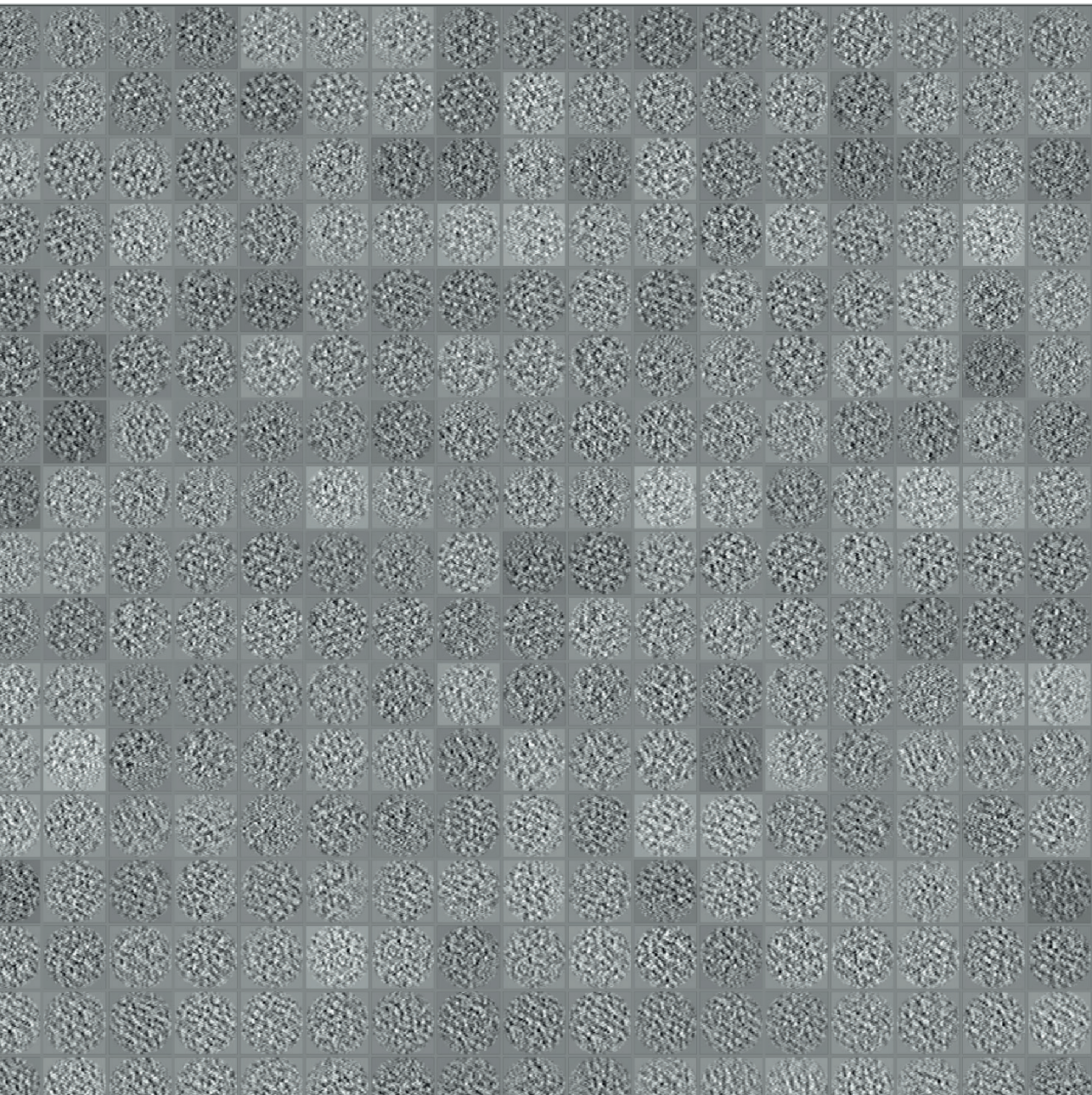
Cryo electron tomography studies of bacterial chemosensory arrays

Yang W.

CHAPTER 7



General Discussion

The research field of biology is an ever-growing collection of knowledge built on the observation of what living things look like and how they function. It has been 127 years since bacterial chemotaxis was first described and we are still searching for answers for two questions, A) what do chemosensory arrays look like and B) how does signal transduction and kinase control work (2, 64). Instead of using traditional methods to understand chemotaxis behavior, such as observing "rings on agar plates" and "bands in capillary tubes", we are now utilizing direct imaging methods to understand the system on a molecular level. Here we apply structural techniques with a high-resolution limit to describe the possible conformational changes in chemoreceptors and the CheA kinase during signal transduction and activation, respectively. The questions about this system have evolved over time, and so have the probes we use to answer them. The studies presented in this thesis, together with numerous other studies published in the past two decades, clearly demonstrate that the method of cryo-electron tomography has provided unprecedented insight into the structure and function of chemoreceptor arrays.

