Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Shawn Walsh April 15, 2015

Investigating Nucleotide Interactions of BdcA, a Rossmann-fold containing protein capable of biofilm dispersal

By

Shawn Walsh

Dr. Emily Weinert
Adviser

Department of Chemistry

Dr. Emily Weinert

Adviser

Dr. Stefan Lutz

Committee Member

Dr. Jose Soria

Committee Member

Dr. Leah Roesch

Committee Member

2015

Investigating Nucleotide Interactions of BdcA, a Rossmann-fold containing protein capable of biofilm dispersal

By

Shawn Walsh

Dr. Emily Weinert
Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Sciences with Honors

Department of Chemistry

2015

Abstract

Investigating Nucleotide Interactions of BdcA, a Rossmann-fold containing protein capable of biofilm dispersal

By Shawn Walsh

Biofilms are a major problem in human health because they confer increased resistance to antimicrobial agents and adhere to medical devices. As a potential therapeutic, the protein BdcA was previously engineered to disperse biofilms, and proposed to function via the sequestering of cyclic dimeric GMP (c-di-GMP).[1] This ubiquitous bacterial second messenger is known to control a variety of cellular processes including biofilm formation and dispersal. Unlike previously characterized c-di-GMP receptors, BdcA contains a Rossmann fold, which typically binds NAD(P)(H). We set out to characterize BdcA as the first of a potentially novel class of c-di-GMP receptors. However, we show that BdcA has no affinity for c-di-GMP. Rather, it binds nicotinamide adenine dinucleotide phosphate (NADPH). We have also created seven rational mutants of BdcA, and initiated isothermal titration calorimetry (ITC) and biofilm dispersal assays in order to correlate NADPH interactions with biofilm dispersal, thereby shedding light on the now enigmatic mechanism behind BdcA-mediated biofilm dispersal.

Investigating Nucleotide Interactions of BdcA, a Rossmann-fold containing protein capable of biofilm dispersal

By

Shawn Walsh

Dr. Emily Weinert
Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Sciences with Honors

Department of Chemistry

2015

Table of Contents

Introduction 1

Results and Discussion 4

1. Work based on c-di-GMP as the substrate for BdcA 4
2. Work based on NADPH-specific binding 11

Conclusion 18

Methods 19

1. Genomic Isolation and Cloning 19
2. Site-directed Mutagenesis 19
3. Protein Expression and Purification 20
4. Phosphodiesterase Activity Assay 20
5. HPLC Analysis of nucleotides 21
6. Sequence and Structure Alignments 22
7. Rapid Equilibrium Dialysis (RED) 22
8. Isothermal Titration Calorimetry 24
9. Biofilm Dispersal Assay 25

Works Cited 26

Supplemental Figures and Tables 28

List of Figures and Tables

Figure 1	Nucleotide Structures	2
Figure 2	3-D BdcA Model	3
Figure 3	BdcA Overlaid with 3OP4	6
Figure 4	Phosphodiesterase Assay Chromatograms	8
Figure 5	C-di-GMP Chromatograms in Various Mobile Phases	9
Figure 6	ITC Results Using c-di-GMP	10
Figure 7	BdcA Overlaid with FabG	12
Figure 8	Mutant Summary	13
Figure 9	ITC Results Using NADPH	14
Figure 10	Biofilm Dispersal Assay	17
Figure S1	Sequence Alignment with DALI Hits	28
Figure S2	Sequence Alignment with 3SJ7 and 3OP4	29
Figure S3	Protein Purification Gel	30
Table S1	Primers for Cloning and Mutagenesis	31
Figure S4	Gel Verification of Cloning	32

Introduction

The vast majority of bacteria in nature live and propagate in complex, multi-species communities attached to a solid surface[2]; these bacterial communities are termed biofilms. Characteristic features of biofilms are radically altered gene expression in comparison to planktonic cells and the production of extracellular polymers that encase the bacteria within a protective barrier. Biofilms are notoriously hard to eliminate, and cause problems in a wide range of settings from the fouling of ship hulls to computer-chip malfunction.[3,4] However, the most alarming consequences of biofilms are those related to human health. Biofilms confer resistance to physical and chemical forces including disinfectants, immune system response, and antibiotics.[5,6] Consequentially, they are involved in 80% of all human infections.[7] They are especially problematic in hospital settings, in which they can contaminate vents and medical devices[8], and even increase the chance of horizontal gene transfer and the resulting spread of antibiotic resistance.[9]

The protein BdcA (biofilm dispersal via c-di-GMP) from *E. coli* was recently discovered as a tool for eliminating biofilms.[1] Subsequent engineering of this protein through site-saturated mutagenesis produced a mutant that showed improved biofilm dispersal capabilities. Expression of BdcA with this single mutation, E50Q, almost completely disperses biofilms in *E coli*. In addition, *bdcA* also shows high conservation among genes in a number of species. When expressed in other gram-negative bacteria, such as *Psuedomonas aeruginosa, P. fluorescens*, and *Rhizobium meliloti*, BdcA retains its ability to disperse biofilms.[10] These findings highlight the potential of BdcA alone to effectively eliminate biofilms formed by a variety of gram-negative bacteria.

BdcA was initially proposed to eliminate biofilms through the sequestering and consequent reduction of the effective concentration of cyclic diguanylate (c-di-GMP) (Figure 1A), which is a ubiquitous small-molecule regulator of biofilms in bacteria. Binding affinity was measured by incubating c-di-GMP with BdcA, separating free c-di-GMP and BdcA-bound c-di-GMP via spin filtering, recovering bound c-di-GMP using a trypsin digest, and finally quantifying c-di-GMP with HPLC. These binding experiments established the binding affinity for BdcA for c-di-GMP as 11.7 μM and showed that the E50Q mutation, which disperses biofilms better than wild-type BdcA, had a higher affinity for c-di-GMP, further supporting the correlation between c-di-GMP sequestering and biofilm dispersal.[1]

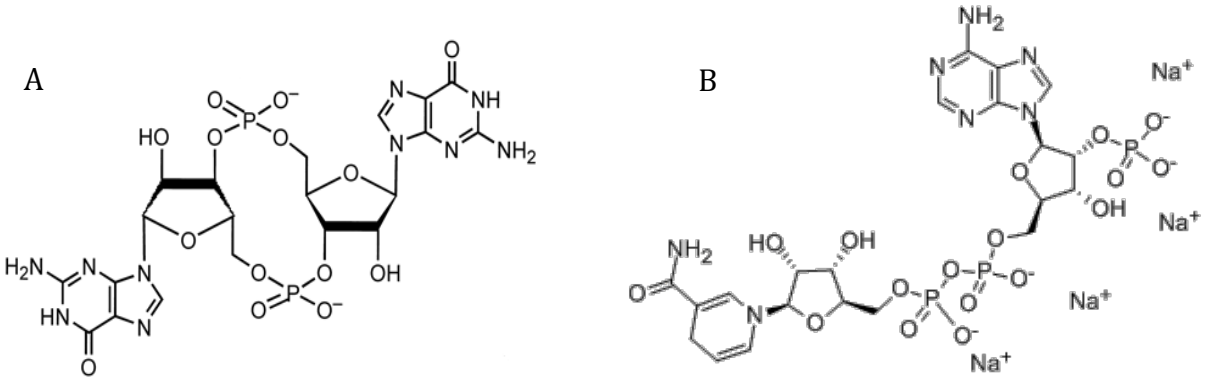


Figure 1. Structures of (A) c-di-GMP and (B) NADPH

Several residues that could be responsible for c-di-GMP binding were identified (Figure 2). Specifically, BdcA contains an EAL sequence, which is necessary within phosphodiesterases that break down c-di-GMP into 5'-pGpG. In addition, some proteins with a catalytically inactive EAL domain function as c-di-GMP receptors. However, it seemed unlikely that this EAL sequence was part of a biologically relevant EAL domain. The motif was found on the exterior of the protein rather than the core, which is inconsistent with other EAL phosphodiesterase structures.[11] Other potentially relevant residues identified in previous work were based on conserved motifs of c-di-GMP binding.[1,12]

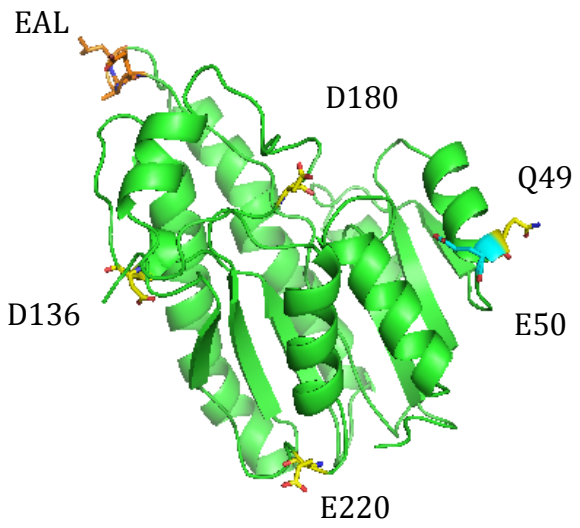


Figure 2. 3-D model of BdcA. The EAL sequence, E50, and other potentially important residues identified are highlighted in orange, blue, and yellow respectively.

