Monovalent cation effects on histidine kinase CheA activity

by

Marie Sophie Balboa

B.S., University of North Carolina, 2012

B.A., University of North Carolina, 2012

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Masters of Science

Department of Chemistry and Biochemistry

2015

UMI Number: 1589936

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

UMI 1589936

Published by ProQuest LLC (2015). Copyright in the Dissertation held by the Author.

Microform Edition © ProQuest LLC.
All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

This thesis entitled:

Monovalent cation effects on histidine kinase CheA activity

written by Marie Sophie Balboa

has been approved for the Department of Chemistry and Biochemistry

---

Dr. Joseph Falke

Date

Dr. Robert Batey

Date

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

ABSTRACT

Balboa, Marie Sophie (M.S., Biochemistry)

Monovalent cation effects on histidine kinase CheA activity

Thesis directed by Professor Joseph J. Falke

CheA is a multi-domain histidine kinase believed to be a central component of all bacterial and archaeal chemosensory arrays. The best studied CheA proteins are those of *Escherichia coli*, *Salmonella typhimurium*, and *Thermotoga maritima*. The P4 catalytic domain of CheA is of particular interest due to its role in ATP binding and auto-phosphorylation of the substrate His residue on the P1 substrate domain, a process which involves Mg^{2+} in the P4 domain ATP-binding pocket. CheA auto-phosphorylation reaction has also long been known to be sensitive to the monovalent cation composition of the reaction buffer. Recent studies in GHKL superfamily members MutL (from *E coli*) and BCK (from *Rattus norvegicus*) demonstrated improved activity in the presence of monovalent K^+. Here experimental and computational analyses assess the effect of monovalent cations on the activity of *Salmonella typhimurium* CheA, and generate a model for a putative monovalent cation binding pocket. The findings show that CheA prefers the physiological K^+ ion over the other monovalent cations tested, indicating the site is specific and tuned to possess the appropriate K^+ affinity for the cytoplasmic environment.

ACKNOWLEDGEMENTS

Once upon a time there was a little bunny who needed to write a masters thesis.

“Carrots” said the mother bunny, “you will need lots and lots of carrots in order to write”

“beer” said the little bunny “works better”

“damn” said the mother bunny, “you are right”

And so they had a beer and wrote a thesis. And it was good.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Two-component signaling in bacteria

Bacteria use two-component signaling pathways to adaptively respond to environmental cues and to control a variety of cellular functions. This array of functions includes bacterial virulence, antibiotic resistance, cell division and taxis. The abundance and importance of the two-component pathways in prokaryotes, along with the conservation of signaling components, has made the pathways a target for studies such as the development of novel antibacterial drugs and the design of engineered biosensors [1]. The chemosensory array found widely in bacterial and archaeal genomes, including those of the enteric bacteria *Escherichia coli*, *Salmonella typhimurium*, and *Thermotoga maritima*, is one such pathway [2, 3].

The two-component chemosensory signaling pathway allows bacteria to travel along the gradient of a chemo-effector, an attractant inducing positive chemotaxis or a repellent leading to negative chemotaxis. The two components are the histidine kinase and the response regulator. The majority of prokaryotes studied have an intracellular signaling pathway consisting of a receptor domain/protein interacting with the histidine kinase domain/protein and an aspartate kinase for the response regulator of the pathway [4].

The receptor-histidine kinase complex can be classified based upon their sensor-kinase architecture. The three main categories of complexes include periplasmic-sensing kinases that detect via a linked extracellular domain, kinases with transmembrane regions that sense stimuli but lack an extracellular domain, and cytoplasmic-sensing kinases that detect either internal, diffusible stimuli or have a separate transmembrane sensor [5]. The chemotactic array found in *E. coli* and *S. typhimurium* belongs to the latter with a separate transmembrane receptor as depicted below [Figure 1.1].

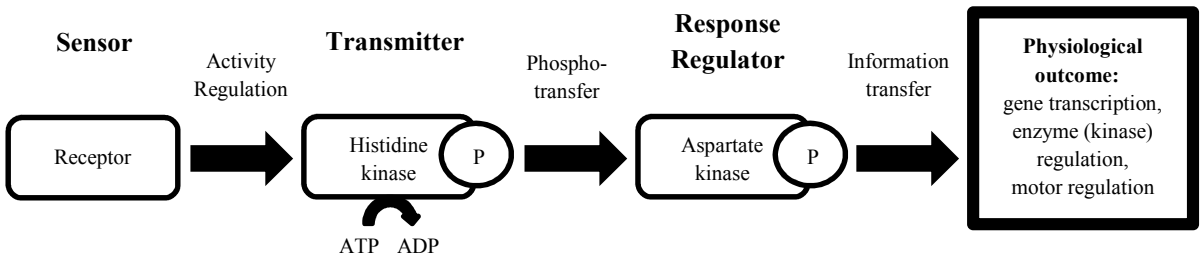
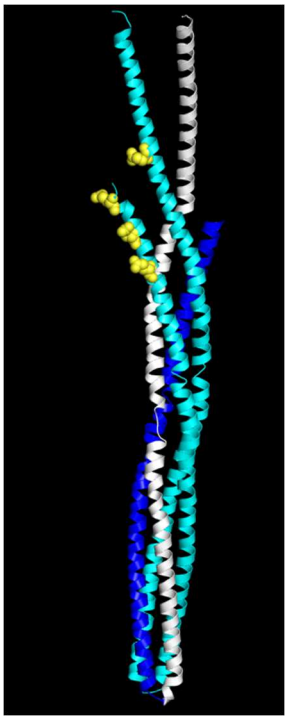


Figure 1.1 **Schematic of the Two-Component Signaling Pathway.** Components are shown in boxes while arrows indicate the information transfer. An additional phosphatase, not shown here and not found in all two-component pathways, deactivates the response regulator by removing the phosphate group [2]**.**

Bacteria sense a wide range of environmental conditions [pH, temperature] and molecular concentrations of external moieties such as amino acids, sugars, peptides, and environmental oxygen. The five transmembrane sensors that accomplish this in *E. coli* are Tar, Tsr, Tap, Trg and Aer, while *S. typhimurium* contains the additional receptor Tcp and a putative chemoreceptor proposed to bind to the cytoplasmic domain of other receptors [6, 7, 8]. The serine receptor Tsr and aspartate receptor Tar are the high abundance receptors, with an estimated tenfold higher cellular concentration than the low abundance receptors [9]. The study herein focuses upon the chemosensory array containing serine receptor, Tsr.

The serine receptor Tsr is classified as a methyl-accepting chemotaxis protein (MCP). Transmembrane serine receptor exists as a stable homo-dimer of helical bundles. The cytoplasmic domain of the serine receptor, and related MCPs, which is responsible for binding and regulating CheA kinase, consists of an N-terminal cytoplasmic extension coupling the second transmembrane helix (TM2) to the HAMP domain, followed by a long helical hairpin [10, 11].

The two identical helical hairpins of the dimer, each composed of two helices (CD1 and CD2) in which the methylation sites (*E. coli* residues 295, 302, 309 and 491) occur, associate to form an extended, anti-parallel 4-helix bundle as shown in Figure 1.2 [12]. The methylation state is controlled by the methyl-transferase protein CheR and methyl-esterase CheB, which methylate and demethylate respectively. Binding of the attractant serine leads to an increased methylation state of the adaptation sites on the



receptor helices [13]. Serine binding reduces the associated histidine kinase activity while methylation drives adaptation to the attractant by increasing kinase activity [14]. The net signal output depends on the balance between attractant binding and methylation of the receptor-kinase population. Inhibition of kinase activity by attractant reduces the production of the tumble signal phospho-CheY, thereby reducing tumble frequency and increasing the length of smooth swimming steps - see below.

Figure 1.2 **PyMOL representation of the transmembrane serine receptor.** Representation of the cytoplasmic domain of Tsr with adaptation sites 491, 295, 302 and 309, shown in gold from top to bottom respectively, on Chain A of the homodimer in light blue. Chain B is shown with CD1 in white and CD2 in dark blue. The accession number is 1QU7. The four-helix bundle shown here is perturbed by the truncations at the upper ends of the helices, causing the bundle to splay apart but disulfide scanning and modeling studies have revealed the native packing of the bundle. [15, 16, 11]**.**

The chemotaxis array and the information pathway

The pathway components form a highly ordered lattice that is ultra-stable and ultra-sensitive. In E. coli the chemotaxis array exhibits a Hill coefficient up to 20 while responding to chemo-attractant quantities over three orders of magnitude [10, 11]. The array consists of core ternary complexes consisting of methyl-accepting chemotaxis protein (MCP) receptors, histidine kinase CheA and adaptor protein CheW. Electron cryo-

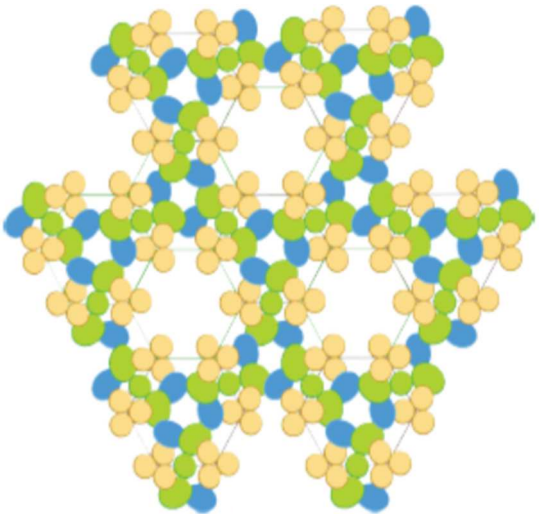


Figure 1.3 **Schematic model with the receptor trimer-of-dimers (tan), dimeric CheA kinase (green), and CheW adaptor (blue) shown as circular objects** [84]**.**

tomography revealed that the receptors, arranged in trimer of dimers, associate with CheA and CheW into hexagonal rings spaced ~12 nm apart and linked together into a higher order hexagonal lattice like that in Figure 1.3 [12].

CheA and CheW form a hetero-hexameric CheA-CheW ring that stabilizes a hexagonal lattice of trans-membrane receptors. The points of contact between the kinase and the adaptor in this ring are believed to aid in lattice formation and transmission of signal, thus contributing to the ultra-sensitivity of the lattice. The lattice components of the array in *Thermotoga maritima* have been crystallized successfully [13, 14, 15]. The sequence conservation was determined for CheA kinase in several bacterial genera, including *S. typhimurium* and *E. coli* [16]. This has been used to create a structural model of *E. coli* and *Salmonella* CheA, CheW and the hetero-hexameric ring by aligning the secondary structural components of the two sequences and threading the primary structure over the resulting tertiary structure using the Protein Homology/analogY Recognition Engine V 2.0 (Phyre2) [17].

The γ-phosphate covalently attached to the substrate histidine on the kinase CheA is quickly transferred in a Mg^{2+} dependent auto-phosphorylation reaction catalyzed by the response regulator CheY, an aspartate kinase, specifically to the substrate aspartate of CheY, D57 [17]. The response regulator then diffuses from the receptor-histidine kinase complex to the flagellar motor. The phosphorylated CheY interacts with the cytoplasmic or C-ring switch of the flagella, with a twenty-fold higher affinity over non-phosphorylated CheY for the C-ring protein FLiM [18, 19].

Phospho-CheY (CheY-P) binds to the N-terminus of the protein FLiM and stabilizes the clockwise rotation of the C-ring [20]. The CheY-P-FLiM binding changes the orientation of FLiG relative to the stationary component MotA from a left-handed to a right-handed helix and changes the movement of the MS-ring and the flagella from counterclockwise to clockwise [21].

The counterclockwise motion of the flagellar motor leads to filament bundling and smooth swimming runs of motion by the bacteria. The clockwise rotation disrupts those bundles leading to reorientation and

tumbling by the bacteria. An increase in serine concentration as a bacterium swims up the gradient lowers the kinase activity and production of the phospho-CheY tumble signal, reducing the tumble frequency and promoting continued smooth swimming up the gradient. Tumbles and runs still occur in a gradient meaning that bacterial movement is best characterized as a biased walk as depicted below in Figure 1.4.

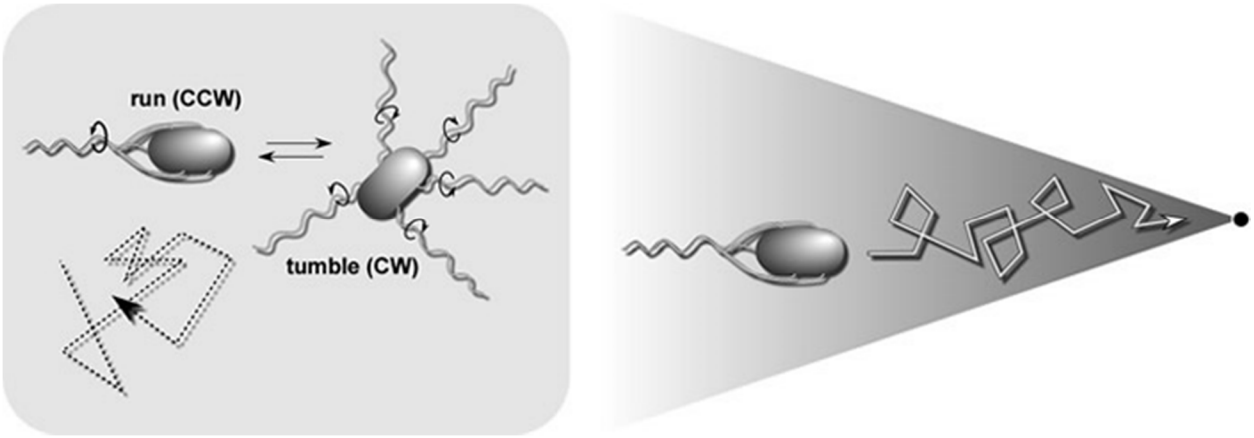


Figure 1.4 **Characteristic movement in the absence and presence of a chemo-effector gradient shown on the left and right respectively.** In the absence of a gradient the bacterial cell will execute a random walking motion of straight line runs alternating with random re-orienting tumbles. In an attractant gradient the runs up the gradient become longer than the runs down the gradient, yielding a biased walk up the gradient [2]. Image was taken from: http://chemotaxis.biology.utah.edu/Parkinson_Lab/projects/ecolichemotaxis/ecolichemotaxis.html.

The structure of histidine kinase CheA

Histidine kinases share five characteristic sequence motifs that define their transmitter domains- the H, N, G1, F, and G2 boxes [22]. The H box contains the substrate histidine domain, while the catalytic and ATP-binding (CA) domain contains the N, G1, F, and G2 boxes [23]. Histidine kinases can be classified into two groups by the organization of the H box respective to the ATP binding domain. In the first class of kinase the histidine phospho-transfer domain is directly adjacent to the CA domain while in second

class of kinases (including CheA) the histidine phospho-transfer domains is distal from the CA domain, separated by a dimerization domain as shown below [Figure 1.5] [24].

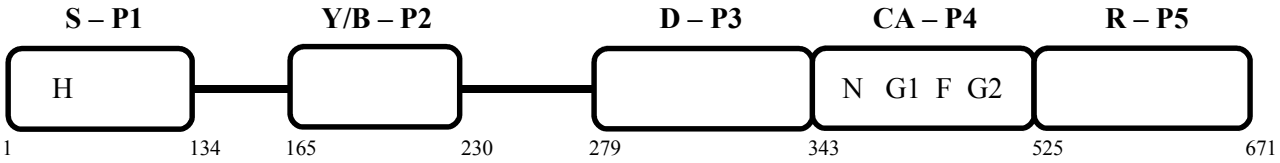


Figure 1.5 **Schematic of the domains of histidine kinase CheA.** Domains P1 through P5 are shown in boxes with the linker regions shown as lines. The abbreviations indicate respectively: S - substrate domain, Y/B – CheY/CheB binding domain, D – dimerization domain, CA - catalytic and ATP-binding domain, and R – regulatory domain. The numbers beneath the domains indicate the amino acid lengths of each while within the domains are the five conserved sequence motifs of the kinase family [25].

His kinase CheA has five domains which are functionally distinctive and autonomous topographically. The H box of CheA is located at the N-terminus of the protein and contains the substrate histidine His48 on a solvent exposed face of one of the five anti-parallel helices comprising the P1 domain [26]. CheA forms stable homo-dimers. The catalytic domain of one CheA transfers the phosphate group onto the substrate domain of its dimeric partner. Isolated P1 domain is inactive, but can be phosphorylated by a separate full length CheA and phosphor-transfer to CheY [27, 23].

The docking site for CheY is P2, which folds into a stable β-sandwich, regardless of the presence of P1 or the catalytic domain. At physiological and assay levels of CheY, phospho-transfer from CheA to CheY is at least tenfold faster than CheA auto-phosphorylation, leaving CheA auto-phosphorylation as the rate limiting step of the pathway [24, 25]. The dimerization of CheA occurs due to the interaction of P3 domains between proteins. The dimerization domains form a four α-helix bundle resulting from two hairpins associating [28].

The fifth domain is the regulatory domain P5, which is responsible for binding to the receptor and the adaptor protein CheW. CheW binding is essential for receptor clustering at cell poles and signaling [29].

The structurally homologous P5 and CheW proteins are each dual-SH3 domain-like folds and there is extensive tight interaction between the proximal barrel of P5, the first sub-domain, and the second sub-domain of CheW [30]. The introduction of mutations to the CheW eliminated or diminished binding to CheA, consequently the receptor could not form activated ternary complexes *in vitro* and cells had aberrant chemotaxis *in vivo* [31].

Multiple studies indicated that the reversible association of *S typhimurium* CheA from inactive monomers into an active dimeric state [32]. Truncated CheA variant lacking the substrate domain restored the kinase activity of active site deficient CheA mutants. *In vivo c*hemotaxis and *in vitro* phosphorylation of CheA and CheY were all restored by the over-expression of the shortened CheA indicating a trans-phosphorylation mechanism within the dimer pairs between the ATP binding site and the substrate histidine in the P1 domain [33, 34]. The atomic phosphorylation site on the histidine was determined in a study with *S typhimurium* using P-31 nuclear magnetic resonance which discerned the chemical shifts of an N-3 phosphohistidinyl residue [35].

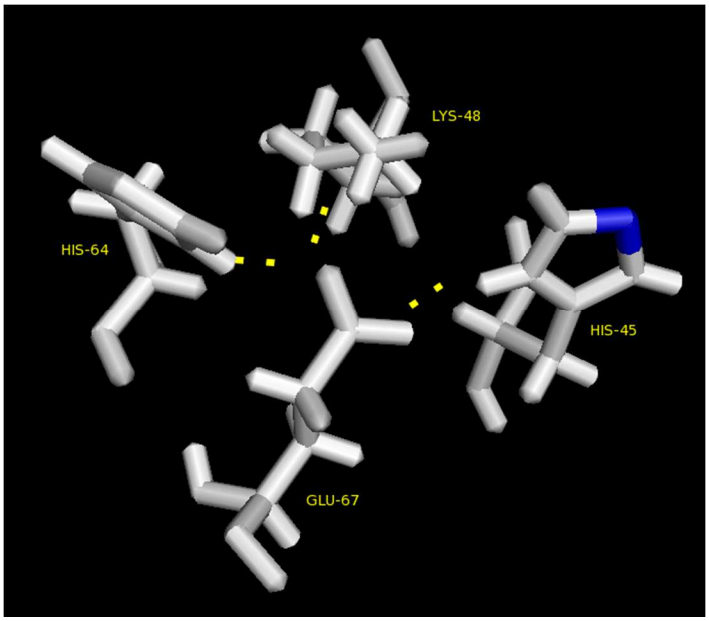


Figure 1.6 **PyMOL representation of histidine activation for phosphotransfer.** The phosphor-accepting histidine, His 48 in *E coli* or here His 45 in *Thermotoga*, is pictured here with the nucleophilic nitrogen that binds the γ-phosphate depicted in blue. This histidine is activated by the general base properties of the nearby glutamate shown hydrogen bonding with dashed yellow lines to the histidine. These amino acids are referred to as the catalytic dyad of the active site. There are further stabilizing hydrogen bonds shown as yellow dashes. The structures accession number is 1TQG [36].

