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.

Mohammed Basel Allaw April 12, 2016

Expression, Purification, Activity, and Crystallization of M.EcoGII, a Sequence Non-specific N6-Adenine Methyltransferase

by

Mohammed Basel Allaw

Xiaodong Cheng, Ph.D.
Adviser

Department of Biology

Xiaodong Cheng, Ph.D.

Adviser

Roger Deal, Ph.D.

Committee Member

Jeremy Weaver, Ph.D.

Committee Member

2016

Expression, Purification, Activity, and Crystallization of M.EcoGII, a Sequence Non-specific N6-Adenine Methyltransferase

By

Mohammed Basel Allaw

Xiaodong Cheng, Ph.D.

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 Biology

2016

Abstract

Expression, Purification, Activity, and Crystallization of M.EcoGII, a Sequence Non-specific N6-Adenine Methyltransferase

By Mohammed Basel Allaw

M.EcoGII, initially identified as a sequence non-specific N6-adenine DNA methyltransferase in pathogenic *Escherichia coli*, is purified to >99% homogeneity using an improved method. This improved method is significantly quicker and results in enhanced adenine methyltransferase activity. M.EcoGII activity is greatest at pH 7.0 under conditions of low ionic strength. M.EcoGII is active on various substrates including double stranded DNA (ds-DNA), RNA/DNA hybrid (ds-RNA/DNA), single stranded DNA or RNA (ss-DNA or ss-RNA), as well as tRNA. M.EcoGII exhibits a preference for these substrates in the order of ds-DNA > ss-DNA > ds-RNA/DNA > ss-RNA. M.EcoGII is active on tRNA, irrespective of whether the secondary or tertiary structure is preserved, in agreement with the observation that the enzyme is active on ss-RNA. Computational homology modeling of M.EcoGII suggested conservation of amino acid residues involved in S-adenosyl-L-methionine (AdoMet) binding. Crystallization of M.EcoGII in ternary complexes with ds-DNA or ss-DNA and S-adenosyl-L-homocysteine (AdoHcy), or in a binary complex with the reaction product AdoHcy, was attempted following each purification. Preliminary crystals diffracted X-rays to 4Å with synchrotron radiation.

Expression, Purification, Activity, and Crystallization of M.EcoGII, a Sequence Non-specific N6-Adenine Methyltransferase

By

Mohammed Basel Allaw

Xiaodong Cheng, Ph.D.

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 Biology

2016

Acknowledgements

I would like to express my gratitude and thanks to Dr. Cheng for giving me the opportunity to partake in such an enriching research experience, Dr. Zhang for her aid in the design of experiments, Dr. Horton for assistance with X-ray diffraction data collection and crystal picking, Dr. Patel and Dr. Hashimoto for their keen insights and advice, and finally, Samuel Hong and John Shanks for their warmth, receptivity, and lab aid. I would also like to thank New England Biolabs for proving me with the initial supply of NEB enzyme and its associated purification protocol.

Table of Contents

List of Figures

List of Tables

Chapter 1: Background and Introduction

One of the most common post-replicative base modifications in bacterial genomes is methylation of adenine at the N6 position (Figure 1; Wion & Casadesús 2006). Adenine methyltransferases (MTases), the enzymes which catalyze the conversion of adenine to N6-methyl-adenine, usually fall into two general classes. Many known adenine MTases are part of classical restriction-modification systems that protect bacterial genomes from viral invasion (Bickle & Kruger 1993). Other adenine MTases, however, termed "orphan or solitary MTases," are not linked to restriction-modification systems (Nagaraja & Vasu 2013). These orphan adenine MTases often play a role in modulating the affinity of specific DNA-binding proteins, such as transcription factors, for specific DNA sequences. One of the most extensively studied "orphan" MTases is Dam, a bacterial DNA adenine MTase which preferentially methylates adenine in a GATC sequence context. In many enteric bacteria, Dam helps imprint an epigenetic code that, when interpreted by effector molecules, promotes either transcriptional repression or activation of a large number of pilus operons (Wion & Casadesús 2006). In this way, adenine methylation plays a significant role in determining not only the transcriptional potential of certain bacterial genetic elements, but also the level of virulence in bacterial pathogens that cause urinary tract infections and diarrheal diseases (Low, *et al.* 2001). For this reason, adenine MTases are clinically relevant, and pharmacological inhibition of these MTases might provide therapeutic benefit.

Most MTases, like Dam, methylate their target base within a specific sequence context: it is rare in nature for MTases to indiscriminately methylate their target base. However, it has been recently shown that M.EcoGII, a Dam ortholog, possesses sequence non-specific adenine MTase activity in pathogenic *Escherichia coli* (Fang, *et al.* 2012). One of the goals of this study is to

gain a structural understanding of why M. EcoGII, unlike Dam and other adenine MTases, is capable of such indiscriminate adenine methylation. In this light, M.EcoGII serves as an important target for functional, enzymatic, and structural studies. Considering M.EcoGII activity may be responsible for conferring pathogenicity to particular strains of *E.coli*, a structural understanding of M.EcoGII will help direct future structure-based drug design.

The catalytic Asp-Pro-Pro-Tyr (DPPY) tetrapeptide motif is absolutely conserved between DNA, RNA, and tRNA adenine methyltransferases, and M. EcoGII is no exception. The methyl donor, S-adenosyl-L-methionine (AdoMet), must be present and strategically located to facilitate the chemical exchange. The target adenine is flipped out of the helix and bound in the enzyme's active site by the DPPY motif in an extra-helical conformation (Goedecke, et al. 2001). The exact base flipping mechanism remains poorly understood, but seems to be mediated, in part, by intercalation of an aromatic ring side chain between consecutive base pairs involving the target adenine (Horton, *et al*. 2006). The nucleophilic N6 nitrogen of the target base participates in hydrogen bonding with the residues in the active site. The first hydrogen bond is with one of the oxygen atoms of the Asp side chain carboxylate group; the second, with the main-chain carbonyl oxygen of the first Pro (Goedecke, *et al.* 2001). These hydrogen-bonding interactions are absolutely critical for effective catalysis. For one, they strategically position the lone pair of the nucleophilic nitrogen so that it directly faces the donor methyl of AdoMet. Second, they cause the target nitrogen atom to acquire negative charge character, which activates it for nucleophilic attack. The aromatic ring of Tyr is positioned face-to-face with the target adenine, which allows it to engage in π-stacking interactions (Cheng 1995). These interactions help stabilize the target adenine within the active site of the enzyme. Following nucleophilic attack, a charged methyl-amine intermediate is formed; however, the methylated amine

intermediate quickly reverts back to its original configuration following deprotonation, most likely by a water molecule located within the active site (Blumenthal, *et al.* 1995; Blumenthal, *et al*. 2003; Newby, *et al.* 2002).

In this present study, the M.EcoGII gene was expressed in *E. coli*, and I purified seven times the full-length, untagged M.EcoGII via chromatography. I attempted to crystallize the protein in complex with only S-adenosyl-L-homocysteine (AdoHcy)—the reaction product—or with AdoHcy together with either ds-DNA or ss-DNA. The purpose of this study is three-fold: (1) To develop a purification method that allows me to purify the enzyme quickly with higher levels of purity and enhanced activity; (2) To examine its selective adenine methyltransferase activity in vitro using various substrates; and (3) To understand via X-ray crystallography the structural characteristics which allow it to methylate adenine non-specifically in a wide variety of nucleic acid substrates.

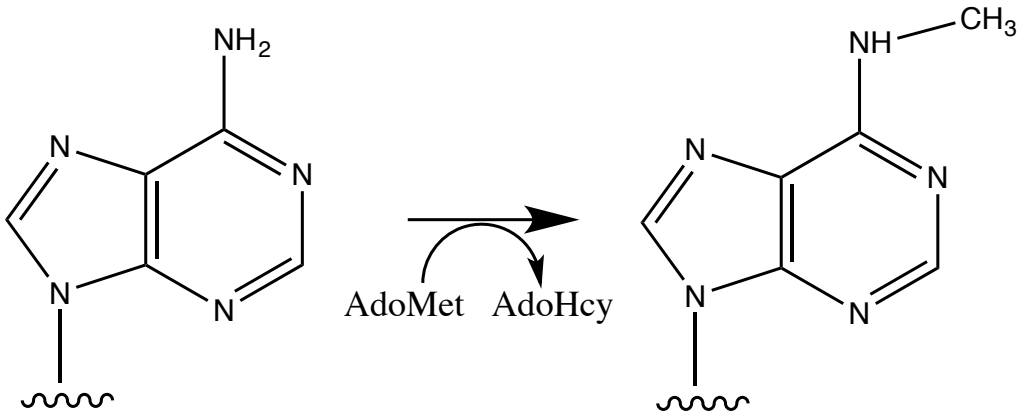


Figure 1. **Schematic detailing the enzyme-catalyzed addition of a methyl group to the N6-position of adenine.**

Chapter 2: Materials and Methods

2.1 Cloning and Overexpression of M.EcoGII:

The untagged M.EcoGII was expressed in *Escherichia coli* cells obtained from New England Biolabs. Protein expression is under the control of the L-arabinose inducible pBAD promoter. A single colony was inoculated in 13 mL of LB containing 100 ug mL^{-1} ampicillin at 37°C overnight while shaking (215 revolutions min^{-1}). Cell aliquots were then used to inoculate larger cultures, usually 6 L in total volume. The cells were cultured in LB medium at 37°C until OD_{600} ~ 0.500-1 was reached. The temperature was kept at 37°C, and 10 mL of 20% w/v L-arabinose were then added per liter of culture to induce expression of the untagged M. EcoGII protein for either 90 or 120 minutes, depending on the preparation.

2.2 Purification of M.EcoGII:

Induced cells from 6X 1L were harvested at 4000g for 20 min. at 4°C. The supernatant was removed, and the pellet was re-suspended in some residual LB liquid. The cells were transferred to 50 mL tubes and underwent another round of centrifugation at 4000g for 20 min. at 4°C. A ~15 mL cell pellet was recovered and frozen O/N at -20°C. The cell pellet was thawed the next day. The cells were lyzed on ice for 10 minutes in 300 mM NaCl, 20 mM Tris-HCl pH 8.0, 5% Glycerol, 0.1 mM EDTA, 1 mM DTT, and 75 uL of PMSF protease inhibitor. Lysis was enhanced by one-second strokes of sonication, followed by 2 sec. of off time.

Crude protein extracts were then recovered by 60 minutes of high-speed centrifugation at 16,500 rpm and 4°C before purification on HiTrap™ **Q** HP Column (GE Healthcare). The column was equilibrated with 30% Buffer B (300 mM NaCl, 20 mM Tris-HCl pH 8.0, 5% Glycerol, 0.1 mM EDTA, and 1 mM DTT). The protein lysate was loaded onto the column, and the flow-through fractions were collected and pooled. The pooled fractions were diluted three-

fold with 0% Buffer B (20 mM Tris-HCl pH 8.0, 5% Glycerol, 0.1 mM EDTA, and 1 mM DTT) so that a final [NaCl] of 100 mM was reached.

The pooled fractions were then loaded onto the HiTrap™ **Heparin** HP Column (GE Healthcare), and all non-specific binding was washed away with 10% Buffer B (100 mM NaCl, 20 mM Tris-HCl pH 8.0, 5% Glycerol, 0.1 mM EDTA, and 1 mM DTT). The proteins bound to the Heparin resin were eluted using a gradient of increasing NaCl concentration from 100 mM to 500 mM. The protein eluted at ~ 300 mM NaCl. The elution fractions were collected and pooled. The pooled fractions were diluted three-fold with 0% Buffer B (20 mM Tris-HCl pH 8.0, 5% Glycerol, 0.1 mM EDTA, and 1 mM DTT) so that a final [NaCl] of 100 mM was reached.

The diluted mix was then loaded onto tandem Q-SP columns [HiTrap™ **Q** HP Column/SP Sepharose™ High Performance HiTrap™ **SP** HP Column (GE Healthcare)]. Most of the protein flowed through the Q column onto the SP column. After all of the protein was loaded onto the columns, the Q column was removed, leaving only the SP column. All non-specific binding was washed away with 10% Buffer B (100 mM NaCl, 20 mM Tris-HCl pH 8.0, 5% Glycerol, 0.1 mM EDTA, and 1 mM DTT). The proteins bound to the SP resin were eluted using an elution gradient of increasing [NaCl] from 100 mM to ~500 mM. The protein eluted at ~ 150 mM NaCl. The elution fractions were collected and pooled. The pooled fractions were diluted 1.5-fold with 0% Buffer B (20 mM Tris-HCl pH 8.0, 5% Glycerol, 0.1 mM EDTA, and 1 mM DTT) so that a final [NaCl] of 100 mM was reached.

The diluted mix was then loaded onto HiTrap™ **Heparin** HP Column (GE Healthcare), and all non-specific binding was washed away with 10% Buffer B (100 mM NaCl, 20 mM Tris-HCl pH 8.0, 5% Glycerol, 0.1 mM EDTA, and 1 mM DTT). The proteins bound to the Heparin resin were eluted using a gradient of increasing NaCl concentration from 100 mM to 500 mM.

The protein eluted at ~ 300 mM NaCl. The elution fractions were collected, pooled, and concentrated in 5K MWCO concentrators at 4°C and ~4000g until a final volume of 2 mL was obtained. The concentrated sample was then loaded onto an S200 16/60 Size Exclusion Column (GE Healthcare). The protein emerged from the column at 0.69 column volumes (CV), which is characteristic of proteins roughly 40 kDa in weight. Final M.EcoGII concentration was determined by absorbance at 280 nm. The elution fractions were collected, pooled, and concentrated in 5K MWCO concentrators at 4°C and ~4000g until the target concentration for crystallization and assays was obtained.

2.3 Preparation of Oligonucleotides:

Nucleic acid substrates were obtained from Integrated DNA Technologies and purified by HiTrap^{TM} Q HP Column (GE Healthcare). Concentrations were determined by measuring absorbance at 260 nm using the molar extinction values calculated by biotools.nubic.northwestern.edu/OligoCalc.html. Oligonucleotides were annealed by heating at 100°C in the presence of 100 mM NaCl, followed by slow cooling to room temperature.

2.4 Radiometric Assay:

The enzymatic activity of M.EcoGII was assessed by addition of a ^{3}H-labelled methyl group from *S*-adenosyl-L-[methyl-^{3}H]methionine to the nucleic acid substrate. Unless stated otherwise, substrate was methylated by M.EcoGII enzymes in the presence of 3.7 μM [methyl-^{3}H]AdoMet (tritium-labeled methyl of AdoMet) and 50 mM Hepes pH 7.0. All enzymatic reactions were carried out at 37ºC and terminated by incorporation of 2.3 mM AdoHcy, a reaction product of methylation that functions as a competitive inhibitor of M.EcoGII. Tubes were subsequently placed in a 4 ºC ice bath to ensure termination. The nucleic acid products

were immobilized on Whatman™ DE81 cellulose filter paper (GE Healthcare). The filter papers were washed twice with 5 mL of 0.2 M NH_4HCO_3, twice again with 5 mL Milli-Q water, and once with 100% (w/v) ethanol. They were subsequently subject to liquid-scintillation counting (PerkinElmer). Bars, lines, and curves were fitted using GraphPad Prism 5.0 software (GraphPad Software, Inc.).