The characterization of biological objects and the understanding of biological events are frequently limited by our ability to see them directly. Cryo-ET is a powerful tool that allows for the visualization of objects in three dimensions in their near-native context (252, 253). Although the basic concept of the 3D reconstruction from 2D projections was established in 1968, the technique has gained exponential popularity in the last two decades after numerous theoretical and experimental limitations in its practical application have been overcome (254-256). Among multiple applications of cryo-ET in life sciences, studies focusing on resolving the molecular architecture of complex macromolecular assemblies have been particularly successful. The bacterial chemosensory array is one such case where cryo-ET studies led to groundbreaking discoveries. Ever since the molecular architecture was discovered in 2012, new insights were continuously gained to produce a better understanding of how chemosensory arrays function on a molecular level and how chemosensory arrays have evolved to generate structural diversity across species (60). All the studies presented in this thesis rely substantially on applying cryo-ET and subtomogram averaging to characterize the chemosensory arrays either in different functional states (**Chapter 5**) or in different bacteria other than the model organism *E. coli* (**Chapter 2-4**). The main scientific observations and conclusions regarding chemosensory arrays have been discussed in the individual chapters. Here, I will discuss the cryo-ET technique itself as well as its shortcomings and limitations that currently pose as obstacles. In turn, this will hopefully help to assess whether cryo-ET can be further exploited and improved for future bacterial chemotaxis studies.

One of the fundamental cryo-ET limitations that hinder visualization of the chemosensory arrays directly at the molecular, or even atomic, level lies in the sample thickness (**Chapter 6**). Even though chemosensory arrays are predicted to commonly exist in motile bacteria, the average cell diameters vary greatly between species. Different cell morphologies such as spherical, rod-shaped or spiral shaped cells differ drastically in diameter and typically range between 0.25 to 2 microns. Among them, only a few bacteria are naturally thin enough to allow cryo-ET imaging for structural analysis of the chemosensory arrays located at the cell pole (42, 50, 51, 55, 225). To overcome the size limitation, various approaches have been applied to make the bacterial cells thinner. For example, several detailed descriptions of the chemosensory arrays at the molecular level were derived from imaging the *E. coli* minicells (43, 44). This approach is still widely used, for example in the latest structural characterization of the chemosensory arrays in *E. coli* (177). An alternative approach to address the thickness issue is to introduce gentle lysis so that after releasing the cytoplasmic content the thickness of cells drastically decreases in the cryo-specimen (54, 58, 59).

Noticeably, reflected in the recent studies, there are a few significant differences between using the lysis treatment and using minicells (**Chapter 4 & 5**). It has been reported that the chemoreceptor lattice in *E. coli* is generally well preserved in lysed cells. However, it has been observed in many studies with the lysis method that instead of one extended receptor lattice per cell, smaller patches of receptor lattice are dominant. In addition, the chemosensory arrays imaged from a severely lysed cell usually shows a loss of packing order towards the membrane-proximal ends of the receptors with no detectable pattern at the periplasmic domains of the receptors. Conversely, in the subtomogram averaging results of arrays from intact minicells, it has been shown that the hexagonal pattern can be detected throughout the full length of the cytoplasmic domain of the receptors and even for the periplasmic domains (177). Together, this implies that although the native architecture of the baseplate is preserved in both ways, the compactness of the receptor lattice is still susceptible to the changes of the geometry of the inner membrane, particularly for the membrane-proximal ends of the arrays. Because of the much-reduced diameter, the inner membrane of a minicell usually exhibits a much more extreme curvature compared to the cell pole of native wild type cells. As a result, it is plausible that the membrane-proximal ends of the receptors splay out more readily and are less likely to have the same level of likelihood to interact with each other. In contrast, lysis treatment is aimed to flatten the cell so that the majority of the inner membrane appears to resemble more or less a 2D sheet. In other words, the curvature of the inner membrane is much reduced so that the possible interaction among the membrane-distal ends of the receptors is encouraged. This geometry difference of the inner membrane is likely to have an impact on the compactness on the membrane-bound chemoreceptor lattice in the averaging results.

Since the receptor lattice is tethered into the hexagonal pattern by the protein network at the baseplate, the compactness of the receptor lattice is likely to exhibit significant changes at its membrane-distal ends. Indeed, the ~ 26 nm distance of the membrane-spanning region of receptor could allow the chemosensory arrays to tolerate a range of packing compactness at the membrane-proximal ends of the receptor lattice. However, such a tolerance of the structural plasticity still depends on baseplate architecture that tethers the receptor lattice in place.