The structural specifics were studied most thoroughly in *Thermotoga maritima*. High resolution crystallography showed the loss of a hydrogen bonding network between His-45, Glu-67, Lys-48 and His-64 upon mutagenesis on these sites that only affected the side chain conformations but did not extend to the overall domain structure [36]. The active site hydrogen bonds in the wild type structure of the protein are depicted in Figure 1.6. The conserved glutamate site deprotonates the histidine and stabilizes an $N^{\delta 1}H$ tautomeric state [26]. The $N^{\varepsilon 2}$ atom is then activated for a nucleophilic substitution with the γ-phosphate.

CheA ATP-binding domain and the ATPase/kinase superfamily

His kinase CheA is a member of the GHKL (gyrase, Hsp90, histidine kinase, MutL) phosphor-transferase superfamily. The recent examination of structural motifs of a variety of proteins led to the identification of this phospho-transferase superfamily due to the unique nucleotide binding fold. The nucleotide binding motif of histidine kinases, the catalytic domain P4 in CheA, conforms to their classification within the GHKL superfamily [37].

The nucleotide binding fold, referred to henceforth as the Bergerat fold, consists of four conserved structural motifs [38]. The four sequence motifs are respectively: Motif I (uubEuuaNouDA), Motif II (uxuxDNGxGuxbaauxxuu), Motif III (uGxxGxouxSxxxuoxbuTuxT), and Motif IV (Tx_nGT) in which u are aromatic residues, o are small side chain residues, a and b are acidic and basic residues. The first three of these motifs are known as the N, G1 and G2 boxes in histidine kinases.

The core structural element of the Bergerat fold is α/β sandwich with three alpha helices and four beta sheets. Two of the alpha helices, α1 and α3, are amphipathic and packed against the hydrophobic face of the beta strands forming the main pocket. The second alpha helix is transverse to the others and forms the lid of the pocket allowing a conical conformation for ATP to enter and interact before phospho-transfer occurs. A general schematic representation is shown below in Figure 1.7.

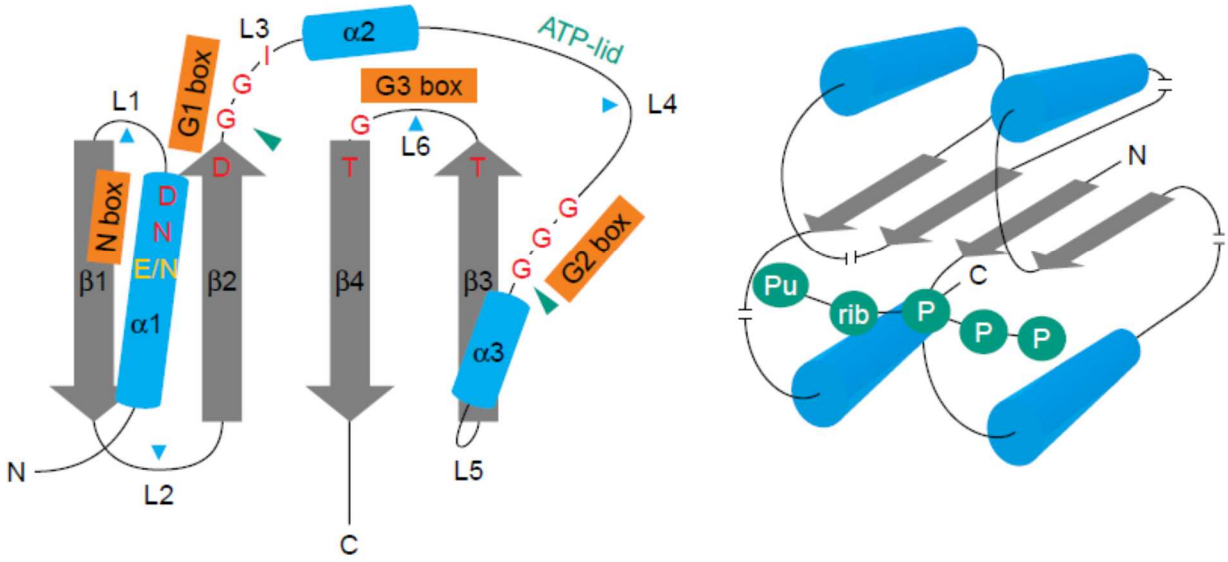


Figure 1.7 **Schematic of the Bergerat fold as determined by structural studies of EnvZ, CheA, DNA gyrase B, Hsp90 and MutL** [39, 37]. Alpha helices are represented as blue cylinders and beta sheets as gray arrows, the loops are the labeled black curves. Important residues are labeled by their single letter abbreviations at their relative sites in the structure. The four conserved motifs (1, II, III, and IV) are shown as orange boxes N, G1, G2 and G3 respectively. Major structural variations between the families within the super group occur at the locations indicated by the blue triangles.

Sequence conservation amongst the GHKL superfamily members occurs mainly in surface loops linking the structural components together. It is these surface loops that create the points of contact with the bound ATP. One notable exception is the conserved asparagine in the first motif (uubEuua**N**ouDA), also known as the N box in histidine kinases. The asparagine coordinates the divalent cation which in turn coordinates the phosphates of the ATP in an octahedral geometry via hydrogen bonds [39].

The most common divalent cation confirmed by crystallography is magnesium, this also being the most relevant divalent cation *in vivo* [4, 23]. The specificity of the pocket for ATP over GTP is thought to occur by hydrogen bonding between the N6 of the adenine and a conserved aspartate in the second motif

or G1 motif. The conserved glycine and threonine of motif IV (TxnGT) indirectly interact through the solvent with the N1 and N6 of adenine.

The primary sub-classification of the GHKL superfamily is characterized by the composition of the ATP-binding pocket lid. These four classifications can be represented by gyraseB, MutL, Hsp90 and EnvZ. The F-box, conserved in histidine kinases such as CheA, is the novel characteristic of the lid of the EnvZ family. The motif is indicated in Figure 1.7 by the blue triangle meant to indicate major structural variations between the super-group families.

The common conformations of the lid of the binding pocket are the lid closed with the nucleotide completely enclosed, partially open and the nucleotide partially exposed, and finally the lid open and ATP completely exposed to nearby solvents. Crystal structures of all these conformations have been resolved and confirmed for all four classifying families indicating that the mobility of the L4 loop is shared between the classes. The interacting residues of the lid differ between proteins and may contribute to the binding orientations and specificities of the nucleotides within the pocket [40]. While a specific function has not been determined it was previously proposed that the mobile loop of EnvZ was able to interact with the phosphorylation site on a separate domain of the histidine kinase and participates in the mechanism either directly or via stabilization of the catalyst [41].

Monovalent cation activation of hydrolysis in the ATP binding pocket

The subsequent hydrolysis after ATP binding is mediated by the conserved glutamate that acts as a base for water activation and nucleophilic attack on the phosphate [42]. BCK hydrolysis was recently determined to be monovalent cation dependent as well as glutamate and divalent cation dependent [43]. BCK, a mitochondrial serine protein kinase in the superfamily, was found to require the presence of monovalent cation and magnesium for ATP binding and hydrolysis [44]. Activity was maintained in the presence of potassium or rubidium, but not sodium, suggesting size dependence in the nucleotide binding

pocket. The preferential selection by BCK for potassium is hypothesized to have developed due to its high abundance [45, 46, 47].

A study of superfamily members MutL and BCK saw that the ATP binding pocket could replace two water molecules with a rubidium ion or one water molecule with a potassium ion, apparently to stabilize the conserved lid of the pocket. Modeling comparisons of BCK and MutL revealed a common monovalent cation-binding site by the lid of the pocket at the glycine motif. The preferred potassium ion was coordinated by the alpha phosphate and various carbonyl oxygens in the backbone [43].

Conformations normally energetically inaccessible to polypeptides in aqueous solutions are stabilized by the presence of monovalent cations via carbonyl oxygen coordination. For this reason monovalent cation binding sites tended to occur near surface exposed loops lacking secondary structure [48, 49]. A similar site was determined in a crystal structure of gyrase B refined by R.J. Lewis and D.B. Wigley, suggesting this monovalent cation coordination may be another unifying feature of the superfamily [50].

The classification of monovalent cation activated enzymes was laid out by Di Cera [51, 52]. This classification uses two criteria- the selectivity of the cation and the mechanism of activation. The selectivity breaks down into two basic groups of classification: K^+ or Na^+. Enzymes dependent on potassium are generally also activated by ammonium and rubidium and not activated as well or at all by cesium or sodium. Enzymes that are sodium selective are poorly activated by any other ion.

Some cases, such as dialkylglycine dehydrogenase and Hsc70, accomplished cation specificity with pockets selective for ionic radius and subsequent coordination geometry that differed in their hydration shells [53]. Monovalent cations of a different size than the selective pocket require reorganization of the active site residues leading to entropic costs and an energy barrier to binding.

The differentiation of mechanism is dependent upon the metals role as a direct co-factor or an allosteric binding partner. The first group, Type I activated enzymes, require both monovalent and divalent cations to be present. The mechanism involves simultaneous coordination from both cations to the ATP before

optimal docking into protein binding pocket tends to occur. The second group or Type II is characterized by the lack of direct contact between the monovalent and the ATP. In this case the cation perturbs the conformation of the active site organizing it into the architecture best suited for ATP binding.

Traditionally, the reaction buffers for CheA auto-phosphorylation include potassium as the dominant monovalent cation [54, 17]. Preliminary results from our lab (Piasta, Turvey and Falke, unpublished) have shown a preference of our ternary complex's histidine kinase activity for buffers containing potassium over sodium. We will continue this study with other monovalent cations to test for size dependence as well as titrating potassium while maintaining constant ionic strength buffer to elucidate the potassium concentration dependence. Subsequently, modeling studies will investigate possible potassium binding site and its ionic interactions with the CheA ATP binding pocket and the G2 motif.

RESULTS AND DISCUSSION

GHKL superfamily sequence alignments and activation classification

Sequence alignments of multiple proteins allow for the prediction of phylogenetic trees and of critical residues. One of the most common methods of alignment to date is the use of a progressive method of pairwise alignments while creating an estimated tree. These alignments were performed with FASTA files created from their respective UniProt entries [see Table 1 for a full list of PDB and UniProtKB entries]. The compiled sequences were run through MUSCLE via the command line on a 64-bit Windows system running Windows 7. The aligned sequences were viewed through SeaView 4.5.2 which was chosen for its availability, ease of use and previous use [55, 56].

K^+-activated Type I	**PDB**	**UniProtKB**
Branched chain α-ketoacid dehydrogenase kinase	1GKZ	P21310
Diol dehydratase	1IWB	Q59471
Glycerol dehydratase	1IWP	O08505
GroEL	1KP8	P0A6F5
MvRadA	1XU4	O73948
Pyruvate kinase	1A3W	P00549
S-Adenosylmethionine synthase	1MXA	P0A817
K^+-activated Type II		
Dialkylglycine decarboxylase	1D7R	P16932
MutL	1NHI	P23367
Ribokinase	1GQT	P0A9J6
Ribokinase	1TYY	Q8ZKR2
Tryptophanase	1AX4	P28796
Tyrosinase	1TPL	P31013

Na^+-activated Type I		
β-Galactosidase	1DP0	P00722
β-Galactosidase	1TG7	Q700S9
Fructose-1,6-bisphosphate aldolase	1B57	P0AB71
Fructose-1,6-bisphosphate aldolase	1RV8	Q9RHA2
Tagatose-1,6-bisphosphate aldolase	1GVF	P0AB74
Na^+-activated Type II		
Trp synthase	1A50	P00929
Trp synthase	1V8Z	Q8U093
Outgroup		
Serine/threonine-protein kinase HipA	3TPD	P23874
Unknown		
Histidine kinase CheA	N/A	P09384

Table 1. **PDB Accession Numbers and UniProtKB Accession Entry for Members of the GHKL superfamily.** CheA does not have a listed PDB entry for reasons discussed in the text.

Within SeaView, a parsimony tree was created by randomizing the sequence order five times, and bootstrapping a hundred times with a 70% minimum level for consensus tree building. This 70% confidence limit was chosen as previous study had determined that ≥70% bootstrap proportion corresponded to a ≥95% of actual clade determination in a family with equal change rates [59]. The parsimony tree used the sequence of the serine/threonine protein kinase HipA from *Escherichia coli* as the outgroup root [Fig 2.1].

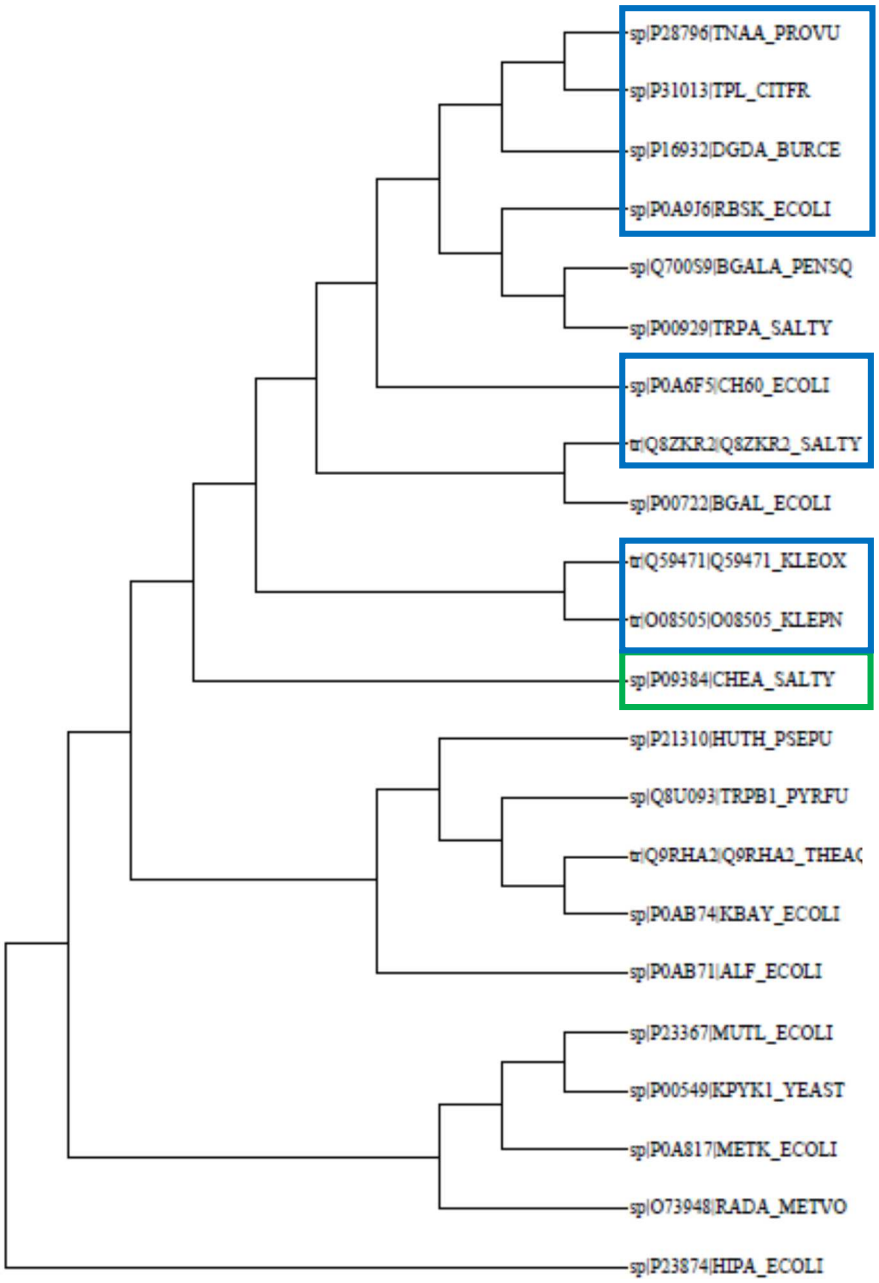


Figure 2.1 **Sequence parsimony tree from SeaView bootstrap analysis of GHKL superfamily members.** Protein identities in the tree are displayed by their UniProt entries which are listed in Table 1. CheA is indicated in green while potassium activated proteins of recent divergence are indicated with dark blue.

The protein activations were categorized by mechanism and cation preference in a review discussing M^+ activation [52]. For enzymes listed with multiple PDB entries listed unicellular organisms were chosen

with preference given to the higher resolution structures. This selection criterion limited the number of sodium Type II activated enzymes due to the predominance of studies upon the medically crucial protein thrombin. In each classification efforts were made to include at least one pair of homologous proteins from divergent organismal species or divergent proteins from species within the same genus.

Eleven other proteins subsequently diverged from the same node as CheA, eight of which were potassium activated. Specifically, the proteins with the closest divergence node to CheA in the phylogeny tree were diol dehydratase and glycerol dehydratase, both potassium Type I activated proteins. The sequence parsimony would suggest that monovalent dependent activity of CheA would likely require potassium.

Studies of MutL were performed in *E. coli* wherein a cleavage pathway is known in which thrombin cleaves the 70kDa protein into a 40kDa amino terminus fragment (henceforth referred to as LN40) and a 30 kDa carboxyl terminus fragment. The ATP-binding pocket occurs in the first domain of LN40 and is representative of the MutL family [57]. Further study resulted in crystallographic confirmation of monovalent cation binding with potassium (accession code 1NHI).

This structure found that the cation was coordinating in the loop following the N-terminal α-helix. The short alpha helix α' is unique to the ATP-lid of the MutL family and partially exposes the nucleotide inside the pocket. There is a conserved arginine at the end of the ATP-lid as part of the third motif that interacts with the alpha and gamma phosphate groups of ATP. The coordinating residues (L70, A71, A73 and A76) were interacting with the cation via their carbonyl oxygen.

The coordinating residues of BCK also chelated the cation via the carbonyl oxygen. Alignments showed the BCK coordinating residues were super-imposable onto the MutL residues. In both BCK and MutL the cation stabilized the ATP lid and in BCK the potassium directly coordinated the α-phosphate of ATPγS. The same potassium carbonyl coordination was seen in GyraseB and conjectured to occur in Hsp90 and CheA.

Structural alignments were performed with the CheA with the MutL structure 1NHI. The structures overlay the main elements of the Mg^{2+}/ATP/K^+ binding pocket reasonably well as shown in Figure 2.2. The putative location of the monovalent cation is occupied by an electron density in the crystallized *Thermotoga* structure that has been assigned as the oxygen of water molecule ten in the PDB. The crystallization conditions of 0.8 M ammonium acetate and 0.085 M sodium acetate in the reservoirs mean that this density may in fact be an ammonium cation.