2.5 Crystallography:

Crystallization of M.EcoGII alone or in complex with nucleic acid was attempted by the sitting-drop and hanging-drop vapor diffusion methods at 16ºC. Various nucleic acid substrates were used for co-crystallization. SAH-DNA-Protein were mixed in 2: 1.1: 1 molar ratios and incubated at 4ºC before crystallization. An aliquot of the complexes (0.2 µL) was mixed with an equal volume of mother liquor by Phoenix (Art Robbins Instruments). A total of 69 screens were performed. Diffracting crystals of macromolecular nature grew overnight at 16ºC. The crystals used for diffraction experiments were obtained from mother liquor containing 0.2 M Zinc Acetate and 20% PEG3350. Crystals were cryoprotected by soaking in mother liquor supplemented with 27% (v/v) ethylene glycol or 22% (v/v) glycerol before plunging in liquid nitrogen

An X-ray diffraction data set was collected at the SER-CAT 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory and processed using HKL2000 (Otwinowski, *et al.* 2003). One data set at 4Å resolution from one crystal was collected at 100K, containing approximately 200 frames of one-degree oscillations at a wavelength of 1Å.

2.6 Structural Modeling:

Computational structure prediction of M.EcoGII was performed via Phyre (Kelley & Sternberg 2015). PyMol was used to visualize the structure of M.EcoGII and analyze the amino acid residues involved in AdoMet binding (The PyMol Molecular Graphics System, Version 1.5.0.4 Shrödinger, LLC).

Chapter 3: Results and Discussion

3.1 Purification of M.EcoGII:

3.1.1 Purifications #1-2:

For the first two purifications, two deviations from the original NEB protocol occurred. For purification #1, the Heparin elution fractions were first loaded onto a Q/SP tandem column. This was in accordance with the NEB protocol. The Q/SP tandem column used by NEB to purify M.EcoGII was such that M.EcoGII flowed through both segments. However, during the Q/SP tandem step of purification #1, no M.EcoGII in the flow-through was detected. It was unclear whether M.EcoGII was binding to the Q or SP portion of the tandem column. M.EcoGII was eluted from the Q/SP tandem column and reloaded onto SP alone. The loading of M.EcoGII onto SP in isolation was the first deviation. M.EcoGII was observed to bind to the SP column and eluted at ~150 mM NaCl. The fact that M.EcoGII flowed through the NEB Source 15S cation exchange column but associated with our SP equivalent is most likely reflected by differences in the resin composition. As can be observed from the gels in Figures 2-3d vs. Figures 4-5d, the presence of the Q column in tandem with the SP is essential. The presence of the Q column in tandem with the SP column separates contaminant species that would otherwise co-elute with M.EcoGII in the SP elution fractions. A second Heparin run was not performed following Q/SP. This was the second deviation. As will be described in Purification #6, the second Heparin

column step is not essential and does not improve purity as long as the Q/SP tandem column precedes it.

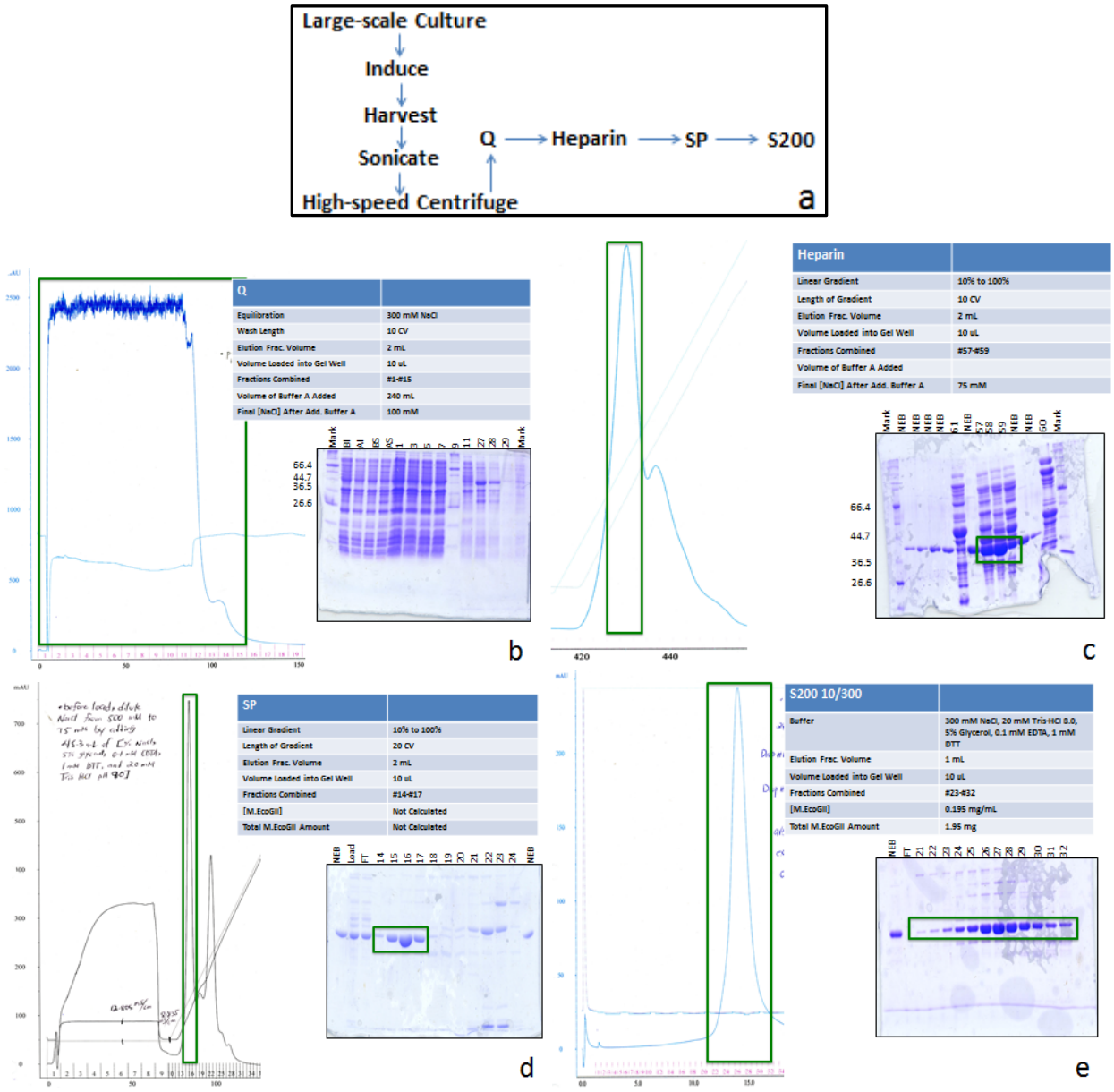


Figure 2. **Stained SDS-PAGE gels and chromatogram peaks corresponding to successive purification stages of preparation #1**. (a) Step-wise expression and purification scheme (b) SDS-PAGE gel, flow-through peak, and purification parameters of Q column step (c) SDS-PAGE gel, elution peak, and purification parameters of Heparin column step (d) SDS-PAGE gel, elution peak, and purification parameters of SP column step (e) SDS-PAGE gel, elution peak, and purification parameters of S200 column step

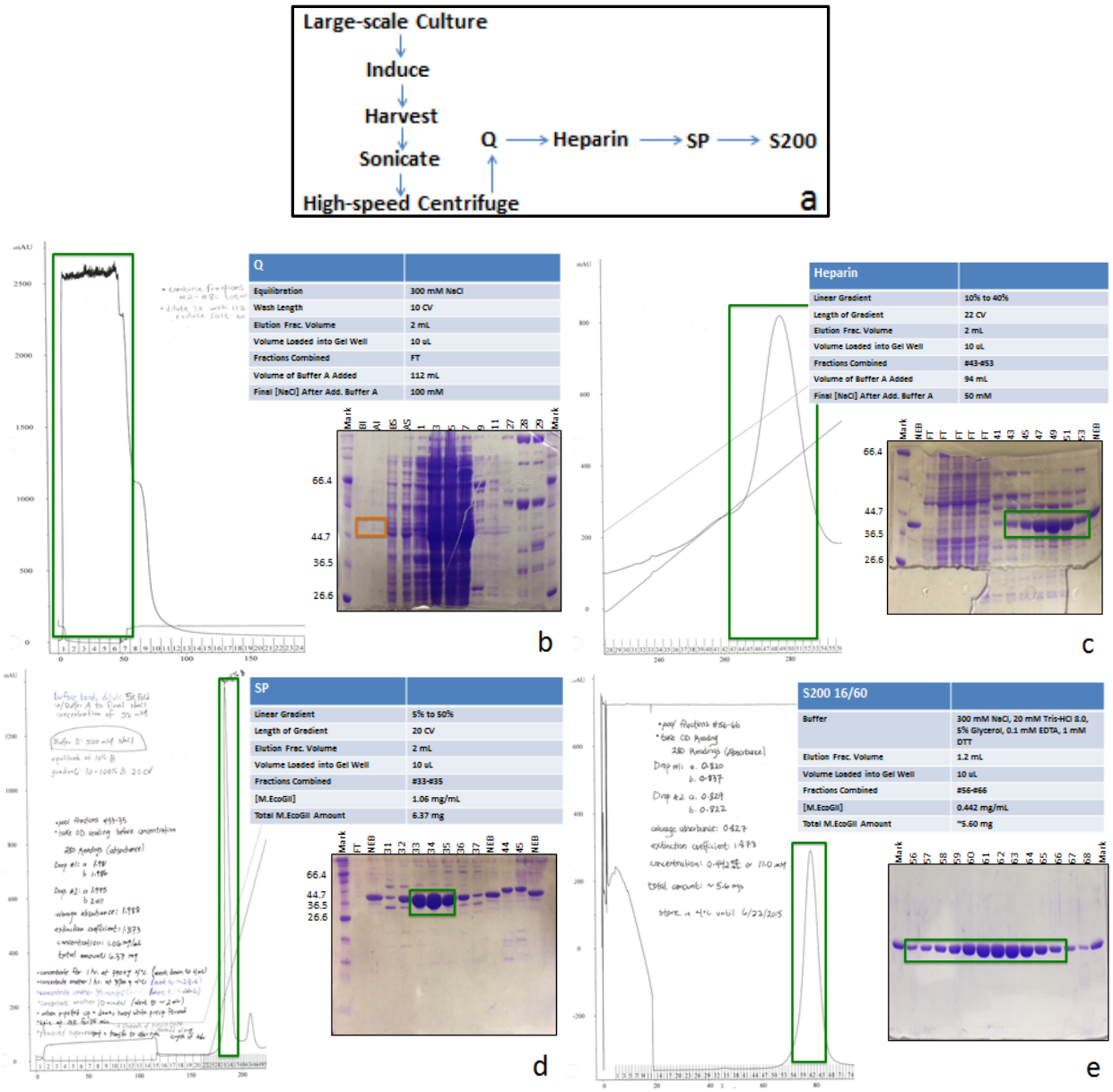


Figure 3. **Stained SDS-PAGE gels and chromatogram peaks corresponding to successive purification stages of preparation #2**. (a) Step-wise expression and purification scheme (b) SDS-PAGE gel, flow-through peak, and purification parameters of Q column step (c) SDS-PAGE gel, elution peak, and purification parameters of Heparin column step (d) SDS-PAGE gel, elution peak, and purification parameters of SP column step (e) SDS-PAGE gel, elution peak, and purification parameters of S200 column step

3.1.2 Purifications #3-5:

Purifications #3-4:

For these purifications, the NEB protocol was faithfully followed. This resulted in a level of purity that was superior to that obtained in purifications #1-2 in which a Q/SP tandem run was not performed (Figures 2-3e. vs. Figures 4-5f).

Purification #5:

In purifications #3-4, although most of the M.EcoGII adhered to the SP portion of the Q/SP tandem column, a small fraction flowed through during the washing step with 100 mM NaCl, which is very close to the salt concentration at which M.EcoGII elutes (Figures 4-5d). This resulted in loss of some M.EcoGII. In order to prevent the flow-through, and thus loss, of M.EcoGII during the Q/SP tandem step of purification #5, M.EcoGII was loaded onto the Q/SP tandem column with 50 mM NaCl instead of the normal 100 mM NaCl. The idea was that by decreasing salt, the affinity of M.EcoGII for the SP portion of the Q/SP tandem column would increase, resulting in less flow-through. For one, lowering the NaCl concentration to 50 mM caused a bit of white precipitation to form and resulted in a slightly decreased yield relative to purifications #3-4 (Table 1). Second, rather than flowing through the Q portion of the tandem column, as was the case for preparations #3-4, M.EcoGII adhered to it at 50 mM NaCl. This is reflected by the fact that little protein eluted from the SP column after the tandem column was separated (Figure 6c). Elution of the Q column revealed that M.EcoGII had adhered to it at 50 mM salt (Figure 6d). This purification was particularly enlightening, as it demonstrated that salt concentration should not be decreased below 75-100 mM salt at any stage during the purification process.

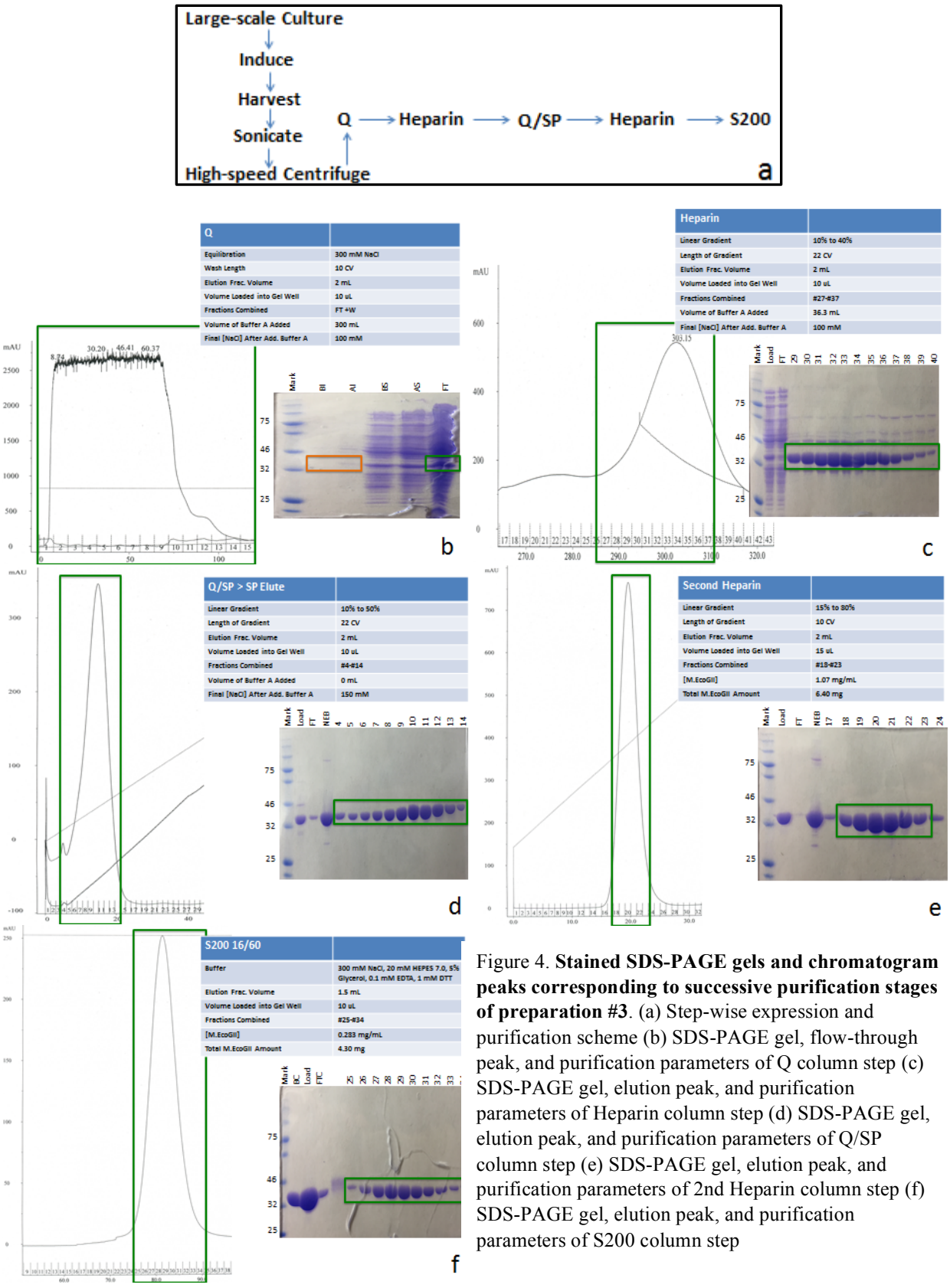


Figure 4. **Stained SDS-PAGE gels and chromatogram peaks corresponding to successive purification stages of preparation #3**. (a) Step-wise expression and purification scheme (b) SDS-PAGE gel, flow-through peak, and purification parameters of Q column step (c) SDS-PAGE gel, elution peak, and purification parameters of Heparin column step (d) SDS-PAGE gel, elution peak, and purification parameters of Q/SP column step (e) SDS-PAGE gel, elution peak, and purification parameters of 2nd Heparin column step (f) SDS-PAGE gel, elution peak, and purification parameters of S200 column step

