2011

Surface Entropy Reduction to Increase the Crystallizability of the Fab-RNA Complex

Priyadarshini Palaniandy Ravindran
University of Central Florida

SURFACE ENTROPY REDUCTION TO INCREASE THE CRYSTALLIZABILITY OF THE FAB-RNA COMPLEX

by

PRIYADARSHINI PALANIANDY RAVINDRAN
B.Tech. Anna University, 2008

A thesis submitted in partial fulfillment of the requirements
for Masters in Molecular and Microbiology
in the Burnett School of Biomedical Sciences
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Summer Term
2011

Major Professor: Jingdong Ye

ABSTRACT

Crystallizing RNA has been an imperative facet and a challenging task in the world of RNA research. Assistive methods such as Chaperone Assisted RNA Crystallography (CARC), employing monoclonal antibody fragments (Fabs) as crystallization chaperones have enabled us to obtain RNA crystal structures by increasing the crystal contacts and providing initial phasing information. Using this technology the crystal structure of ΔC209 P4-P6 RNA (an independent folding domain of the self-splicing *Tetrahymena* group I intron) complexed to Fab2 (high affinity binding Fab) has been resolved to 1.95 Å (1). Although the complexed class I ligase ribozyme has also been crystallized using CARC (2), in practice, it has been found that the crystallization of, large RNA-Fab complex remains a confrontation. The possible reason for this difficulty is that Fabs have not been optimized for crystallization when complexed with RNA. Here we have used the Surface Entropy Reduction technique (SER) for the optimization process. Candidate residues for mutations were identified based on combining results from visual inspection of ΔC209 P4-P6/Fab2 crystal structure complex using pyMOL software and a web-based SER software. The protruding lysine and glutamate residues were mutated to a set of alanine (Super Mutant Alanine SMA) and serine (Super Mutant Serine SMS) mutant clones. Filter binding assay studies confirmed that the mutant clones bind to ΔC209 P4-P6 with similar binding affinities as that of the parent Fab2. Large scale expression of the mutants, parent clone and ΔC209 P4-P6 RNA were optimised. Crystal trays for ΔC209 P4-P6 complexed with Fab2, Fab2SMA and Fab2SMS were set-up side-by-side using Hampton crystal screen kits and ~600 conditions including temperature as a variable condition were screened. Crystal screening shows significantly higher crystal-forming ratios for the mutant complexes. As the chosen SER residues

are far away from the CDR regions of the Fab, the same set of mutations can be potentially applied to other Fabs binding to a variety of ribozymes and riboswitches to improve the crystallizability of the Fab-RNA complex.

ACKNOWLEDGMENTS

I take this opportunity to express my heartfelt gratitude to my mentor Dr. Jingdong Ye, PhD for being my guiding beacon throughout the project. I sincerely thank my committee members Dr. James Turkson, Dr. Sean Moore and Dr. Shaw Li for their unflagging support and guidance. I extend my thanks to my lab mate Breena Stoner for constructing the VCIII binding parent Fabs. I wish to thank Eileen, Joe, Jackie, RJ, Sean, Neesha, Cody, Mohammed, Galal and Brian for helping me in the course of the project and for being a great work group. I thank my friends Maha, Prabhu, Soumya, Akansha, Anusha, Gowri, Nithya, Vidusha, Priya, Ramya, Arthi, Sathish, Supreeth, Bihag, Sai, Pavi, CK, Jagadeesh, Suny, Bharathi, Pranav, Swetha, Pankaj, Guatam, Bala, Vineet, Vanathy, Sathya, and Rucha from the bottom of my heart for their undying affection and moral support. I thank Shiva and Vani for their constant support and encouragement. I thank my family for being the protective shield in every walk of my life. Above all I thank the Lord almighty for showering me with his countless blessings.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 INTRODUCTION

With every new discovery of a functional RNA, there is a sprint to obtain its crystal structure and elucidate the structure-function relationship. The dynamic RNA research is mired by the difficulty in obtaining RNA crystals due to its structural properties. The negatively charged phosphate backbone of RNA disfavors good crystal lattice formation. Compared to proteins, RNAs have fewer functional groups available to make crystal contacts. In addition, the generally weaker and flexible tertiary structure of RNAs may lead to inter-domain movements within RNA, making crystal packing all the more difficult (3). Assistive methods for crystallization are in place to alleviate these complications (4,5). Chaperone Assisted RNA Crystallography is one such assistive method where a monoclonal antibody fragment binds to RNA of interest and aids in crystallization. Antigen-binding fragments Fabs selected from a synthetic phage display library can bind to large functional RNAs. These Fabs with large surface area participate extensively in crystal contacts formation and provide good initial phasing information (1,2).