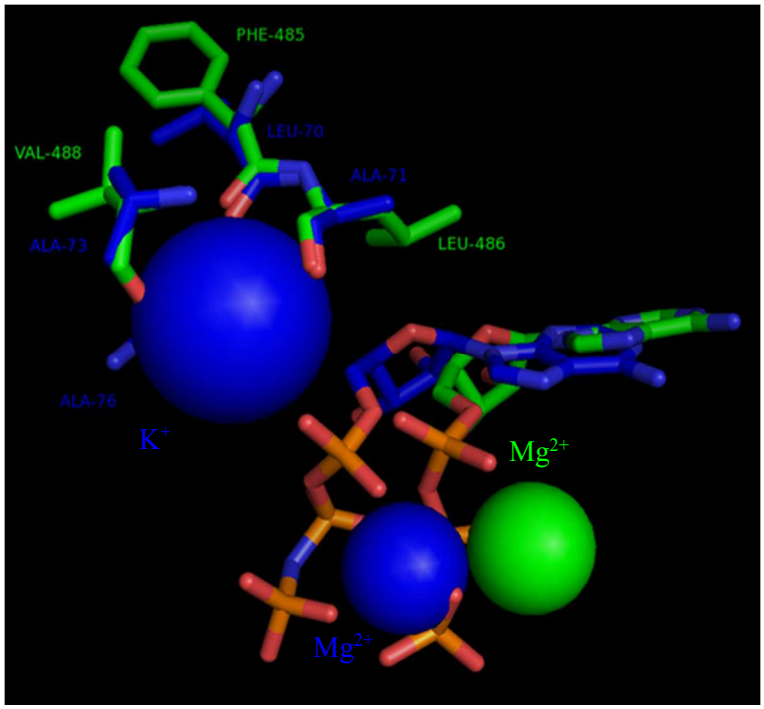


Figure 2.2 **Alignment of MutL (1NHI structure shown in dark blue) active site and *Thermotoga* CheA (1I59 structure shown in light green).** The protein alignment was produced via alignment of the oxygen atoms of the labeled residues which coordinated the potassium through their carbonyl oxygen. The magnesium ions are also shown colored by their respective original crystal structure as are the ATP analogues used in the crystallography conditions.

The Mg^{2+}/ATP/K^+ complex aligns into the turn of the G2 and F motifs. The lid is partially open and the ligands are partially exposed to solvent and thus available for the reaction with the P1 domain. With this initial alignment the carbonyl oxygen sites of two residues, L470 and I471, in CheA are roughly three and two angstroms from the center of the potassium ion respectively. The exact coordination of a putative potassium site in CheA would require further investigation and crystallographic study to ascertain however.

Experimental CheA monovalent cation preference

A correlation between monovalent cation and CheA specific activity was first observed by a collaborator with the kinase activity buffer (for full description see Materials and Methods) containing sodium or potassium cation (Piasta, Turvey and Falke, unpublished). Those preliminary tests observed a ten-fold specific activity increase when the reaction buffer contained K^+ rather than Na^+, in each case at 150 mM. The increased specific activity was the result of an increase in the kinase activity without significant effect on the incorporation of CheA into the chemotactic array as can be seen in Figure 2.3.

The chemotactic array is a hexagonal lattice of receptor oligomers and hexameric rings of the histidine kinase and adaptor protein. Each ring within the lattice consists of three core units, each containing two receptor oligomers, a CheA dimer and two CheW proteins. The receptor oligomers are trimers of dimers; meaning in optimally formed lattices the CheA incorporation can reach a ratio between 1:9 and 1:6, or roughly 11 to 16% uptake. These optimal saturation ratios have been confirmed by another lab experimentally [58]. The uptake ratio of 1:6 rarely occurs because of receptor zippers that are unavailable to bind CheA.

The monovalent cation conditions consisted of five monovalent cations of the alkali metal group: lithium, sodium, potassium, rubidium, and cesium. The majority of uptake ratios shown here were within error of 150 mM potassium (used as a standard condition for comparison herein) regardless of the monovalent cation identity. The samples with 150 mM rubidium were marginally outside error for the standard condition but within error of other conditions and did not seem to indicate a significant trend. The trials for 300 mM potassium showed a higher average uptake than others but were still within error of the standard condition. Furthermore, this increased uptake may have been the result of solubility properties rather than actual incorporation differences, a phenomenon discussed later in more depth.

The samples with 300 mM potassium were slightly cloudy by visual examination after the addition of 2X Laemmli buffer. The buffer contains 4% sodium dodecyl sulfate (SDS) by weight/volume, approximately

139 mM SDS. While the Krafft point of SDS is 10°C, the Krafft point of potassium dodecyl sulfate (KDS) is 45°C and is partially insoluble at the experimental condition of 22°C [59]. It is likely that the uptake quantity was the result of variability caused by this precipitation.

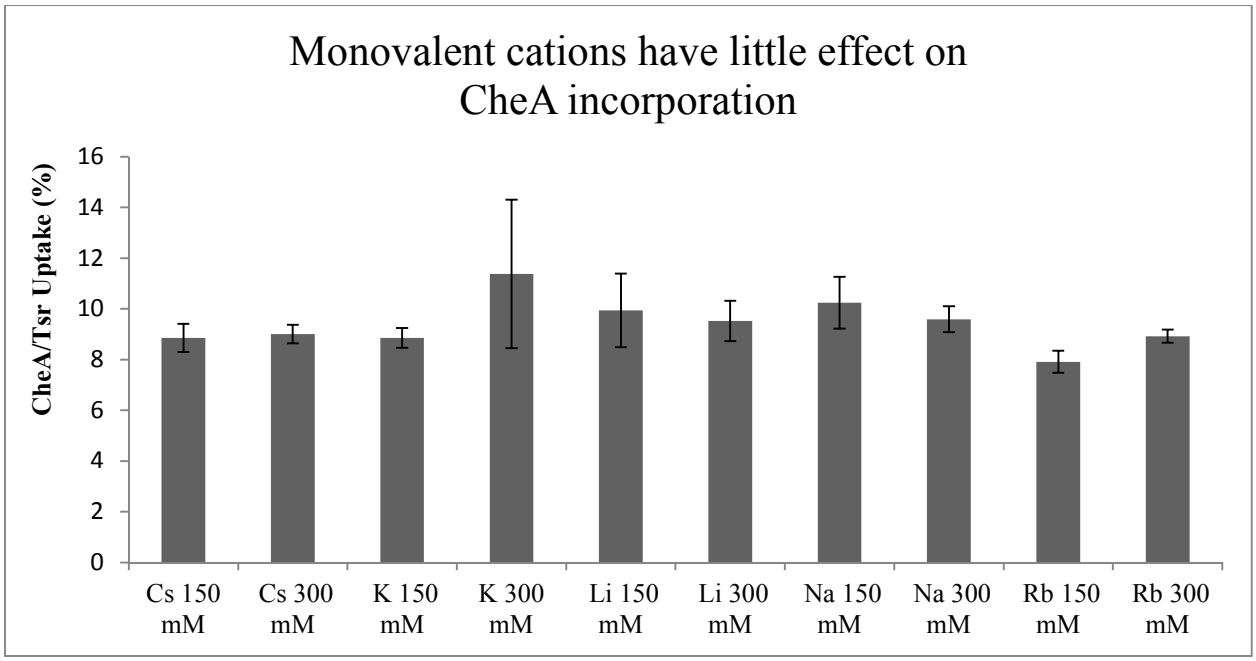


Figure 2.3 **Monovalent cation effects on CheA incorporation into the chemotactic array.** Results shown are from triplicate data sets (*n=3*) with standard error of the mean as the error bars.

Tests with the stable members of the alkali metal group indicated that CheA specific activity is sensitive to the type and concentration of monovalent cation. The highest specific activity was found with 150 mM potassium. The ten-fold decrease in activity with the same concentration (150 mM) sodium observed in the preliminary work indicated above was confirmed with an activity ratio of 0.09 ± 0.05 normalized to 150 mM potassium.

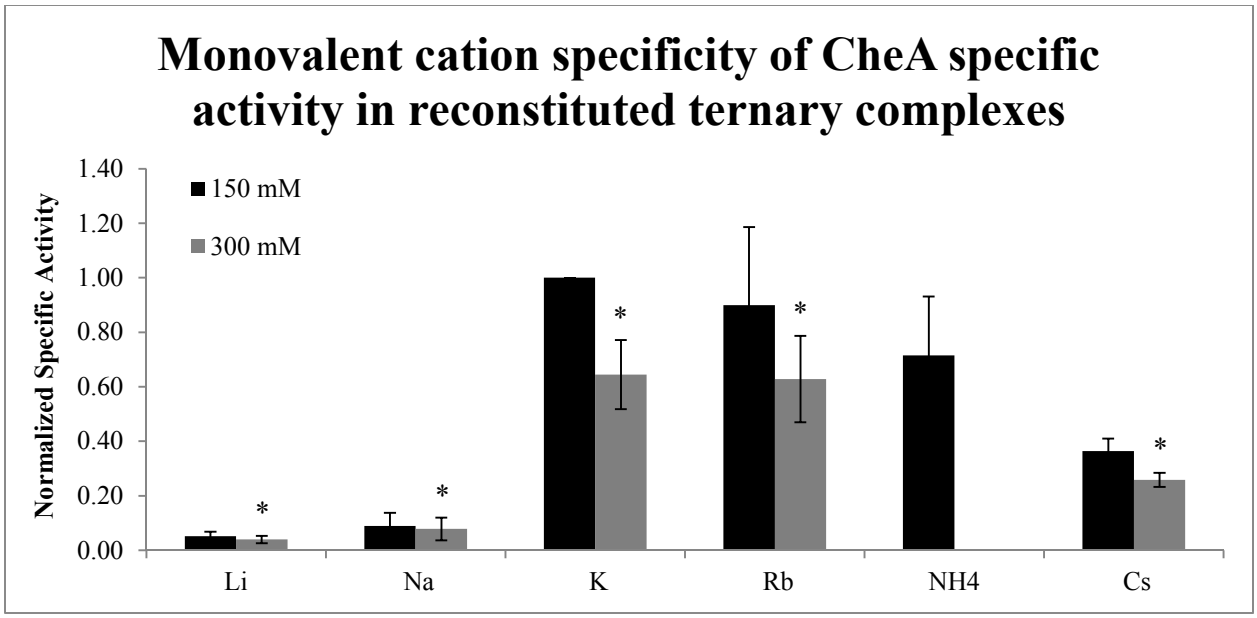


Figure 2.4 **Monovalent cation effects on CheA specific activity.** Results shown are from triplicate data sets (*n=3*) with standard error of the mean as the error bars. Asterisks designate the 300 mM trials in the graph. No 300 mM trials were performed with ammonium chloride.

M^+ activated enzymes have been classified with remarkable specificity by their preferred monovalent cation (Na^+ or K^+) and their relative activities with other salts predicted fairly accurately. Potassium activated enzymes are generally also activated by ammonium and rubidium but not activated or poorly activated by lithium, sodium and cesium. CheA specific activity demonstrated its categorization as a potassium activated enzyme as had been predicted by its parsimony to other GHKL superfamily members activated by potassium. The specific activity of CheA with rubidium and ammonium holds with the trend predicted for potassium activated enzymes, namely a lower but appreciable activation seen with these cations.

The trend seen in Figure 2.4 and the values reported in Table 2 also confirmed the relative activations with other monovalent cations that had been predicted for a potassium activated enzyme. Lithium demonstrated essentially no histidine kinase activity at either concentration while reduced but discernible

activity was seen with both rubidium and cesium relative to potassium. Rubidium is within error of potassium level of specific activity while cesium is significantly lower.

	Normalized Specific Activity		Normalized Specific Activity
Lithium 150 mM	0.05 ± 0.02	Lithium 300 mM	0.04 ± 0.01
Sodium 150 mM	0.09 ± 0.04	Sodium 300 mM	0.08 ± 0.04
Potassium 150 mM	1*	Potassium 300 mM	0.64 ± 0.13
Rubidium 150 mM	0.90 ± 0.29	Rubidium 300 mM	0.63 ± 0.16
Ammonium 150 mM	0.72 ± 0.22		
Cesium 150 mM	0.36 ± 0.05	Cesium 300 mM	0.26 ± 0.03

Table 2. **Normalized specific activity of M^+ at 150 mM and 300 mM.** The error values are the standard error of the mean of triplicate data sets (*n=3*). No error value is shown for 150 mM potassium since this was the condition used to normalize the data sets.

Lithium and rubidium trials with both 150 mM and 300 mM concentrations showed no significant difference between the two concentrations. Each was within error of the other as determined by the standard error of the mean. Potassium, sodium and cesium saw a significant decrease in activity with an increased salt concentration as seen by the t-test results in Table 3. This trend resembles that seen in a previous study on the monovalent cation dependence of MutL hydrolysis in which the wild-type protein had the same or lower activity with monovalent salt concentrations 200 mM and higher [43].

	Unpaired t – test between 150 mM and 300 mM conditions						
	t	df	P-value	Mean Difference	Standard Error of Difference	95% Confidence Interval	
						Lower	Upper
Lithium	1.057	16	0.306	0.013	0.012	-0.013	0.038
Sodium	3.202	16	0.006*	0.070	0.022	0.024	0.117
Potassium	3.805	16	0.002*	0.409	0.107	0.181	0.637
Rubidium	1.568	16	0.136	0.279	0.178	-0.098	0.656
Cesium	2.471	16	0.045*	0.111	0.045	0.016	0.206

Table 3. **Unpaired t-test results of M^+ at 150 mM and 300 mM.** Triplicate data sets containing triplicate samples (*n=9*) were entered for each salt condition after normalization to 150 mM potassium as previously described. The two-tailed P-values were calculated and those deemed significant with a 95% confidence interval are marked with an asterisk. Values were calculated with the GraphPad QuickCalcs server.

The K_M of ATP in the phosphorylation reaction of CheA is 260 ± 50 μM (Unpublished data from Buehler and Falke). The cellular concentration of ATP is approximately 1 mM, the concentration used in the experimental conditions of the kinase assay [60]. The nearly 4-fold ratio of ATP concentration to K_M of the reaction ensures that ATP binding is close to saturation, so that ATP occupancy of CheA is likely to be relatively insensitive to ionic strength. Thus the lower kinase activity observed at high ionic strength may result from a shift of the on-off switching equilibrium of the complex towards the off-state, or from a perturbation of structure or dynamics of the complex.

The kinase reaction of the reconstituted complex was then tested more carefully to determine its dependence on potassium concentration at constant ionic strength (150 mM). The potassium concentrations were: 0, 12.5, 25, 50, 75, 100 or 150 mM. Sodium was used as a potassium replacement to maintain ionic strength, capitalizing on its significantly lower kinase activation. The results were

normalized to 150 mM potassium within each data set and the standard error of the mean at each concentration determined [See Figure 2.5].

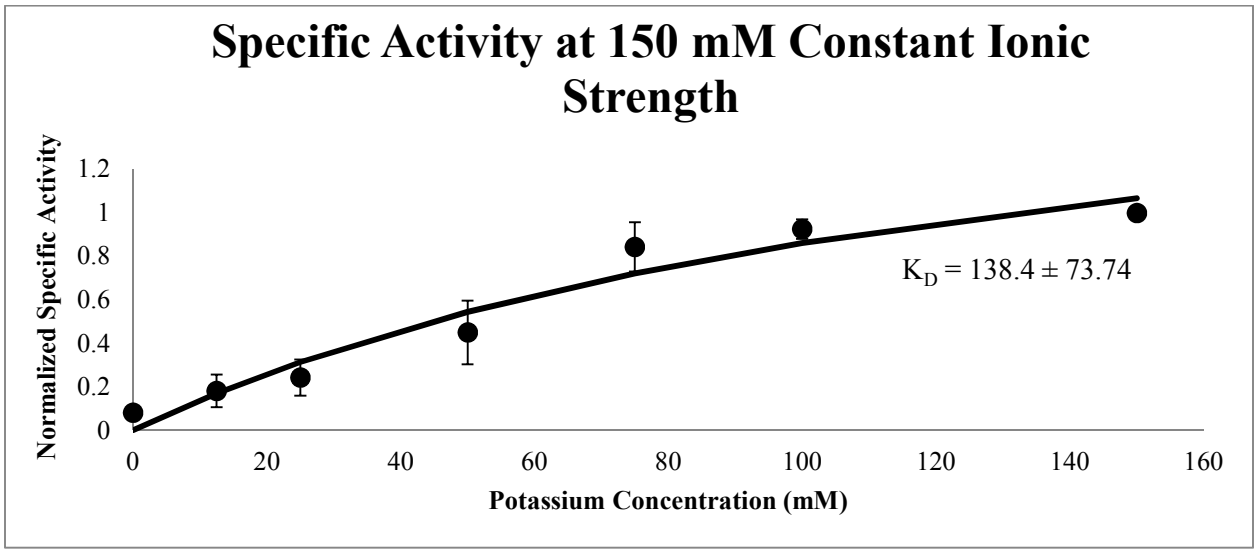


Figure 2.5 **Monovalent cation concentrations versus CheA specific activity independent of ionic strength.** Results shown are from triplicate data sets (*n=3*) with standard error of the mean as the error bars. The equilibrium binding constant is reported as determined within GraphPad with analysis of one site specific binding with an un-weighted fit.

The values for the normalized specific activities over this range of concentrations were analyzed with GraphPad (Version 6.04). The results of analysis are in Table 3 and are well fit by a single, independent site model with no cooperativity (Hill coefficient ~ 1.0) indicating there is a saturated binding site for the cation. The standard error for the best fit values and the correlation coefficient of the line fit is reported.

Best Fit Values	
B_{max}	2.052 ± 0.6460
K_d	138.4 ± 73.74

Table 4. **Saturation binding constants with single site binding without cooperativity.** The correlation coefficient of the fit was 0.9500 and the standard errors for each value were reported (*n = 7*).

Further work is needed to confirm whether the monovalent cation coordinates the ATP in a Type I interaction as a directly interacting cofactor or functions as a Type II allosteric effector that perturbs the active site residue conformations. One test that should help resolve the models is the measurement of the effect of K^+ on the ATP K_M. Type I predicts that 150 mM K^+ should stabilize ATP binding by directly coordinating phosphate, and thus should yield a lower Km than observed in 150 mM Na^+. If K^+ and Na^+ yield the same ATP K_M, this would support the Type II model in which the monovalent cation does not directly contact ATP.

Molecular Simulation and Structural Analysis of CheA

In the absence of direct experimental data, simulations of the structural molecular dynamics were run to explore the role of the monovalent cation in the active site. Previous sequence alignments of CheA from over sixty genera have shown remarkable conservation at the motif locations [61]. Alignment of CheA from Salmonella typhimurium, Thermotoga maritima, and Escherichia coli was performed on the European Molecular Biology Laboratory – European Bioinformatics Institute server. The alignment is shown in Appendix A.

The sequences were run with MUSCLE (Version 3.8.31) and analyzed for their conservation score assigned within the BLOSUM62 matrix [62, 63, 64]. Salmonella CheA run within MUSCLE showed appreciable conservation, roughly 84% of residues. The BLOSUM62 matrix was chosen for its range of applicable evolutionary distances and its results within previous rounds of assessment [65, 66].

Performance assessments of multiple alignment matrices using known structures were run within MULTALIN. Homology search assessments were run with LSHRRAT, RTARAT, and UL33$HCMVA queries within BLAST, FASTA and Smith-Waterman. BLOSUM62 consistently performed best of the matrices not only amongst the substitution matrices but also as the optimal sequence identity level.

Figure 2.6 **The sequence alignment of CheA species created with MUSCLE and colored by the BLOSUM62 score matrix.** The alignment of *Escherichia coli*, *Salmonella typhimurium* and *Thermotoga maritima* CheA sequences was created with MUSCLE Version 3.8.31 run through the EMBL-EBI server. The coloring is assigned by the average BLOSUM62 score of the letters within the column- light blue is a

score ≥ 3, light gray is a score ≥ 0.2 and scores lower than that are not colored. The results can be accessed at muscle-E20150127-024910-0643-83901533-oy.

The amino acid sequence of Salmonella typhimurium CheA was threaded onto the secondary structural elements of crystallized Thermotoga maritima with the Phyre2 server. The modifications are detailed in the Materials and Methods section. The resulting structure has been colored by the conserved motifs and important structural elements of the GHKL super-family [See Figure 2.7].