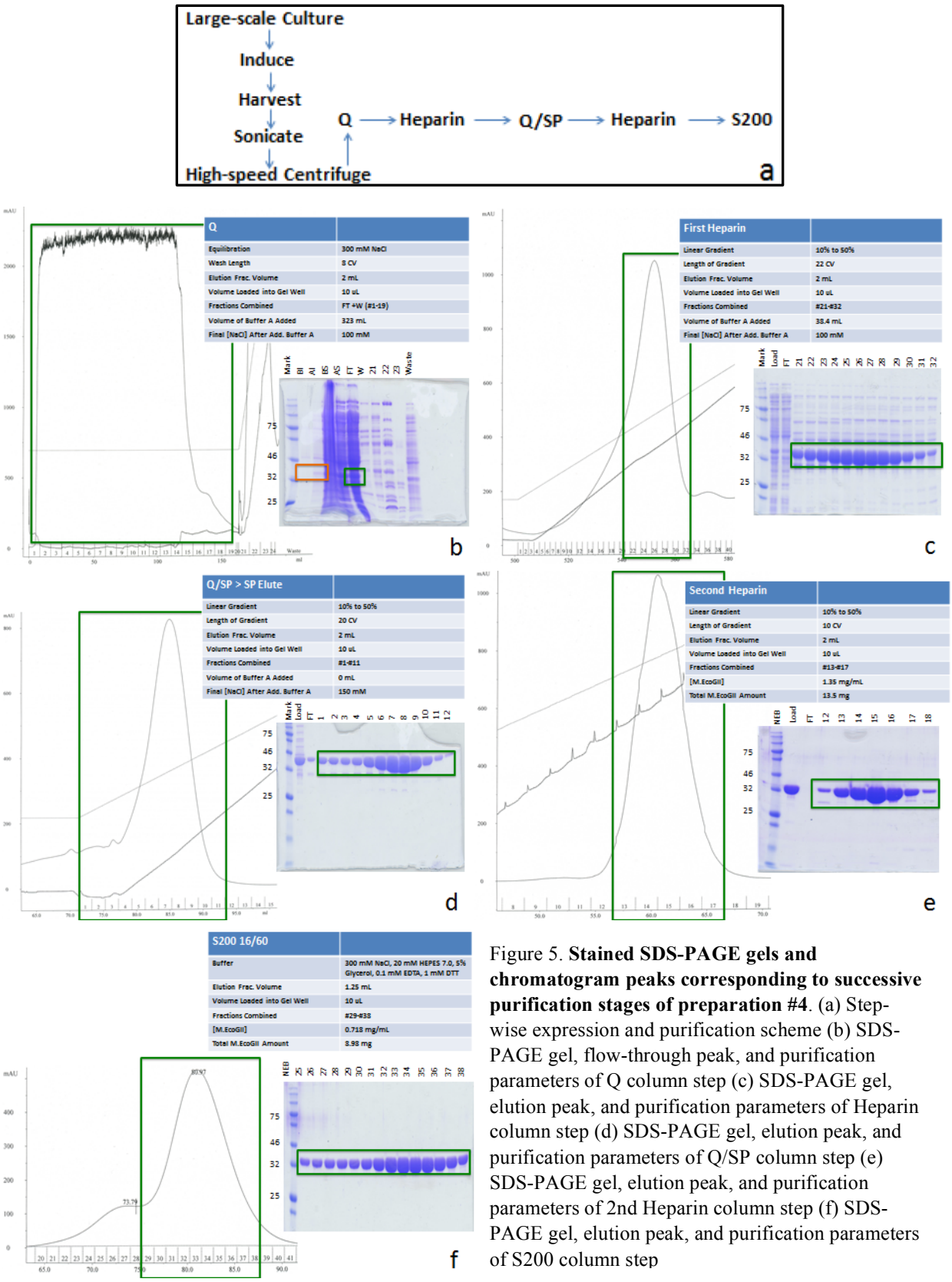


Figure 5. **Stained SDS-PAGE gels and chromatogram peaks corresponding to successive purification stages of preparation #4**. (a) Step-wise expression and purification scheme (b) SDS-PAGE gel, flow-through peak, and purification parameters of Q column step (c) SDS-PAGE gel, elution peak, and purification parameters of Heparin column step (d) SDS-PAGE gel, elution peak, and purification parameters of Q/SP column step (e) SDS-PAGE gel, elution peak, and purification parameters of 2nd Heparin column step (f) SDS-PAGE gel, elution peak, and purification parameters of S200 column step

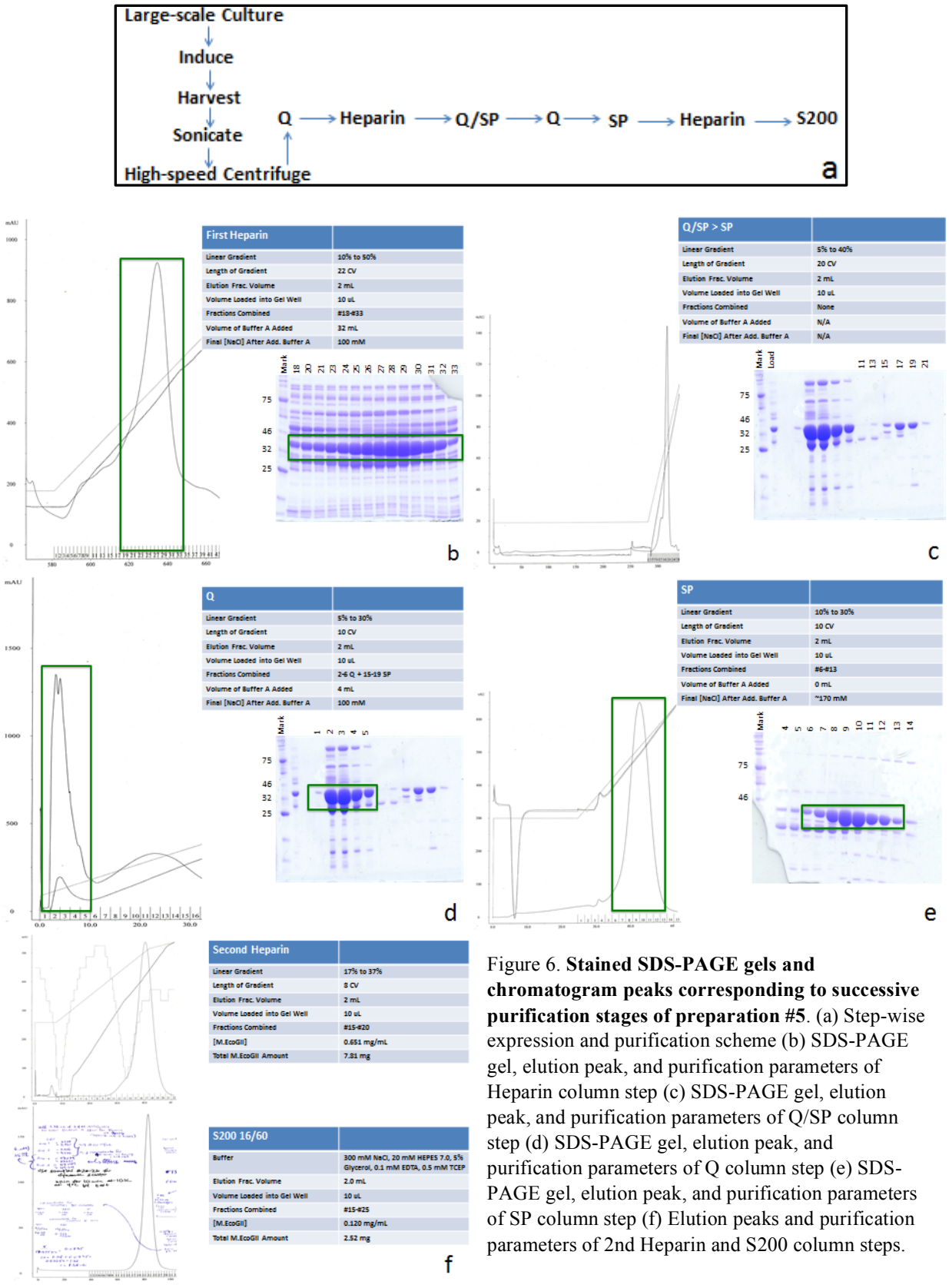


Figure 6. **Stained SDS-PAGE gels and chromatogram peaks corresponding to successive purification stages of preparation #5**. (a) Step-wise expression and purification scheme (b) SDS-PAGE gel, elution peak, and purification parameters of Heparin column step (c) SDS-PAGE gel, elution peak, and purification parameters of Q/SP column step (d) SDS-PAGE gel, elution peak, and purification parameters of Q column step (e) SDS-PAGE gel, elution peak, and purification parameters of SP column step (f) Elution peaks and purification parameters of 2nd Heparin and S200 column steps.

3.1.3 Purifications #6-7:

Purification #6:

The NEB purification protocol calls for a second Heparin column run following Q/SP. However, a very high level of purity was already achieved following Q/SP tandem; no contaminant species can be visualized via SDS-Page gel (Figure 7d). For this reason, a second Heparin column run was deemed unnecessary, and the Q/SP elution fractions were pooled and directly loaded onto the size-exclusion S200 16/60 Superdex column. The final M.EcoGII contained a level of purity that was comparable to earlier preparations #2-5. The low yield of purification #6 relative to preparations #3-5 was due to the shorter induction time (90 minutes vs. 120 minutes).

Purification #7:

The NEB purification protocol calls for the use of 20 mM Tris-HCl pH 8.0 for the Q/SP tandem step. The elution fractions collected and pooled following Q/SP must be concentrated to ~ 2 mL before loading onto the S200. At pH 8.0, concentration is slow, taking ~ 5 hours to concentrate a sample of 22 mL to 2 mL. Therefore, to accelerate concentration and save time, a change to the Q/SP buffer was made. Specifically, the buffer was changed to 20 mM Hepes pH 7.0. The decrease in pH from 8.0 to 7.0 significantly accelerated concentration. It took ~ 1 hour to concentrate a sample of 8 mL to 1.9 mL.

The buffer change, however, altered the binding properties of M.EcoGII to the SP column. The isoelectric point of M.EcoGII is 8.57. M.EcoGII has a greater net positive charge at pH 7.0 than at pH 8.0; therefore, it is expected that M.EcoGII binds more tightly to the negatively-charged SP column under these new pH conditions. This is indeed what was

observed. Previously, under pH 8.0 conditions, M.EcoGII eluted at roughly 150 mM NaCl (Figure 7d). Under pH 7.0 conditions, M.EcoGII eluted at 250 mM NaCl (Figure 8d). This increase in salt reflects increased binding to the SP column. However, in comparison with the Q/SP gels from earlier preparations, more contamination can be observed (Figures 4-5d, Figure 7d, and Figure 8d). The bound proteins were eluted using a gradient of increasing salt from 100 to 800 mM NaCl. A wide gradient was used because it was unknown at which [NaCl] M.EcoGII would elute in pH 7.0. Next time, a narrower gradient, from 100 to 400 mM NaCl will be used. This will help separate the major, 30 kDa contaminant species visible in fractions #C9-C12 (Figure 8d.). Higher molecular-weight contamination (~ 90-100 kDa) can also be observed (Figure 8d.). This contamination appears to co-elute with M.EcoGII. This contamination did not appear in Q/SP gels of earlier preparations; however, size-exclusion chromatography separated this contaminant from M. EcoGII (Figure 8e.).

The low yield of purification #7 relative to preparations #3-5 is due to the shorter induction time (90 minutes vs. 120 minutes).

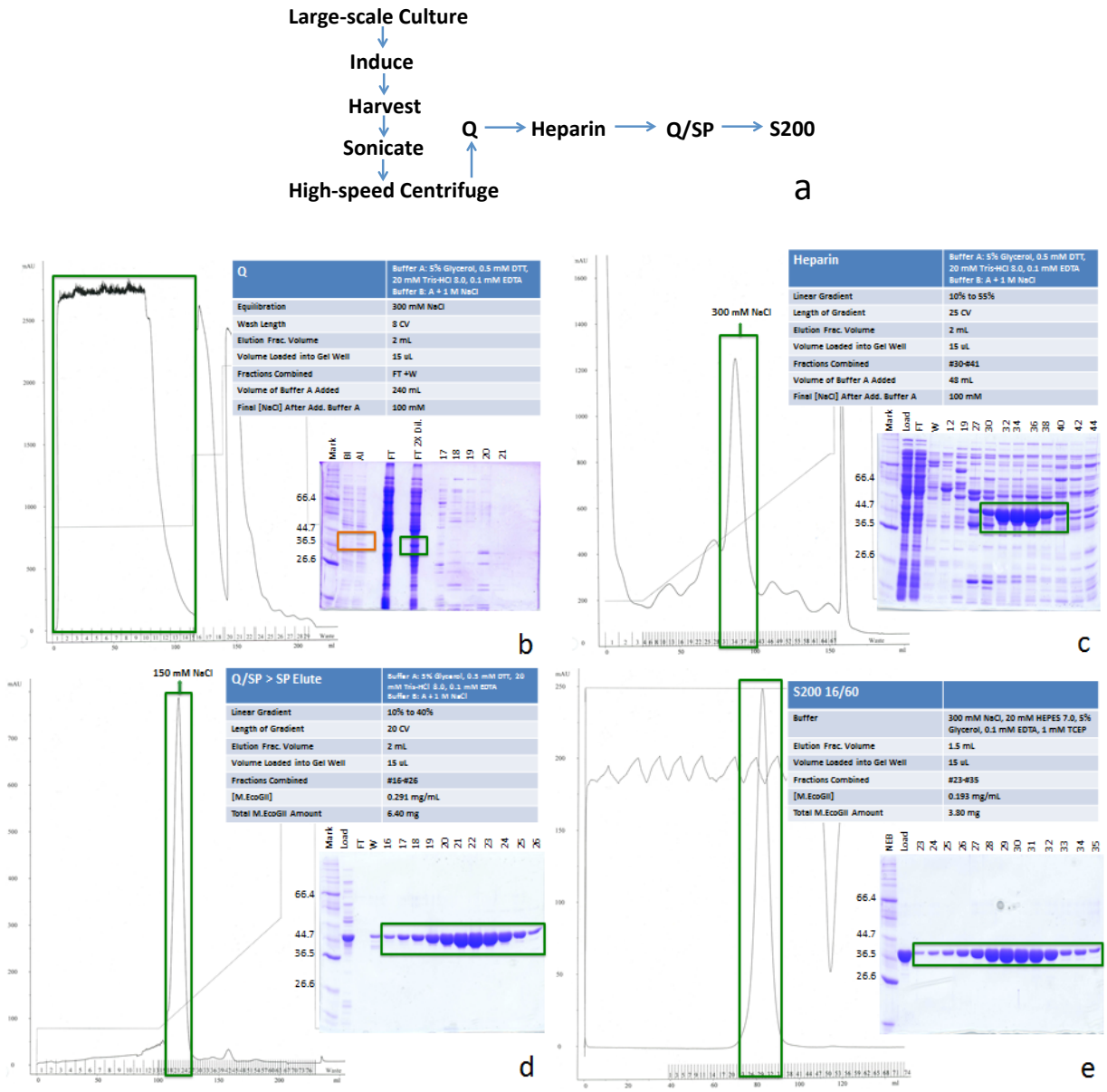


Figure 7. **Stained SDS-PAGE gels and chromatogram peaks corresponding to successive purification stages of preparation #6**. (a) Step-wise expression and purification scheme (b) SDS-PAGE gel, flow-through peak, and purification parameters of Q column step (c) SDS-PAGE gel, elution peak, and purification parameters of Heparin column step (d) SDS-PAGE gel, elution peak, and purification parameters of Q/SP column step (e) SDS-PAGE gel, elution peak, and purification parameters of S200 column step