YSGX synthetic superlibrary is a reduced genetic codon library, which contains randomized CDRs especially enriched in tyrosines and serines (6). This was used to obtain the first RNA-Fab complex crystal structure (1.95 Å) of ΔC209 P4-P6 binding to Fab2 (1). ΔC209 P4-P6 RNA is a single site mutant form of the P4-P6 of *Tetrahymena* group I intron which catalyses its own excision from precursor RNA. The single site mutation increases the tertiary stability by reducing the conformational flexibility of the P4 helix. This 159 nt mutant structure was resolved to 2.25 Å vs the wild type of 2.8 Å (7,8). Using ΔC209 P4-P6 RNA as a proof-of-concept model

system, Ye and coworkers have selected specific anti-RNA Fabs and demonstrated the co-crystallization of ΔC209 P4-P6 RNA in complex of Fab2. This crystal structure further improved the resolution of the ΔC209 P4-P6 RNA to 1.95 Å. Fab chaperone showed great phasing capacity. The electron density map generated by molecular replacement using Fab solution alone can be used to build ΔC209 P4-P6 models. The crystal structure of ΔC209 P4-P6/Fab2 revealed extensive crystal contacts involving Fab-Fab and Fab-RNA intermolecular interactions. Fab contributed to 61% of the total surface area buried by crystal lattice interactions. These features validate the Fabs as good crystallization chaperones.

Class I ligase ribozyme is another functional RNA which has been crystallized in complex of the Fab chaperone. Its cognate Fab BL3-6, was selected against the class I ligase using YSGRX synthetic Fab library with reduced codon enriched in tyrosines, serines, glycines and arginines (2). The structure of the class I ligase/BL3-6 complex was resolved to 3.1 Å. The crystal structure reveals that binding of Fab to the class I ligase doesn't change its catalytic function. Molecular replacement using Fab co-ordinates provided sufficient initial phasing information to resolve the RNA structure. Fab BL3-6 also involved extensively in the crystal contact with the RNA. Together with the Fab-RNA contacts of the original complex, Fab BL3-6 mediated crystal contacts account for 78.7% of buried surface area in the structure.

In both Fab mediated ΔC209 P4-P6 and class I ligase ribozyme crystal structures, the cognate Fabs showed great phasing value and extensive crystal contact participation, corroborating Fabs as general crystallization chaperones for RNA targets.

Despite these favorable features of Fabs as vital crystallization chaperones, crystallizing large RNAs is still an impediment. For instance, although ΔC209 P4-P6/Fab2 complex has been

crystallized at 4 ^{o}C, there were no crystal hits at 20 ^{o}C using Hampton Crystallization Screening kits. A possible reason is that although Fab in general is a great crystallization module, its crystallization capability was not optimized when in complex with RNA molecules. The nucleation event and crystal growth kinetics of Fab-RNA complexes may differ substantially from that of Fab alone or RNA alone. As an empirical experience, it is observed that Fab-RNA complexes in general are more soluble than RNA alone and less soluble than Fab alone. As a consequence, Fab-RNA complexes precipitate out more frequently than Fab alone and less frequently than RNA alone. It is reasoned that by providing large surface area to bind to large RNA antigen molecule, much of the Fab surface is inaccessible to make crystal contacts either directly or indirectly masked by RNA molecules through steric hindrance. Therefore, optimizing the crystallizability of the Fab module in complex of RNA may provide a solution to the low crystal hit rate during crystal screening of the Fab-RNA complexes.

Developed by the Derewenda group, surface entropy reduction generates 'low-entropy' surface patches through site-directed mutagenesis to develop crystallizable variants (9-12). To better explain this technique, it is important to understand the thermodynamics of crystallization. Crystallization is a thermodynamically costly process. $\Delta G = \Delta H - T(\Delta S_{protein} + \Delta S_{solvent})$ represents the free energy change that drives crystallization. Enthalpy ΔH is considered to be a very small value, when compared to the entropy changes of solvent and protein. Protein molecules incorporated in lattices confers a negative entropy term disfavoring crystallization. The release of structured water molecules around the hydrophilic and hydrophobic patches on protein surface can contribute to positive entropy and favor protein crystallization. Amino acids with large flexible and charged side-chains (such as lysines and glutamates) are detrimental to the

crystallization process for two reasons. Large residues have high conformational entropies which will be lost if they are directly involved in the crystal contact. In addition, large residues are less likely to organize solvent molecules whose release is the major driving force of the crystallization process (9,11,13-16). Thus, surface entropy reduction enhances the crystallizability of proteins by rational mutagenesis of surface residues to decrease the surface entropy. The choice of lysines and glutamates pertains to the fact that they are residues found to be over-represented as exposed on protein surface. About 6% of lysine residues are found to be buried in proteins while 68% is exposed with its amino group. The rest of the 26% is found to be partly exposed. Only 12% of glutamic acids are found to be buried. Both the amino acids account for a considerable proportion of total amino acid content of many proteins (17). Their side chains have high conformational entropy, ~2 kcal/mol (under normal conditions) supporting the fact that high entropy side chain residues remain on the protein surface (18,19). In addition, the atomic structure of protein-protein interfaces show a stark depletion of lysine and glutamate (20). By replacing surface lysines and glutamates to alanines, surface entropy reduction has been used successfully to improve the crystallizability of many proteins and complexes (9,11).