Surprisingly, subsequent work from the same group identified BdcA as a short-chain dehydrogenase/reductase (SDR) that binds nicotinamide adenine dinucleotide phosphate (NADPH) (Figure 1B) specifically, rather than c-di-GMP (the authors could not detect c-di-GMP binding using isothermal titration calorimetry (ITC)), which discredited the established hypothesis.[13] As BdcA binding to c-di-GMP has now been refuted, the mechanism of BdcA-dependent biofilm dispersal has become an enigma. There are surprisingly few examples of SDRs that have been shown to influence any kind of quorum sensing phenotypes.[14] Dissecting correlations between NADPH dependence of BdcA on biofilm dispersal would provide valuable insights into its mechanism and potentially validate a link between these two disparate fields.

In this work, we initially created seven mutants to investigate c-di-GMP binding to BdcA. This protein contain a Rossmann fold, which was entirely new in a c-di-GMP receptor and suggested that BdcA might employ a previously uncharacterized mechanism of c-di-

GMP binding. The Rossmann fold is the common motif of SDRs that typically binds to NAD(P)(H) and, although the nearly 47,000 distinct SDRs known show low sequence similarity[14], the Rossmann fold can be identified by its characteristic structure; a central beta sheet with 6-7 strands flanked on each side by 3-4 alpha-helices.[15] The Rossmann fold contains an integral glycine-rich region for binding the pyrophosphate moiety of NADPH. A conserved Asn-Ser-Tyr-Lys tetrad is responsible for oxidoreductase activity. Following the publication of the second paper refuting c-di-GMP binding, the panel of mutants has been used to investigate determinants of NADPH binding and anti-biofilm activity.

Our results validate NADPH binding to BdcA and the lack of c-di-GMP affinity. Furthermore, we hope to shed light on the mechanism through which BdcA exerts its remarkable ability to disperse biofilms, which until recently had a seemingly simple explanation. The rational mutations that were incorporated within BdcA are in prime positions to establish the connection between NADPH binding affinity, potential oxidoreductase ability, and biofilm dispersal.

Results and Discussion

1. Work based on c-di-GMP as the substrate for BdcA

Initially, we focused on characterizing BdcA as a novel receptor for c-di-GMP due to the report of BdcA as a novel c-di-GMP binding protein.[1] Although several conserved binding domains of c-di-GMP are well known, the search for novel effectors that may help explain the complexity of c-di-GMP signaling is ongoing. Prior to work showing that BdcA bound c-di-GMP,[1] there were no examples of c-di-GMP receptors that contained a Rossmann fold. We reasoned that BdcA might be the first identified protein within a novel

class of c-di-GMP receptors, and that by understanding how c-di-GMP bound to BdcA we could identify other new receptors with a similar binding domain.

Towards this end, we performed a homology search to identify potential homologs based on structure. We choose to focus on structural homologs, as opposed to identifying proteins based on sequence similarity, for two reasons. First, proteins containing a Rossmann fold are notorious for exhibiting low sequence similarity (typically 20-30%)[16], whereas they can be quite structurally similar.[14] Second, validating c-di-GMP binding in several proteins with structural similarity to BdcA but a low sequence homology would provide valuable insight into the essential residues of the c-di-GMP binding domain involved.

Lacking a crystal structure of BdcA, we created a 3-D model using an online structure homology-modeling server, SWISS-MODEL, which automatically identified a dehydrogenase/reductase from *Sinorhizobium meliloti* 1021 (PBD ID: 3V2G) as a template.[17–20] We then used this model to search for structural homologs on the Protein Data Bank (PDB) using the Dali Network Service.[21] A sequence alignment with BdcA for the top results of this search is shown in Figure S1. Based on these results, it seemed unlikely that we would be able to identify promising candidates to test for c-di-GMP affinity because none of the hits contained the EAL sequence motif or the small collection of residues that the previous group to work with BdcA had identified as possibly being involved in c-di-GMP binding. Specifically, these residues were D136, D180, Q49, and E220.[1]

Without clear c-di-GMP binding homologs that could be expressed for direct comparison to BdcA, we decided to begin probing c-di-GMP binding through the use of mutagenesis. We created the mutations E90A, D180A, and D136T to investigate whether

these residues were really involved in c-di-GMP binding. In addition, we recreated the E50Q mutant, which had been previously shown to affect c-di-GMP binding and biofilm dispersal.[1] Departing from previous work, it seemed that the residues within the Rossmann fold that had known roles in NAD(P)(H) binding were also attractive candidates for mutagenesis. NADH has a structure similar to c-di-GMP (Figure 1), especially in that they both contain a purine-sugar-phosphate moiety. This suggested that c-di-GMP might bind within the Rossmann fold itself. With this possibility in mind, we aligned BdcA with the crystal structure of a 3-ketoacyl reductase (PBD ID: 3OP4) in complex with NADP+, which DALI had identified as a potential homolog (Figure 3).

A B

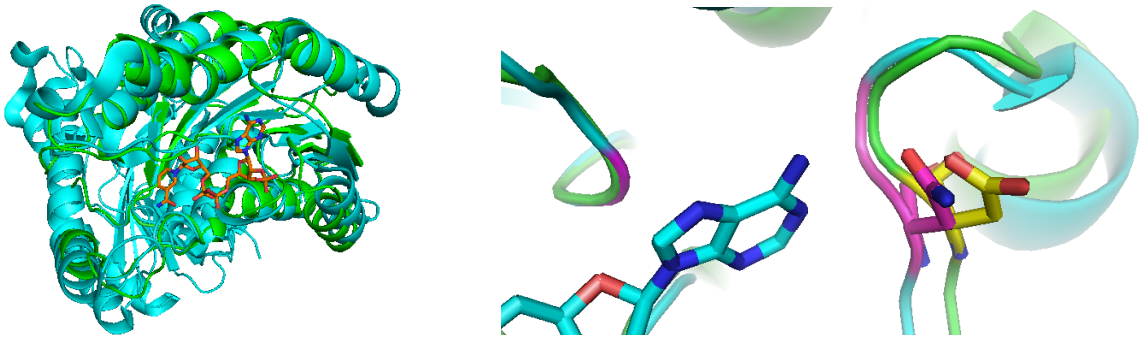


Figure 3. (A) BdcA (green) overlaid with 3OP4 (cyan). (B) Close-up of N63 (yellow) on 3OP4 and D59 (magenta) on BdcA next to the purine of NADP+

In 3OP4, the Asparagine at position 63 (Figure 3B, magenta) is involved in hydrogen bonding to the amine group of the purine ring on NADP+. When aligned with 3OP4, BdcA shows an aspartic acid (Figure 3B, yellow) in roughly the same location, a residue that is also capable of hydrogen bonding. Therefore, we created the mutation D59A based on the rationale that it would disrupt c-di-GMP binding if c-di-GMP bound to BdcA with a similar orientation to NADPH binding within the Rossmann fold. Finally, we created the mutations

G14V and K150I. The former is within a well-established glycine-rich region of Rossmann fold proteins that is involved in NADPH binding, and the mutation was designed to disrupt c-di-GMP binding if binding occurred in a manner similar to NADP+ binding to 3OP4. The K150I mutation was incorporated because the residue is essential within the catalytic tetrad required for the redox chemistry of Rossmann fold proteins. A representative gel of the protein purification can be found in S3.

Before conducting binding studies of the BdcA mutants, we wanted to verify that BdcA showed no catalytic activity towards c-di-GMP. The protein carries an EAL sequence motif, which is characteristic of phosphodiesterases that cleave c-di-GMP specifically into 5' pGpG. Catalytic activity of EAL proteins depends on coordination of Mg^{2+} or Mn^{2+} ions.[22] EAL domains that lack metal coordination sites are catalytically inactive and can function as c-di-GMP effectors.[23]

Although previous work reported that BdcA lacked phosphodiesterase activity, these experiments were carried out in solution with EDTA.[1] Therefore, lack of c-di-GMP turnover in these experiments could plausibly have been a result of catalytically inactive protein or simply the sequestering of metal ions required for catalysis. Therefore, the first step in characterization of BdcA was testing the phosphodiesterase activity.

The BdcA gene was cloned from the genomic DNA of *E. coli* K12 MG1655 strain. Cloning of the gene into the pET28a vector was verified by agarose gel electrophoresis (Fig. S3) and sequencing. Protein expression and purification were performed as described in the Methods section; a representative gel from the purification can be found in Fig. S4.

We observed no catalytic activity of BdcA towards c-di-GMP, even in the absence of metal chelators (Figure 4). Unfortunately, the chromatograms from this experiment

showed an erratic baseline that made accurate c-di-GMP quantification impossible. It is interesting to note that the chromatograms we obtained resembled those from previous phosphodiesterase activity assays for BdcA.[1] In light of the second study that did not observe c-di-GMP binding, the results based on HPLC chromatograms from the Ma, *et al* paper also may have been artifacts.