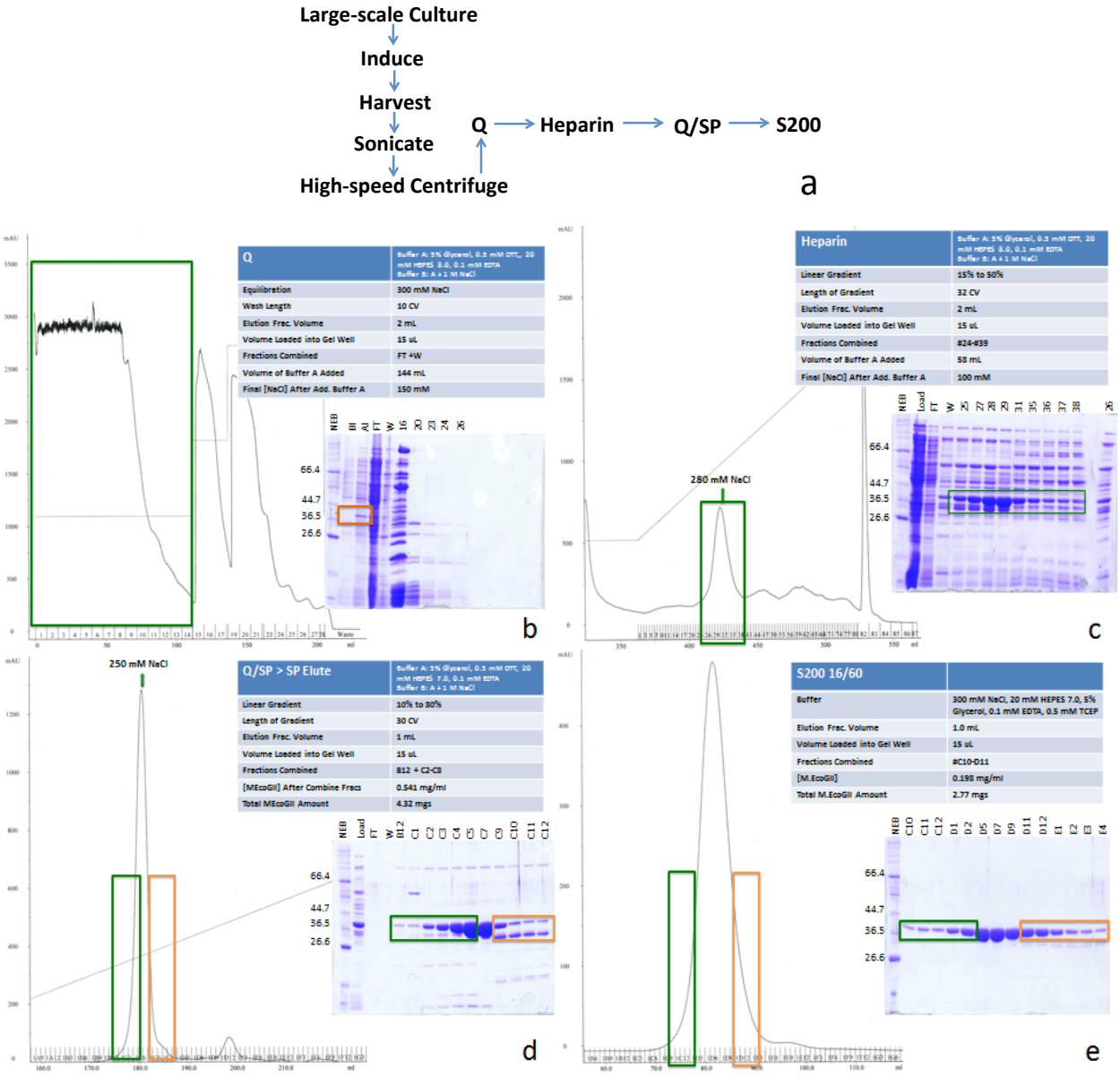


Figure 8. **Stained SDS-PAGE gels and chromatogram peaks corresponding to successive purification stages of preparation #7**. (a) Step-wise expression and purification scheme (b) SDS-PAGE gel, flow-through peak, and purification parameters of Q column step (c) SDS-PAGE gel, elution peak, and purification parameters of Heparin column step (d) SDS-PAGE gel, elution peak, and purification parameters of Q/SP column step (e) SDS-PAGE gel, elution peak, and purification parameters of S200 column step

3.1.4 Summary of How Original NEB Purification Can Be Optimized:

The NEB protocol calls for 90 minutes of induction time. The NEB protocol states that induction of M.EcoGII exceeding 120 minutes is toxic to the cells. For preparations #3-5, the

total induction time was 120 minutes; for preparations #6-7, 90 minutes. The yields for preparations #3-5 were 1 mg/L, 0.9 mg/L, and 0.6 mg/L, respectively (Table 1). For preparations #6-7, the yields were respectively 0.6 mg/L and 0.5 mg/L (Figure 10). The difference in average yield between preparations #3-5 and preparations #6-7 (0.8 mg/L vs. 0.6 mg/L) might be due to the difference in induction time. Allowing induction to proceed for a full 2 hours permitted additional production of M.EcoGII in preparations #3-5. Therefore, it would be prudent to prolong total induction time to 120 minutes in the future. 120 minutes of induction maximizes M.EcoGII expression, but falls just short of the threshold whereby M.EcoGII expression becomes toxic.

Additionally, purification #6 demonstrates that the 2nd Heparin run is superfluous since all contaminant species were removed in the preceding Q/SP tandem step (Figure 7d). Purification #7 also demonstrates that the 2nd Heparin is unneeded since the size-exclusion step removed almost all of the contaminants that co-eluted with M.EcoGII during the Q/SP tandem step. (Figures 8d-e). The ability to forego the 2nd Heparin run accelerates the purification process and confers a valuable time saving advantage. With any purification, there is invariably a small percentage of the loaded protein that is lost at a given column step. The ability to forego the 2nd Heparin run, thus, minimizes the amount of M.EcoGII lost during the purification process, resulting in an increased overall yield. The ability to forego the 2nd Heparin run also minimizes the total time that M.EcoGII is stored at 4ºC during the purification process, which helps protect against degradative processes that may abolish M.EcoGII activity and binding properties. Purification #7 demonstrates that the use of 20 mM Hepes pH 7.0 in place of 20 mM Tris-HCl pH 8.0 as the Q/SP tandem column buffer accelerates the concentration time before loading onto the S200 column. The switch from pH 7.0 buffer to pH 8.0 buffer at this junction during the

preparation also aids in accelerating the purification process and confers an additional time saving advantage. With these modifications, approximately 5-6 hours of lab time can be saved, which is significant considering how involving M.EcoGII purification is.

Preparation	Yield (mg M.EcoGII / L culture)
#1	0.300
#2	0.900
#3	1.00
#4	0.900
#5	0.600
#6	0.600
#7	0.500

Table 1. **M.EcoGII Yield from Each Purification**. Yield is calculated as mg of M.EcoGII recovered following the final purification step per L of culture grown**.**

3.2 Enzymatic Activity of M.EcoGII

3.2.1 M.EcoGII Activity is pH and NaCl Dependent:

In these experiments, the effect of pH and NaCl on M.EcoGII activity was investigated. 7.5 μM ss-DNA (5'-ACCGGCACGCGC-3') was used for all pH and NaCl experiments. To investigate the enzyme's NaCl dependence, ss-DNA was incubated with 1 μM and 0.5 μM M.EcoGII at 4 different NaCl concentrations: 0, 12.5, 25, and 50 mM. The results can be modeled as one phase exponential decay (Figure 9b). M.EcoGII activity was greatest at a NaCl concentration of 0 mM. As the NaCl concentration increased, M.EcoGII activity decayed and eventually plateaued. 50% inhibition was attained at 14.3 mM NaCl for the experiment using 1 μM enzyme, and 5.5 mM NaCl for the experiment using 0.5 μM enzyme. At 50 mM NaCl, 85% inhibition was attained for the experiment using 1 μM enzyme, and 67% inhibition was attained for the experiment using 0.5 μM enzyme. This suggests that high levels of NaCl inhibit enzymatic activity by interacting with the DNA substrate and/or M.EcoGII, and these

interactions collectively reduce the binding affinity of M.EcoGII such that a tight complex cannot form. Alternatively, large amounts of NaCl could disrupt catalysis. These findings are in agreement with previous studies that report NaCl's inhibitory effect on methyltransferase activity (Adams, *et. al.* 1979; Jones & Taylor 1982).

M.EcoGII activity was tested at the following pH values: 6.5, 7.0, 7.5, 8.0, and 8.5 (Figure 9a). M.EcoGII activity rose by 10% as pH was increased from 6.5 to 7.0, at which point a plateau was reached. After this plateau, M. EcoGII activity steadily declined as pH progressively increased, eventually reaching a minimum level of activity at the most basic pH value tested, 8.5. There was a 33% reduction in M.EcoGII activity as pH rose from 7.0 to 7.5, a 35% reduction in activity from pH 7.5 to 8.0, and finally, a 61% reduction in activity from pH 8.0 to 8.5. This amounts to an average 43% reduction in M.EcoGII activity for every 0.5 unit increase in pH starting at pH 7.0. The pH shift from 7.0 to 8.5 represents an 83% decline in M.EcoGII methyltransferase activity. The decline in enzyme efficiency at basic pHs likely reflects a change in the charge of residues involved in substrate binding and/or catalysis.

Collectively, these results helped guide and optimize conditions for subsequent experiments and crystallization trials. For subsequent experiments, NaCl concentration was reduced so as to approach 0 mM. All subsequent radiometric experiments were also performed under pH 7.0 conditions. For the final four purifications, M.EcoGII was eluted from the S200 size exclusion column (GE Healthcare) using pH 7.0 buffer.

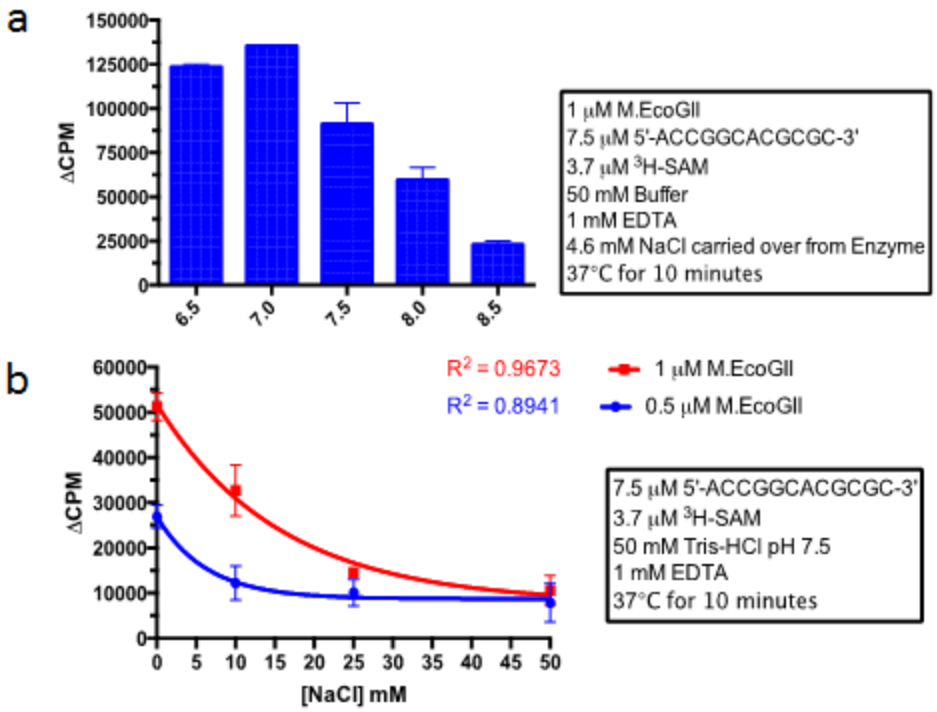


Figure 9. **pH and NaCl Dependence of M.EcoGII** (a) M.EcoGII activity is greatest at a pH of 7.0. M.EcoGII activity decreases as pH progressively increases past pH 7.0. (b) M.EcoGII activity decays with increasing NaCl.

3.2.2 M.EcoGII Activity on ds-DNA vs. ss-DNA:

In this experiment, M.EcoGII activity on ds-DNA was compared with its activity on ss-DNA. The ds-DNA and ss-DNA oligonucleotides used for the enzymatic comparison possessed matching sequences (Figure 10a). Each oligonucleotide was 12 bases in length and contained one internal adenine, flanked on the 5' end by GCCGGC, and by CGCGC on the 3' end. Nucleic acid concentration was held constant at 2 μM. Four different M.EcoGII concentrations were used: 0, 31.25, 62.5, 125, and 250 nM.

Both substrate and cofactor concentration greatly exceeded enzyme concentration; therefore, the amount of product formed should increase linearly with enzyme concentration. The

R^2 values for ds-DNA (0.9953) and ss-DNA (0.9797) nearly approach 1.0 (Figure 10b), which confirms that this theoretical premise indeed holds in practice and lends credence to the results obtained. The slope for ds-DNA was 337 ± 8 ΔCPM/nM M.EcoGII compared to 284 ± 14 ΔCPM/nM M.EcoGII for ss-DNA. The ΔCPM/nM M.EcoGII for ds-DNA is approximately 1.2-fold greater the ΔCPM/nM M.EcoGII for ss-DNA (Figure 10b). The percent difference of CPM between ds-DNA and ss-DNA, averaged across M.EcoGII concentrations tested, is 36%. These results indicate the M.EcoGII exhibits a slight preference for ds-DNA over ss-DNA.

A review of the literature indicates that M.EcoGII is not the only adenine methyltransferase capable of methylating both ds and ss-DNA. M.EcoP1I, which belongs to the Type III restriction-modification system and methylates the internal adenine within an AGACC sequence context, has been shown to exhibit activity on both ds and ss-DNA substrates (Roberts, *et al.* 2010; Sistla, *et al.* 2004). M.EcoP1I has been shown to methylate ds and ss-DNA with almost equal efficiency (Bheemanaik, *et al.* 2010). Some cytosine N4 methyltransferases, such as Bcn1, have been also been reported to exhibit activity on both ds and ss-DNA substrates (Merkiene, *et al.* 1998).