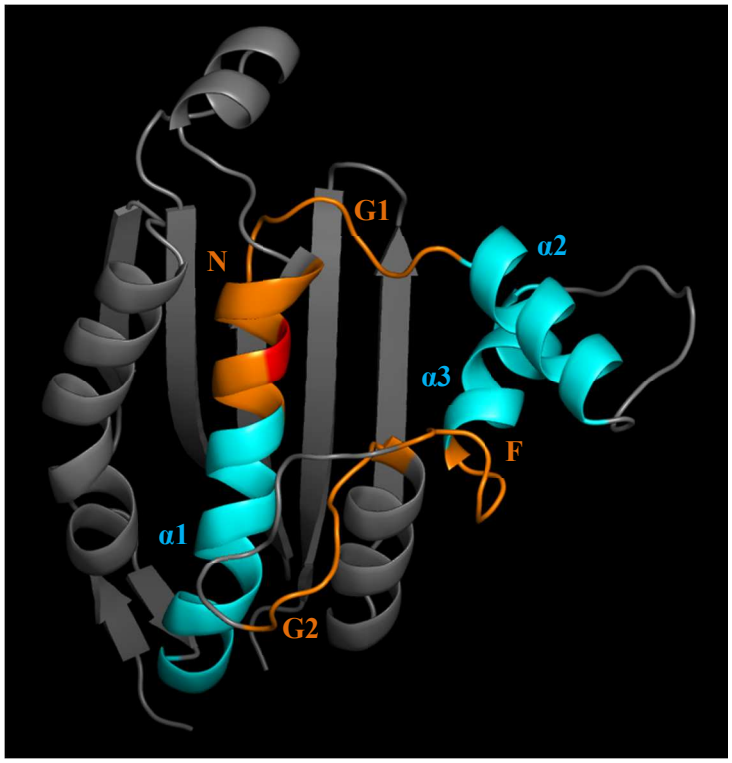


Figure 2.7 **PyMOL representation of CheA depicted like the Bergerat fold schematic within Figure 1.7.** Alpha helices 1-3 (labeled next to motifs N, G1 and G2) are represented in blue and the conserved coordinating asparagine, N397, is colored red. The four conserved motifs (1, II, III, and IV) are shown as orange boxes N, G1, G2 and F respectively.

The Mg^{2+}/ATP/K^{+} complex was added manually to the protein, resulting in location of the monovalent cation near the G2 and F motifs. This location was chosen by homology to monovalent cation coordinating sites found in previous structural studies [43, 52]. The threaded *Salmonella* structure produced with the Phyre server had good homology to the *Thermotoga* crystal structure. The Mg^{2+} and ATP stitched into the *Salmonella* structure overlaid with those of the crystal structure well as depicted below.

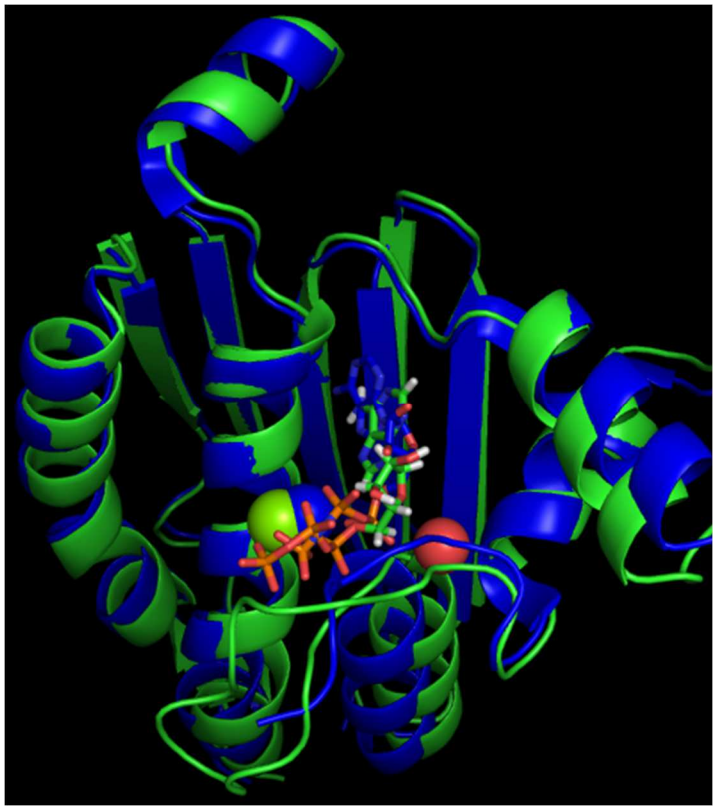


Figure 2.8 **Alignment of *Thermotoga* CheA (1I59 structure shown in dark blue) active site and *Salmonella* CheA (structure shown in light green).** The protein structure for *Salmonella* CheA was created with the Phyre server and the Mg^{2+}/ATP/K^{+} complex was added manually. The magnesium ions are also shown colored by their respective original crystal structure as are the ATP analogues used in the crystallography conditions. The putative location of the monovalent previously assigned water molecule ten is shown as a red sphere.

Dr. Margaret Hurley, at the Army Research Laboratory, set up multiple dynamics simulations in which the site of water molecule ten was used as an initial location for the potassium cation with degrees of similarity to the coordinates of the Mg^{2+}/ATP/K^{+} of the MutL crystal structure (PDB 1NHI). Her initial dynamics runs were unsuccessful. While the Mg^{2+}/ATP complex was quite stable, the K^{+} cation detached quickly and became solvated; a representative image is shown below [Figure 2.9].

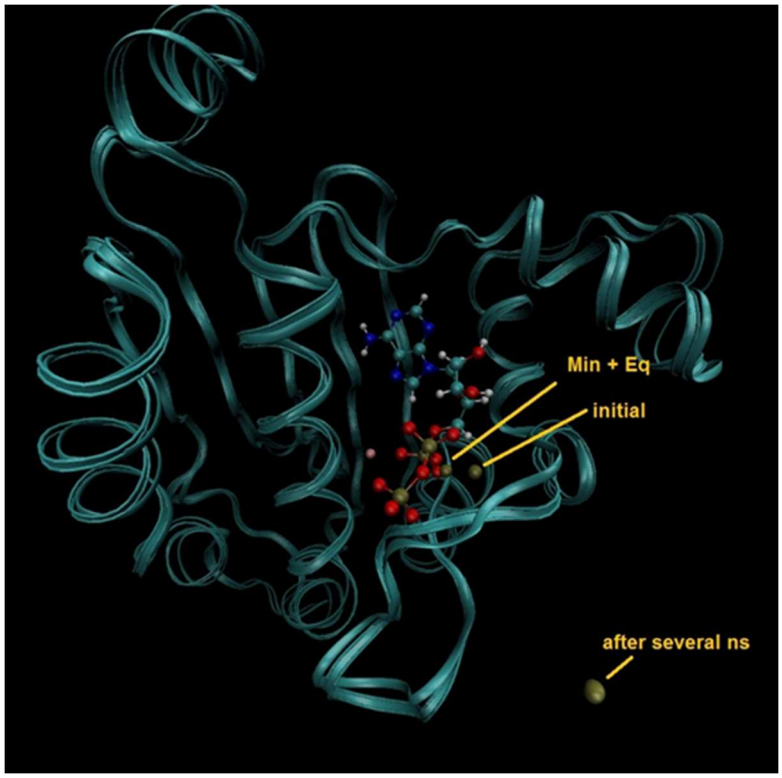


Figure 2.9 **Visual Molecular Dynamics (VMD) representation of CheA with potassium cation before, during, and after dynamics simulation.** The protein structure is a ribbon model while the ATP is a ball and stick representation. Magnesium is shown here as a pink sphere. The locations of ATP/Mg^{2+} did not change significantly during the dynamics simulation. The potassium ion locations are indicated by a golden sphere with the respective time within the dynamics labeling the figure. The locations are shown for the original structure, after minimization and some equilibration, and finally after further dynamics. This figure was generously provided during personal communication with Dr. Hurley.

Accordingly, new simulations were initiated mimicking the Mg^{2+}/ANP complex structure of the 1I59 *Thermotoga* crystal structure. The runs attempted placing the potassium ion closely within various portions of the lid, location B, and by minimizing into the complex (location C). These sites, along with site A – the location of water ten as assigned in 1I59, are shown [Figure 2.10].

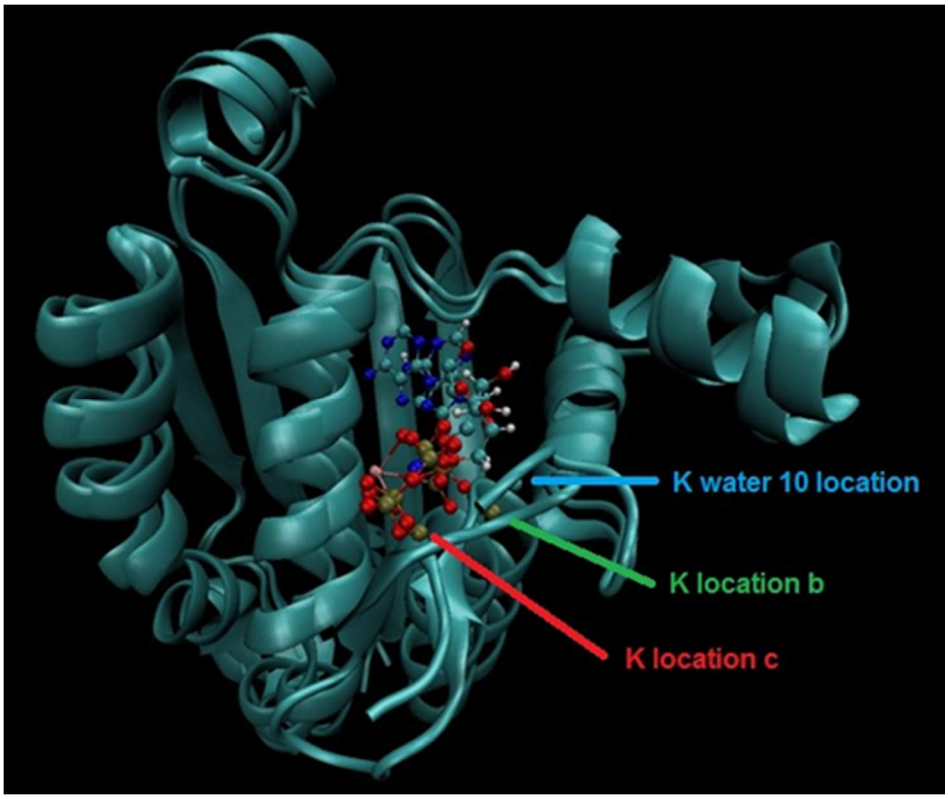


Figure 2.10 **Visual Molecular Dynamics (VMD) representation of CheA with potassium locations indicated.** Structures with the monovalent cations initially located well away from the α-phosphate of ATP minimized to location C, or detached from the protein complex entirely and moved into solution during dynamics. Further exploration of this phenomena remains to be done. This figure was also generously provided during personal communication with Dr. Hurley.

It should be noted that upon performing simulations with the former two cation locations (location A- the site of water 10 and location B), both resulted in a stable Mg^{2+}/ATP complex but a detached, solvated monovalent cation. Simulations performed with the cation placed in location C showed excellent stability for the potassium ion in particular, with varying degrees of stability upon substitution for other cations (Rb^{+}, Na^{+}, Li^{+}, Cs^{+}). It is quite possible that location C is not a global energy minimum, and further structural study is required.

In collaboration with Dr. Hurley a solvated structural system was built iteratively with Visual Molecular Dynamics (VMD Version 1.9.1) by padding the protein with 15 Å of water molecules on all sides [61],

resulting in a 78.3 x 79.25 x 68.75 angstrom simulation cell. Magnesium ions were added to neutralize the total system charge after the initial addition of the respective monovalent cation. Using this model as a template, a series of further models were generated by modification of the monovalent cation to Li^+, Na^+, Rb^+, and Cs^+. A further model was generated by complete deletion of the monovalent cation near the ATP complex [Figure 2.11].

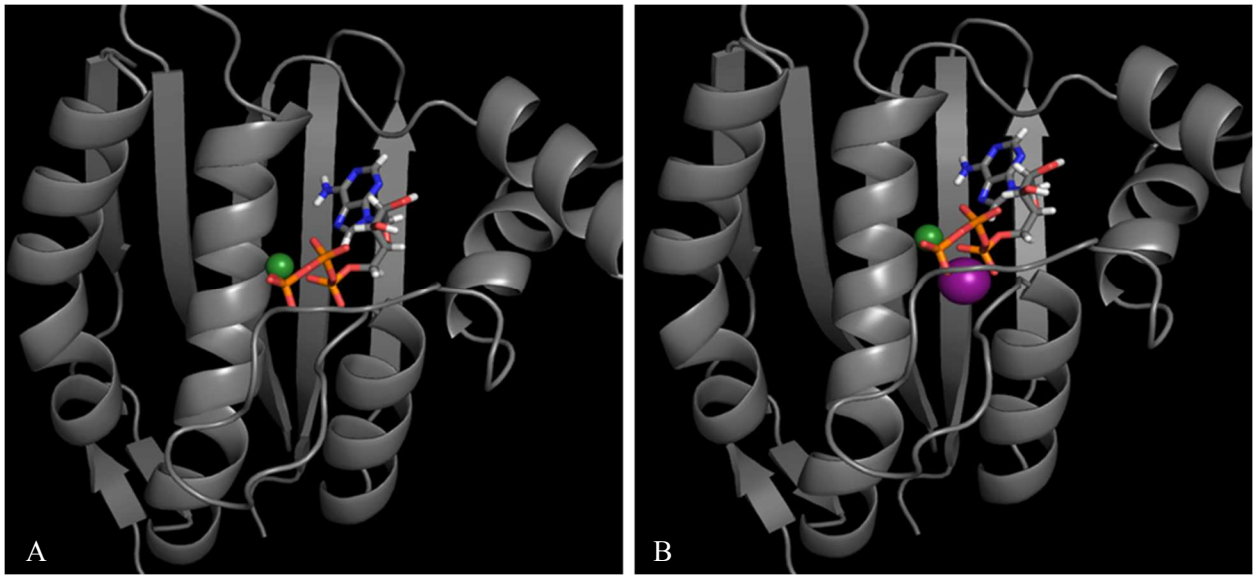


Figure 2.11 **PyMOL representation of the initial CheA structures depicting the monovalent/divalent cation-ATP complex.** Figure A on the left is the initial apo structure constructed for the dynamics simulation while Figure B on the right contains the Mg^{2+}/ATP/K^+ complex that was used for all initial monovalent cation dynamics runs subsequently. It should be noted again that the complex was created using the relative coordinates from the MutL crystal file (accession code 1NHI). The apo structure lacks the monovalent cation but retains the Mg^{2+}/ATP complex. The Mg^{2+}/ATP/K^+ complex is colored violet for the alkali metal and dark green for the alkali earth metal while ATP uses CPK conventions.

Molecular dynamics simulations were performed by Dr. Margaret Hurley. The system was energy minimized and equilibrated for 10ns at constant particle number, pressure and temperature (NPT), before accumulating data for an additional 30 ns. Simulations were performed with the Charmm36 force field with CMAP correction and TIP3P water within the Nanoscale Molecular Dynamics program (NAMD

Version 2. 9) [62, 63, 64, 65, 66] Force-field parameters for monovalent cations were taken from Roux et al [67]. Force-field parameters for ATP were taken from the Charmm General Force field v2b9 (CGenFF) [68, 69, 70].

Within the course of these simulations, the monovalent cation (MVC) drifted away from the Mg^{2+}/ATP complex for the Na^{+} and Rb^{+} model systems and joined the neutralizing ions within the solvent. Models containing Li^{+}, K^{+} and Cs^{+} all maintained a stable Mg^{2+}/ATP/MVC complex during the course of this 30ns trajectory. The K^{+} model was simulated for an additional 80ns, maintaining a stable complex. A meaningful discussion and quantification of the exact role of MVC within the ATP complex and protein pocket would require full quantum mechanical analysis which is beyond the scope of the current work. However, an analysis of the protein movement with and without the presence of the monovalent cation does provide some interesting insight.

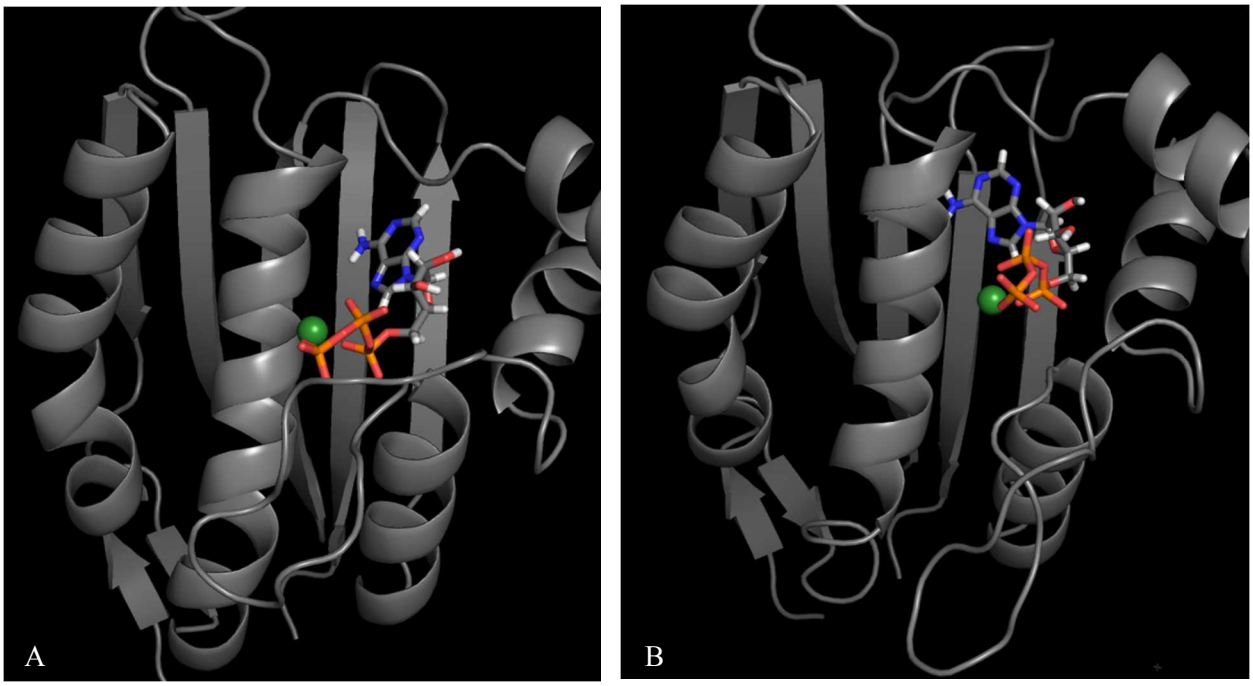


Figure 2.12 **PyMOL representations of the initial and final CheA structures depicting the divalent cation-ATP complex.** Figure A on the left is the initial apo structure constructed for the dynamics

simulation while Figure B on the right is a representative structure from the various frames deposited during the trajectory of the dynamics run as described further below. The Mg^{2+}/ATP complex is colored dark green for the alkali earth metal while ATP uses CPK conventions.

Sixteen frames were taken from the trajectory file and the (RMSD) calculated for all the α-carbons of the protein. One such frame configuration, represented with PyMOL in Figure 2.12B, was taken during the apo dynamics simulation. The RMSD values of the structures were calculated using the algorithm described fully in the Methods and Materials. A graphic representation of the RMSD values over the span of the P4 domain with/out monovalent cation can be seen in Figure 2.15 comparing the structural movements between apo and potassium specifically.