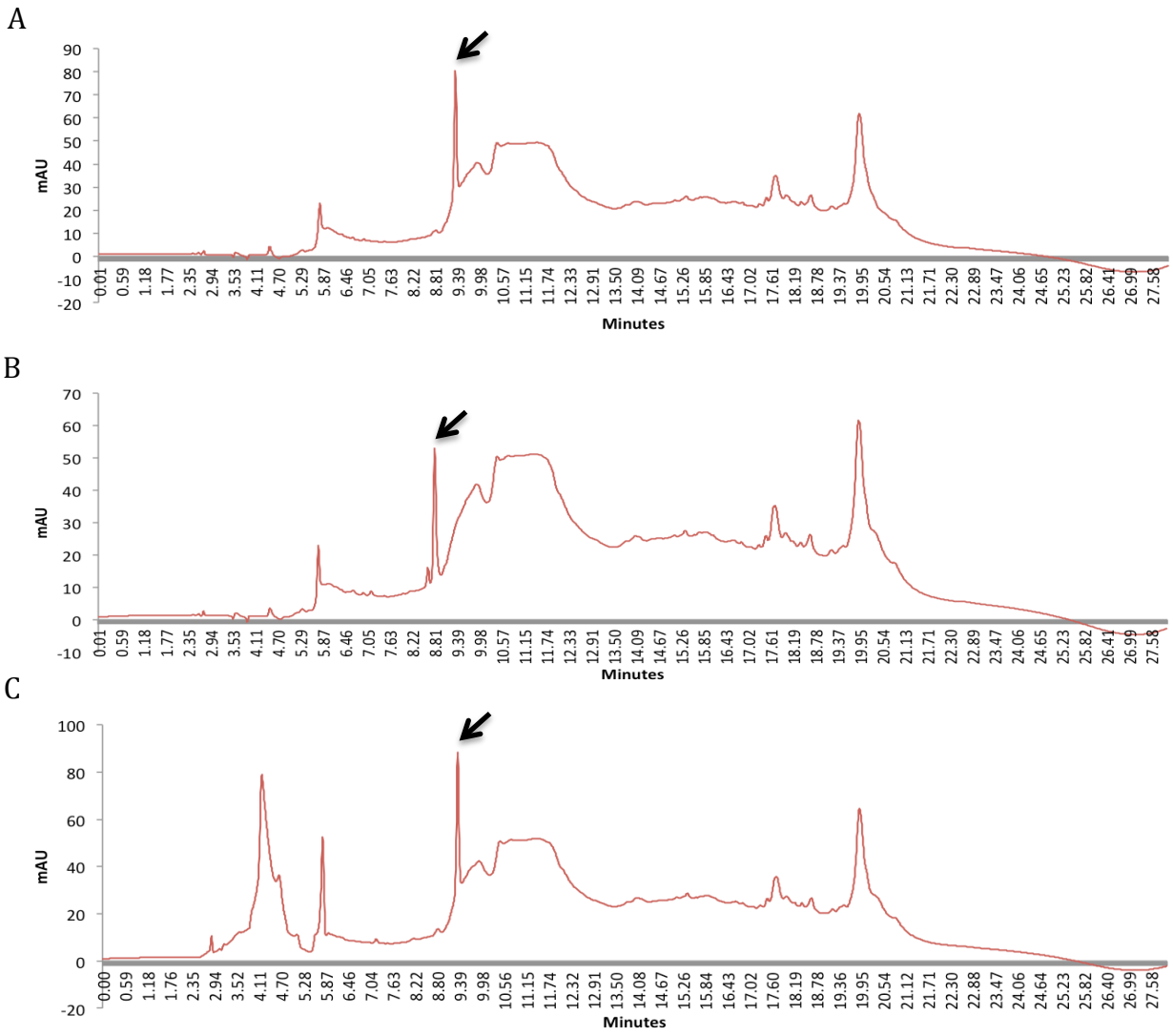


Figure 4. Chromatograms taken at 250nm for (A) pGpG standard, (B) c-di-GMP incubated with BdcA under PDE assay conditions, and (C) c-di-GMP standards reveal that no c-di-GMP degradation occurred in the PDE assay. C-di-GMP eluted at 9.34 minutes and pGpG at 8.81 minutes (indicated by arrows).

Despite several attempts to remove apparent contamination from the system, including several methods of washing, several new mobile phase combinations, and a new column, problems with accurately quantifying c-di-GMP persisted (Figure 5). These problems may have arisen in part due to contaminants in a failing HPLC seal.

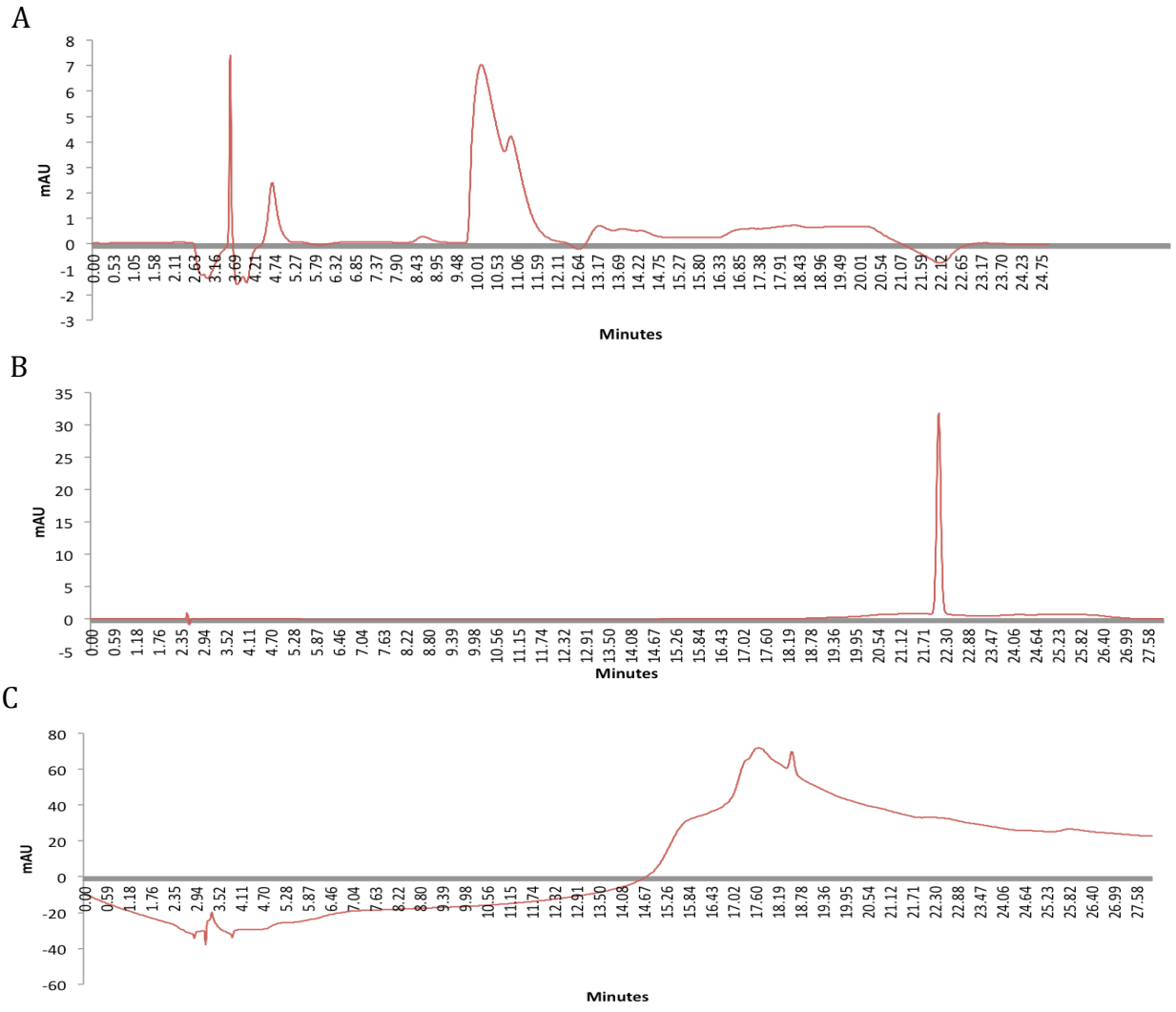


Figure 5. C-di-GMP chromatograms measured at 250nm under various conditions. (A) Mobile phases of methanol with 0.1% formic acid (buffer A) and water with 0.1% formic acid (buffer B) lead to reduced absorbance for c-di-GMP samples (10.45 minutes) that makes quantification imprecise, especially for low concentrations. (B) Mobile phases of 0.1M KH_2PO_4, pH=6 (buffer A) and 90% buffer A with 10% methanol (buffer B) showed a c-di-GMP peak at 22.13 minutes, but lead to problems with the HPLC. Specifically, pressure spiked and rose above tolerable limits. (C) Mobile phases of 0.1M triethyl amine, pH=6 (buffer A) and 50% buffer A with 50% methanol led to an uneven baseline, although a small peak could be seen for c-di-GMP at 18.6 minutes.

The initial binding studies of mutants using rapid equilibrium dialysis (RED) and analysis via HPLC were inconclusive because of these problems reliably quantifying c-di-GMP. In retrospect, these issues were likely exacerbated due to the lack of c-di-GMP binding by BdcA. However, as the only report in the literature at the time had quantified c-di-GMP binding to BdcA with a K_d of ~12 μM,[1] the lack of observed c-di-GMP binding was attributed to issues with the HPLC. Therefore, to sidestep potential HPLC issues, c-di-GMP was quantified using UV-Vis analysis following RED. Surprisingly, these experiments showed that BdcA had no affinity for c-di-GMP (data not shown). In order to verify these results, c-di-GMP affinity was also measured via isothermal titration calorimetry (ITC, Figure 6). ITC was chosen because it measures binding directly by quantifying heat of binding, as opposed to techniques that correlate a measurement of ligand concentration to binding affinity, and thereby reduces the potential for other factors of an experimental setup to skew the results.