Besides changes to the sample itself, generating minicells or introducing lysis also results in differences in how to process the different data. For the most ideal lysis results, the cell envelope will be completely flattened and preserved as a sheet oriented perpendicular to the electron beam. Due to the limited image sampling range of cryo-ET, this would readily cause the majority of the subtomograms to exhibit a stronge preference of the orientation. In contrast, minicells largely preserve the curvature of the membrane and the chemosensory arrays attaching to it. Thus, imaging minicells offers a higher chance of resulting in a homogeneous spatial sampling of the chemosensory arrays compared to imaging the lysed cells. The sampling differences may strongly influence sub-tomogram averaging (STA) data processing and results. An averaged density map generated mainly, if not exclusively, from receptors lattice parallel to the electron beam direction will reflect the incomplete sampling perpendicular to the lattice plane. The most obvious defects in the EM maps from arrays with such preferred orientation is an artificial elongation of the receptors and severely underrepresented baseplate components. This is particularly problematic for resolving the baseplate components in the arrays. Despite this drawback, averages from flat arrays have proven to be most sensitive to capture structural differences in the packing order of the receptor lattice.

The sample thickness is a universal bottleneck for achieving a higher resolution of biological structures *in situ*. The most informative model building for the chemosensory arrays has therefore been based on the cryo-ET studies of a recombinant chemosensory array sample *in vitro* (45). For the vast majority of questions remaining for the chemosensory complex structure, an atomic resolution structural description is urgently needed. For example, we still need to understand how exactly the receptor binds to the kinase and the coupling protein, and how it controls CheA enzymatic activity. With the current imaging and image analysis technique, it is unlikely that minicells or the lysis treatment would be sufficient to offer samples that permit a resolution needed to answer these questions. Instead, single particle analysis would be a more suitable tool to address this issue if ternary complex comprising chemoreceptor trimers-of-dimers, kinase CheA and CheW could be engineered and proven suitable for cryo-EM imaging (205). An atomic-resolution description of the binding interfaces within such a complex, as well as revealing the spatial arrangement of different domains within the

kinase CheA, would undoubtedly be a major milestone towards understanding signal transduction in chemotaxis (257).

In the meantime, cryo-ET has its unparalleled advantages in exploring the diversity of the chemosensory arrays by imaging them *in situ* in a variety of bacteria. This undoubtedly broadens our understanding and appreciation of diversity of the chemosensory system (**Chapter 2**). Given that the chemoreceptor lattice is one of the most frequently observed macromolecular machines in the tomographic database ETDB (https://etdb.caltech.edu) of bacterial cells, there is a rich abundance of image data openly available for analyzing the appearance of the chemosensory arrays (258). The current consensus still stands that the hexagonal packing of the receptors in the chemotaxis arrays is universal in bacteria and archae (51, 55). However, a potential structural diversity may have been overlooked so far.

Functional studies have put a substantial amount of effort in illustrating a "vertical" signal transduction mechanism that is proposed occur through conformational changes passing through the receptor and propagating to the kinase and coupling protein. Meanwhile, the "horizontal" signal transduction, which has been less extensively characterized, is proposed to take place across the baseplate through long-distance allosteric dynamics in the coupling protein and kinase networks. This type of horizontal signal transduction is thought to support signal amplification and signal cooperation. This implies that the chemosensory arrays function not only as a nose, but also as a brain that evaluates different input signals to ultimately control swimming trajectory. In order to understand the nature of this horizontal signaling transduction, the kinase distribution and the molecular architecture of the baseplate are crucially relevant. Emerging research suggests that this distribution is likely to be species-dependent. This is in contrast to the packing order of the chemoreceptor lattice that continues to be a universal feature.

Aside from the number and arrangement of kinase proteins in the array, compositional diversity in the baseplate commonly exists in bacteria outside *E. coli*. In this model organism, which represents the most extensively investigated system among all bacteria, the ternary signaling core unit might just be the simplest case (**Chapter 3**). CheV, a hybrid protein containing a CheW-like domain, can substitute for CheW in the chemosensory arrays. Although not as widely distributed as CheW in the known genomes of bacteria and achaea, in some bacterial species it can even function as the only coupling protein in the chemotaxis system (70, 91, 108, 110, 259, 260). Studies in *V. cholerae* have shown that, in addition to CheV, ParP is another alternative coupling protein that is directly involved in chemosensory array formation and localization (112, 146).

