; Wolf, E.; et al. GraFix: Sample Preparation for Single-Particle Electron Cryomicroscopy. *Nat. Methods* **2008**, *5*, 53–55. [CrossRef] [PubMed]
17. Schmidt, C.; Urlaub, H. Combining Cryo-Electron Microscopy (cryo-EM) and Cross-Linking Mass Spectrometry (CX-MS) for Structural Elucidation of Large Protein Assemblies. *Curr. Opin. Struct. Biol.* **2017**, *46*, 157–168. [CrossRef] [PubMed]
18. Manglik, A.; Kobilka, B.K.; Steyaert, J. Nanobodies to Study G Protein–Coupled Receptor Structure and Function. *Annu. Rev. Pharmacol. Toxicol.* **2017**, *57*, 19–37. [CrossRef] [PubMed]
19. Billesbølle, C.B.; Azumaya, C.M.; Kretsch, R.C.; Powers, A.S.; Gonen, S.; Schneider, S.; Arvedson, T.; Dror, R.O.; Cheng, Y.; Manglik, A. Structure of Hepcidin-Bound Ferroportin Reveals Iron Homeostatic Mechanisms. *Nature* **2020**, *586*, 807–811. [CrossRef] [PubMed]
20. Ilca, S.L.; Kotecha, A.; Sun, X.; Poranen, M.M.; Stuart, D.I.; Huiskonen, J.T. Localized Reconstruction of Subunits from Electron Cryomicroscopy Images of Macromolecular Complexes. *Nat. Commun.* **2015**, *6*, 8843. [CrossRef]
21. Zhang, K.; Julius, D.; Cheng, Y. A Step-by-Step Protocol for Capturing Conformational Snapshots of Ligand Gated Ion Channels by Single-Particle Cryo-EM. *STAR Protoc.* **2022**, *3*, 101732. [CrossRef] [PubMed]
22. Nakane, T.; Scheres, S.H.W. Multi-Body Refinement of Cryo-EM Images in RELION. *Methods Mol. Biol.* **2021**, *2215*, 145–160. [CrossRef] [PubMed]
23. Morris, K.L.; Jones, J.R.; Halebian, M.; Wu, S.; Baker, M.; Armache, J.-P.; Avila Ibarra, A.; Sessions, R.B.; Cameron, A.D.; Cheng, Y.; et al. Cryo-EM of Multiple Cage Architectures Reveals a Universal Mode of Clathrin Self-Assembly. *Nat. Struct. Mol. Biol.* **2019**, *26*, 890–898. [CrossRef] [PubMed]

24. Coscia, F.; Estrozi, L.F.; Hans, F.; Malet, H.; Noirclerc-Savoye, M.; Schoehn, G.; Petosa, C. Fusion to a Homo-Oligomeric Scaffold Allows Cryo-EM Analysis of a Small Protein. *Sci. Rep.* **2016**, *6*, 30909. [CrossRef] [PubMed]
25. Liu, Y.; Gonen, S.; Gonen, T.; Yeates, T.O. Near-Atomic Cryo-EM Imaging of a Small Protein Displayed on a Designed Scaffolding System. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 3362–3367. [CrossRef] [PubMed]
26. Yeates, T.O.; Agdanowski, M.P.; Liu, Y. Development of Imaging Scaffolds for Cryo-Electron Microscopy. *Curr. Opin. Struct. Biol.* **2020**, *60*, 142–149. [CrossRef] [PubMed]
27. Mittl, P.R.; Ernst, P.; Plückthun, A. Chaperone-Assisted Structure Elucidation with DARPins. *Curr. Opin. Struct. Biol.* **2020**, *60*, 93–100. [CrossRef] [PubMed]
28. King, N.P.; Bale, J.B.; Sheffler, W.; McNamara, D.E.; Gonen, S.; Gonen, T.; Yeates, T.O.; Baker, D. Accurate Design of Co-Assembling Multi-Component Protein Nanomaterials. *Nature* **2014**, *510*, 103–108. [CrossRef] [PubMed]
29. Liu, Y.; Huynh, D.T.; Yeates, T.O. A 3.8 Å Resolution Cryo-EM Structure of a Small Protein Bound to an Imaging Scaffold. *Nat. Commun.* **2019**, *10*, 1864. [CrossRef] [PubMed]
30. Castells-Graells, R.; Meador, K.; Arbing, M.A.; Sawaya, M.R.; Gee, M.; Cascio, D.; Gleave, E.; Debreczeni, J.É.; Breed, J.; Leopold, K.; et al. Cryo-EM Structure Determination of Small Therapeutic Protein Targets at 3 Å-Resolution Using a Rigid Imaging Scaffold. *Proc. Natl. Acad. Sci. USA* **2023**, *120*, e2305494120. [CrossRef] [PubMed]
31. Jeong, W.H.; Lee, H.; Song, D.H.; Eom, J.H.; Kim, S.C.; Lee, H.S.; Lee, H.; Lee, J.O. Connecting two proteins using a fusion alpha helix stabilized by a chemical cross linker. *Nat. Commun.* **2016**, *7*, 11031. [CrossRef] [PubMed]
32. Zivanov, J.; Nakane, T.; Forsberg, B.O.; Kimanius, D.; Hagen, W.J.; Lindahl, E.; Scheres, S.H. New Tools for Automated High-Resolution Cryo-EM Structure Determination in RELION-3. *eLife* **2018**, *7*, e42166. [CrossRef] [PubMed]
33. Zhang, K. GCTF: Real-Time CTF Determination and Correction. *J. Struct. Biol.* **2016**, *193*, 1–12. [CrossRef] [PubMed]
34. Gautomatch: An Efficient and Convenient Gpu-Based Automatic Particle Selection Program. Available online: https://www2.mrc-lmb.cam.ac.uk/download/gautomatch-053/ (accessed on 8 December 2023).
35. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. [CrossRef] [PubMed]
36. Asarnow, D.; Palovcak, E.; Cheng, Y. UCSF pyem v0.5. *Zenodo* **2019**. [CrossRef]
37. Meng, E.C.; Goddard, T.D.; Pettersen, E.F.; Couch, G.S.; Pearson, Z.J.; Morris, J.H.; Ferrin, T.E. UCSF ChimeraX: Tools for Structure Building and Analysis. *Protein Sci.* **2023**, *32*, e4792. [CrossRef] [PubMed]
38. Liebschner, D.; Afonine, P.V.; Baker, M.L.; Bunkóczi, G.; Chen, V.B.; Croll, T.I.; Hintze, B.; Hung, L.W.; Jain, S.; McCoy, A.J.; et al. Macromolecular Structure Determination Using X-rays, Neutrons and Electrons: Recent Developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **2019**, *75*, 861–877. [CrossRef] [PubMed]
39. Emsley, P.; Lohkamp, B.; Scott, W.G.; Cowtan, K. Features and Development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 486–501. [CrossRef] [PubMed]
40. Chen, V.B.; Arendall, W.B., 3rd; Headd, J.J.; Keedy, D.A.; Immormino, R.M.; Kapral, G.J.; Murray, L.W.; Richardson, J.S.; Richardson, D.C. MolProbity: All-Atom Structure Validation for Macromolecular Crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 12–21. [CrossRef] [PubMed]
41. *The PyMOL Molecular Graphics System*, Version 1.2r3pre; Schrödinger, LLC.: New York, NY, USA, 2010.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.