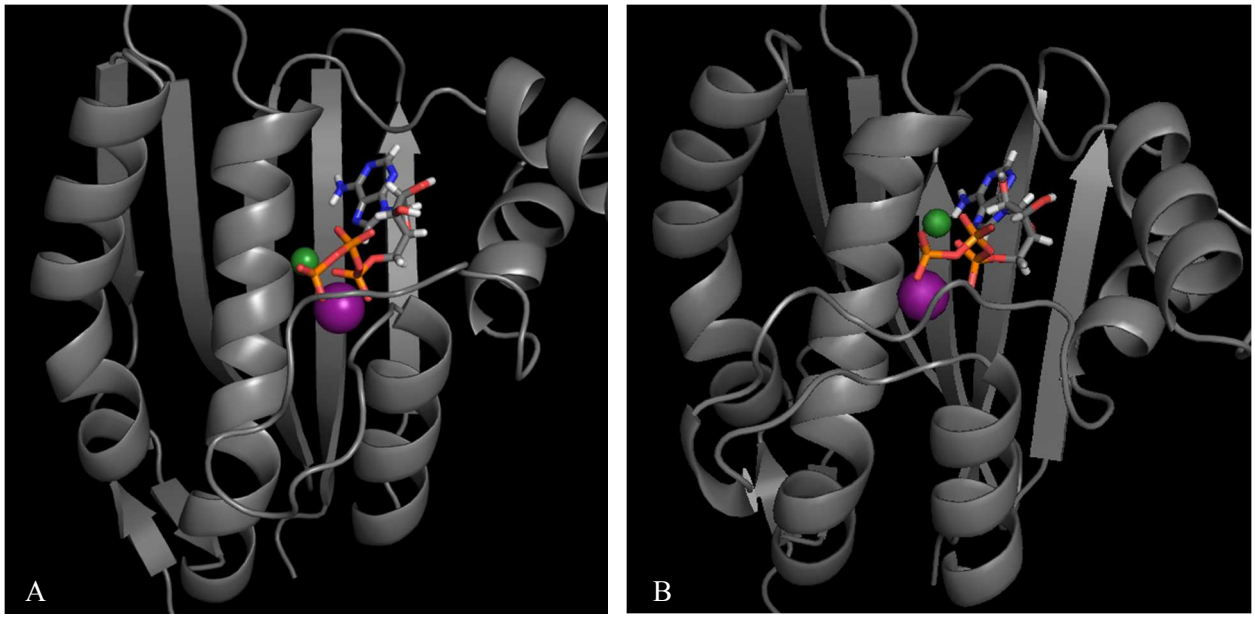


Figure 2.13 **PyMOL representations of the initial and final CheA structures depicting the potassium monovalent/divalent cation-ATP complex.** Figure A on the left is the initial potassium containing structure constructed for the dynamics simulation while Figure B on the right is a representative structure from the various frames deposited during the trajectory of the dynamics run. The Mg^{2+}/ATP/K^{+} complex

are colored violet for the alkali metal and dark green for the alkali earth metal while ATP uses CPK conventions.

BCK and MutL had revealed a common monovalent cation-binding site by the lid of the pocket at the glycine motif. The VMD representations herein indicate that the mobility of the domain structure is diminished in the presence of cation relative to the trajectory seen without cation. These sites of mobility are relatively localized to the G2 and F motifs at the hinges of the ATP binding pocket lid, with the greatest displacement occurring in the structure that lacks a monovalent cation in this region. This mobility can be noted in Figure 2.12B in which the loop has swung out further from the pocket over the course of the dynamics run.

The mobility of this ATP lid was previously seen in structural studies of CheA determined with ATP-analog, ADP or nucleotide-free in which this loop varied significantly [23]. This variation along with the NMR structure of EnvZ suggested that the loop remains highly mobile in the absence of nucleotides. When the structure of the P4 domain is determined with the ADPCP–Mg^{2+} complex the α2-α3 loop turns into a short helical section bordering the domain pocket.

This structural feature is discernible in Figure 2.13B as the loop remains near the pocket and gains secondary structure in multiple frames saved during dynamics. The previous study noted that the groove resulting from this conformation surrounds the γ-phosphate and may aid in P1 binding to P4 [71].

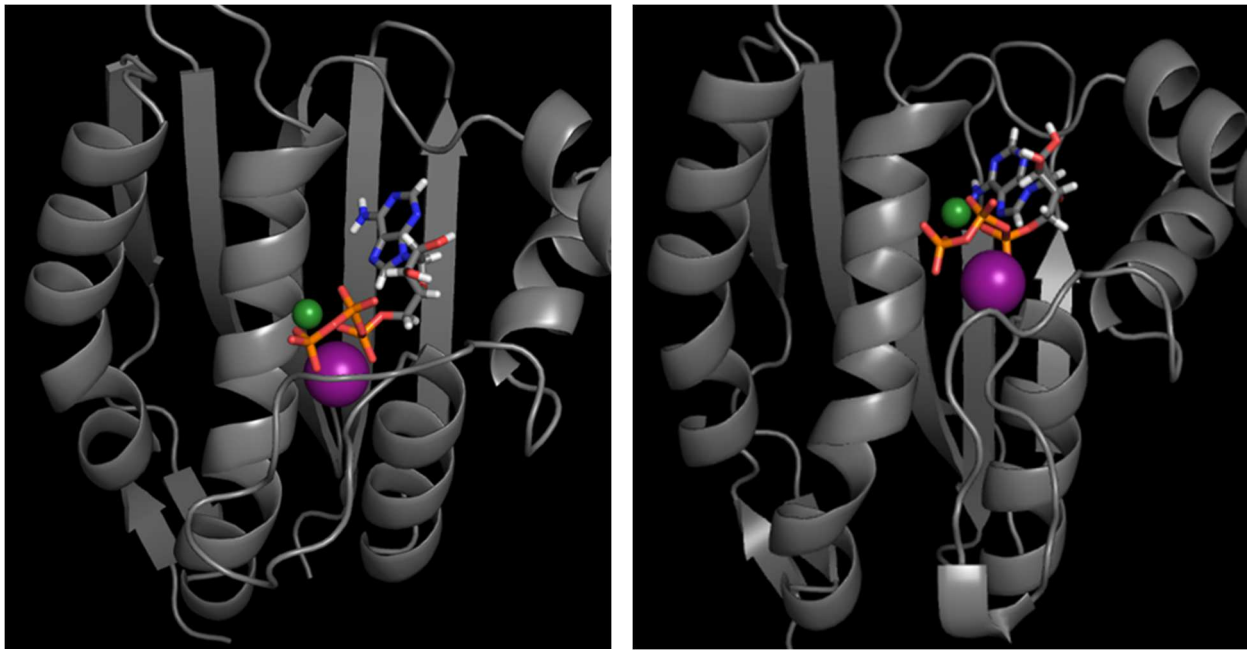


Figure 2.14 **PyMOL representations of the initial and final CheA structures depicting the cesium monovalent/divalent cation-ATP complex.** Figure A on the left is the initial cesium containing structure constructed for the dynamics simulation while Figure B on the right is a representative structure from the various frames deposited during the trajectory of the dynamics run. The Mg^{2+}/ATP/Cs^{+} complex are colored violet for the alkali metal and dark green for the alkali earth metal while ATP uses CPK conventions.

The dynamics simulation of the P4 domain determined with the Mg^{2+}/ATP/Cs^{+} complex showed that the α2-α3 loop was capable of changing configuration into the same short helical section. The helical loop section can be seen in Figure 2.13B (the potassium structure) and 2.14B (the cesium representation shown). However this structural component maintained less contact with the pocket and its ligands since the loop had now swung out in a fashion reminiscent of the apo dynamics configurations.

The potassium structure was focused upon due to this particular monovalent cation having the highest experimental activity as shown earlier and the categorization of the enzyme as a potassium activated enzyme of unknown mechanism. The respective backbone mobility as determined by the RMSD was

compared over the entire range of P4 amino acids (residues 341 to 525) between the structures with/out the potassium cation within the ATP lid pre-equilibration. These values, calculated as a distance of angstroms, are depicted graphically in Figure 2.15 and displayed in the table in Appendix A.

The root-mean-square displacement average of the α-carbon chain in the catalytic domain was calculated and the standard deviation determined using the values reported in the appendix. The average displacement in the absence of cation was 1.70 ± 1.15 Å. The average displacement in the presence of potassium cation was 1.64 ± 1.10 Å. Significant displacements were determined as values greater than a 95% confidence interval from these averages. The upper limits for significant movement therefore were 3.95 and 3.79 Å for monovalent lacking and potassium respectively.

While the average displacements are well within error of each for the structures the maximum displacements are not. Also of some interest was the observation that the locations of these mobile regions differ between the two structures. The cation lacking structure has mobility in the stretch of residues 478-486 with 9 Å of displacement. The structure with potassium however moves 7.27 Å near amino acids 457-462.

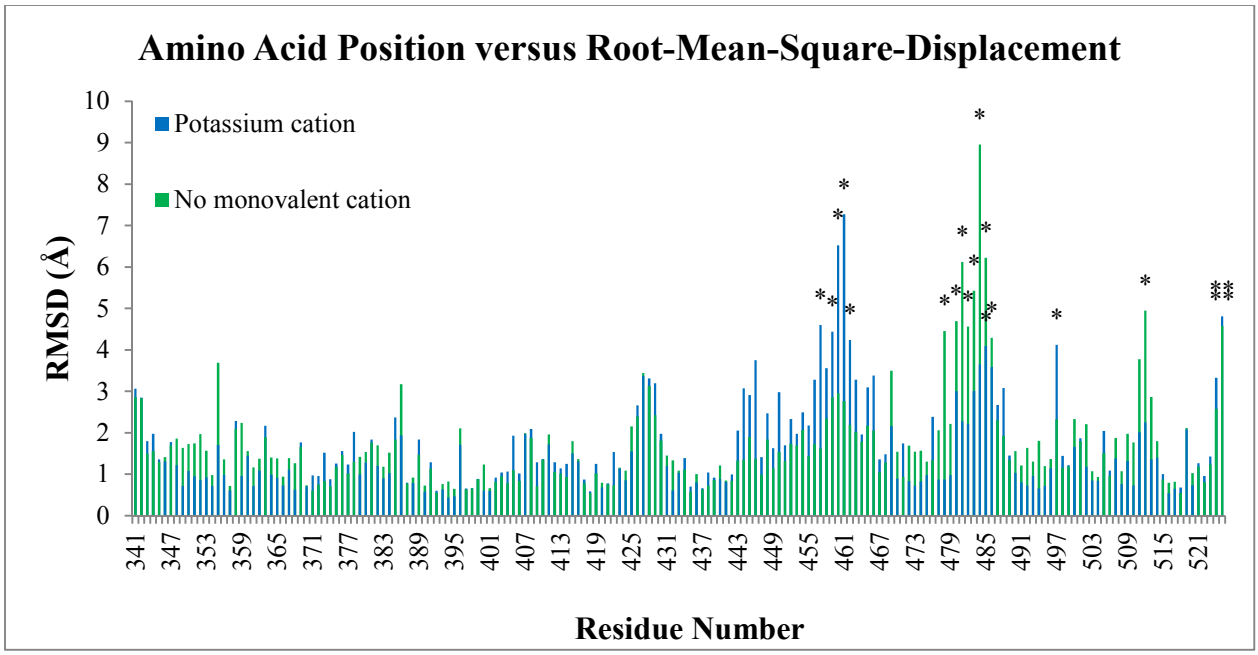


Figure 2.15 **Root mean square displacements of the α-carbons over the entire P4 domain.** RMSD values were calculated with and without monovalent cation, significantly mobile residues have asterisks marking them. The values can also be seen in Appendix A; the entries for the significantly mobile residues contain asterisks in the table as well.

A total of 80 ns of dynamics resulted in a structure in which the turn of the loop between G2 and the F motif was stable. The cause of this stability at residues 478-486 relative to the monovalent lacking run seems related to the location of the monovalent cation. The potassium sits within the partially open ATP lid that allows the ATP to be partially solvated and the P1 domain to make contact. Within three angstroms of the potassium post dynamics are four oxygen atoms available for coordinating the cation. The side-chain oxygen of S477 and the carbonyl oxygen of V490 are 2.8 and 2.9 Å respectively. The ATP is close enough to coordinate with the cation via O_2A and O_3G oxygen.

The implication of this potential coordination is that the monovalent and divalent cations would contribute to the stabilization of the ATP molecule within the P4 domain without interfering with the P1-P4 interaction required for kinase activity. Crystallography and studies of the effect of monovalent cation

on the K_M of ATP are needed confirm this putative site while further analysis would determine the binding order and mechanism of this interaction.

The spherical atomic radial distribution functions were calculated using the trajectory files of the dynamics runs with and without the potassium cation as part of the Mg^{2+}/ATP/MVC complex. The radial distribution functions between the phosphorus atom of the γ-phosphate and the oxygen atoms of the solvent molecules were determined with the default settings of the *g(r)* plug in with the Visual Molecular Dynamics introduced in Version 1.3 [Figure 2.16]. The algorithm is described further in the Methods and Materials.

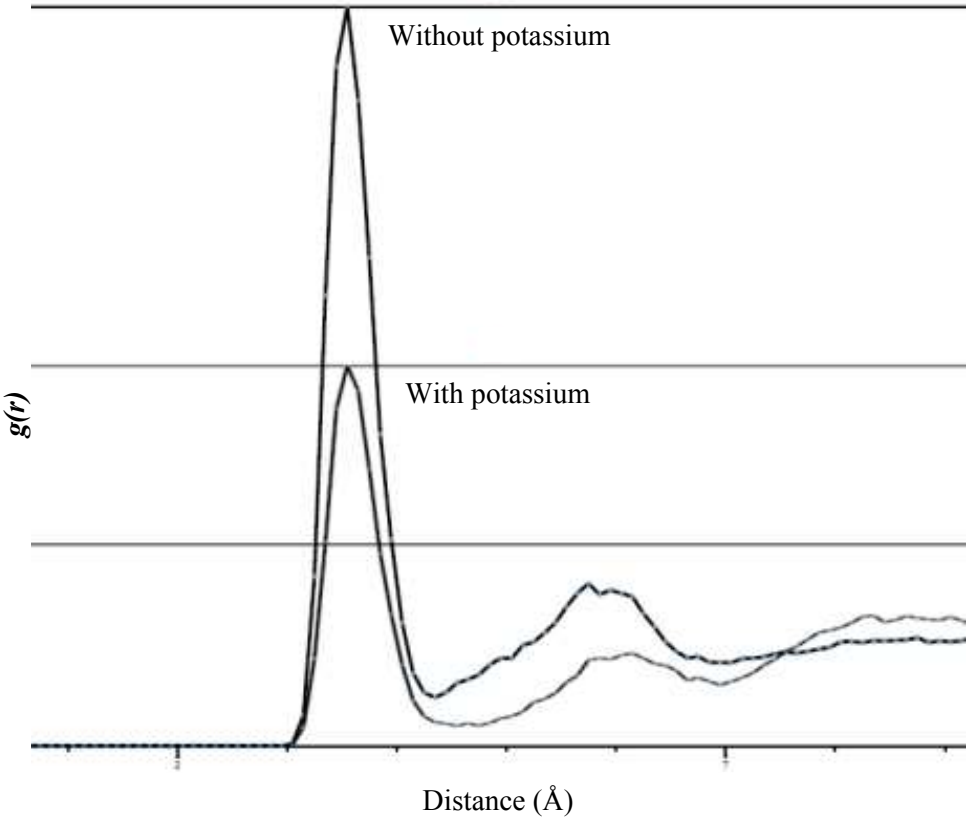


Figure 2.16 **Overlaid radial distribution functions calculated with/out potassium.** The representation herein shows the radial distribution functions calculated over roughly ten angstroms of distance from the atomic center of the γ-phosphate phosphorus. These radial distributions were created with the plug-in within VMD and overlaid using the GNU Image Manipulation Program (GIMP) freeware. The trajectory files resulting from the dynamics run with and without potassium cation were used to produce the graphs seen and the numerical values can be seen in Appendix B.

The radial distribution of the complex without the monovalent cation had a more extensive network of water surrounding the γ-phosphate. The shell apparent at roughly 3.5 Å from the center of the phosphorus atom was approximately twice the quantity of water molecules without the monovalent cation as that with the potassium cation. The value of the cumulative water network was also larger, as shown in Table 5. The complete set of radial distribution values (reported as the distance, the value of the radial distribution function at that distance and the cumulative integral value of the radial distribution function) are available in Appendix B.

	With potassium	**Without monovalent cation**
Cumulative after 10.05 Å	**62.4798**	**76.8247**

Table 5. **Cumulative sum of potential energies from the radial distribution function with/out potassium cation.** The energies were calculated for the solvated structures containing the monovalent/divalent cation-ATP complex over a span of roughly ten angstroms.

The ATP γ-phosphate seems to be less solvent accessible with the presence of a monovalent cation. This indicates that the accessibility is likely not the cause of the enhanced kinase activity achieved with potassium cation. The structural changes seen previously in crystallographic and NMR studies were apparent here within the dynamics results. The transition of the α2-α3 loop into a short helical section bordering the domain pocket was a unifying result of past studies and the current work herein. The structural transition then possibly freezes the binding pocket lid into a configuration that aids in the binding of the P1 domain to the P4 domain ATP pocket.

MATERIALS AND METHODS

Materials

All chemicals were highly pure reagent-grade. Chemicals were obtained from Sigma Aldrich with the following exceptions: [γ-^{32}P] ATP was from Perkin-Elmer, DTT from Research Products International, Ni-NTA agarose resin from Qiagen, and the bicinchoninic acid assay (BCA) reagents were from Biorad.

GHKL superfamily member alignment with MUSCLE

The MUSCLE program was chosen for its availability to the public as freeware, benchmark speed on a desktop or relative computational complexity, ease of use, biological accuracy over CLUSTALW, Progressive POA, T-Coffee and MAFFT script FFTNSI [72, 73]. Accuracy during the program comparisons was determined by the benchmarks BAliBASE, SABmark, SMART and PREFAB.

MUSCLE functions by calculating the sub-sequence distances, or kmer distances, of the amino acid pairs. The kmer counts of distance are determined by a scoring function derived from the probabilities of both amino acids from the 240 PAM VTML matrixes. The function of distance score, or log-expectation (LE) score has a normalized frequency and is scaled by the occupancy of the sequence to account for binary penalization for gaps.

$$\text{Binary approximation: } F^{Binary} = \frac{\sum_{\tau}\delta_{XY}(\tau)}{[\min(L_X, L_Y) - k + 1]} \qquad \text{Equation 2}$$

In the above approximation two sequence lengths, X and Y, are compared over a kmer span of τ. The function δ is either one or zero depending on the mutual sequence occupation, 1 for amino acid presence in both sequences and 0 in case of indel or gap insertion. The PAM VTML matrix uses a maximum likelihood (ML) approach suited to small scale divergent sequence alignments. The matrix stems from Markov chains with exchange rates of evolution in which one amino acid could mutate to any other. This divergence in sequence is referred to as the point accepted mutation unit or PAM [74, 75].

The kmer distances are compiled into a distance matrix that is clustered by the un-weighted pair-group method using arithmetic averages (UPGMA) or neighbor-joining (NJ) [76]. This creates a binary tree that is aligned by branching order of the subsequently divergent sequences before ancestral producing a multiple alignment file at the predicted phylogenetic root. In order to avoid bias due to approximations of the kmer distance re-estimations are made within each round using Kimura distance measured from the previous multiple alignment file.