Such discoveries of compositional diversity in the baseplate are inspiring for a further illustration of alternative molecular architecture of the chemosensory arrays other than arrays in *E. coli*. Yet, precisely distinguishing the different components in the baseplate can be a formidable task. To identify a structure of interest, particularly a specific protein, has always been a major challenge with cellular tomography. So far, there is not a universally applicable EM-visible tag that can readily serve as an indicator similar to a fluorescent tag for light microscopy. Immunolabelling using gold-conjugated antibodies could work for binding to proteins exposed on the cell surface or restricted to the section surfaces, but such labels could not reliably access the baseplate in the cytoplasm without extensive cell lysis. A few attempts have been made to design cloneable, protein-specific, electron-dense tags, among which is the ion-loaded ferritin-based label that was tested on chemosensory arrays in *E. coli* (261-263). Yet, instead of being fused to any protein component that directly integrates into the chemosensory arrays, the ferritin label was fused to the CheY protein that co-localizes with arrays through association with kinase CheA. Given the relatively large size of the ferritin label (~ 12 nm), it is hard to imagine that the chemosensory arrays can be efficiently or correctly assembled when either CheV or ParP proteins are fused with such a label. On the other hand, tags with a significantly smaller size might not perturb the local cellular environment but might be difficult to recognize in cellular tomograms. Such labeling would eventually suffer from low fidelity and efficiency, in particular if the fused protein (for example ParP) is predicted to be sparse by nature.

In the absence of a reliable tag, an alternative approach for identifying components in supermolecular machinery *in situ* with cellular tomography data is to compare STA maps derived from a set of mutant strains that each has a single component of the structure of interest knocked out. This method has proven to be powerful through numerous studies that revealed for example the molecular architecture of the bacterial secretion system (264, 265). However, this approach would not be equally effective when applied to identify CheV or ParP in the chemosensory arrays. Since CheV, ParP, CheW, and their homologs all share the SH3-domain-like topology, they are all predicted to be capable of serving as scaffolding proteins in the baseplate. In fact, it has been shown through fluorescence microscopy studies that CheV2, which does not normally integrate into arrays in *V. cholerae* under standard growth conditions, increasingly localizes with the F6 arrays in the absence of CheA (102). Thus, in arrays with a high degree of variability in the baseplate, mutations can be easily overcome by substitution. In contrast, overexpressing either CheW or deletion of CheA in *E. coli*, where there is little variability in the baseplate, severely impairs array assembly (266). In short, it is not feasible to identify some components in the baseplate by mapping the missing density in different STA maps.

Would it be possible to distinguish the baseplate components exclusively through i*n silico* analyses, namely based on the classification during the STA analysis (**Chapter 4**)? The cellular tomographic image provides unprecedented structural information that is unfortunately accompanied by an enormous amount of noise. This low signal to noise ratio impedes a reliable interpretation of the structure of interest in the tomographic data directly. Low signal to noise ratio also hinders the accuracy of defocus determination and the precision of subunit position and orientation in the STA alignment, which eventually limits the attainable resolution in final EM maps. The implementation of dose-symmetric tilt scheme and accurate 3D contrast transfer function correction will be an essential step to an overall improvement of attainable resolution in the STA reconstruction (267, 268). This will be true for various samples including chemosensory arrays. Additionally, using a more sophisticated classification algorithm will also improve the achievable resolution. For example, a new algorithm has recently been developed that has been tailored to distinguish the structural differences of 3D maps derived from STA analysis (269, 270). With those technical developments, identifying structures from cryo-ET data *in silico* is becoming more realistic (271-274). However, the successful applications of such analyses are still rather rare and limited to protein complexes with the size and abundance at a similar level of ribosomes. In contrast, distinguishing the proteins located in the baseplate of *V. cholerae*, such as CheV (36kDa) and ParP (42kDa), from CheW (18kDa) would be a formidable task due to their small and similar size. Not only do the sizes of the CheV and ParP present a challenge for analysis, but in addition both proteins are also extremely sparse in the arrays. It is difficult to predict what amount of tomographic data would be required to provide sufficient repeats so that the classification results hold fidelity. The sparseness of such proteins poses a difficulty for applying the correlative light electron microscope approach, given that it would be a challenge for reaching both the fluorescent detection level and the resolution required for the localization accuracy. Altogether, although cryo-ET and STA are progressing towards a molecule-based cellular landscape of macromolecular complexes with their native spatial coordinates, directly identifying chemotaxis proteins *in silico* for the baseplate is unlikely to be feasible in the foreseeable future.