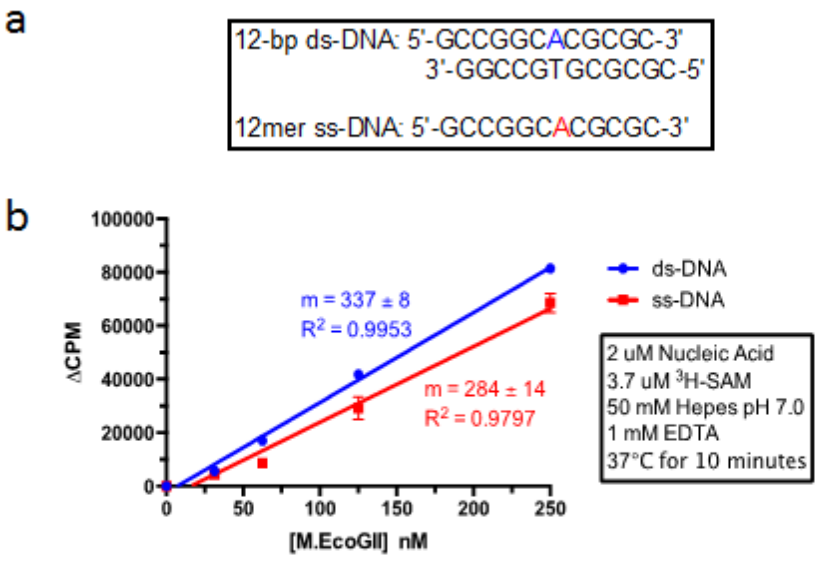


Figure 10. **M.EcoGII Activity on ds-DNA vs. ss-DNA** (a) Synthetic oligonucleotides used for the comparison. Both oligonucleotides contain only one internal adenine. (b) M.EcoGII exhibits a preference for ds-DNA.

3.2.3 M.EcoGII Activity on a Wide Variety of Nucleic Acid Substrates:

In this experiment, M.EcoGII activity on ss-DNA was compared with its activity on ss-RNA. The ss-DNA and ss-RNA oligonucleotides used for this enzymatic comparison possessed matching sequences (Figure 11a). Each oligonucleotide was 27 bases in length and contained nine internal adenine sites; two of these internal adenines were positioned consecutively, while all others had intervening guanosines, cytosines, and in the case of RNA, uracils as well. Additionally, M. EcoGII activity on ds-DNA was compared with its activity on a ds-RNA/DNA hybrid. The ds-DNA and ds-RNA/DNA hybrid used for this enzymatic comparison possessed matching sequences (Figure 11a). Each oligonucleotide was 27 bases in length, and contained a total of fourteen adenine sites, nine of which were in the top strand and five on the bottom. The ds-RNA/DNA hybrid consisted of a nine-adenine RNA top strand and a five-adenine DNA bottom strand. The RNA top strand used to form the hybrid was the same ss-RNA used for the ss-RNA vs. ss-DNA comparison. Both the ds-DNA and ds-RNA/DNA hybrid possessed the same five-adenine DNA bottom strand.

In order for the comparisons to be valid, the concentration of each substrate greatly exceeded its respective K_m (the [substrate] at which $1/2V_{max}$ is achieved). In other words, M.EcoGII catalyzes the methyl transfer to its target adenine at maximum velocity for each substrate. The average ΔCPM for ss-DNA was 26,000 ± 500, whereas for ss-RNA it was 3500 ± 300 (Figure 11b). M.EcoGII's activity on ss-DNA was approximately 7-fold greater than its activity on ss-RNA. This data indicate that M.EcoGII exhibits a preference for ss-DNA over ss-RNA. The average ΔCPM for ds-DNA was 53000 ± 3000, whereas for ds-RNA/DNA it was 8410 ± 20 (Figure 11b). M.EcoGII's activity on ds-DNA was approximately 6-fold greater than

its activity on ds-RNA/DNA. These data indicate that M.EcoGII exhibits a preference for ds-DNA over ds-RNA/DNA.

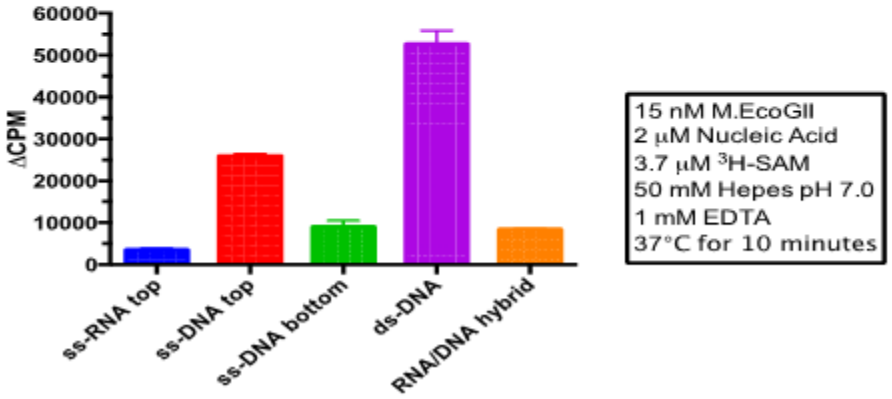
Additional experiments confirm the results described above. In Figure 11c, M.EcoGII activity on ss-DNA was compared with its activity on ss-RNA. The same oligonucleotides were used for this comparison and the comparison illustrated in Figure 11b; the only differences between the two are an increased [M.EcoGII] and [substrate] in the second experiment. The average ΔCPM for ss-DNA was 290,000 ± 20,000, whereas for ss-RNA it was 62,000 ± 3,000 (Figure 11c). M.EcoGII's activity on ss-DNA was approximately 5-fold greater than its activity on ss-RNA. These data indicate that M.EcoGII exhibits a preference for ss-DNA over ss-RNA and confirms the findings of the experiment illustrated in Figure 11b.

In this experiment, M. EcoGII activity on ds-DNA was compared with its activity on a ds-RNA/DNA hybrid. The same oligonucleotides were used for this comparison and the comparison illustrated in Figure 11.b; the only difference is that three different M.EcoGII concentrations were used: 0, 3, and 15 nM. The slope for ds-DNA was 2310 ± 90 ΔCPM/nM M.EcoGII compared to 610 ± 20 ΔCPM/nM M.EcoGII for ds-RNA/DNA (Figure 11d). The ΔCPM/nM M.EcoGII for ds-DNA is approximately 4-fold greater the ΔCPM/nM M.EcoGII for ds-RNA/DNA. The percent difference of CPM between ds-DNA and ds-RNA/DNA, averaged across M.EcoGII concentrations tested, is 36%. These results indicate the M.EcoGII exhibits a preference for ds-DNA over ds-RNA/DNA and is consistent with the findings of the previous experiment (Figure 11.b).

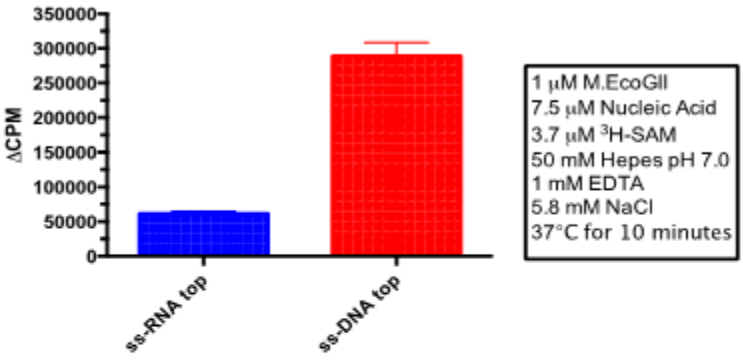
a


ss-RNA top: 5'-UGACAUGAACACAGGUGCUCAGAUAGC-3'
ss-DNA top: 5'-TGACATGAACACAGGTGCTCAGATAGC-3'
ss-DNA bot: 5'-GCTATCTGAGCACCTGTGTTCATGTCA-3'
ds-DNA: 5'-TGACATGAACACAGGTGCTCAGATAGC-3'
3'-ACTGTACTTGTGTCCACGAGTCTATCG-5'
RNA/DNA: 5'-UGACAUGAACACAGGUGCUCAGAUAGC-3'
3'-ACTGTACTTGTGTCCACGAGTCTATCG-5'

b


ΔCPM
60000
50000
40000
30000
20000
10000
0
ss-RNA top
ss-DNA top
ss-DNA bottom
ds-DNA
RNA/DNA hybrid
15 nM M.EcoGII
2 μM Nucleic Acid
3.7 μM ³H-SAM
50 mM Hepes pH 7.0
1 mM EDTA
37°C for 10 minutes

c


ΔCPM
350000
300000
250000
200000
150000
100000
50000
0
ss-RNA top
ss-DNA top
1 μM M.EcoGII
7.5 μM Nucleic Acid
3.7 μM ³H-SAM
50 mM Hepes pH 7.0
1 mM EDTA
5.8 mM NaCl
37°C for 10 minutes

d

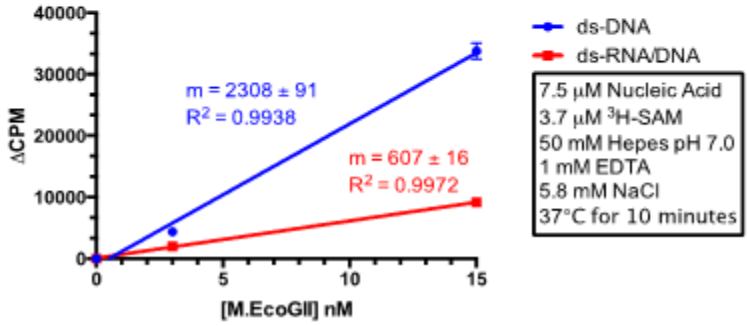

ΔCPM
40000
30000
20000
10000
0
0
5
10
15
[M.EcoGII] nM
m = 2308 ± 91
R² = 0.9938
m = 607 ± 16
R² = 0.9972
ds-DNA
ds-RNA/DNA
7.5 μM Nucleic Acid
3.7 μM ³H-SAM
50 mM Hepes pH 7.0
1 mM EDTA
5.8 mM NaCl
37°C for 10 minutes

Figure 11. **M.EcoGII activity on wide variety of nucleic acid substrates** (a) Synthetic oligonucleotides used for the comparison (b) M. EcoGII's activity on ss-DNA was approximately 7-fold greater than its activity on ss-RNA. M. EcoGII's activity on ds-DNA was approximately 6-fold greater than its activity on ds-RNA/DNA (c) Results of an additional experiment confirming M.EcoGII's preference for ss-DNA over ss-RNA (d) Results of an additional experiment confirming M.EcoGII's preference for ds-DNA over ds-RNA/DNA

3.2.3 M.EcoGII May Exhibit Sequence Preference:

The original purpose of these experiments was to investigate whether M.EcoGII functioned in a processive or distributive manner (Figure 12b-c). In these two experiments, M.EcoGII activity was tested on 3 ss-DNA substrates. One substrate possessed a total of 9 adenines (27-bp), another of 5 (27-bp), and the final of one (29-bp) (Figure 12a). Oligonucleotide concentration was adjusted so that the final adenine concentration was held constant. In one experiment, adenine concentration was held constant at 1 μM, and in the other, at 40 nM. The rationale was that if M.EcoGII functioned in a processive manner, then oligonucleotides possessing a greater number of adenines would exhibit a higher count relative to those with a fewer number; conversely, if M.EcoGII functioned in a distributive manner, then all oligonucleotides, irrespective of their total number of adenines, would exhibit a similar count.

The average ΔCPM for the 9A ss-DNA was 96,000 ± 9,000, 98,000 ± 6,000 for the 5A ss-DNA, and 184,000 ± 8,000 for the 1A ss-DNA (Figure 12b). M.EcoGII's activity on the 9A ss-DNA and 5A ss-DNA was nearly identical (only 2.6% difference). However, M.EcoGII's activity on the 1A ss-DNA was approximately two-fold higher than its activity on both the 9A and 5A ss-DNAs. In Figure 12c, the average ΔCPM for the 9A ss-DNA was 1,870 ± 70, 1,948 ± 290 for the 5A ss-DNA, and 4,010 ± 190 for the 1A ss-DNA. Once again, M.EcoGII's activity on the 9A ss-DNA and 5A ss-DNA was nearly identical (only 4.1% difference). However,

M.EcoGII's activity on the 1A ss-DNA was approximately two-fold higher than its activity on both the 9A and 5A ss-DNAs.

These unexpected results suggest one of two possibilities: (1) that M.EcoGII, while methylating adenines in a sequence non-specific manner, possesses a slight preference for adenines that reside within GCAGC fragments, which is absent in the 9A or 5A substrates, or (2) that M.EcoGII can better latch onto the 29-bp 1A substrate than it can to the 27-bp 5A and 9A substrates. Previous studies have shown that C5 cytosine methyltransferases methylate CpGs at a higher rate when the CpGs reside within longer DNA fragments. (Bestor and Ingram 1983). Taken together, this might indicate that M.EcoGII exhibits a higher activity on longer oligonucleotides and that the level of adenine methyltransferase activity is dependent on DNA length. However, while the substrates are not exactly identical in length (29-bp vs. 27-bp), a 2-bp difference in length is most likely not responsible for the observed difference in activity. The target adenine in the 29-bp oligonucleotide resided within a GCAGC fragment, and this fragment is unique to the 29-bp oligonucleotide (Figure 12a). Therefore, the observed difference in activity is most likely reflected by differences in sequence preference. In order to confirm this, however, binding studies need to be performed, and specific K_D values must be calculated. If GCAGC is indeed preferred by M.EcoGII, then 12-bp oligonucleotides containing this sequence can be synthesized and used for co-crystallization.

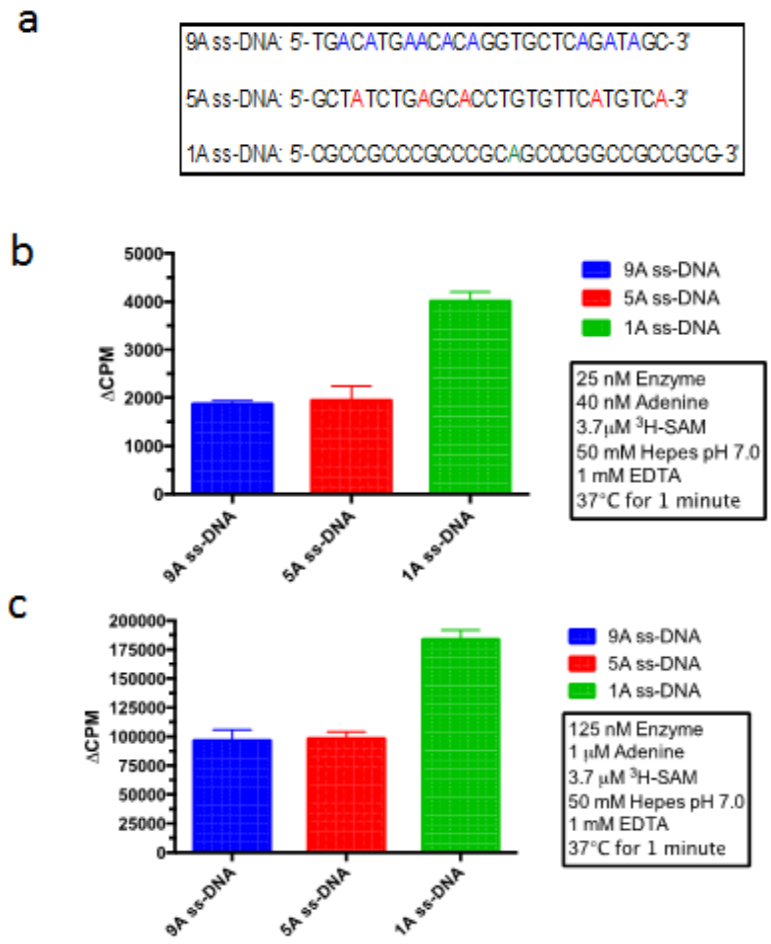


Figure 12. **M.EcoGII Activity on Oligonucleotides with Different Number of Adenines** (a) Synthetic oligonucleotides used for the experiment (b) [Adenine] is 40 nM and [M.EcoGII] is 25 nM (c) [Adenine] is 1 uM and [M.EcoGII] is 125 nM.