The Kimura distances are created by the Kimura two-parameter model in which the genetic correlation of a pair decreases exponentially with increasing distance [77]. The Kimura distances are clustered like before and used to produce a set of subtrees to recreate the alignment tree with branching orders relative to the original tree. The subtrees are separated by deleting edges of the aligned tree starting from the root and progressing distally. With each subtree progression the sum of pairs is computed with improved pairs being kept and others being discarded until a convergence or pre-defined parameters are met. The Kimura distance depends upon the fractional identity that is determined by a mutational distance computed by the number of mutations between the historic mutational paths between two amino acids as decided by the most recent common ancestor. Thus as sequences diverge the probability of multiple mutations at a single site increase [73].

The comparative benchmarks used to determine the optimal program to use were the reference alignment sets for versions MUSCLE 3.2, CLUSTALW 1.82, T-Coffee 1.37 and MAFFT 3.82. Cursory tests were made with MUSCLE-p which incorporates time and space complexity parameters but lack refinement steps. Comparative attempts were made with Align-m 1.0 but the program self-aborted or was impractically slow.

The sets are considered indicative as they covered a variety of alignment conditions such as repeats, inversion, transmembrane helices, super-families, signaling domains and a proprietary set using structural methods but not sequence similarities selected from the FSSP database [78, 79, 80]. The accuracy

measurements were produced with three scores: Q (quality or the number of correctly aligned residue pairs), TC (total column score or the correctly aligned columns divided by the number of columns), and APDB (a derived score from structure that is not dependent upon the reference sequence alignment). The benchmark speed was determined by the CPU times.

Protein expression and purification

S. typhimurium CheA kinase and CheW adaptor cysteine-less pET vectors were expressed with histidine tags at the N-termini in *E. coli* strain BL21 [81]. *E. coli* CheY pVS vectors were expressed with histidine tags at the C-termini in *E. coli* strain M15 containing pREP4 plasmid [82]. CheA, CheW, and CheY proteins were purified using standard Ni-NTA agarose affinity chromatography [83]. Protein concentrations were calculated from the UV absorbance at 276-280 nm using the respective extinction coefficients previously determined [84, 58].

E. coli transmembrane serine receptor was overexpressed in *E. coli* strain UU1581, lacking all chemotactic proteins [85]. Inner membrane bacterial vesicles containing Tsr were isolated as previously described with minor modification [81, 86]. The total concentration of membrane associated proteins was determined with a standard BCA assay and the receptor concentration calculated by the receptor fraction of total protein found by densitometry on SDS-PAGE [87].

Functional array formation

Arrays were reconstituted as previously described by combining 6.7 μM *E. coli* Tsr containing inner membrane vesicles, 5 μM *S. typhimurium* CheA, and 10 μM *S. typhimurium* CheW in 1X physiological kinase buffer (140 mM KCl, 15mM NaCl, 5 mM $MgCl_2$, 0.5 mM EDTA, 50 mM Tris, pH 7.5) with 0.5 mg/ml BSA, 2mM PMSF, and 2 mM TCEP [88, 89, 90, 47]. This mixture was incubated for 45 min at 22 °C, after which membranes with bound components were pelleted by centrifugation at 13200*g* for 7 min at 4 °C. Pellets were re-suspended in tenfold their original volume of the various 1X monovalent cation buffers, re-pelleted and re-suspended, then finally re-suspended in 15μL of their respective 1X

monovalent cation buffer. The chemotaxis proteins of *S. typhimurium* and *E. coli* are functionally interchangeable, and chemotaxis arrays generated using the protocol described herein are active and regulated by serine attractant [58].

The monovalent cation buffers were 5 mM $MgCl_2$, 0.5 mM EDTA, 50 mM Tris, pH 7.5 with 150 mM monovalent cations. These included: LiCl, NaCl, KCl, NH_4, RbCl, and CsCl. There was an additional gradient of solutions containing 12.5 mM, 25 mM, 50 mM, 75 mM, and 100 mM KCl with compensatory amounts of NaCl for a concentration of 150 mM total cation.

Quantification of functional array specific activity

The relative specific activities of the arrays were determined as previously described with minor modifications [81, 84]. Kinase reaction mixtures with a final volume of 10μL contained 3.4 μM *E. coli* Tsr with associated CheA and CheW, 40 μM CheY, 1X monovalent cation buffer, and 1 mM cold ATP - [γ-^{32}P] ATP (0.1 μM) mixture. The attractant regulation was tested by setting up parallel reactions with the addition of 2mM serine in the mixture. The reaction was initiated by ATP addition and quenched after ten seconds by the addition of 30μL 2X Laemmli sample loading buffer (125 mM Tris, 4% w/v SDS, 30% w/v glycerol, 0.005% bromo-phenol blue, pH 6.8) containing 40 mM EDTA. The samples were resolved with denaturing urea SDS-PAGE that was dried and imaged on a phosphor-imager screen and quantified with ImageJ.

The specific activity was calculated by determining the CheA densitometry on 12.5% acrylamide SDS-PAGE. Of the 15μL final re-suspension during array formation, 5μL was mixed with 10μL 2 X Laemmli samples loading buffer containing 40 mM DTT and resolved with Coomassie Brilliant Blue R-250. The image was acquired with a LAS4000 imagers and the density of the CheA band determined with ImageJ. The relative CheY-P^{32} levels were divided by the relative CheA levels to find the specific kinase activity that was then normalized to the free cysless CheA loading controls to control for inter-gel staining variability.

Data and error analysis

All data points shown are averages of "*n*" replicates in which each set includes three technical replicates of each experimental condition. The error bars indicate standard error of the mean calculated with "*n*" indicated in the figure. Results are normalized to 150 mM KCl for ease of comparison.

De novo structural modeling with Phyre2

Various algorithms have been developed recently and the Critical Assessment of Structure Prediction, or the biennial CASP assessment, has emerged as a premier benchmark for this field of computational modeling [91]. For the purposes of this study, we chose the Phyre2 server for its availability on a public webserver, benchmark speed, ease of use, acceptable results within previous rounds of the CASP assessment, and use previously in the lab [90].

In the Phyre2 server three independent algorithms (PSIPRED, SSpro3, and Jpred 3) are used for a three-state secondary structure prediction of alpha helix, beta strand and coil states along with a confidence interval that is averaged between the three and used to create a consensus prediction [92, 93, 94]. The consensus prediction, along with a two-state prediction of protein dis-order from the DISOPRED server, is put into a set of profile-profile algorithms to determine alignment scores and construct a three dimensional model from the ten highest E-value alignments [95]. The E-value is the scoring function which determines the average number of tests that would obtain the extremes of the observed statistics assuming a null hypothesis.

$$E = K * m * n * e^{-S*\lambda} \qquad \text{Equation 1}$$

The E-value is dependent upon the lengths of the sequences (m and n) and the assigned score S of the high-scoring segment pairs as determined by a modified Smith-Waterman algorithm for local alignment. The scoring algorithm used determines the parameters K and λ, which are scales for search size and score system respectively. The profile-profile algorithms used were PDP (Probability Dot Product), BASIC

(Bilaterally Amplified Sequence Information Comparison), B-DHIP (Bi-Directional Heterogeneous Inner Product) and PCorr (Pearson's Correlation Coefficient of Probability) as described previously [96].

A current structural model of the kinase in complex with its associated adaptor protein CheW is not available for *Salmonella typhimirium* but is for *Thermotoga maritima* [97]. However, previous studies have shown that *S. typhimirium*, *Escherichia coli,* and *T. maritima* all have extensive sequence and structural homology with greater conservation occurring between *S. typhimirium* and *E. coli* [98, 88, 89]. Any structural references herein refer to a structural model of the *S. typhimirium* amino acid sequence (UniProtKB entry P09384) threaded onto the *T. maritima* crystal structure (Accession number 1I59) with the Protein Homology/analogY Recognition Engine V 2.0 (Phyre2) modeling server [99].

In silico dynamics modeling

The monovalent cation protein models were generated with the protocol described above with the Phyre2 server. Minor modification to histidine 393 was made using the PyMOL Molecular Graphics System (Version 1.3) with the threaded structures. The Crystallographic Object-Oriented Toolkit (WinCoot Version 0.7.2) confirmed a suitable electron density map for rotamer alternatives for this histidine.

A solvated structural system was then built with Visual Molecular Dynamics (VMD Version 1.9.1) by padding the protein with 15 Å of water molecules on all sides [61]. Magnesium ions were added to neutralize the charge and the periodic boundary conditions of constant particle numbers, a constant pressure of 1 atmosphere and the constant temperature of 300 K (NPT) set up.

Molecular dynamics simulations were then performed by a collaborator, Dr. Margaret Hurley at the Army Research Laboratory. The system minimized and equilibrated for 10ns at constant particle number, pressure and temperature (NPT), before accumulating data for an additional 30 ns. Simulations were performed with the CHARMM36 bio-molecular force field with CMAP correction and TIP3P water within the Nanoscale Molecular Dynamics program (NAMD Version 2.9) [62, 63, 64, 65, 66] Force-field

parameters for monovalent cations were taken from Roux et al [67]. Force-field parameters for ATP were taken from the CHARMM General Force Field for organic moleculesv2b9 (CGenFF) [68, 69, 70].

The system was minimized for 5000 steps then ran time-steps of 0.2 fs over a time course of 10 nanoseconds with the aforementioned periodic boundary conditions. A further 30 ns trajectory was then performed for analysis to generate the results shown herein. The trajectory files from the simulation were then compared to the original PDB file created as described above. Sixteen frames were taken from the trajectory file and run with the tool command language file "coloring.tcl" made freely available by the Theoretical and Computational Biophysics Group. This script colors the structure based upon the displacement. The root mean square deviation of the atomic positions was calculated for all α-carbons of the proteins by the script "residue_rmsd.tcl". The equation is as follows:

$$RMSD = \sqrt{\frac{\sum_{i=1}^{N\ atoms}\left(r_{i,1} - r_{i,2}\right)^2}{N_{atoms}}} \qquad \text{Equation 2}$$

$$N_{atoms} = total\ number\ of\ atoms\ whose\ positions\ are\ compared$$

Calculation of the radial distribution function

The *g(r)* GUI Plugin, Version 1.3 was chosen for its availability- namely its integration into the VMD open source program that had been previously used, benchmark speed, ease of use and the multithreaded implementation for standard desktop GPU use. The trajectory files of the dynamics runs obtained as described above, with and without the potassium cation as part of the Mg^{2+}/ATP/MVC complex, were used to determine the spherical atomic radial distribution functions. The functions were found specifically for the initial selection, the phosphorus atom of the γ-phosphate, and the oxygen atoms of the solvent molecules as the secondary selection. The default binning setting of the *g(r)* plugin with the Visual Molecular Dynamics program was used to create the histograms [100].

The radial distribution function *g(r)* is a measure of the probability of finding two particles a distance *r* apart. It may be expressed as Equation 3, where the numerator *p(r)* is the average number of a specified atomic pair found within a distance '*r'* and '*r + dr'* of each other, and *V* is the system volume. This is approximated by histogram with the default width *dr* of 0.1 Å, which is deemed sufficient for the detail shown in Figure 2.16. The results were output as the distance, the value of the radial distribution function at that distance and the cumulative integral value of the radial distribution function.

$$g(r) = \lim_{dr \to 0} \frac{p(r)}{4\pi(\frac{N_{pairs}}{V})r^2 dr} \qquad \text{Equation 3}$$

REFERENCES

1. Barrett, J. F. and Hoch, J. A. "**Two-Component Signal Transduction as a Target for Microbial Anti-Infective Therapy**" *Antimicrobial Agents and Chemotherapy* 42, no. 7 (1998): 1529–1536. Available at http://www.ncbi.nlm.nih.gov/pubmed/9660978

2. Falke, J. J., Bass, R. B., Butler, S. L., Chervitz, S. A., and Danielson, M. A. "**THE TWO-COMPONENT SIGNALING PATHWAY OF BACTERIAL CHEMOTAXIS: A Molecular View of Signal Transduction by Receptors, Kinases, and Adaptation Enzymes**" *Annual Review of Cellular and Developmental Biology* 13, (1997): 457–512. doi:10.1146/annurev.cellbio.13.1.457.THE, Available at http://www.ncbi.nlm.nih.gov/pubmed/9442881

3. Falke, J. J. and Piasta, K. N. "**Architecture and Signal Transduction Mechanism of the Bacterial Chemosensory Array: Progress, Controversies, and Challenges.**" *Current opinion in structural biology* 29, (2014): 85–94. doi:10.1016/j.sbi.2014.10.001, Available at http://www.ncbi.nlm.nih.gov/pubmed/25460272

4. Stock, a M., Robinson, V. L., and Goudreau, P. N. "**Two-Component Signal Transduction.**" *Annual review of biochemistry* 69, (2000): 183–215. doi:10.1146/annurev.biochem.69.1.183, Available at http://www.ncbi.nlm.nih.gov/pubmed/10966457

5. Mascher, T., Helmann, J. D., and Unden, G. "**Stimulus Perception in Bacterial Signal-Transducing Histidine Kinases.**" *Microbiology and molecular biology reviews* 70, no. 4 (2006): 910–38. doi:10.1128/MMBR.00020-06, Available at http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1698512&tool=pmcentrez&rendertype=abstract

6. Hazelbauer, G. L., Engström, P., and Harayama, S. "**Methyl-Accepting Chemotaxis Protein III and Transducer Gene Trg.**" *Journal of bacteriology* 145, no. 1 (1981): 43–9. Available at http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=217242&tool=pmcentrez&rendertype=abstract

7. Weerasuriya, S. "**Chimeric Chemoreceptors in Escherichia Coli: Signaling Properties of Tar-Tap and Tap-Tar Hybrids**" *Journal of bacteriology* 180, no. 4 (1998): 914–920. Available at http://jb.asm.org/content/180/4/914.short

8. Frye, J., Karlinsey, J., and Felise, H. "**Identification of New Flagellar Genes of Salmonella Enterica Serovar Typhimurium**" *Journal of bacteriology* 188, no. 6 (2006): 2233–2243. doi:10.1128/JB.188.6.2233, Available at http://jb.asm.org/content/188/6/2233.short

9. Li, M. and Hazelbauer, G. L. "**Cellular Stoichiometry of the Components of the Chemotaxis Signaling Complex**" *Journal of Bacteriology* 186, no. 12 (2004): 3687–3694. doi:10.1128/JB.186.12.3687-3694.2004, Available at http://jb.asm.org/content/186/12/3687.short

10. Kim, S.-H., Wang, W., and Kim, K. K. "**Dynamic and Clustering Model of Bacterial Chemotaxis Receptors: Structural Basis for Signaling and High Sensitivity.**" *Proceedings of the National Academy of Sciences of the United States of America* 99, no. 18 (2002): 11611–5. doi:10.1073/pnas.132376499, Available at http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=129317&tool=pmcentrez&rendertype=abstract

11. Kim, K. K., Yokota, H., and Kim, S. H. "**Four-Helical-Bundle Structure of the Cytoplasmic Domain of a Serine Chemotaxis Receptor.**" *Nature* 400, no. 6746 (1999): 787–92. doi:10.1038/23512, Available at http://www.ncbi.nlm.nih.gov/pubmed/10466731

12. Falke, J. J. and Hazelbauer, G. L. "**Transmembrane Signaling in Bacterial Chemoreceptors.**" *Trends in biochemical sciences* 26, no. 4 (2001): 257–65. doi:10.1016/S0968-0004(00)01770-9, Available at http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2895674&tool=pmcentrez&rendertype=abstract

13. Levit, M. N. and Stock, J. B. "**Receptor Methylation Controls the Magnitude of Stimulus-Response Coupling in Bacterial Chemotaxis.**" *The Journal of biological chemistry* 277, no. 39 (2002): 36760–5. doi:10.1074/jbc.M204325200, Available at http://www.ncbi.nlm.nih.gov/pubmed/12119291

14. Rice, M. S. and Dahlquist, F. W. "**Sites of Deamidation and Methylation in Tsr, a Bacterial Chemotaxis Sensory Transducer.**" *The Journal of biological chemistry* 266, no. 15 (1991): 9746–53. Available at http://www.ncbi.nlm.nih.gov/pubmed/2033064

15. Falke, J. J. and Piasta, K. N. "**Architecture and Signal Transduction Mechanism of the Bacterial Chemosensory Array: Progress, Controversies, and Challenges.**" *Current opinion in structural biology* 29, (2014): 85–94. doi:10.1016/j.sbi.2014.10.001, Available at http://dx.doi.org/10.1016/j.sbi.2014.10.001

16. Falke, J. J. and Kim, S. H. "**Structure of a Conserved Receptor Domain That Regulates Kinase Activity: The Cytoplasmic Domain of Bacterial Taxis Receptors**" *Current Opinion in Structural Biology* 10, no. 4 (2000): 462–469. doi:10.1016/S0959-440X(00)00115-9, Available at http://www.ncbi.nlm.nih.gov/pubmed/10981636

17. Bourret, R. B., Hess, J. F., and Simon, M. I. "**Conserved Aspartate Residues and Phosphorylation in Signal Transduction by the Chemotaxis Protein CheY.**" *Proceedings of the National Academy of Sciences of the United States of America* 87, no. 1 (1990): 41–5. Available at http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=53195&tool=pmcentrez&rendertype=abstract

18. McEvoy, M. M., Bren, a, Eisenbach, M., and Dahlquist, F. W. "**Identification of the Binding Interfaces on CheY for Two of Its Targets, the Phosphatase CheZ and the Flagellar Switch Protein fliM.**" *Journal of molecular biology* 289, no. 5 (1999): 1423–33. doi:10.1006/jmbi.1999.2830, Available at http://www.ncbi.nlm.nih.gov/pubmed/10373376

19. Dyer, C. M. and Dahlquist, F. W. "**Switched or Not?: The Structure of Unphosphorylated CheY Bound to the N Terminus of FliM.**" *Journal of bacteriology* 188, no. 21 (2006): 7354–63. doi:10.1128/JB.00637-06, Available at http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1636273&tool=pmcentrez&rendertype=abstract

20. Bai, F., Branch, R. W., Nicolau, D. V, Pilizota, T., Steel, B. C., Maini, P. K., and Berry, R. M. "**Conformational Spread as a Mechanism for Cooperativity in the Bacterial Flagellar Switch.**" *Science (New York, N.Y.)* 327, no. 5966 (2010): 685–9. doi:10.1126/science.1182105, Available at http://www.ncbi.nlm.nih.gov/pubmed/20133571

21. White, D. "**The Physiology and Biochemistry of Prokaryotes / David White**" (2007):

22. Swanson, R., Alex, L., and Simon, M. "**Histidine and Aspartate Phosphorylation: Two-Component Systems and the Limits of Homology**" *Trends in biochemical sciences* no. 21 (1994): 485–490. doi:10.1016/0968-0004(94)90135-X, Available at http://www.sciencedirect.com/science/article/pii/096800049490135X

23. Bilwes, A. M., Alex, L. A., Crane, B. R., and Simon, M. I. "**Structure of CheA, a Signal-Transducing Histidine Kinase.**" *Cell* 96, no. 1 (1999): 131–41. doi:10.1016/S0092-8674(00)80966-6, Available at http://www.ncbi.nlm.nih.gov/pubmed/9989504

24. Zhou, H., Lowry, D. F., Swanson, R. V, Simon, M. I., and Dahlquist, F. W. "**NMR Studies of the Phosphotransfer Domain of the Histidine Kinase CheA from Escherichia Coli: Assignments, Secondary Structure, General Fold, and Backbone Dynamics.**" *Biochemistry* 34, no. 42 (1995): 13858–70. doi:10.1021/bi00042a018, Available at http://www.ncbi.nlm.nih.gov/pubmed/7577980