Last but not least, although all biological subjects are inherently dynamic, the sample preparation of cryo-ET captures only a single time point of any given biological process. In other words, cryo-ET fundamentally acquires static 3D snapshots instead of recording biological events in time. In cases where the biological processes of interest occur slowly, the different specimens frozen at the same time point could readily be captured in different stages of such a biological process. By imaging a sufficient amount of different specimens, it is feasible to eventually assembly all the snapshots along the timeline. This would allow for the description of the entire biological process

of interest even though the observation is not carried out strictly on a single specimen. However, the heterogeneities among the specimens could also bring ambiguity into the interpretation of the cryo-ET observation. The bacterial flagellar motors can be used as an example for such a case. In stationary phase cells of gamma-proteobacteria, cryo-ET of the cells reveals incomplete motor structures called 'relics' that are embedded in the cell envelopes. When judged solely from the tomographic data, it is impossible to distinguish whether a relic is an intermediate structure before a flagellar motor is fully assembled or the remaining structure after the flagella has been ejected from the motor (20, 275).

This time-course related ambiguity also affects both structural and functional studies of the chemosensory arrays by cryo-ET. The cell lysis treatment, which was intended to facilitate the higher-resolution imaging by flattening the cells, resulted in a high degree of heterogeneity among specimens that were frozen and imaged at a random point during lysis of the cells (**Chapter 3 & 4**). Although most of the receptor arrays in *V. cholerae* were found completely disordered, there were indeed rare cases in which the hexagonal packing order was retained in lysed *V. cholerae* cells. One could easily suspect that the lattice order was preserved because this cell was frozen the moment it was lysed so there was not enough time for the baseplate to disassociate. Alternatively, it could also be the case that the array patch retaining order upon lysis happens to have a high CheA occupancy so that the baseplate structure is more resilient to the lysis process and may hold up to the structural integrity particularly well. Due to the lack of time-course information, we could not depict how the chemosensory array structure changes during the cell lysis process and characterize it in greater detail to see how resilient the different baseplate structures may be. Compared to a gradual event such as lysis, the switching between different signaling states of the chemosensory complexes within the arrays is thought to be inherently dynamic on a much faster time scale. As a result, such functional and conformational dynamics cannot simply be captured by freezing at different time points before and after exposing chemosensory arrays to the target compounds (57). Instead, mutations of either the receptor or the kinase are required to strongly bias the signaling states of the complexes (**Chapter 5**). Only combined with extensive biochemistry characterization of the kinase activity of such mutants, structural information derived from cryo-ET data can be conclusive for illustrating the conformational dynamics related to the signaling process.

With its unparalleled advantages for visualizing arrays *in situ*, cryo-ET will certainly remain the best tool to explore the structural diversity of chemosensory arrays (**Chapter 2**). As insights into this diversity accumulate, interest will increase for understanding the physiological relevance of different types of chemosensory arrays. Cryo-ET will continue to be an essential tool to carry out the most direct and conclusive assessment

on whether any type of chemosensory array is assembled at a given culturing condition, as well as in any chemotaxis protein mutant strain. Moreover, observing the appearance and the occurrence of the arrays can help to reveal the strategies bacteria have evolved for dealing with complex environmental inputs. Cryo-ET may also assist in understanding the possible benefits of the crosstalk among different arrays or the necessity of segregation among different chemosensory pathways. For bacteria, both phenomena are essential for maintaining an efficient and accurate perception of the immediate chemical environment and thus is essential for their survival. In the case of *V. cholerae*, the chemosensory system is capable of converting signal inputs from 32 different 40H chemoreceptors into the net flux of phosphor-CheY to bias its swimming direction (278). Still, much is yet to be learned about how this sophisticated sensory system sorts and simplifies information.

Although it is developing faster than ever, cryo-ET has not yet advanced to a point where we can unambiguously determine individual chemotaxis protein structure directly in the native cellular context. To interpret the structural description of the arrays made by cryo-ET, we often require other lines of evidence and analysis derived from other characterization methods. Given that cryo-ET is very costly, time consuming, and not readily accessible outside highly specialized labs, the samples need to first be finely engineered or characterized through molecular biology and biochemistry studies. In addition, computational methods such as proteomics analysis on chemoreceptors and other chemotaxis proteins can shed insights into what the composition of the arrays might be. Moreover, compared to any other method, genomics-based bioinformatics analysis can most efficiently compare various chemosensory systems on a broad spectrum and make a commonly applicable prediction of the functional specificities of chemosensory arrays. Lastly, genetics analysis has contributed greatly to predicting the functionality of chemotaxis proteins and how they might selectively interact with chemoreceptors. In combination with gentics analysis, cryo-ET will remain as an essential tool to solve the open questions in understanding bacterial chemotaxis.