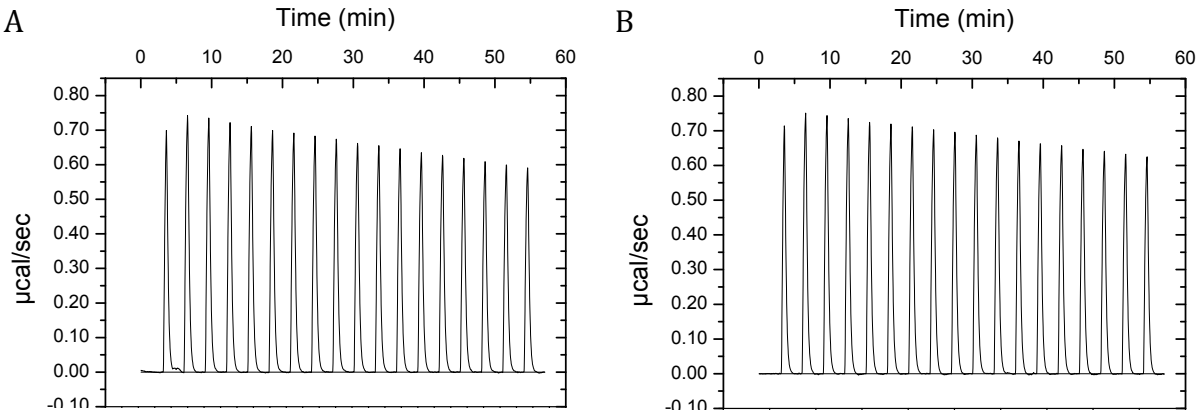


Figure 6. Raw ITC data for BdcA E50Q titrated with c-di-GMP (A) and buffer (B) are equivalent.

Contrary to previous results, ITC also showed that BdcA lacked affinity for c-di-GMP. Testing under a variety of conditions, including the use of high-salt buffers, produced the same result. Although our efforts were aimed toward characterizing a novel c-di-GMP

receptor, our data suggested that BdcA has no affinity for c-di-GMP and is not a novel c-di-GMP receptor.

2. Work based on NADPH-specific binding

A second paper characterizing BdcA recently reported binding to NADPH specifically, with no observable affinity of BdcA for c-di-GMP.[13] The lack of c-di-GMP affinity corroborated with our own results, and in light of these new findings we were forced to reevaluate our aims. Fortunately, some mutants we created to probe c-di-GMP binding were also relevant to exploring the correlation between NADPH interactions and biofilm dispersal mediated by BdcA. Since c-di-GMP sequestering is no longer a valid hypothesis for BdcA's biofilm dispersal capabilities, we turned our attention towards understanding this phenomenon.

Among the proteins identified by DALI during the bioinformatics analysis, a β-ketoacetyl-CoA reductase from *Staphylococcus aureus* termed FabG (PBD: 3SJ7) that has been shown to bind to NADPH shares a 32% sequence identity with BdcA.[24] Alignment of one asymmetric unit of FabG with the BdcA model shows how NADPH might bind to BdcA (Figure 7A).

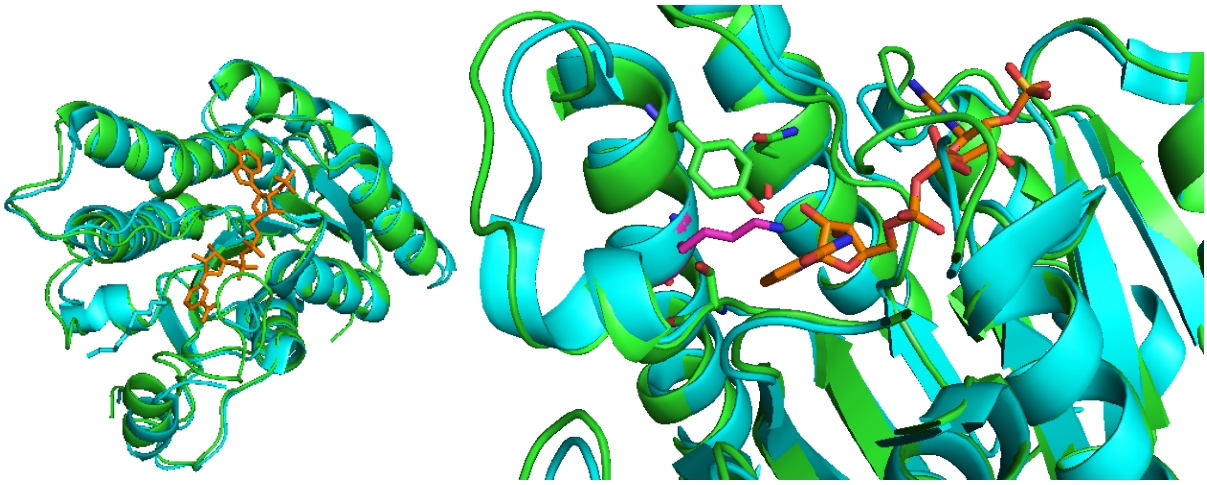


Figure 7. BdcA model (green) aligned with FabG asymmetric unit (cyan). NADPH is shown in orange. (A) Overall protein structures align to a high degree. (B) The lysine at position 150 (magenta) in BdcA is in prime position to engage in catalytic activity. Side chains for the rest of the potentially catalytic triad are shown in green

FabG oxidizes NADPH using the characteristic Asn-Ser-Tyr-Lys catalytic tetrad of proteins that contain a Rossman fold.[24] The model of BdcA contains these same residues in the same positions when it is structurally aligned with FabG (Figure 7B). The catalytic role of these residues was explored over twenty years ago, and it is well known that mutating either the tyrosine or lysine residues totally abolishes catalytic activity in otherwise active proteins, except for rare substitutions that exhibit a fraction of the original activity (substitution with cysteine or arginine).[25] Therefore, the K150I mutation should have no catalytic activity towards NADPH, whether or not WT BdcA does. Any difference in biofilm dispersal between this mutant and WT BdcA would provide evidence as to the role of catalysis in biofilm dispersal. The placement of the other mutations that were generated on BdcA are shown in Figure 8, in which BdcA was aligned with FabG bound to NADPH and then FabG was hidden.

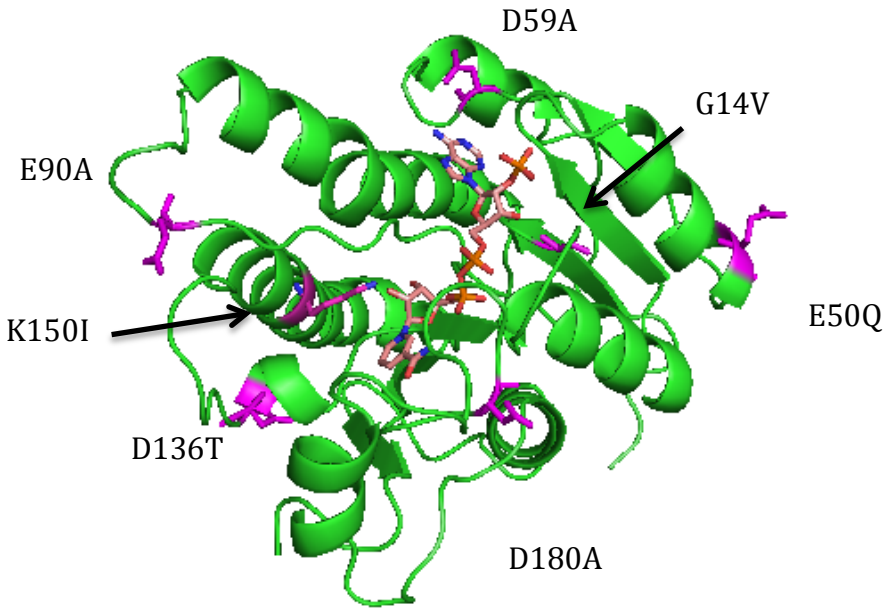


Figure 8. Model of BdcA aligned with FabG (hidden) complexed to NADPH with mutated residues highlighted in magenta

It is interesting to note that the glutamic acid residue at position 50 is extremely distal to the likely NADPH binding pocket, but that mutating this residue to glutamine in the previous Ma, *et al* study dramatically improved BdcA's ability to disperse biofilms.[1] Although E50 may affect protein dynamics or stability, leading to changes in biofilm formation due to altered BdcA-dependent catalysis, it also raises the possibility that BdcA may somehow disperse biofilms independently of its interaction with NADPH.

ITC was used to determine the NADPH affinity of BdcA. Previously, this technique was used to establish a K_d of 25.9±4.1 μM for WT BdcA.[13] Our preliminary data suggest a value for NADPH binding that falls within the error range of this literature value. In a single trial, WT BdcA was found to bind NADPH with a K_d of 27.5±2.6 μM (Figure 9A). Further trials are necessary to determine the quality of this data, but as a qualitative measure it is sufficient to verify NADPH binding.