25. Levit, M. N., Liu, Y., and Stock, J. B. "**Mechanism of CheA Protein Kinase Activation in Receptor Signaling Complexes.**" *Biochemistry* 38, no. 20 (1999): 6651–8. doi:10.1021/bi982839l, Available at http://www.ncbi.nlm.nih.gov/pubmed/10350484

26. Zhou, H. and Dahlquist, F. W. "**Phosphotransfer Site of the Chemotaxis-Specific Protein Kinase CheA as Revealed by NMR.**" *Biochemistry* 36, no. 4 (1997): 699–710. doi:10.1021/bi961663p, Available at http://www.ncbi.nlm.nih.gov/pubmed/9020767

27. Swanson, R. V, Schuster, S. C., and Simon, M. I. "**Expression of CheA Fragments Which Define Domains Encoding Kinase, Phosphotransfer, and CheY Binding Activities.**" *Biochemistry* 32, no. 30 (1993): 7623–9. doi:10.1021/bi00081a004, Available at http://www.ncbi.nlm.nih.gov/pubmed/8347572

28. Park, S. Y., Ham, S. W., Kim, K. Y., and Crane, B. R. "**Crystallization and Preliminary X-Ray Crystallographic Analysis of Escherichia Coli CheA P3 Dimerization Domain.**" *Acta crystallographica Section F: Structural biology and crystallization communications* 67, no. 6 (2011): 662–664. doi:10.1107/S1744309111010335, Available at http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3107136&tool=pmcentrez&rendertype=abstract

29. Gestwicki, J. E. and Kiessling, L. L. "**Inter-Receptor Communication through Arrays of Bacterial Chemoreceptors.**" *Nature* 415, no. 6867 (2002): 81–4. doi:10.1038/415081a, Available at http://www.ncbi.nlm.nih.gov/pubmed/22729033

30. Park, S.-Y., Borbat, P. P., Gonzalez-Bonet, G., Bhatnagar, J., Pollard, A. M., Freed, J. H., Bilwes, A. M., and Crane, B. R. "**Reconstruction of the Chemotaxis Receptor-Kinase Assembly.**" *Nature structural & molecular biology* 13, no. 5 (2006): 400–7. doi:10.1038/nsmb1085, Available at http://www.ncbi.nlm.nih.gov/pubmed/16622408

31. Boukhvalova, M. S., Dahlquist, F. W., and Stewart, R. C. "**CheW Binding Interactions with CheA and Tar. Importance for Chemotaxis Signaling in Escherichia Coli.**" *The Journal of biological chemistry* 277, no. 25 (2002): 22251–9. doi:10.1074/jbc.M110908200, Available at http://www.ncbi.nlm.nih.gov/pubmed/11923283

32. Gegner, J. and Dahlquist, F. "**Signal Transduction in Bacteria: CheW Forms a Reversible Complex with the Protein Kinase CheA**" *Proceedings of the National Academy of Sciences* 88, no. February (1991): 750–754. Available at http://www.pnas.org/content/88/3/750.short

33. Wolfe, A. and Stewart, R. "**The Short Form of the CheA Protein Restores Kinase Activity and Chemotactic Ability to Kinase-Deficient Mutants.**" *Proceedings of the National Academy of Sciences* 90, no. February (1993): 1518–1522. doi:10.1073/pnas.90.4.1518, Available at http://www.pnas.org/content/90/4/1518.short

34. Levit, M., Liu, Y., Surette, M., and Stock, J. "**Active Site Interference and Asymmetric Activation in the Chemotaxis Protein Histidine Kinase CheA**" *Journal of Biological Chemistry* 271, no. 50 (1996): 32057–32063. doi:10.1074/jbc.271.50.32057, Available at http://www.jbc.org/content/271/50/32057.short

35. Surette, M. G., Levit, M., Liu, Y., Lukat, G., Ninfa, E. G., Ninfa, a., and Stock, J. B. "**Dimerization Is Required for the Activity of the Protein Histidine Kinase CheA That Mediates Signal Transduction in Bacterial Chemotaxis**" *Journal of Biological Chemistry* 271, no. 2 (1996): 939–945. doi:10.1074/jbc.271.2.939, Available at http://www.jbc.org/cgi/doi/10.1074/jbc.271.2.939

36. Quezada, C. M., Hamel, D. J., Gradinaru, C., Bilwes, A. M., Dahlquist, F. W., Crane, B. R., and Simon, M. I. "**Structural and Chemical Requirements for Histidine Phosphorylation by the Chemotaxis Kinase CheA.**" *The Journal of biological chemistry* 280, no. 34 (2005): 30581–5. doi:10.1074/jbc.M505316200, Available at http://www.ncbi.nlm.nih.gov/pubmed/15994328

37. Dutta, R., Qin, L., and Inouye, M. "**Histidine Kinases: Diversity of Domain Organization**" *Molecular microbiology* 34, (1999): 633–640. doi:10.1046/j.1365-2958.1999.01646.x, Available at http://onlinelibrary.wiley.com/doi/10.1046/j.1365-2958.1999.01646.x/full

38. Bergerat, A., Massy, B. de, and Gadelle, D. "**An Atypical Topoisomerase II from Archaea with Implications for Meiotic Recombination**" *Nature* 386, no. 6623 (1997): 414–417. doi:10.1038/386414a0, Available at http://www.nature.com/nature/journal/v386/n6623/abs/386414a0.html

39. Dutta, R. and Inouye, M. "**GHKL, an Emergent ATPase/kinase Superfamily**" *Trends in biochemical sciences* 25, no. 1 (2000): 24–28. doi:10.1016/S0968-0004(99)01503-0, Available at http://www.sciencedirect.com/science/article/pii/S0968000499015030

40. Guarnieri, M. T., Blagg, B. S. J., and Zhao, R. "**A High-Throughput TNP-ATP Displacement Assay for Screening Inhibitors of ATP-Binding in Bacterial Histidine Kinases.**" *Assay and drug development technologies* 9, no. 2 (2011): 174–183. doi:10.1089/adt.2010.0289, Available at http://www.ncbi.nlm.nih.gov/pubmed/21050069

41. Tanaka, T., Saha, S. K., Tomomori, C., Ishima, R., Liu, D., Tong, K. I., Park, H., Dutta, R., Qin, L., Swindells, M. B., Yamazaki, T., Ono, a M., Kainosho, M., Inouye, M., and Ikura, M. "**NMR Structure of the Histidine Kinase Domain of the E. Coli Osmosensor EnvZ.**" *Nature* 396, no. 6706 (1998): 88–92. doi:10.1038/23968, Available at http://www.ncbi.nlm.nih.gov/pubmed/9817206

42. Jackson, A. and Maxwell, A. "**Identifying the Catalytic Residue of the ATPase Reaction of DNA Gyrase**" *Proceedings of the National Academy of Sciences* 90, no. 23 (1993): 11232–11236. Available at http://www.pnas.org/content/90/23/11232.short

43. Hu, X., Machius, M., and Yang, W. "**Monovalent Cation Dependence and Preference of GHKL ATPases and Kinases**" *FEBS Letters* 544, no. 1-3 (2003): 268–273. doi:10.1016/S0014-5793(03)00519-2, Available at http://linkinghub.elsevier.com/retrieve/pii/S0014579303005192

44. Machius, M., Chuang, J. L., Wynn, R. M., Tomchick, D. R., and Chuang, D. T. "**Structure of Rat BCKD Kinase: Nucleotide-Induced Domain Communication in a Mitochondrial Protein Kinase.**" *Proceedings of the National Academy of Sciences of the United States of America* 98, no. 20 (2001): 11218–23. doi:10.1073/pnas.201220098, Available at http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=58710&tool=pmcentrez&rendertype=abstract

45. Czyz, A., Szewczyk, A., Nałecz, M. J., and Wojtczak, L. "**The Role of Mitochondrial Potassium Fluxes in Controlling the Protonmotive Force in Energized Mitochondria.**" *Biochemical and biophysical research communications* 210, no. 1 (1995): 98–104. doi:10.1006/bbrc.1995.1632, Available at http://www.ncbi.nlm.nih.gov/pubmed/7741755

46. Rosen, B. P. "**Ion Transport in Prokaryotes**" (1987):

47. Neidhardt, F. C. and Curtiss, R. "**Escherichia Coli and Salmonella : Cellular and Molecular Biology**" (1996):

48. Gunasekaran, K., Ma, B., and Nussinov, R. "**Is Allostery an Intrinsic Property of All Dynamic Proteins?**" *Proteins* 57, no. 3 (2004): 433–43. doi:10.1002/prot.20232, Available at http://www.ncbi.nlm.nih.gov/pubmed/15382234

49. Harding, M. M. "**The Architecture of Metal Coordination Groups in Proteins.**" *Acta crystallographica. Section D, Biological crystallography* 60, no. Pt 5 (2004): 849–59. doi:10.1107/S0907444904004081, Available at http://www.ncbi.nlm.nih.gov/pubmed/15103130

50. Wigley, D., Davies, G., and Dodson, E. "**Crystal Structure of an N-Terminal Fragment of the DNA Gyrase B Protein**" *Nature* 351, (1991): 624–629. doi:10.1038/351624a0, Available at http://www.nature.com/nature/journal/v351/n6328/abs/351624a0.html

51. Page, M. J. and Cera, E. Di. "**Role of Na+ and K+ in Enzyme Function.**" *Physiological reviews* 86, no. 4 (2006): 1049–92. doi:10.1152/physrev.00008.2006, Available at http://www.ncbi.nlm.nih.gov/pubmed/17015484

52. Cera, E. Di. "**A Structural Perspective on Enzymes Activated by Monovalent Cations.**" *The Journal of biological chemistry* 281, no. 3 (2006): 1305–8. doi:10.1074/jbc.R500023200, Available at http://www.ncbi.nlm.nih.gov/pubmed/16267046

53. Collins, K. D. "**Charge Density-Dependent Strength of Hydration and Biological Structure.**" *Biophysical journal* 72, no. 1 (1997): 65–76. doi:10.1016/S0006-3495(97)78647-8, Available at http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1184297&tool=pmcentrez&rendertype=abstract

54. Bourret, R. B., Drake, S. K., Chervitz, S. A., Simon, M. I., and Falke, J. "**Activation of the Phosphosignaling Protein CheY II. Analysis of Activated Mutants by 19 F NMR and Protein Engineering**" *Journal of Biological Chemistry* 268, no. 18 (1993): 13089–13096. Available at http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2892986

55. Gouy, M., Guindon, S., and Gascuel, O. "**SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building.**" *Molecular biology and evolution* 27, no. 2 (2010): 221–4. doi:10.1093/molbev/msp259, Available at http://www.ncbi.nlm.nih.gov/pubmed/19854763

56. Galtier, N., Gouy, M., and Gautier, C. "**SEA VIEW and PHYLO_ WIN: Two Graphic Tools for Sequence Alignment and Molecular Phylogeny**" *Computer applications in the biosciences* 12, no. 6 (1996): 543–548. Available at http://www.ncbi.nlm.nih.gov/pubmed/9021275

57. Ban, C. and Yang, W. "**Crystal Structure and ATPase Activity of MutL: Implications for DNA Repair and Mutagenesis**" *Cell* 95, (1998): 541–552. doi:10.1016/S0092-8674(00)81621-9, Available at http://www.sciencedirect.com/science/article/pii/S0092867400816219

58. Levit, M. N., Grebe, T. W., and Stock, J. B. "**Organization of the Receptor-Kinase Signaling Array That Regulates Escherichia Coli Chemotaxis.**" *The Journal of biological chemistry* 277, no. 39 (2002): 36748–54. doi:10.1074/jbc.M204317200, Available at http://www.ncbi.nlm.nih.gov/pubmed/12119289

59. Missel, P. J., Mazer, N. A., Carey, M. C., and Benedek, G. . "**Thermodynamics of the Sphere-to-Rod Transition in Alkyl Sulfate Micelles**" *Solution Behavior of Surfactants* (1982): 373–388. doi:10.1007/978-1-4613-3491-0_19, Available at http://link.springer.com/chapter/10.1007%2F978-1-4613-3491-0_19

60. Yaginuma, H., Kawai, S., Tabata, K. V, Tomiyama, K., Kakizuka, A., Komatsuzaki, T., Noji, H., and Imamura, H. "**Diversity in ATP Concentrations in a Single Bacterial Cell Population Revealed by Quantitative Single-Cell Imaging**" *Scientific Reports* 4, no. 6522 (2014): 1–7. doi:10.1038/srep06522, Available at http://www.ncbi.nlm.nih.gov/pubmed/25283467

61. Humphrey, W., Dalke, A., and Schulten, K. "**VMD: Visual Molecular Dynamics.**" *Journal of molecular graphics* 14, no. 1 (1996): 33–8, 27–8. doi:10.1016/0263-7855(96)00018-5, Available at http://www.ncbi.nlm.nih.gov/pubmed/8744570

62. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kalé, L., and Schulten, K. "**Scalable Molecular Dynamics with NAMD.**" *Journal of computational chemistry* 26, no. 16 (2005): 1781–802. doi:10.1002/jcc.20289, Available at http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2486339&tool=pmcentrez&rendertype=abstract

63. Best, R., Zhu, X., Shim, J., Lopes, P. E. M., Mittal, J., Feig, M., and MacKerell, A. D. "**Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone Φ, Ψ and Side-Chain χ1 and χ2 Dihedral Angles**" *Journal of chemical ...* 8, no. 9 (2012): 3257–3273. doi:10.1021/ct300400x, Available at http://pubs.acs.org/doi/abs/10.1021/ct300400x

64. MacKerell, A. and Bashford, D. "**All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins**" *The Journal of ...* 5647, no. 97 (1998): 3586–3616. doi:10.1021/jp973084f, Available at http://pubs.acs.org/doi/abs/10.1021/jp973084f

65. MacKerell, A. D., Feig, M., and Brooks, C. L. "**SUPPORTING INFORMATION: Improved Treatment of the Protein Backbone in Empirical Force Fields.**" *Journal of the American Chemical Society* 126, (2004): 698–9. doi:10.1021/ja036959e, Available at http://www.ncbi.nlm.nih.gov/pubmed/14733527

66. Buck, M., Bouguet-Bonnet, S., Pastor, R. W., and MacKerell, A. D. "**Importance of the CMAP Correction to the CHARMM22 Protein Force Field: Dynamics of Hen Lysozyme.**" *Biophysical journal* 90, (2006): L36–L38. doi:10.1529/biophysj.105.078154, Available at http://www.ncbi.nlm.nih.gov/pubmed/16361340

67. Beglov, D., Roux, B., and Hc, C. "**Finite Representation of an Infinite for Computer Simulations Bulk System : Solvent Boundary Potential**" *Journal of Medical Physics* 100, no. 1994 (1994): 9050–9063. doi:10.1063/1.466711, Available at http://link.aip.org/link/?JCPSA6/100/9050/1

68. Steffen, C., Thomas, K., Huniar, U., Hellweg, A., Rubner, O., and Schroer, A. "**CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields**" *Journal of computational chemistry* 31, no. 4 (2010): 2967–2970. doi:10.1002/jcc, Available at http://www.ncbi.nlm.nih.gov/pubmed/19575467

69. Vanommeslaeghe, K. and MacKerell, a. D. "**Automation of the CHARMM General Force Field (CGenFF) I: Bond Perception and Atom Typing**" *Journal of Chemical Information and Modeling* 52, no. 1 (2012): 3144–3154. doi:10.1021/ci300363c, Available at http://www.ncbi.nlm.nih.gov/pubmed/23146088

70. Vanommeslaeghe, K. and MacKerell, a. D. "**Automation of the CHARMM General Force Field (CGenFF) II: Assignment of Bonded Parameters and Partial Atomic Charges**" *Journal of Chemical Information and Modeling* 52, (2012): 3144–3154. doi:10.1021/ci300363c, Available at http://www.ncbi.nlm.nih.gov/pubmed/23145473

71. Bilwes, a M., Quezada, C. M., Croal, L. R., Crane, B. R., and Simon, M. I. "**Nucleotide Binding by the Histidine Kinase CheA.**" *Nature structural biology* 8, no. 4 (2001): 353–360. doi:10.1038/86243, Available at http://www.ncbi.nlm.nih.gov/pubmed/11276258

72. Edgar, R. C. "**MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput.**" *Nucleic acids research* 32, no. 5 (2004): 1792–7. doi:10.1093/nar/gkh340, Available at http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=390337&tool=pmcentrez&rendertype=abstract

73. Edgar, R. C. "**MUSCLE: A Multiple Sequence Alignment Method with Reduced Time and Space Complexity.**" *BMC bioinformatics* 5, (2004): 113. doi:10.1186/1471-2105-5-113, Available at http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=517706&tool=pmcentrez&rendertype=abstract

74. Müller, T., Spang, R., and Vingron, M. "**Estimating Amino Acid Substitution Models: A Comparison of Dayhoff's Estimator, the Resolvent Approach and a Maximum Likelihood Method**" *Molecular biology and evolution* 19, no. 1 (2002): 8–13. doi:10.1093/oxfordjournals.molbev.a003985, Available at http://mbe.oxfordjournals.org/content/19/1/8.short

75. Gonnet, G., Cohen, M., and Benner, S. "**Exhaustive Matching of the Entire Protein Sequence Database**" *Science* 256, no. 5062 (1992): 1443–1445. doi:10.1126/science.1604319, Available at http://www.sciencemag.org/cgi/doi/10.1126/science.1604319

76. Backeljau, T., Bruyn, L. De, Wolxt, H. De, Dongen, S. Van, Winnepenninckx, B., and Jordaens, K. "**Multiple UPGMA and Neighbor-Joining Trees and the Performance of Some Computer Packages**" *Molecular biology and evolution* 13, no. 2 (1996): 309–313. doi:10.1093/oxfordjournals.molbev.a025590, Available at http://mbe.library.arizona.edu/data/1996/1302/2back.pdf

77. Kimura, M. and Weiss, G. "**The Stepping Stone Model of Population Structure and the Decrease of Genetic Correlation with Distance**" *Genetics* 49, no. 480 (1964): 561–576. doi:10.1093/oxfordjournals.molbev.a025590, Available at http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1210594/

78. Walle, I. Van, Lasters, I., and Wyns, L. "**Align-M--a New Algorithm for Multiple Alignment of Highly Divergent Sequences.**" *Bioinformatics (Oxford, England)* 20, no. 9 (2004): 1428–35. doi:10.1093/bioinformatics/bth116, Available at http://www.ncbi.nlm.nih.gov/pubmed/14962914

79. Brenner, S. E., Koehl, P., and Levitt, M. "**The ASTRAL Compendium for Protein Structure and Sequence Analysis**" *Nucleic acids research* 28, no. 1 (2000): 254–256. doi:doi: 10.1093/nar/28.1.254, Available at http://www.ncbi.nlm.nih.gov/pmc/articles/PMC102434/

80. Holm, L. and Sander, C. "**Touring Protein Fold Space with Dali/FSSP**" *Nucleic acids research* 26, no. 1 (1998): 316–319. doi:10.1093/nar/26.1.316, Available at http://www.ncbi.nlm.nih.gov/pubmed/9399863

81. Miller, A. S., Kohout, S. C., Gilman, K. A., and Falke, J. J. "**CheA Kinase of Bacterial Chemotaxis : Chemical Mapping of Four Essential Docking Sites**" *Biochemistry* 45, no. 29 (2006): 8699–8711. doi:10.1021/bi060580y, Available at http://www.ncbi.nlm.nih.gov/pubmed/16846213

82. Miller, A. and Falke, J. "**Side Chains at the Membrane-Water Interface Modulate the Signaling State of a Transmembrane Receptor**" *Biochemistry* 43, no. 7 (2004): 1763–1770. doi:10.1021/bi0360206, Available at http://pubs.acs.org/doi/full/10.1021/bi0360206