3.2.5 $tRNA^{Asp}$ Activity:

In this experiment, M.EcoGII activity was tested on $tRNA^{Asp\text{-}GTC}$ (Figure 13). The effect of secondary and tertiary structure of $tRNA^{Asp\text{-}GTC}$ on M.EcoGII activity was also investigated. Human DNMT2, which methylates the cytosine-38 position of $tRNA^{Asp\text{-}GTC}$ was used as a

control (Goll, M.G., *et al.* 2006)—as a positive control when secondary and tertiary structures were maintained, and a negative control when disrupted. Secondary structure of $tRNA^{Asp\text{-}GTC}$ was disrupted by heating at 98°C for 10 minutes, followed by rapid cooling on ice for an additional 10 minutes. Magnesium is essential for the maintenance of tRNA tertiary structure, as it neutralizes the repulsive forces between negatively-charged phosphates and allows for folding into a compact and stable L conformation (Draper, D. 2004; Römer & Hach 1975). Therefore, 5 mM $MgSO_4$ was used to preserve tertiary $tRNA^{Asp\text{-}GTC}$ structure, and the absence of $MgSO_4$ was used to disrupt tertiary structure.

As expected, human DNMT2 exhibits robust activity on $tRNA^{Asp\text{-}GTC}$ when both secondary and tertiary structures are maintained (Figure 13). Surprisingly, M.EcoGII also exhibits activity when both secondary and tertiary structures are maintained, albeit 2.8-fold less than human DNMT2 (Figure 13). Human DNMT2 activity is almost entirely abolished when both secondary and tertiary structures are disrupted. Unlike human DNMT2, however, M.EcoGII retains residual activity. There is an approximate 2.4-fold reduction in M.EcoGII activity when both secondary and tertiary $tRNA^{Asp\text{-}GTC}$ structures are disrupted relative to when they are preserved (Figure 13). tRNAs possess 4 separate base-paired stems. Having already observed M.EcoGIIs preference for double-stranded nucleic acid substrates (Figure 10b), this relative decrease in activity is possibly due to the disruption of these ds-RNA stem structures.

A review of the literature indicates that most tRNA adenine methyltransferase modify adenine at the N-1 position. To my knowledge, only one characterized methyltransferase (TrmM) modifies tRNA adenine at the N6 position (Dunin-Harkowitz, *et al*. 2006). TrmM methylates the adenine-37 position of $tRNA^{Val}$ (Golovina, *et al.* 2009). This enzyme possesses an NPPY tetra-peptide catalytic motif rather than the consensus DPPY (Dunin-Harkowitz, *et al*.

2006). The modified base m^6A37 of tRNA, located within the anticodon stem-loop (ASL) of $tRNA^{Val}$, participates in a cross-strand stack with G1 of the mRNA codon. The addition of the methyl group to A37 is thought to expand the ring surface area, thus enhancing this cross-strand stacking interaction (Weixlbaumer, *et al*., 2007). It has been postulated that m^6A37 may contribute to anti-codon loop stability (Weixlbaumer, *et al*. 2007). A future objective will be to investigate whether M.EcoGII is active on all tRNAs, not solely $tRNA^{Asp\text{-}GTC}$. It will be of additional interest to investigate whether M.EcoGII is capable of methylating A37, considering m^6A37's possible role in stabilizing codon-anticodon interactions.

Bacterial ribosomes consist of two subunits: the 50S and 30S subunits (Wittmann-Leibold, B. 1986). The 50S subunit, which is responsible for catalyzing peptide formation, is composed of two RNA molecules, termed 23S and 5S rRNA, and 31 distinct proteins (Wittmann-Leibold, B. 1986; Ban, *et al*. 2000). The macrolide-lincosamide streptogramin B (MLS) antibiotics bind to the 50S subunit and inhibit protein translation (Brock & Brock 1959; Mao & Putterman 1969). Adenine 2058 of 23S rRNA resides within the binding site of these antibiotics. Dimethylation of adenine 2058 in 23S rRNA confers resistance to MLS antibiotics by reducing their affinity for the 50S subunit. (Lai & Weisblum 1973; Skinner, *et al.* 1983). To date, numerous 23S rRNA adenine methyltransferase have been discovered and characterized. ErmAM (*S. pneumoniae*), ErmBC (*E. coli*), and ErmC' (*B. subtilis*) all dimethylate adenine 2058 in 23S rRNA and confer resistance to MLS antibiotics (Denoya & Dubnau 1989; Leclercq & Courvalin 1991; Stanislaw, *et al*. 2006; Weisblum, *et al*. 1979). A future objective will be to investigate whether M.EcoGII is capable of methylating adenine 2058 in 23S rRNA. If so, then M.EcoGII may play a role in conferring resistance to macrolide-lincosamide streptogramin B

antibiotics. A structural understanding of M.EcoGII could pave the way for rational design of inhibitors that serve as antibiotic treatment for resistant strains.

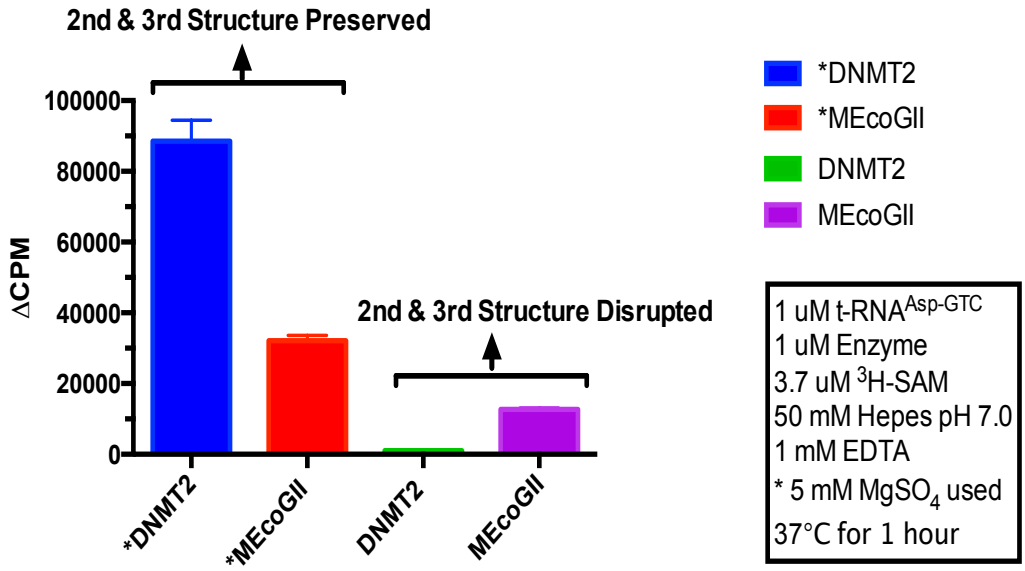


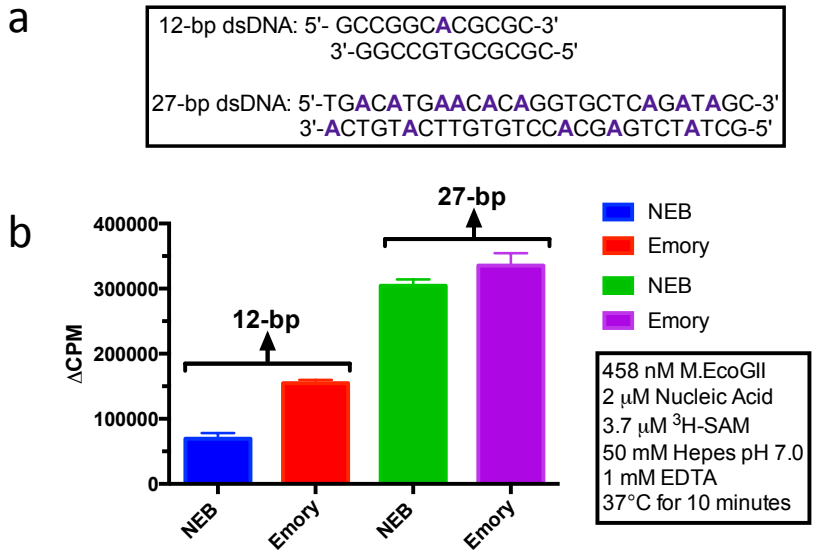
Figure 13. **M.EcoGII Activity on tRNA$^{Asp\text{-}GTC}$.** M.EcoGII exhibits activity on tRNA$^{Asp\text{-}GTC}$ irrespective of secondary and tertiary structure. However, M.EcoGII exhibits greater activity when both secondary and tertiary tRNA$^{Asp\text{-}GTC}$ structure are maintained.

3.2.6 Comparison of Activities of Emory-Purified M.EcoGII to NEB-Purified M.EcoGII:

Using the Emory protocol I developed, a higher level of purity was attained earlier during the purification process. The eluted fractions from both Q/SP and the 2nd Heparin runs were more pure than the NEB-purified enzyme following S200 (Figure 14c). For NEB-purified enzyme, numerous contaminants that span the length of the gel(s) can be visualized. The most conspicuous contaminants include an ~85-kDa and ~29 kDa species. For Emory-purified enzyme, these same contaminants are absent in fractions following both Q/SP and 2nd Heparin

runs. This increased level of purity corresponded to an enhanced level of adenine methyltransferase activity. Methyltransferase activity of Emory-purified M.EcoGII was compared to that of NEB-purified M.EcoGII on two substrates: 12-bp ds-DNA and 29-bp ds-DNA (Figure 14a). The 12-bp ds-DNA substrate possessed one internal adenine site, while the 29-bp ds-DNA possessed 14 total adenines.

The average ΔCPM for Emory M.EcoGII on the 12-bp ds-DNA was 155,000 ± 5,000, compared to 69,000 ± 9,000 for the NEB M.EcoGII (Figure 14b). This amounts to an approximate 2-fold difference in activity. The average ΔCPM for Emory M.EcoGII on the 29-bp ds-DNA was 335,000 ± 20,000, compared to 300,000 ± 10,000 for the NEB M.EcoGII (Figure 14b). This amounts to an approximate 1.1-fold difference in activity. The fold-difference in activity on the 29-bp ds-DNA was not as pronounced as it was for the 12 bp ds-DNA because methyl-accepting sites were approaching saturation.



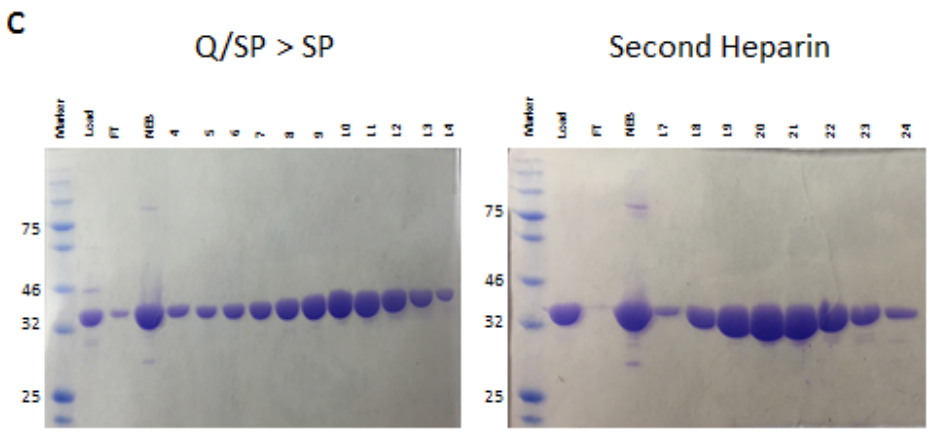


Figure 14. **Comparison of Emory-Purified M.EcoGII to NEB-Purified M.EcoGII** (a) Synthetic oligonucleotides used for the comparison (b) Emory-purified exhibited greater activity on both the 12-bp and 29-bp substrates than did NEB-purified M.EcoGII (c) The eluted fractions from both Q/SP and the 2nd Heparin runs were more pure than the NEB-purified enzyme following S200

3.3 Computational Structural Modeling of AdoMet Binding Pocket:

The modeling study indicated that the methyl donor, AdoMet, could be accommodated in the active site of M.EcoGII. The NH_2 group of the AdoMet moiety interacts via hydrogen bonding with the carbonyl of the aspartic acid side chain in the DPPY catalytic motif (Figure 15b). The hydroxyl group of both S312 and S314, and the amino group of S312 participate in hydrogen bonding with the carboxylic end of the AdoMet molecule (Figure 15b-d). E330 and R335 also interact with the ribose component of AdoMet (Figure 15c, 15e-f). Additionally, the carboxylic acid of D16 interacts with the adenine ring system of SAM at both the N1 and N6 positions (Figure 15b, 15g). In all, these electrostatic interactions help stabilize the AdoMet molecule within the binding pocket of M.EcoGII, thus allowing for transfer of the methyl group CH_3 to the target adenine base of the substrate.