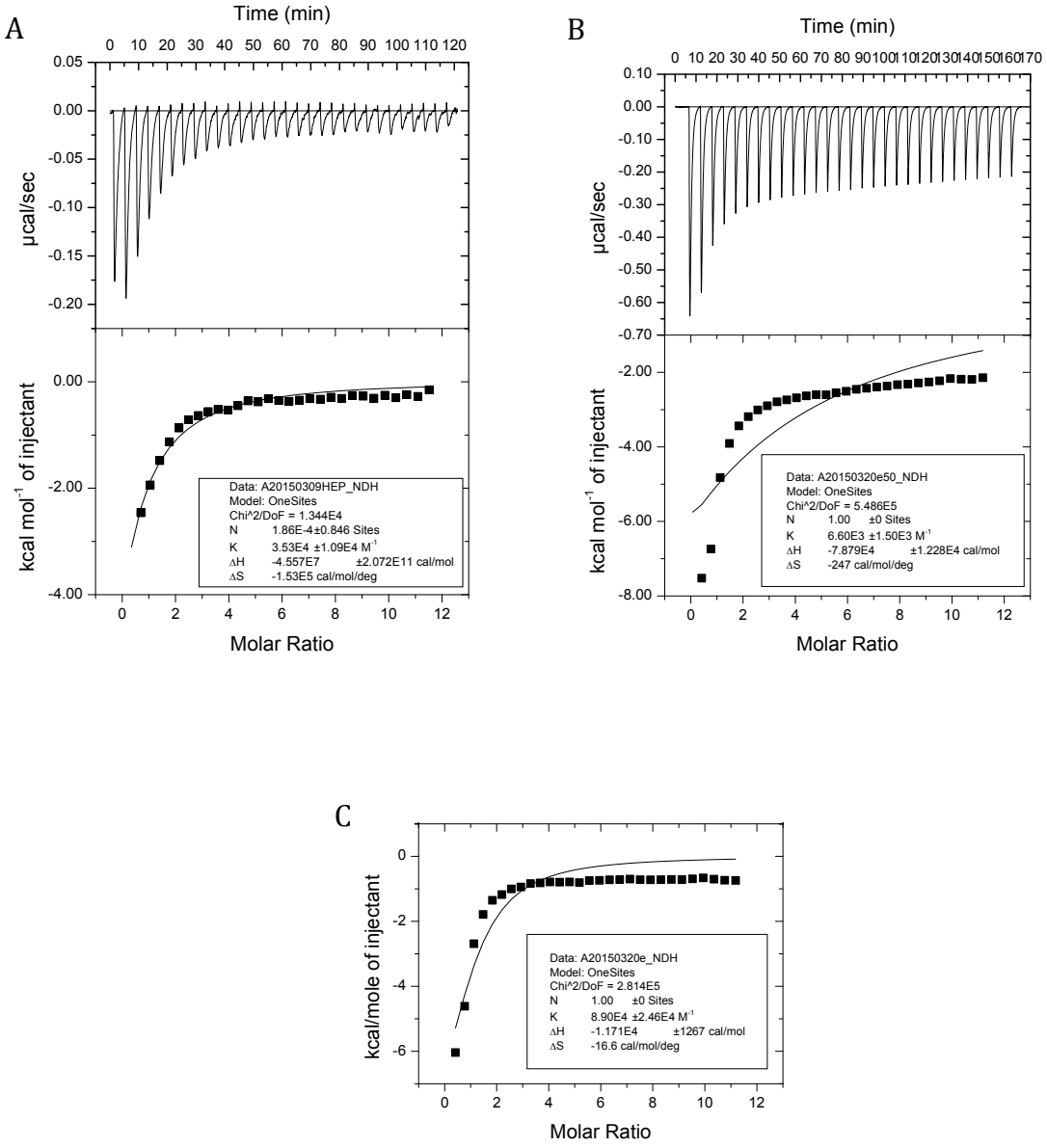


Figure 9. Binding isotherms for (A) WT BdcA, (B) Bdca E50Q, and (C) BdcA E50Q with reference data from the negative control subtracted. Because it is constructed by subtracting one set of data from another, there is no raw binding isotherm to show.

Although BdcA was engineered to better disperse biofilms by means of an E50Q mutation, it is unknown how this mutation affects cofactor binding. Any differences in NADPH affinity for BdcA E50Q compared to WT BdcA would have implications for the method by which BdcA disperses biofilms. Initials trials to determine BdcA E50Q affinity

for NADPH are inconclusive (Figure 9). ITC experiments thus far have yielded binding isotherms that exhibit a high heat of dilution, which manifests in peaks that level off well below 0 kcal per mol of injectant. In other words, injections of NADPH still produce relatively large amounts of heat even after BdcA has become saturated or almost completely saturated with NADPH. This makes the data difficult to fit, because the model for data fitting is built on the assumption that the addition of ligand into a solution of fully saturated protein will results in a negligible change in the heat content of the system. As a result, the fitting in Figure 9B is much worse than 9A, which have Chi^2 values that differ by more than an order of magnitude. Varying the time between injections from 330 seconds to 200 seconds resulted in no significant difference in data fitting (data not shown), further strengthening our hypothesis that the ITC issues are due to heats of dilution.

Fitting for the E50Q binding isotherm can be marginally improved by subtracting the reference data of buffer titrated into protein without NADPH (Figure 9C). The resulting isotherm yields a K_d value of 11.2±3.1 μM, which is a promising result for correlating cofactor binding and biofilm dispersal. However, several issues suggest that this value could be inaccurate. The Chi^2 still differs substantially from data fitting to the WT BdcA binding isotherm in Figure 9A. In addition, it is unsettling that Figure 9C is the result of a binding isotherm from which reference data has been subtracted, and no reference data has been subtracted from 9A. Consequent ITC trials all resulted in heats of dilution comparable to those obtained for E50Q, suggesting that the single binding isotherm for WT BdcA used to calculate its affinity for NADPH might be anomalous (data not shown), possibly due to better buffer matching between BdcA and NADPH in the first ITC trial.

In future experiments, it will be worthwhile devote effort to standardizing the technique and obtaining reliable values for the heat of dilution. For figure 9C, reference data used resulted from the titration of buffer into protein, but it will also be worthwhile to determine the effect of NADPH titration into a solution of buffer. Finally, reducing the concentration of titrated NADPH may help to resolve the issue, because a greater portion of peaks in the binding isotherm will show heat release that is truly dependent on binding rather than dilution effects.

The E50Q mutant of BdcA has previously been shown to disperse biofilms better than WT BdcA.[1] However, it is not known how any of the other mutants used in this work affect biofilm dispersal relative to WT BdcA. Any trends between NADPH affinity and biofilm dispersal capability could provide valuable insights into the mechanism of BdcA-dependent biofilm dispersal. Additionally, differences in the ability of the K150I mutant to disperse biofilms would suggest that catalytic ability for NADPH is important in its mode of function.

Expression of BdcA in *E. coli* during biofilm dispersal assays did not affect cell growth compared to uninduced controls (Figure 10A). Beyond this, it would be difficult to extract any conclusions from the data obtained from biofilm dispersal assays. The crystal-violet staining method of biofilm quantification is highly sensitive to small variations in technique. The method involves washing of the assay plate in order to remove planktonic cells and then later washing again to remove crystal violet dye, but washing that is too rigorous will disrupt biofilms within the wells of the plate.

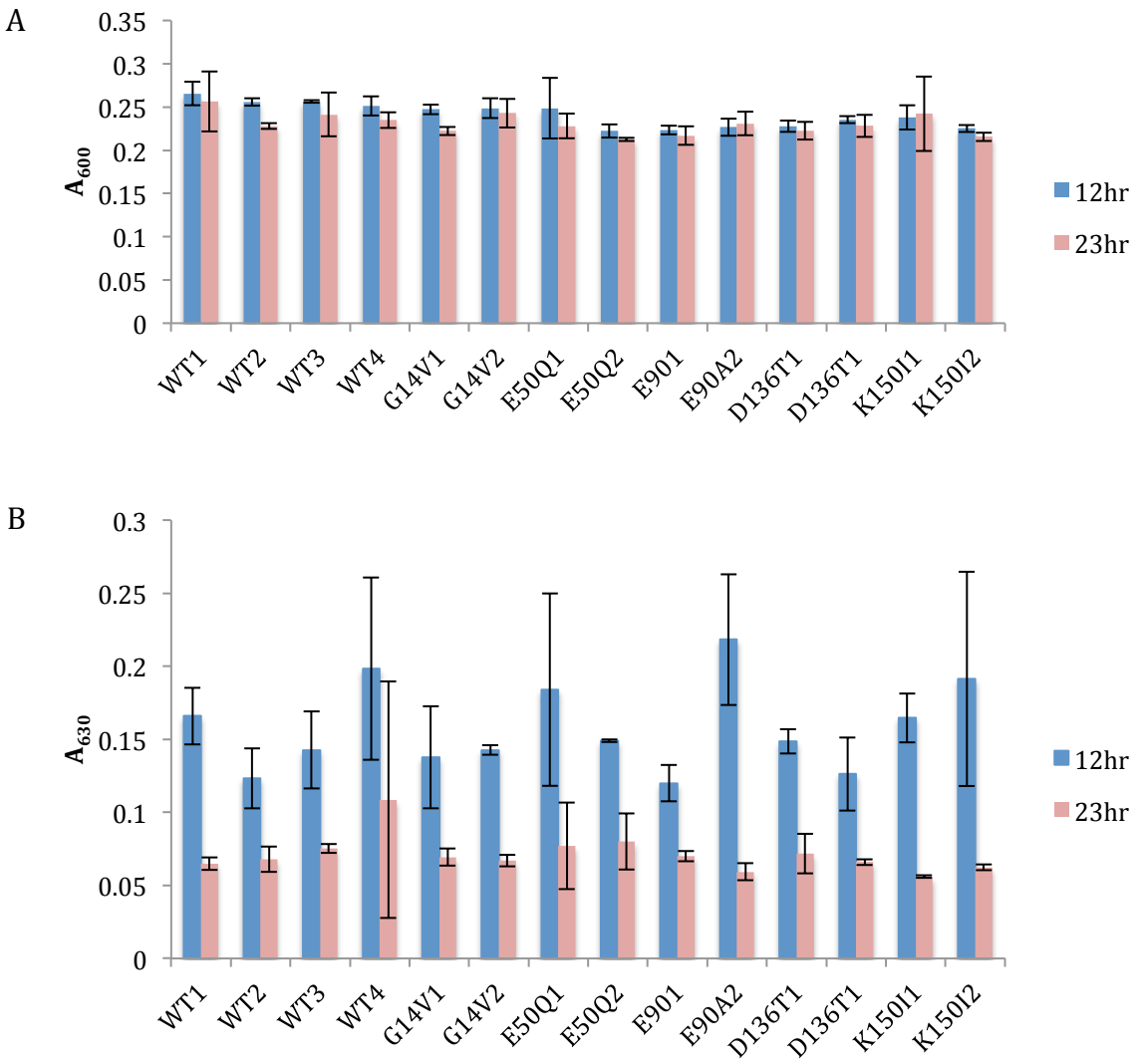


Figure 10. Biofilm dispersal assay. (A) A_{600} readings of biofilm assay wells before staining with crystal violet at 12 and 23 hours after IPTG induction show that BdcA expression has no significant effect on cell density. (B) A_{630} readings of assay wells after staining with crystal violet at 12 and 23 hours post-induction. Results are the average of three wells for each colony, with two colonies used for WT and each mutant. WT3 and WT4 are negative controls that were not induced with IPTG. Errors shown are one standard deviation.