83. Hengen, P. "**Purification of His-Tag Fusion Proteins from Escherichia Coli**" *Trends in biochemical sciences* 20, (1995): 8–9. doi:10.1016/S0968-0004(00)89045-3, Available at http://www.sciencedirect.com/science/article/pii/S0968000400890453

84. Piasta, K. N., Ulliman, C. J., Slivka, P. F., Crane, B. R., and Falke, J. J. "**Defining a Key Receptor–CheA Kinase Contact and Elucidating Its Function in the Membrane-Bound Bacterial Chemosensory Array: A Disulfide Mapping and TAM-IDS Study**" *Biochemistry* 52, no. 22 (2013): 3866–3880. doi:10.1021/bi400385c, Available at http://www.ncbi.nlm.nih.gov/pubmed/23668882

85. Parkinson, J. and Houts, S. "**Isolation and Behavior of Escherichia Coli Deletion Mutants Lacking Chemotaxis Functions.**" *Journal of bacteriology* 151, no. 1 (1982): 106–113. Available at http://jb.asm.org/content/151/1/106.short

86. Li, G. and Weis, R. M. "**Covalent Modification Regulates Ligand Binding to Receptor Complexes in the Chemosensory System of Escherichia Coli.**" *Cell* 100, no. 3 (2000): 357–65. doi:10.1016/S0092-8674(00)80671-6, Available at http://www.ncbi.nlm.nih.gov/pubmed/10676817

87. Smith, P., Krohn, R., and Hermanson, G. "**Measurement of Protein Using Bicinchoninic Acid**" *Analytical biochemistry* 150, no. 1 (1985): 76–85. doi:10.1016/0003-2697(85)90442-7, Available at http://www.sciencedirect.com/science/article/pii/0003269785904427

88. Slivka, P. F. and Falke, J. J. "**Isolated Bacterial Chemosensory Array Possesses Quasi- and Ultrastable Components: Functional Links between Array Stability, Cooperativity, and Order.**" *Biochemistry* 51, no. 51 (2012): 10218–28. doi:10.1021/bi301287h, Available at http://www.ncbi.nlm.nih.gov/pubmed/23186266

89. Erbse, A. H. and Falke, J. J. "**The Core Signaling Proteins of Bacterial Chemotaxis Assemble to Form an Ultrastable Complex.**" *Biochemistry* 48, no. 29 (2009): 6975–87. doi:10.1021/bi900641c, Available at http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2766635&tool=pmcentrez&rendertype=abstract

90. Natale, A., Duplantis, J., Piasta, K., and Falke, J. "**Structure, Function, and On–Off Switching of a Core Unit Contact between CheA Kinase and CheW Adaptor Protein in the Bacterial Chemosensory Array: A Disulfide**" *Biochemistry* 040731, no. 303 (2013): 1–40. doi:10.1021/bi401159k, Available at http://pubs.acs.org/doi/abs/10.1021/bi401159k

91. Moult, J., Fidelis, K., Kryshtafovych, A., and Tramontano, A. "**Critical Assessment of Methods of Protein Structure Prediction (CASP)--Round IX.**" *Proteins* 79 Suppl 1, (2011): 1–5. doi:10.1002/prot.23200, Available at http://www.ncbi.nlm.nih.gov/pubmed/21997831

92. Pollastri, G. and Przybylski, D. "**Improving the Prediction of Protein Secondary Structure in Three and Eight Classes Using Recurrent Neural Networks and Profiles**" *Proteins: Structure, ...* 235, no. July 2001 (2002): 228–235. doi:10.1002/prot.10082, Available at http://onlinelibrary.wiley.com/doi/10.1002/prot.10082/full

93. McGuffin, L. J., Bryson, K., and Jones, D. T. "**The PSIPRED Protein Structure Prediction Server.**" *Bioinformatics (Oxford, England)* 16, no. 4 (2000): 404–405. doi:10.1093/bioinformatics/16.4.404, Available at http://www.ncbi.nlm.nih.gov/pubmed/10869041

94. Cole, C., Barber, J. D., and Barton, G. J. "**The Jpred 3 Secondary Structure Prediction Server.**" *Nucleic acids research* 36, no. Web Server issue (2008): W197–201. doi:10.1093/nar/gkn238, Available at http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2447793&tool=pmcentrez&rendertype=abstract

95. Ward, J. J., McGuffin, L. J., Bryson, K., Buxton, B. F., and Jones, D. T. "**The DISOPRED Server for the Prediction of Protein Disorder.**" *Bioinformatics (Oxford, England)* 20, no. 13 (2004): 2138–9. doi:10.1093/bioinformatics/bth195, Available at http://www.ncbi.nlm.nih.gov/pubmed/15044227

96. Bennett-Lovsey, R. M., Herbet, A. D., Sternberg, M. J. E., and Kelley, L. A. "**Exploring the Extremes of Sequence/structure Space with Ensemble Fold Recognition in the Program Phyre**" *Proteins: Structure, ...* no. April (2008): 611–625. doi:10.1002/prot, Available at http://onlinelibrary.wiley.com/doi/10.1002/prot.21688/full

97. Li, X., Fleetwood, A. D., Bayas, C., Bilwes, A. M., Ortega, D. R., Falke, J. J., Zhulin, I. B., and Crane, B. R. "**The 3.2 Å Resolution Structure of a Receptor:CheA:CheW Signaling Complex Defines Overlapping Binding Sites and Key Residue Interactions within Bacterial Chemosensory Arrays**" *Biochemistry* 52, no. 22 (2013): 3852–3865. doi:10.1021/bi400383e, Available at http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3705729&tool=pmcentrez&rendertype=abstract

98. Briegel, A., Ortega, D. R., Tocheva, E. I., Wuichet, K., Li, Z., Chen, S., Müller, A., Iancu, C. V, Murphy, G. E., Dobro, M. J., Zhulin, I. B., and Jensen, G. J. "**Universal Architecture of Bacterial Chemoreceptor Arrays.**" *Proceedings of the National Academy of Sciences of the United States of America* 106, no. 40 (2009): 17181–6. doi:10.1073/pnas.0905181106, Available at http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2761316&tool=pmcentrez&rendertype=abstract

99. Kelley, L. a and Sternberg, M. J. E. "**Protein Structure Prediction on the Web: A Case Study Using the Phyre Server.**" *Nature protocols* 4, no. 3 (2009): 363–71. doi:10.1038/nprot.2009.2, Available at http://www.ncbi.nlm.nih.gov/pubmed/19247286

100. Levine, B. G., Stone, J. E., and Kohlmeyer, A. "**Fast Analysis of Molecular Dynamics Trajectories with Graphics Processing Units-Radial Distribution Function Histogramming**" *Journal of Computational Physics* 230, no. 9 (2011): 3556–3569. doi:10.1016/j.jcp.2011.01.048, Available at http://dx.doi.org/10.1016/j.jcp.2011.01.048

APPENDIX A

Residue Number	No MVC	K^+	Residue Number	No MVC	K^+
341	2.86	3.06	**387**	0.78	0.80
342	2.84	2.85	**388**	0.92	0.79
343	1.49	1.80	**389**	1.47	1.84
344	1.56	1.98	**390**	0.73	0.58
345	1.30	1.36	**391**	1.13	1.29
346	1.41	1.31	**392**	0.56	0.60
347	1.70	1.78	**393**	0.76	0.63
348	1.86	1.22	**394**	0.83	0.44
349	1.64	0.72	**395**	0.65	0.47
350	1.73	1.08	**396**	2.10	1.71
351	1.75	0.95	**397**	0.63	0.66
352	1.97	0.86	**398**	0.66	0.67
353	1.57	0.93	**399**	0.88	0.88
354	0.98	0.71	**400**	1.23	0.60
355	3.69	1.71	**401**	0.67	0.60
356	1.36	0.98	**402**	0.82	0.91
357	0.72	0.61	**403**	0.95	1.04
358	2.10	2.28	**404**	0.79	1.07
359	2.24	0.96	**405**	1.09	1.93
360	1.56	1.44	**406**	0.84	1.02
361	1.17	0.71	**407**	1.99	1.97

362	1.37	1.09	**408**	1.87	2.09
363	1.90	2.17	**409**	0.71	1.29
364	1.41	0.99	**410**	1.35	1.37
365	1.38	0.91	**411**	1.96	1.72
366	0.94	0.72	**412**	1.05	1.29
367	1.39	1.11	**413**	1.01	1.14
368	1.26	0.63	**414**	0.93	1.25
369	1.66	1.77	**415**	1.80	1.51
370	0.73	0.71	**416**	1.37	1.31
371	0.61	0.97	**417**	0.78	0.87
372	0.75	0.95	**418**	0.54	0.58
373	0.83	1.52	**419**	1.02	1.25
374	0.71	0.88	**420**	0.79	0.68
375	1.17	1.26	**421**	0.75	0.78
376	1.47	1.56	**422**	0.73	1.54
377	1.02	1.24	**423**	1.16	1.14
378	1.31	2.02	**424**	1.09	0.86
379	1.42	1.00	**425**	2.15	1.55
380	1.54	1.28	**426**	2.39	2.66
381	1.77	1.84	**427**	3.45	3.37
382	1.70	1.21	**428**	3.12	3.31
383	1.18	0.90	**429**	2.42	3.19
384	1.52	1.03	**430**	1.81	1.98
385	1.84	2.37	**431**	1.45	1.20
386	3.17	1.94	**432**	1.33	0.61

Residue Number	No MVC	K^+	Residue Number	No MVC	K^+
433	1.09	1.03	**481**	6.12*	2.28
434	1.11	1.39	**482**	4.57*	2.21
435	0.58	0.70	**483**	5.43*	3.01
436	1.00	0.80	**484**	8.96*	3.65
437	0.63	0.66	**485**	6.22*	4.09*
438	0.72	1.04	**486**	4.29*	3.59
439	0.86	0.92	**487**	2.30	2.67
440	1.21	0.88	**488**	1.93	3.08
441	0.85	0.81	**489**	1.34	1.45
442	0.84	1.00	**490**	1.56	1.03
443	1.33	2.06	**491**	1.21	0.81
444	1.35	3.07	**492**	1.64	0.72
445	1.90	2.91	**493**	1.31	0.96
446	1.38	3.75	**494**	1.80	0.66
447	0.99	1.42	**495**	1.20	0.72
448	1.84	2.47	**496**	1.36	1.14
449	1.14	1.63	**497**	2.32	4.13*
450	1.53	2.98	**498**	1.16	1.44
451	1.32	1.70	**499**	1.18	1.22
452	1.73	2.33	**500**	2.33	1.66
453	1.70	1.98	**501**	1.80	1.87
454	2.07	2.50	**502**	2.21	1.18
455	1.44	2.18	**503**	1.07	0.85

456	1.73	3.28	**504**	0.94	0.84
457	1.62	4.60*	**505**	1.52	2.04
458	2.34	3.56	**506**	0.97	1.09
459	2.85	4.44*	**507**	1.87	1.38
460	2.95	6.53*	**508**	1.07	0.76
461	2.76	7.27*	**509**	1.97	1.32
462	2.20	4.24*	**510**	1.77	0.74
463	2.02	3.28	**511**	3.78	2.02
464	1.79	1.96	**512**	4.95*	2.25
465	2.17	3.09	**513**	2.86	1.36
466	2.06	3.38	**514**	1.80	1.40
467	1.05	1.36	**515**	0.87	1.01
468	1.28	1.48	**516**	0.80	0.54
469	3.50	2.16	**517**	0.82	0.65
470	1.55	0.90	**518**	0.55	0.68
471	0.93	1.74	**519**	2.11	2.08
472	1.69	0.83	**520**	1.03	0.73
473	1.55	0.72	**521**	1.18	1.27
474	1.57	0.83	**522**	0.81	0.96
475	1.31	0.98	**523**	1.24	1.43
476	1.34	2.38	**524**	2.58	3.33
477	2.06	0.87	**525**	4.58	4.81
478	4.45*	0.87			
479	2.21	0.98			
480	4.70*	3.00			

APPENDIX B

The results for the apo dynamics trajectory are the left half of the table while the potassium dynamics run results are on the right side of the table.

Distance (Å)	*g(r)*	$\int g(r)$	Distance (Å)	*g(r)*	$\int g(r)$
0.0500	0.0000	0.0000	0.0500	0.0000	0.0000
0.1500	0.0000	0.0000	0.1500	0.0000	0.0000
0.2500	0.0000	0.0000	0.2500	0.0000	0.0000
0.3500	0.0000	0.0000	0.3500	0.0000	0.0000
0.4500	0.0000	0.0000	0.4500	0.0000	0.0000
0.5500	0.0000	0.0000	0.5500	0.0000	0.0000
0.6500	0.0000	0.0000	0.6500	0.0000	0.0000
0.7500	0.0000	0.0000	0.7500	0.0000	0.0000
0.8500	0.0000	0.0000	0.8500	0.0000	0.0000
0.9500	0.0000	0.0000	0.9500	0.0000	0.0000
1.0500	0.0000	0.0000	1.0500	0.0000	0.0000
1.1500	0.0000	0.0000	1.1500	0.0000	0.0000
1.2500	0.0000	0.0000	1.2500	0.0000	0.0000
1.3500	0.0000	0.0000	1.3500	0.0000	0.0000
1.4500	0.0000	0.0000	1.4500	0.0000	0.0000
1.5500	0.0000	0.0000	1.5500	0.0000	0.0000
1.6500	0.0000	0.0000	1.6500	0.0000	0.0000
1.7500	0.0000	0.0000	1.7500	0.0000	0.0000
1.8500	0.0000	0.0000	1.8500	0.0000	0.0000
1.9500	0.0000	0.0000	1.9500	0.0000	0.0000
2.0500	0.0000	0.0000	2.0500	0.0000	0.0000
2.1500	0.0000	0.0000	2.1500	0.0000	0.0000
2.2500	0.0000	0.0000	2.2500	0.0000	0.0000
2.3500	0.0000	0.0000	2.3500	0.0000	0.0000
2.4500	0.0000	0.0000	2.4500	0.0000	0.0000
2.5500	0.0000	0.0000	2.5500	0.0000	0.0000
2.6500	0.0000	0.0000	2.6500	0.0000	0.0000
2.7500	0.0000	0.0000	2.7500	0.0000	0.0000
2.8500	0.0000	0.0000	2.8500	0.0000	0.0000
2.9500	0.0000	0.0000	2.9500	0.0000	0.0000
3.0500	0.0117	0.0043	3.0500	0.0089	0.0030
3.1500	0.1515	0.0640	3.1500	0.0958	0.0373

3.2500	0.8395	0.4159	3.2500	0.4360	0.2036
3.3500	2.2307	1.4095	3.3500	1.0124	0.6138
3.4500	3.3685	3.0010	3.4500	1.6772	1.3346
3.5500	3.6638	4.8337	3.5500	1.8902	2.1946
3.6500	3.1981	6.5248	3.6500	1.7652	3.0437
3.7500	2.4209	7.8760	3.7500	1.3776	3.7431
3.8500	1.5458	8.7854	3.8500	0.9589	4.2562
3.9500	1.0240	9.4195	3.9500	0.6566	4.6261
4.0500	0.6096	9.8164	4.0500	0.4063	4.8667
4.1500	0.3607	10.0630	4.1500	0.2288	5.0090
4.2500	0.2640	10.2522	4.2500	0.1502	5.1070
4.3500	0.2395	10.4322	4.3500	0.1210	5.1896
4.4500	0.2649	10.6405	4.4500	0.1137	5.2709
4.5500	0.3115	10.8964	4.5500	0.1034	5.3482
4.6500	0.3289	11.1786	4.6500	0.1148	5.4379
4.7500	0.3524	11.4942	4.7500	0.1047	5.5232
4.8500	0.4022	11.8697	4.8500	0.1173	5.6228
4.9500	0.4348	12.2926	4.9500	0.1405	5.7471
5.0500	0.4339	12.7318	5.0500	0.1513	5.8864
5.1500	0.4983	13.2562	5.1500	0.1740	6.0530
5.2500	0.5148	13.8194	5.2500	0.2203	6.2722
5.3500	0.5602	14.4558	5.3500	0.2370	6.5172
5.4500	0.6114	15.1766	5.4500	0.2669	6.8034
5.5500	0.6901	16.0203	5.5500	0.3203	7.1596
5.6500	0.7572	16.9797	5.6500	0.3545	7.5681
5.7500	0.8047	18.0357	5.7500	0.4305	8.0820
5.8500	0.7504	19.0550	5.8500	0.4385	8.6238
5.9500	0.7713	20.1386	5.9500	0.4330	9.1773
6.0500	0.7562	21.2373	6.0500	0.4542	9.7774
6.1500	0.7394	22.3472	6.1500	0.4632	10.4099
6.2500	0.6560	23.3642	6.2500	0.4393	11.0293
6.3500	0.5896	24.3079	6.3500	0.4116	11.6285
6.4500	0.5167	25.1609	6.4500	0.4038	12.2349
6.5500	0.4679	25.9577	6.5500	0.3776	12.8197
6.6500	0.4348	26.7208	6.6500	0.3287	13.3446
6.7500	0.4426	27.5212	6.7500	0.3434	13.9094
6.8500	0.4142	28.2926	6.8500	0.3236	14.4575
6.9500	0.4137	29.0856	6.9500	0.3059	14.9910
7.0500	0.4117	29.8977	7.0500	0.3233	15.5711
7.1500	0.4323	30.7747	7.1500	0.3421	16.2026
7.2500	0.4345	31.6811	7.2500	0.3660	16.8970

7.3500	0.4405	32.6255	7.3500	0.4049	17.6868
7.4500	0.4564	33.6308	7.4500	0.4334	18.5551
7.5500	0.4656	34.6841	7.5500	0.4725	19.5275
7.6500	0.4598	35.7521	7.6500	0.4984	20.5805
7.7500	0.4683	36.8684	7.7500	0.5325	21.7351
7.8500	0.4792	38.0403	7.8500	0.5659	22.9940
7.9500	0.4884	39.2656	7.9500	0.5774	24.3116
8.0500	0.5155	40.5915	8.0500	0.5910	25.6941
8.1500	0.5173	41.9550	8.1500	0.6145	27.1676
8.2500	0.5118	43.3376	8.2500	0.6455	28.7537
8.3500	0.5217	44.7811	8.3500	0.6497	30.3892
8.4500	0.5201	46.2549	8.4500	0.6175	31.9810
8.5500	0.5233	47.7731	8.5500	0.6280	33.6385
8.6500	0.5253	49.3329	8.6500	0.6483	35.3895
8.7500	0.5362	50.9623	8.7500	0.6380	37.1529
8.8500	0.5122	52.5545	8.8500	0.6312	38.9377
8.9500	0.5201	54.2079	8.9500	0.6272	40.7514
9.0500	0.5150	55.8820	9.0500	0.6441	42.6558
9.1500	0.5202	57.6105	9.1500	0.6254	44.5462
9.2500	0.5364	59.4319	9.2500	0.6029	46.4085
9.3500	0.5404	61.3069	9.3500	0.5960	48.2896
9.4500	0.5662	63.3136	9.4500	0.5992	50.2213
9.5500	0.5626	65.3502	9.5500	0.5814	52.1356
9.6500	0.5877	67.5225	9.6500	0.5920	54.1260
9.7500	0.6019	69.7934	9.7500	0.5805	56.1183
9.8500	0.5924	72.0746	9.8500	0.5823	58.1579
9.9500	0.6074	74.4612	9.9500	0.5978	60.2946
10.0500	0.5896	76.8247	10.0500	0.5993	62.4798