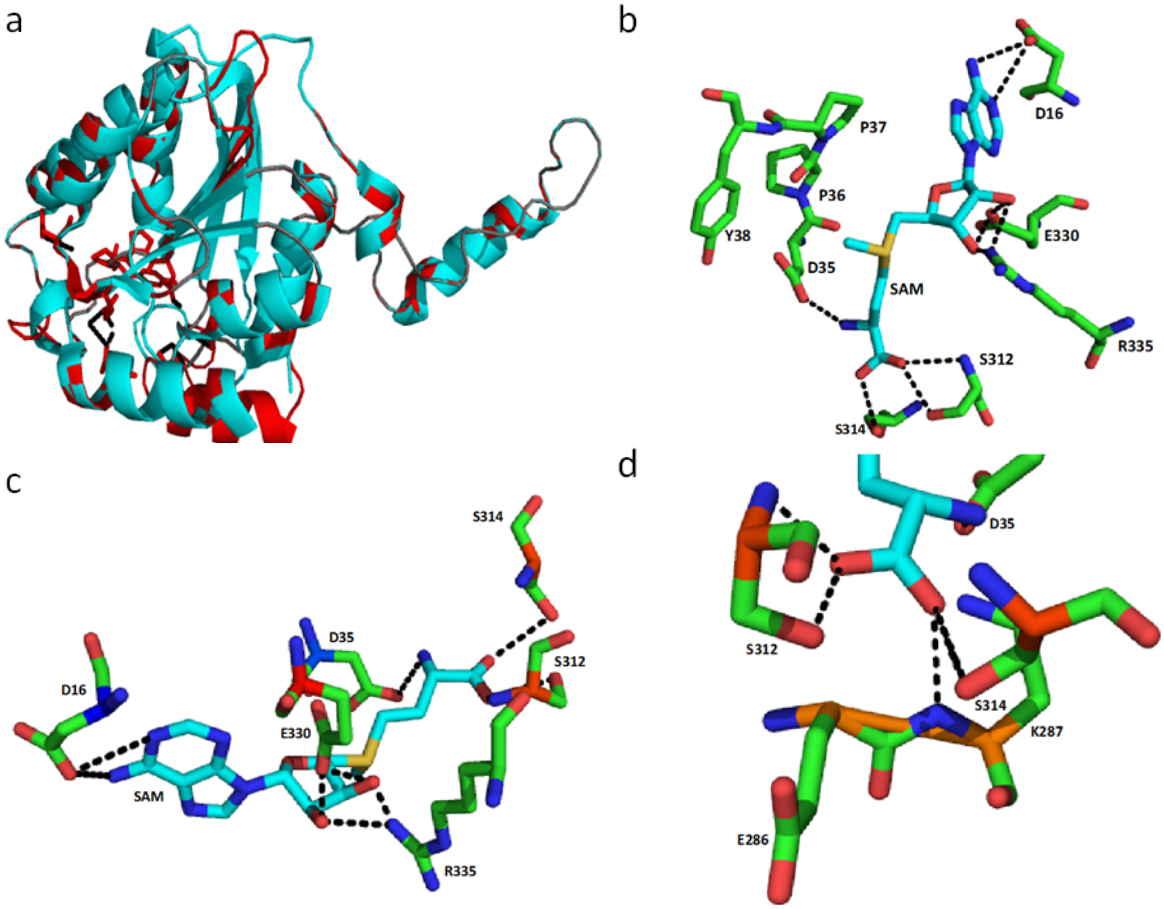


Figure 15. **Structural Modeling of AdoMet/AdoHcy Binding Pocket of M.EcoGII.** The M.EcoGII amino acid residues involved in AdoMet/AdoHcy recognition and binding are shown in green, AdoHcy is shown in cyan, and the hydrogen bonding interactions are represented by black dashes. (a) Superimposition of Moraxella bovis specific adenine methyltransferase (cyan) on M.EcoGII (red) (b) Expanded view of AdoMet-binding pocket with conserved DPPY motif included (c) Expanded view of AdoMet-binding pocket with amino acid residues involved in binding (d) Close view of amino acid residues involved in binding carboxylic acid of AdoMet terminal end

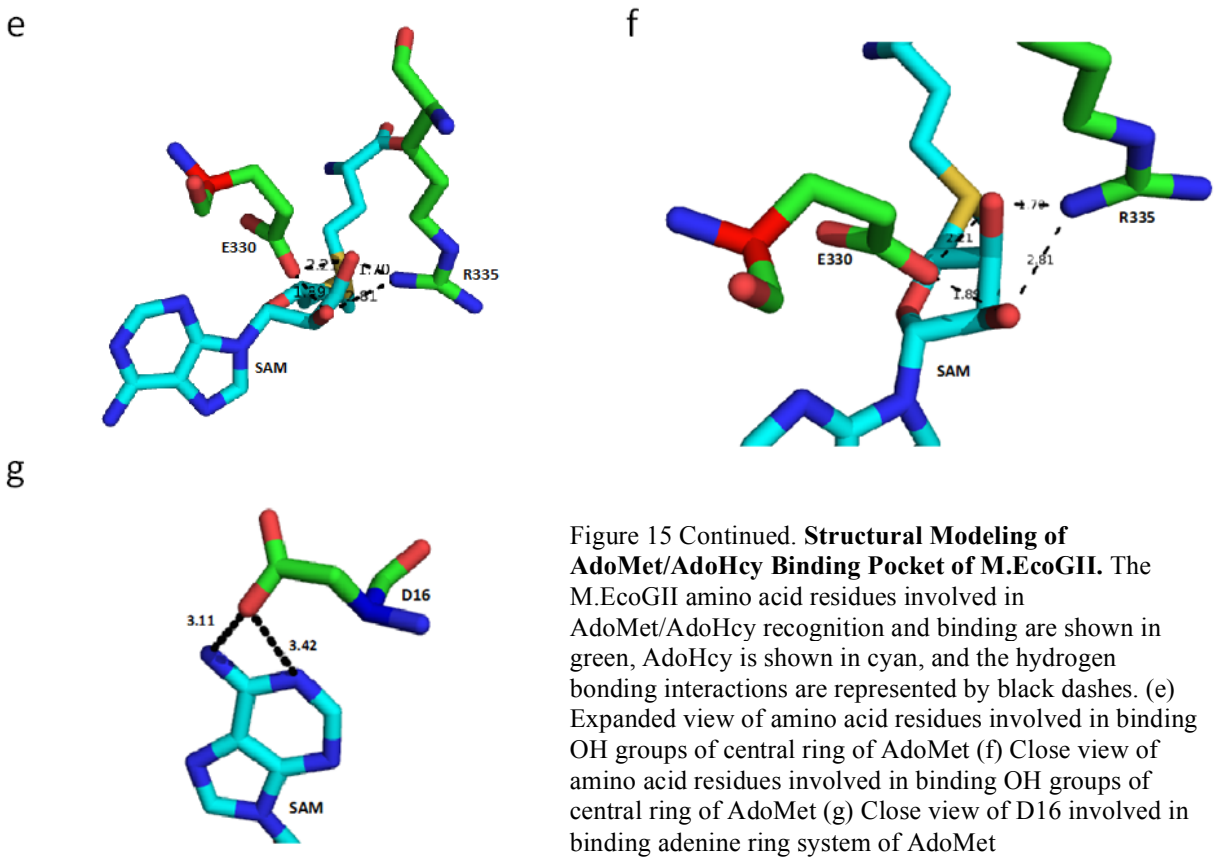


Figure 15 Continued. **Structural Modeling of AdoMet/AdoHcy Binding Pocket of M.EcoGII.** The M.EcoGII amino acid residues involved in AdoMet/AdoHcy recognition and binding are shown in green, AdoHcy is shown in cyan, and the hydrogen bonding interactions are represented by black dashes. (e) Expanded view of amino acid residues involved in binding OH groups of central ring of AdoMet (f) Close view of amino acid residues involved in binding OH groups of central ring of AdoMet (g) Close view of D16 involved in binding adenine ring system of AdoMet

3.4 Crystallization Trails:

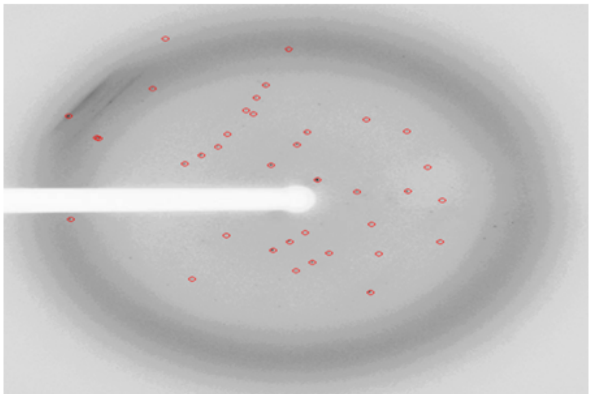
The only macromolecular crystals that grew were obtained from drops containing a complex of M.EcoGII, 29-bp ds-DNA, and AdoHcy (Figure 16a-d). These crystals consistently formed under 0.2 M zinc acetate and 20% (w/v) PEG3350 conditions. These crystals were relatively small and highly birefringent. Each sitting-drop contained in excess of > 20 individual crystals (Figure 16c). The crystals were replicated in both sitting and hanging-drops. Only one data set at 4Å resolution was obtained (Figure 16b). The unit cell dimensions correspond to either a C2 or H3 space group (Table 2). The space group is a description of the underlying unit-cell symmetry.

In 1968 and 1976, Matthews analyzed the solvent content of available crystal structures and observed that it was most common for 43% of the crystal volume to be occupied by solvent (Matthews, B.W. 1968; Matthews, B.W. 1976). This finding led Matthews to develop a novel parameter, V_m, defined as the crystal volume per unit of protein molecular weight (Matthews, B.W. 1968; Kantardjieff and Rupp 2003). Matthews found that the possible fractional volume of solvent within a unit cell exists within a very narrow range for a crystal containing a macromolecule of a given molecular weight. This allowed for one to deduce the probability that a given macromolecule exists within a unit cell based on the predicted solvent content.

I utilized a program developed by Kantardjieff and Rupp applying these basic principles to determine the most probable molecular content of the crystal unit cell (Kantardjieff and Rupp 2003). The probabilities that each of the following macromolecules exist within either the C2 or H3 space group are tabulated (Table 3 and Figure 17). As can be seen in Table 3 and Figure 17, the crystals obtained are most likely either of a binary complex consisting of M.EcoGII and SAH in a C2 space group, or two molecules of 29-bp ds-DNA in a C2 space group. The fact that the crystals were small and highly birefringent is characteristic of DNA. However, the nature of the crystals can only be confirmed after solving the structure via the molecular replacement method.

a

b



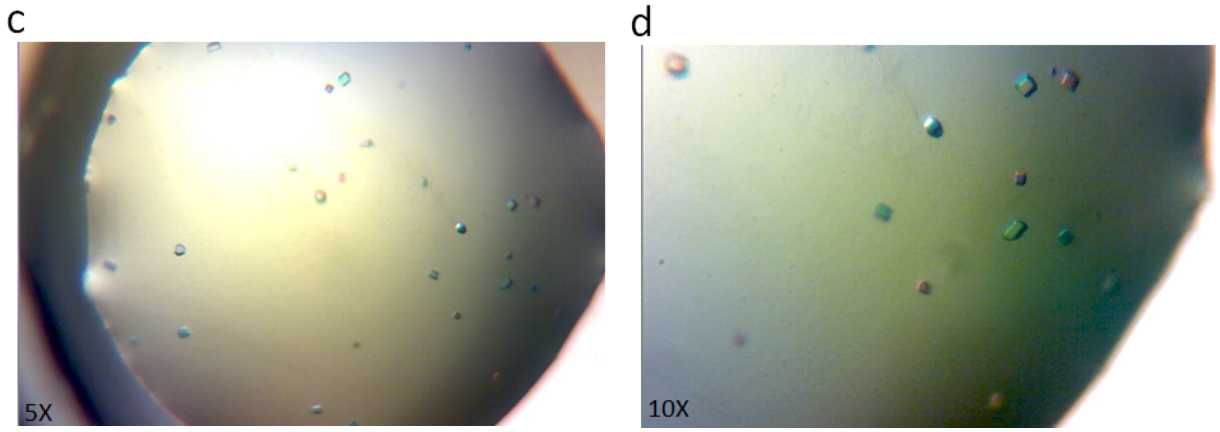


Figure 16. **Crystal Images and Diffraction Pattern** (a) Crystals obtained from 0.2 M zinc acetate and 20% PEG3350 using hanging-drop method. These crystals were used to collect the data set. (b) Diffraction pattern of crystals. (c) 5X magnification of crystals grown in 0.2 M zinc acetate and 20% PEG3350 using sitting-drop method. (d) 10X magnification of crystals grown in 0.2 M zinc acetate and 20% PEG3350 using sitting-drop method.

Space Group	a	b	c	α	β	γ
C2	80.3 Å	46.3 Å	93.7 Å	90°	106.3°	90°
H3	46.3 Å	46.3 Å	269.4 Å	90°	90°	120°

Table 2. **X-ray Diffraction Data Set Collected from SER-CAT 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory and processed using HKL2000**

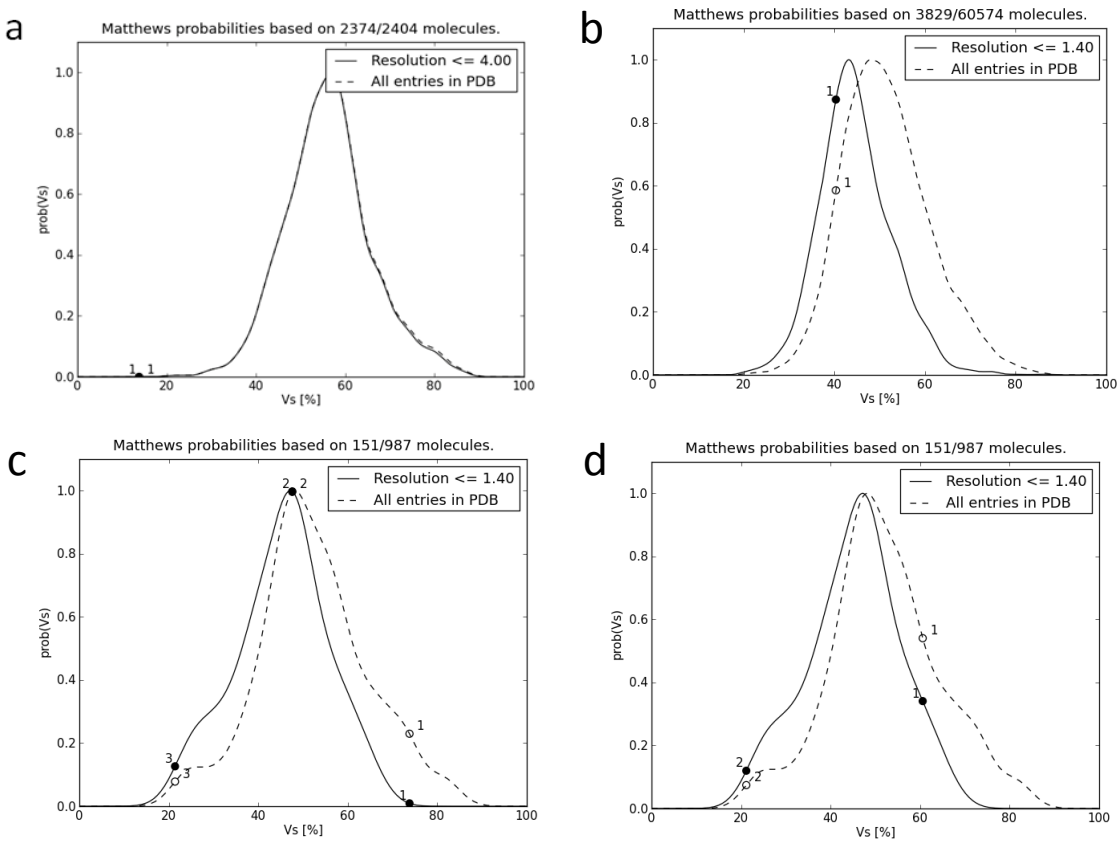


Figure 17. **Matthews Probability Plots** (a) One molecule of M.EcoGII + 29-bp ds-DNA + SAH arranged in C2 space group (b) One molecule of M.EcoGII + SAH arranged in C2 space group (c) Number 1 on the plot corresponds to one molecule of 29-bp ds-DNA arranged in C2 space group, number 2 on the plot corresponds to two molecules of 29-bp ds-DNA arranged in C2 space group, and number 3 on the plot corresponds to three molecules of 29-bp ds-DNA arranged in a C2 space group (d) Number 1 on the plot corresponds to one molecule of 29-bp ds-DNA arranged in an H3 space group, and number 2 on the plot corresponds to two molecules of 29-bp ds-DNA arranged in an H3 space group

Space Group	Macromolecular Content of Unit Cell	Fraction of Unit Cell Volume Occupied by Solvent	Probability
C2	One Molecule of M.EcoGII + 29-bp ds-DNA + SAH	14.15%	~0.0
C2	One Molecule of M.EcoGII + SAH	40.35%	~0.88
C2	One Molecule of 29-bp ds-DNA	73.80%	~0.0
C2	Two Molecules of 29-	47.59%	~0.98

	bp ds-DNA		
C2	Three Molecule of 29-bp ds-DNA	21.39%	~0.10
R3	One Molecule of 29-bp ds-DNA	60.58%	~0.35
R3	Two Molecules of 29-bp ds-DNA	21.17%	~0.10

Table 3. **Matthews Probabilities Based on Fraction of Crystal Volume Occupied By Solvent**.