It is likely that the differences observed between 12 and 23 hours post-induction are a result of different amounts of washing rather than a drastic reduction in overall biofilm levels, especially considering that these results reflect biofilm levels from a single trial. In one subsequent experiment, washing was so vigorous that biofilms were removed

completely before crystal staining was carried out (data not shown). In future experiments, the technique will require refinement to reduce error in the data, and multiple trials will be necessary to determine the variation among different washes within otherwise identical samples.

Conclusion

Initially, we focused on investigating BdcA, a protein that disperses biofilms in a variety of gram-negative bacteria,[10] as a novel c-di-GMP receptor. Although the enzymes involved in making and breaking down c-di-GMP are relatively well characterized, less is known about c-di-GMP receptors, and the search for such binding partners is ongoing. BdcA contains a Rossmann fold, which had never been identified in a c-di-GMP receptor, and suggested that it bound to c-di-GMP through an uncharacterized mechanism. Just as other c-di-GMP binding domains are conserved, we hypothesized that the c-di-GMP binding domain within BdcA might also be conserved, and set out to identify potential homologs with a similar binding domain. We used structural similarity rather than sequence identity to identify potential homologs because proteins with a Rossmann fold show low sequence similarity. No obvious homologs emerged, and therefore we turned to mutagenesis to initially probe the determinants of c-di-GMP binding. We found that BdcA did not interact with c-di-GMP at all, either by degrading or binding it, and recent literature corroborated our findings.[13]

We then turned to probing the mechanism by which BdcA disperses biofilms, since the only explanation behind this process was refuted. We demonstrated that BdcA binds to NADPH. Several of the mutants that we created to investigate c-di-GMP binding are relevant to studies with NADPH. Based on comparison of BdcA with other Rossmann fold

proteins, these mutations should interfere with either binding or catalysis specifically. In addition, the mutation E50Q has previously been shown to enhance biofilm dispersal, but its effect on NADPH affinity is unknown.[1] ITC experiments to determine affinity of BdcA mutants are ongoing. Large heats of dilution in mutant isotherms, which render data inconclusive, have hampered current efforts. Assays to correlate NADPH interaction to biofilm dispersal capability were undertaken. Before producing reliable data, the technique by which we quantify biofilms must be improved.

Methods

1. Genomic isolation and cloning

E. coli K-12 MG1655 was cultured overnight and genomic DNA was extracted using the Wizard® Genomic DNA Purification Kit as specified in the protocol supplied by Promega. The primers used to amplify *bdcA* from the genomic DNA via PCR and incorporate restriction sites for BamHI and XhoI are listed in table S1. Amplified DNA and pET28a were digested using BamHI and XhoI. Ligation was carried out using the Fast-Link™ DNA Ligation Kit. Ligation product was transformed into *E. coli* DH5α cells. These cells were cultured overnight and plasmid was isolated using the GenElute™ Plasmid Miniprep Kit according to the protocol from Sigma-Aldrich Co. Cloning success was verified via PCR amplification followed by gel electrophoresis (Figure S3).

2. Site-directed mutagenesis

Mutations were incorporated into the constructed plasmid using the Quickchange® Lightning Site Directed Mutagenesis Kit as specified by the corresponding instruction

manual. DNA sequencing then confirmed successful mutation. Primers used and mutants created are listed in table S1.

3. Protein expression and purification

pET28a plasmid containing the BdcA gene was transformed into *E. coli* Tuner pLysS. Cultures were induced with 0.1mM IPTG after reaching an OD_{600} of 0.6. All growth media was supplemented by 40 μg/mL kanamycin and 30 μg/mL chloramphenicol. Cells were harvested by centrifugation after 24 hours, homogenized, and purified via affinity chromatography using a Ni-NTA column. Purified protein was immediately dialyzed into the appropriate buffer (specific buffers for each experiment are given in subsequent sections), flash frozen, and stored at -80°C. The viability of this protocol was verified via gel electrophoresis (Figure S4).

4. Phosphodiesterase Activity Assay

BdcA that had been dialyzed into PDE Assay Buffer[1] (50 mM Tris, 5 mM $MgCl_2$, 50 mM NaCl, pH=6) immediately following affinity chromatography purification was incubated with 10 μM c-di-GMP (BioLog) for 1 hour. EDTA was added and samples were heated at 95°C for 5 minutes to denature protein. Supernatant was collected and analyzed via HPLC. Traces were compared to c-di-GMP and pGpG standards (BioLog), as well as standards containing a mixture of the two.

5. HPLC analysis of nucleotides

HPLC was performed using an Agilent 1260 Infinity system with photodiode array detector on reverse phase C18 columns. Flow rates were 1 mL/min and samples were run at room temperature. A variety of conditions were used to optimize conditions for quantifying c-di-GMP.

For runs with a supelcosil (Sigma-Aldrich) LC-18 column (25 cm × 4.6 mm, 5 µL particle size): Gradients of buffer A (50 mM triethylamine, pH 5.2 with acetic acid) and buffer B (60% buffer A and 40% acetonitrile) were used as follows. Starting with 100% A, the composition was changed to 65% A at 10 minutes with a constant gradient, 0% A at 16 minutes, 100% A at 22 minutes, and maintained at 100% A until 25 minutes. These specifications were used for the phosphodiesterase assays. Gradients of buffer A (water with 0.1% formic acid) and buffer B (methanol with 0.1% formic acid) were also tested on this column with the following gradients: 100% A at 0 minutes, 98.5% A at 0.1 minutes, 98.5%A at 4 minutes, 98% A at 4.1 minutes, 98% A at 8 minutes, 92% A at 10 minutes, 92% A at 12 minutes, 85% A at 14 minutes, 85% A at 16 minutes, 98.5% A at 18 minutes, and 98.5% A at 25 minutes.

For runs on a Microsorb-MV 100 C18 column: Gradients of buffer A (0.1M KH_2PO_4, pH=6) and buffer B (90%A, 10% methanol) were used as follows: 100% A at 0 minutes, 100% A at 9 minutes (the flow was increased to 1.3 mL/minute at this time point and maintained for the remainder of the run), 75% A at 15 minutes, 10% A at 17.5 minutes, 0% A from 19 to 23 minutes, and 100% A at 24 minutes. In addition, runs with gradients of buffer A (0.1M triethylamine, pH=6) and buffer B (50% A and 50% methanol) were performed with the following specification. Starting with 100% A, the percentage of A was

transitioned to 50% with a constant gradient until 25 minutes, and then transitioned to 100% A until 30 minutes.

6. Sequence and structure alignments

BdcA FASTA sequence was submitted to NCBI's Basic Local Alignment Search Tool (BLAST), specifically using the protein blast and excluding *Escherichia coli* from the search results to obtain superfamily data. The same sequence was submitted to the SWISS-MODEL server using an automatically generated template and an unspecified template.[17–20] Structural homologs were found using the Dali server, with specifications as detailed previously.[21] MacPyMOL was used to align FabG and BdcA and create a variety of corresponding images.

7. Rapid Equilibrium Dialysis (RED)

Single-Use Rapid Equilibrium Dialysis (RED) Plate with Inserts, 8K MWCO (Pierce) were used to measure c-di-GMP (BioLog) affinity for BdcA. In all experiments, buffers used were the same as those that the protein had been dialyzed into.

For c-di-GMP binding to WT BdcA and analysis via HPLC: Sample chambers were filled with 10 μM c-di-GMP and BdcA in PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8mM KH_2PO_4). Buffer chambers were filled with PBS only. Control experiments were performed with no c-di-GMP added. Each experiment was performed in triplicate using multiple wells on the same plate. Plates were covered with sealing tape and incubated at 37°C for four hours. Samples were collected and analyzed via HPLC.

For c-di-GMP binding to WT BdcA and analysis via UV-Vis spectroscopy: Sample chambers were filled with 10 μM BdcA in PDE assay buffer and c-di-GMP sufficient to reach 20 μM total across both sample and buffer chambers. Control experiments were performed with no c-di-GMP added. Each experiment was performed in triplicate using multiple wells on the same plate. Plates were covered with sealing tape and incubated at 30°C for four hours. Absorbance at 251nm was measured using a Cary 100 UV-Vis Spectrophotometer.

For c-di-GMP binding to BdcA E50Q in sample chambers: Sample chambers were filled w 7.5 μM BdcA E50Q. Buffer chambers were filled with c-di-GMP (kerafast) in PDE assay buffer sufficient to reach 10 in μM both sample and buffer wells. Control experiments were performed with no c-di-GMP added. Each experiment was performed in triplicate using multiple wells on the same plate. Plates were covered with sealing tape and incubated at 30°C for four hours. Absorbance at 251nm was measured using a Cary 100 UV-Vis Spectrophotometer.

For c-di-GMP binding to BdcA E50Q in buffer chambers: Buffer chambers were filled with 7.5 μM BdcA E50Q in triplicate. Corresponding sample chambers were filled with 2.5, 5, or 7.5 μM of c-di-GMP (total concentration between sample and buffer wells) in PDE assay buffer to establish a gradient. Sample and buffer wells were filled in tandem with no BdcA to serve as a negative control. Plates were covered with sealing tape and incubated at 30°C for four hours. Absorbance at 251nm was measured using a Cary 100 UV-Vis Spectrophotometer

8. Isothermal titration calorimetry

A VP-ITC (GE Healthcare) ITC was used to titrate ligand into samples of protein under constant stirring. Bio-Rad Protein Assay was used to determine protein concentration. All samples were degassed using a Microcal Thermovac before loading.

For binding studies of c-di-GMP to E50Q BdcA: Protein samples that had been dialyzed in PDE assay buffer immediately following purification and c-di-GMP in PDE assay buffer were loaded at concentrations of 10 μM and 300 μM, respectively. Each titration was carried out at 20°C and with stirring at 270 rpm, 18 injections of μL. Protein was omitted in one negative control, and c-di-GMP in another. In one experiment, BdcA was dialyzed into high salt buffer (50mM Tris, 0.5mM EDTA, 25mM $MgCl_2$, 250mM NaCl) immediately after purification, and this buffer was also used in place of PDE assay buffer for the rest of the experiment.

For binding studies of NADPH: Immediately after protein purification via affinity chromatography, buffer used during dialysis (20mM HEPES) was collected for each protein in order to create specific NADPH solutions that matched each protein solution exactly. Concentrations of NADPH solutions were verified via absorbance at 340nm. NADPH at 600 uM was titrated into a 12uM sample of protein at 25°C under constant stirring at 310rpm. Each titration was carried out using 30 NADPH injections of 10uL with 200 seconds between injections. After a single titration with WT BdcA verified NADPH binding, the first injection was changed to 2 uL for subsequent experiments. In a single trial that was a repeat of a previous trial in every other aspect, the spacing between injections was changed to 330 seconds.

9. Biofilm dispersal assay

Overnight cultures of WT BdcA and mutants were diluted in LB to 0.05AU at 600nm in 96-well polysterene plates. After 19 hours, sample wells were inoculated with 0.1mM IPTG and positive controls were left undisturbed. Plates were shaken for 1 minute at 150 rpm before incubating further. Crystal Violet endpoint assays[26] were used to quantify biofilms at 12 and 23 hours post-induction.

Works Cited

1. Ma, Q., Yang, Z., Pu, M., Peti, W. & Wood, T. K. Engineering a novel c-di-GMP-binding protein for biofilm dispersal. *Environ. Microbiol.* **13,** 631–42 (2011).

2. Donlan, R. M. Biofilm formation: a clinically relevant microbiological process. *Clin. Infect. Dis.* **33,** 1387–92 (2001).

3. Blenkinsopp, S. A. & Costerton, J. W. Understanding bacterial biofilms. **9,** (1991).

4. Potera, C. Biofilms Invade Microbiology. *Science (80-.).* **273,** (1996).

5. Flemming, H.-C. & Wingender, J. The biofilm matrix. *Nat. Rev. Microbiol.* **8,** 623–33 (2010).

6. Park, A., Jeong, H.-H., Lee, J., Kim, K. P. & Lee, C.-S. Effect of shear stress on the formation of bacterial biofilm in a microfluidic channel. *BioChip J.* **5,** 236–241 (2011).

7. US National Institutes of Health. No Title.

8. Kostakioti, M., Hadjifrangiskou, M. & Hultgren, S. J. Bacterial Biofilms: Development, Dispersal, and Therapeutic Strategies in the Dawn of the Postantibiotic Era. *Cold Spring Harb. Perspect. Med.* 1–24 (2013).

9. Worthington, R. J., Richards, J. J. & Melander, C. Non-Microbicidal Control of Bacterial Biofilms with Small Molecules. *Anti-Infective Agents* **12,** 120–138 (2014).

10. Ma, Q., Zhang, G. & Wood, T. K. Escherichia coli BdcA controls biofilm dispersal in Pseudomonas aeruginosa and Rhizobium meliloti. *BMC Res. Notes* **4,** 447 (2011).

11. Robert-Paganin, J., Nonin-Lecomte, S. & Réty, S. Crystal Structure of an EAL Domain in Complex with Reaction Product 5′-pGpG. *PLoS One* **7,** 1–13 (2012).

12. Spurbeck, R. R., Tarrien, R. J. & Harry, L. T. Enzymatically Active and Inactive Phosphodiesterases and Diguanylate Cyclases Are Involved in Regulation of Motility or Sessility in Escherichia coli CFT073. (2012). doi:10.1128/mBio.00307-12.Editor

13. Lord, D. M., Baran, A. U., Wood, T. K., Peti, W. & Page, R. BdcA, a protein important for Escherichia coli biofilm dispersal, is a short-chain dehydrogenase/reductase that binds specifically to NADPH. *PLoS One* **9,** e105751 (2014).

14. Bijtenhoorn, P. *et al.* A novel metagenomic short-chain dehydrogenase/reductase attenuates Pseudomonas aeruginosa biofilm formation and virulence on Caenorhabditis elegans. *PLoS One* **6,** e26278 (2011).

15. Lesk, a M. NAD-binding domains of dehydrogenases. *Curr. Opin. Struct. Biol.* **5,** 775–83 (1995).

16. Kallberg, Y., Oppermann, U. & Persson, B. Classification of the short-chain dehydrogenase/reductase superfamily using hidden Markov models. *FEBS J.* **277,** 2375–2386 (2010).

17. Biasini, M. *et al.* SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **42,** 252–258 (2014).

18. Kiefer, F., Arnold, K., Künzli, M., Bordoli, L. & Schwede, T. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res.* **37,** 387–392 (2009).

19. Guex, N., Peitsch, M. C. & Schwede, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis* **30,** 162–173 (2009).

20. Arnold, K., Bordoli, L., Kopp, J. & Schwede, T. The SWISS-MODEL workspace: A web-based environment for protein structure homology modelling. *Bioinformatics* **22,** 195–201 (2006).

21. Holm, L. & Rosenström, P. Dali server: Conservation mapping in 3D. *Nucleic Acids Res.* **38,** 545–549 (2010).

22. Schmidt, A. J., Ryjenkov, D. A. & Gomelsky, M. The Ubiquitous Protein Domain EAL Is a Cyclic Diguanylate-Specific Phosphodiesterase : Enzymatically Active and Inactive EAL Domains The Ubiquitous Protein Domain EAL Is a Cyclic Diguanylate-Specific Phosphodiesterase : Enzymatically Active and Inactive E. *Society* **187,** 4774–4781 (2005).

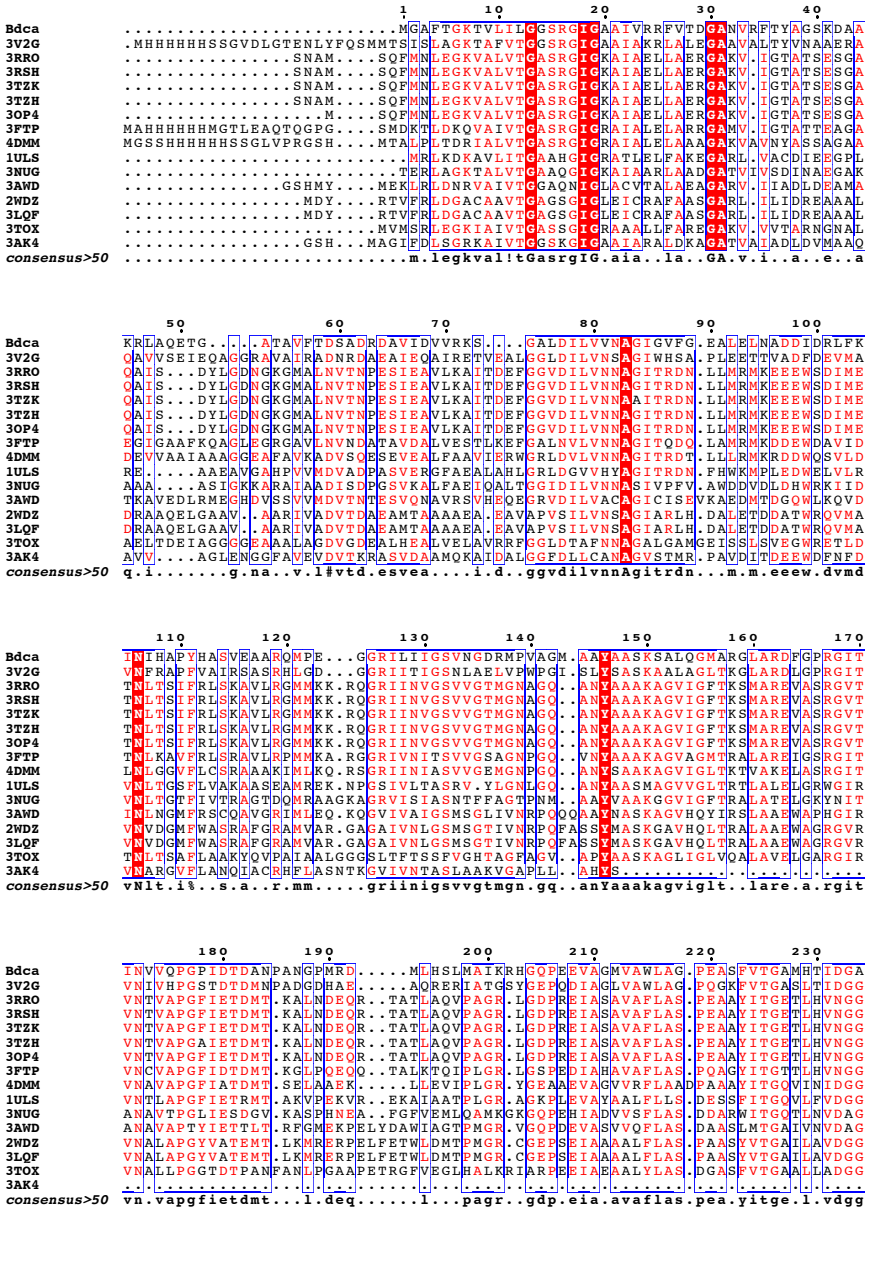
23. Sundriyal, A. *et al.* Inherent Regulation of EAL Domain-catalyzed Hydrolysis of Second Messenger Cyclic di-GMP. *J. Biol. Chem.* **289,** 6978–6990 (2014).