Appendix

A.1 Need for Additional Assay to Assess M.EcoGII Methyltransferase Activity:

The supply of Whatman™ DE81 cellulose filter papers used to assess methyltransferase activity was exhausted. Additional filter papers could not be purchased because their production was discontinued. Therefore, an alternative assay was needed. Previous reports have demonstrated the use of fluorescence-based, enzyme-coupled assays to quantify thiol reactivity and lysine methylation (Gao, *et al.* 2013; Hudec, *et al.* 2012). Adapting the parameters of these assays to meet the needs of this current study was attempted.

A.2 Theoretical Description of Enzyme-coupled SAHH Fluorescence Assay:

AdoMet (or SAM) is converted into AdoHcy (or SAH) upon completion of the enzyme-catalyzed methyl transfer. S-adenosyl-L-homocysteine hydrolase (SAHH) catalyzes the breakdown of AdoHcy (SAH) into homocysteine (Hcy) and adenosine (Ado) (De La Haba & Cantoni 1959; Turner, *et al.* 2000). Adenosine is an inhibitor of SAHH. Adenosine deaminase converts the liberated adenosine into ionosine (Long, E.R. 1913). The breakdown of adenosine into ionosine serves two purposes. (1) The inhibitory effect of adenosine on SAHH is abated and

(2) The removal of adenosine helps drive the activity of SAHH. The CME contains a thiol-reactive maleimide. The –SH group of Hcy participates in a nucleophilic attack of the maleimide. This Michael addition yields a covalently modified compound with altered photo-physical properties (Yi, *et al.* 2009). The photo-physical properties are altered such that the modified compound can emit at 460 nm upon excitation with 380 nm.

A.3 Cloning and Overexpression of S-adenosyl-L-homocysteine hydrolase (SAHH):

A sequence encoding S-adenosyl-L-homocysteine hydrolase (SAHH) was obtained and cloned into the pET22b vector downstream of a His_6-TEV tag. The His_6-TEV-fused SAHH in the pET22b vector was expressed in *Escherichia coli* BL21 (DE3). Protein expression is under the control of the lactose/isopropyl-β-D-1-thiogalactopyranoside (IPTG) inducible T7 promoter. A single colony was inoculated in 5 mL of LB containing 100 ug mL^{-1} ampicillin at 37°C overnight shaking (215 revolutions min^{-1}). Cell aliquots were then used to inoculate a larger culture, 2 L in total volume. The cells were cultured in LB medium at 37°C until OD_{600} ~ 0.500 was reached. The temperature was then reduced to 16°C, and 2 mL of 0.1M IPTG (0.2 mM) were then added per liter of culture to induce expression of the His_6-TEV-SAHH fusion protein overnight.

A.4 Purification of S-adenosyl-L-homocysteine hydrolase (SAHH):

Induced cells from 2X 1L were harvested at 4000g for 20 min. at 4°C. The supernatant was removed, and the pellet was re-suspended in some residual LB liquid. The cells were transferred to 50 mL tubes and underwent another round of centrifugation at 4000g for 20 min. at 4°C. An OD_{600} of 6.5 cell pellet was recovered and frozen overnight at -20°C. The cell pellet was

thawed, and the cells were lyzed on ice for 10 minutes in 500 mM NaCl, 20 mM Tris-HCl pH 7.5, 5% Glycerol, 10 mM imidazole, 0.5 mM TCEP, and 0.1 mM PMSF protease inhibitor. Lysis was enhanced by one-second strokes of sonication, followed by 2 sec. of off time.

Crude protein extracts were then recovered by 60 minutes of high-speed centrifugation at 16500 rpm and 4°C before purification on **His**Trap™ HP Column (GE Healthcare). The **His**Trap™ HP Column (GE Healthcare) was equilibrated with 2% Buffer B (500 mM NaCl, 20 mM Tris-HCl pH 7.0, 5% Glycerol, 10 mM imidazole, and 0.5 mM TCEP). The protein lysate was loaded onto the HiTrap™ Q HP Column (GE Healthcare), and elution fractions (#22-47) were collected and pooled. TEV protease (500 uM) and NADH (500 uM) were added to the pooled fractions. The pooled fractions were dialyzed overnight at room temperature in 20 mM Tris-HCl pH 7.5, 100 mM NaCl, 1 mM EDTA, 1mM DTT, and 10 mM imidazole and then loaded once again onto an **His**Trap™ HP Column (GE Healthcare). Flow-through fractions (#5-25) were pooled and concentrated in 5K MWCO concentrators at 4°C and ~4000g until a final volume of 2 mL was obtained. The concentrated sample was then loaded on S200 16/60 Size Exclusion Column (GE Healthcare) in 20 mM Tris-HCl pH 7.5, 100 mM NaCl, 1 mM EDTA, and 10 mM imidazole. Final SAHH concentration was determined by absorbance at 280 nm. The elution fractions were collected, pooled, and concentrated in 5K MWCO concentrators at 4°C and ~4000g until an SAHH concentration of 210 uM was obtained.

A.5 Enzyme-coupled SAHH Fluorescence Assay:

The target substrate was incubated with M.EcoGII in the presence of 50 μM coumarin maleic acid ester (CME), 0.5 unit/mL adenosine deaminase (ADA), 10 μM S-adenosyl-L-homocysteine hydrolase (SAHH), 4 mM AdoMet, and 50 mM Hepes pH 7.0. In some

experiments, 10% (v/v) DMSO was also present. Fluorescence measurements were read at room temperature using the Synergy 4 microplate reader (BioTek) with 380 nm excitation and 460 nm emission filters. Data indicate that cysteine residues of M.EcoGII may be interacting with CME, indicating that this assay may not be amenable to methylation reactions dependent on the cofactor AdoMet.

References

Adams, R.L.P., McKay, E.L., Craig, L.M., and Burdon, R.H. (1979). "Mouse DNA Methylase: Methylation of Native DNA." Biochimica et Biophysica Acta **561**: 345-357.

Ban, N., Nissen, P., Hansen, J., Moore, P.B., and Steitz, T.A. (2000). "The Complete Atomic Strcuture of the Large Ribosomal Subunit at 2.4 Å Resolution." Science **289.5481**: 905-920.

Bestor, T.H. and Ingram, V.M. (1983). "Two DNA methyltransferases from murine erythroleukemia cells: purificiation, sequence specificity, and mode of interaction with DNA." Proceedings of the National Academy of Sciences of the United States of America **80.18**: 5559-5563.

Bheemanaik, S., Sistla, S., Krishnamurthy, V., Arathi, S., and Desirazu, Narasimha R. (2010). "Kinetics of Methylation by EcoP1I DNA Methyltransferase." Enzyme Research **Article ID 302731**: 14 pages.

Bickle, T.A. and Kruger, D.H (1993). "Biology of DNA Restriction." Microbiological Reviews **57.2** 434-450.

Blumenthal, Robert M., Cheng, Xiaodong, and Schubert, Heidi L. (2003). "Many paths to methyltransfer: a chronicle of convergence." Trends in Biochemical Sciences **28.6:** 329-335.

Blumenthal, Robert M., Cheng, Xiaodong, and Malone, Thomas. (1995). "Structure-guided Analysis Reveals Nine Sequence Motifs Conserved among DNA Amino-methyl-transferases, and Suggests a Catalytic Mechanism for these Enzymes." Journal of Molecular Biology **253.4**: 618-632

Brock, T.D., and Brock, M.L. (1959). "Similarity in the mode of action of chloramphenicol and erythromycin." Biochimica et Biophysica Acta **33**: 274-275.

Wion, Didier, and Casadesús, Josep. (2006). "N^6-methyl-adenine: an epigenetic signal for DNA-protein interactions." Nature Reviews Microbiology **4**:183-192.

Cheng, Xiaodong (1995). "Structure and Function of DNA Methyltransferases." Annual Review of Biophysics and Biomolecular Structure **24**: 293-318.

De La Haba, G., and Cantoni, G.L. (1959). "The Enzymatic Synthesis of *S*-Adenosyl-L-homocysteine from Adenosine and Homocysteine." The Journal of Biological Chemistry **234**: 603-608.

Denoya, Claudio, and Dubnau, David (1989). "Mono- and Dimethylating Activities and Kinetic Studies of the *ermC* 23 S rRNA Methyltransferase." The Journal of Biological Chemistry **264**: 2615-2624.

Draper, David E. (2004). "A guide to ions and RNA structure." RNA **10.3**: 335-343.

Fang, G., Munera, D., Friedman, D.I., Mandlik, A., Chao, M.C., Banarjee, O., Feng, Z., Losic, B., Mahajan, M.C., Jabado, O.J., Deikus, G., Clark, T.A., Luong, K., Murray, I.A., Davis, B.M., Keren-Paz, A., Chess, A., Roberts, R.J., Korlach, J., Turner, S.W., Kumar, V., Waldor, M.K., and Schadt, E.E. (2012). "Genome-wide mapping of methylated adenine residues in pathogenic *Escherichia coli* using single-molecule real-time sequencing." Nature Biotechnology **30**: 1232-1239.

Gao, T., Yang, C., and Zheng, Y.G. (2013). "Comparative Studies of Thiol-Sensitive Fluorogenic Probes for HAT Assays." Analytical and Bioanalytical Chemistry **405.4**: 1361-1371.

Goedecke, K., Pignot, M., Goody, Roger S., Scheidig, Axel J., and Weinhold, E. (2001). "Structure of the N6-adenine DNA methyltransferase M.*Taq*I in complex with DNA and a cofactor analog." Nature Structural Biology **8**: 121-125.

Goll, M., Kirkepekar, F., Maggert, K., Yoder, J., Hseih, C., Zhang, X., Golic, K., Jacobsen, S., and Bestor, T. (2006). "Methylation of $tRNA^{Asp}$ by the DNA methyltransferase homolog Dnmt2." Science **311**: 395-398.

Golovina, Anna Y., Sergiev, Petr V., Golovin, Andrey V., Serebryakova, Marina V., Demina, Irina., Govorun, Vadim M., and Dontsova, Olga A. (2009). "The *yfiC* gene of *E. coli* encodes an adenine-N6 methyltransferase that specifically modifies A37 of $tRNA_1^{Val}(cmo^5UAC)$." RNA **15**: 1134-1141.

Horton, John R., Liebert, K., Bekes, M., Jeltsch, A., and Cheng, X. (2006). "Structure and Substrate Recognition of the *Escherichia coli* DNA Adenine Methyltransferase." Journal of Molecular Biology **358.2**: 559-570.

Hudec, R., Hamada, K., and Mikoshiba, K. (2012). "A fluorescence-based assay for the measurement of *S*-adenosylhomocysteine hydrolase activity in biological samples." Analytical Biochemistry **433**: 95-101.

Jones, Peter A. and Taylor, Shirley M. (1982). "Mechanism of Action of Eukaryotic DNA Methyltransferase." Journal of Molecular Biology **162**: 679-692.

Kantardjieff, Katherine and Rupp, Bernhard (2003). "Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals." Protein Science **12.9**: 1865-1871.

Lai, Ching-Juh, and Weisblum, Bernard (1973). "Alteration of 23 S ribosomal RNA and erythromycin-induced resistance to lincomycin and spiramycin in *Staphylococcus aureus*." Journal of Molecular Biology **74.1**: 67-72.

Leclercq, R. and Courvalin, P. (1991). "Bacterial resistance to macrolide, lincosamide, and streptogramin antibiotics by target modification." Antimicrobial Agents and Chemotherapy **35.7**: 1267-1272.

Long, E.R. (1913). "The Purines And Purine Metabolism Of Some Tumors In Domestic Animals." The Journal of Experimental Medicine **18.5**: 512-526.

Low, David A., Mahan, Michael J., and Weyand, Nathan J. (2001). "Roles of DNA Adenine Methylation in Regulating Bacterial Gene Expression and Virulence." Infection and Immunity **69.12**: 7197-7204.

Mao, J.C.H., and Putterman, M. (1969). "The intermolecular complex of erythromycin and ribosome." Journal of Molecular Biology **44.2**: 347-361.

Matthews, B.W. (1968). "Solvent content of protein crystals." Journal of Molecular Biology **33.2**: 491-497.

Matthews, B.W. (1976). "X-ray crystallographic studies of proteins." Annual Review of Physical Chemistry **27**: 493-523.

Merkiene, E., Vilkaitis, G., Klimasauskas, S. (1998). "A pair of single-strand and double-strand DNA cytosine-*N4* methyltransferases from *Bacillus centrosporus*." Europe PMC **379**: 569-571.

Nagaraja, Valakunja, and Vasu, Kommireddy. (2013) "Diverse Functions of Restriction-Modification Systems in Addition to Cellular Defense." Microbiology and Molecular Biology Reviews **77.1**: 53-72.

Newby, Zachary .E.R, Lau, Edmond Y., Bruice, Thomas C. (2002). "A theoretical examination of the factors controlling the catalytic efficiency of the DNA-(adenine-N^6)-methyltransferase from *Thermus aquaticus*." Proceedings of the National Academy of Sciences of the United States of America **99.12**: 7922-7927.

Otwinowski, Z., Borek, D., Majewski, W., and Minor, W. (2003). "Multiparametric scaling of diffraction intensities." Acta Crystallographica Section A **59**: 228-234.

Kelley, L.A. and Sternberg, M.J.E. (2015). "Protein structure prediction on the web: a case study using the Phyre server." Nature Protocols **4**: 363-371.

Roberts, Richard J., Vincze, T., Posfai, J., and Macelis, D. (2010). "REBASE—a database for DNA restriction and modification: enzymes, genes and genomes." Nucleic Acids Research **38**: D234-D236.

Römer, Roland, and Hach, Renate. (1975). "tRNA Conformation and Magnesium Binding. European Journal of Biochemistry **55**: 271-284.

Sistla, S., Krishnamurthy, V., and Rao, Desirazu N. (2004). "Single-stranded DNA binding and methylation by EcoP1I DNA methyltransferase." Biochemical and Biophysical Research Communications **314.1**: 159-165.

Skinner, R., Cundliffe, E., and Schmidt, F.J. (1983). "Site of action of a ribosomal RNA methylase responsible for resistance to erythromycin and other antibiotics." The Journal of Biological Chemistry **258**: 12702-12706.

Stanislaw, D., Czerwoniec, A., Gajda, Michael J., Feder, M., Grosjean, H., and Bujnicki, Janusz, M (2006). "MODOMICS: a database of RNA modification pathways." Nucleic Acids Research **34.1**: D145-D149.

Turner, M.A., Yang, X., Yin, D., Kuczera, K., Borchardt, R.T., and Howell, P.L. (2000). "Structure and Function of S-Adenosylhomocysteine Hydrolase." Cell Biochemsitry and Biophysics **33**: 101-125.

The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.

Weisblum, B., Holder, S.B., and Halling, S.M. (1979). "Deoxyribonucleic acid sequence common to staphylococcal and streptococcal plasmids which specify erythromycin resistance." Journal of Bacteriology **138.3**: 990-998.

Weixlbaumer, A., Murphy, Frank V., IV, Dziergowska, A., Andrzej, M., Vendeix, Franck A P., Agris, Paul F., and Ramakrishnan, V. (2007). "Mechanism for expanding the decoding capacity of transfer RNAs by modification of uredines." Nature Structural & Molecular Biology **14**: 498-502.

Wittman-Leibold, B. (1986). "Ribosomal Proteins: Their Structure and Evolution." New York: Springer. p.326-361.

Yi, L., Li, H., Sun, L., Liu, L., Zhang, C., and Xi, Z. (2009). "A Highly Sensitive Fluorescence Probe for Fast Thiol-Quantification Assay of Glutathione Reductase." Angewandte Chemie **48.22**: 4034-4037.