24. Dutta, D., Bhattacharyya, S. & Das, A. K. Crystal structure and fluorescence studies reveal the role of helical dimeric interface of staphylococcal fabg1 in positive cooperativity for NADPH. *Proteins Struct. Funct. Bioinforma.* **80,** 1250–1257 (2012).

25. Joernvall, H., Persson, B. & Krook, M. Short-Chain Dehydrogenases/Reductases (SDR). *Biochemistry* **34,** (1995).

26. Jackson, D. W. *et al.* Biofilm Formation and Dispersal under the Influence of the Global Regulator CsrA of Escherichia coli. *J. Bacteriol.* **184,** 290–301 (2002).

Supplementary



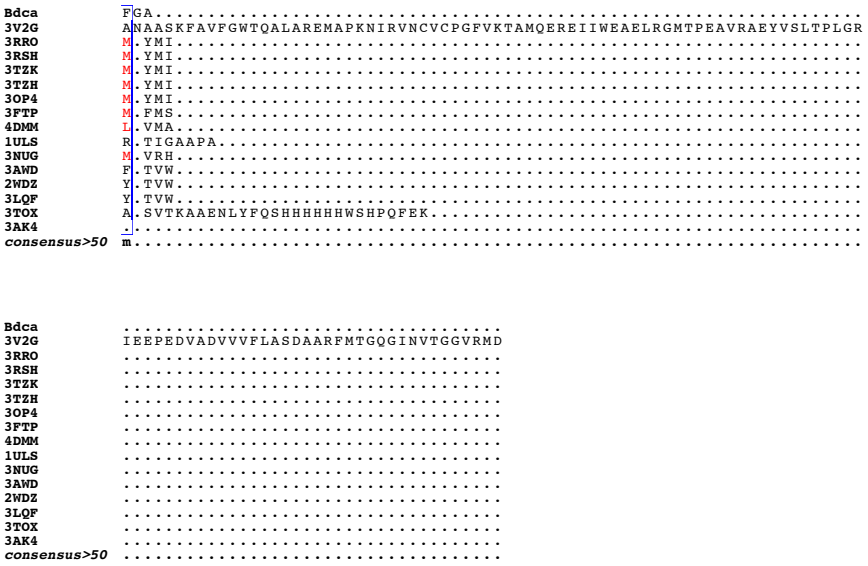


Figure S1. Sequence alignment of BdcA with top hits from DALI[21] search.

Figure S2. Alignment of BdcA with structurally similar proteins show a 32% identity with 3SJ7 and a 34% identity with 3OP4

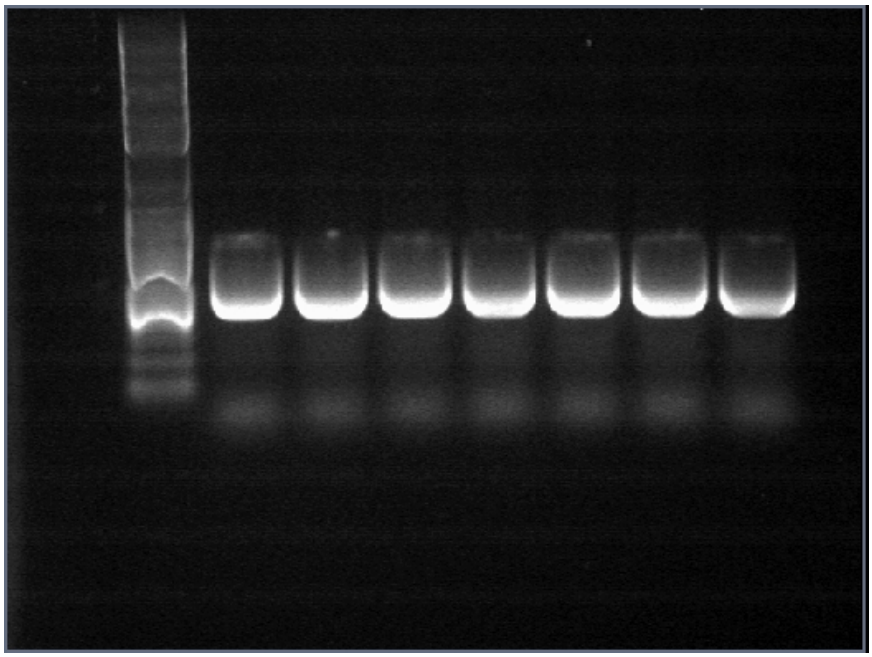


Figure S3. PCR amplification of the BdcA gene from final plasmid product followed by gel electrophoresis verifies that the gene was successfully cloned. Multiple aliquots are shown.

Table S1. Primers used for cloning and site-directed mutagenesis, listed 5' to 3'

Cloning	Forward	GACTGGATCCATGGGCGCTTTTACAGGTAAGACA
	Reverse	GACTCTCGAGCGTAGTCGGTTATGCGCCAAA
G14V	Forward	CAGTTCTCATCCTCGGTGTCAGTCGTGGTATCGGTG
	Reverse	CACCGATACCACGACTGACACCGAGGATGAGAACTG
E50Q	Forward	CTAAACGCCTGGCACAACAGACTGGAGCGACAGC
	Reverse	GCTGTCGCTCCAGTCTGTTGTGCCAGGCGTTTAG
D59A	Forward	CGACAGCAGTATTCACAGCTAGTGCTGACAGAGACGC
	Reverse	GCGTCTCTGTCAGCACTAGCTGTGAATACTGCTGTCG
E90A	Forward	GTATTGGCGTCTTTGGCGCGGCCCTGGAATTAAATG
	Reverse	CATTTAATTCCAGGGCCGCGCCAAAGACGCCAATAC
D136T	Forward	GGCTCCGTGAATGGCACTCGTATGCCTGTTGC
	Reverse	GCAACAGGCATACGAGTGCCATTCACGGAGCC
K150I	Forward	GCTTATGCCGCCAGCATATCTGCCCTGCAAGG
	Reverse	CCTTGCAGGGCAGATATGCTGGCGGCATAAGC
D180A	Forward	GCCAGGGCCAATTGCTACCGACGCTAATC
	Reverse	GATTAGCGTCGGTAGCAATTGGCCCTGGC

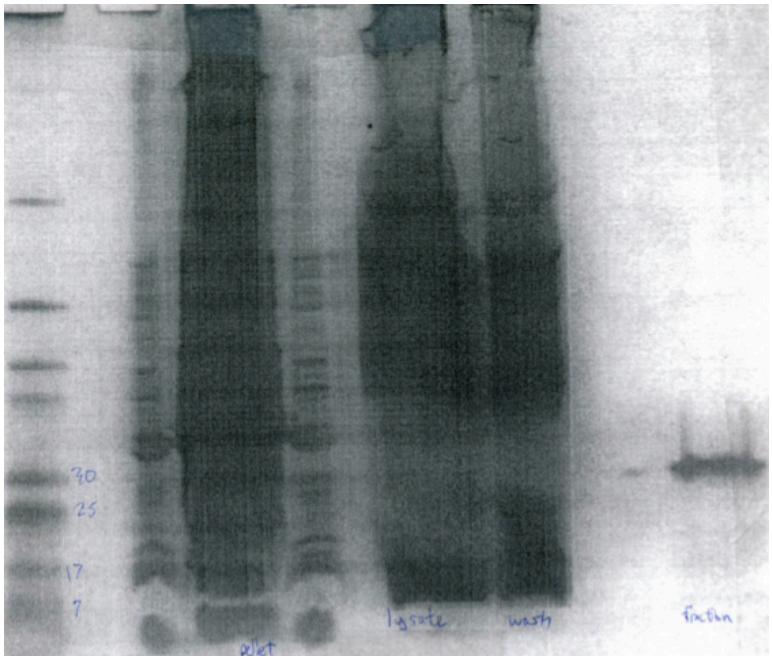


Figure S4. Products from each step of the protein purification process. Lane 1: Ladder; Lane 3: pellet after large cell culture; Lane 5: cell lysate after homogenization; Lane 6: Column wash; Lane 8: collected fraction from the column containing purified protein.