We will discuss the competing methods of engineering crystal contacts and SER approach. SER has been applied in this study to improve the crystallizabity of Fab/RNA complex. The expression of the mutated Fab proteins was optimized and the purified Fabs were complexed with ΔC209 P4-P6 and screened for crystallization. Greater crystal forming rate was observed with the mutant Fabs obtained through the surface entropy reduction method.

CHAPTER 2 MATERIALS AND METHODS

2.1 Single strand DNA synthesis

Double stranded Fab2 plasmid was electroporated into CJ236 cells (uracil deglycosidase deficient *E.coli*) on a micropulser electroporator (Bio-Rad). Single colony from the transformation plate was used to inoculate 1 ml 2YT starting culture with 100 μg/ml ampicillin and 10 μg/ml chloramphenicol and this culture was shaken at 37 ^{o}C for 6 hrs. M13 helper phage (~10^{10} pfu/ml) was added to the slightly cloudy culture, and shaken for 10 min at 37 ^{o}C. 300 μl of the mixture was transferred to 30 ml 2YT with 100 μg/ml ampicillin and 0.25 μg/ml uridine and shaken for 18 hrs at 37 ^{o}C. The phage solution was purified and the uracil-containing ssDNA was isolated with the Qiagen M13 purification kit. The concentration of the eluted ssDNA was determined using a spectrophotometer reading and an agarose gel was run to verify the ssDNA band against a positive control.

2.2 Kunkel Mutagenesis

The following Kunkel primers were designed for construction of Fab2SMA: CTTCATCTTCCCGCCATCTGATGCCCAGTTGAAATCTGGAACTGC (for LC-E123A), AGCAGACTACGAGAAACACGCCGTCTACGCCTGCGAAGTA (for LC-K190A), AGCAACACCAAGGTCGACGCCGCCGTTGCCCCCAAATCTTGTGACAAAACTCACAC ATAGGGCCGGCCCTCTGGTTCC (for HC-K217A, HC-K218A, HC-E220A, and HC-C230stop). Underlined are mutation sites, with GCC for alanine and TAG for amber stop codon. For Fab2SMS, the primers are essentially the same except that the underlined GCC was replaced by AGC to code for serine. The three primers were phosphorylated individually using T4

Polynucleotide kinase (NEB, 10U/µl) at 37 ^{o}C for ~1hr. The primers were annealed to the single strand template at 90 ^{o}C for 1 min, 50 ^{o}C for 3 min and placed on ice. The daughter strand extension was performed at 37 ^{o}C for 1.5 hrs using T7 DNA polymerase and T4 DNA ligase. PCR cleanup using Qiagen kit was performed and the Kunkel DNA was transformed into XL-1blue cells (uracil glycosidase containing strain) by heat shock transformation. Small scale DNA extraction was performed on the XL-1blue cells and sent out for DNA sequencing (Genewiz Inc.). Stop codon was incorporated into the ΔC209 P4-P6/Fab2 gene using Kunkel mutagenesis to express parent ΔC209 P4-P6/Fab2 which will be used as a positive control in crystallization. Kunkel mutagenesis was confirmed using colony PCR.

2.3 Fab expression of mutant clones and parent clone

Following the optimized protocol a starter culture was inoculated with Fab plasmid containing *E coli* 34B8 strain cells at 30 ^{o}C and 300rpm O/N in 25 ml 2YT and 100 µg/ml ampicillin. Spectrophotometeric reading OD at 600nm was measured and glycerol stock was prepared. The 5 ml of starter culture was used to inoculate 500 ml of a specialised media CRAP (27 mM $(NH_4)_2SO_4$, 2.4 mM sodium citrate-$2H_2O$, 14 mM KCl, 5.4 g/liter yeast extract, 5.4 g/liter HyCase SF Casein, 0.11 M Mops– NaOH buffer (pH 7.3), 0.55% (w/v) glucose and 7 mM $MgSO_4$) supplemented 100 µg/ml of ampicillin and grown at 30 ^{o}C and 300 rpm for 24 hrs. OD@600nm was checked (should be >5 with good observable foaming) and the cells were spun down at 6000~8000 rpm for 10 min at 4 ^{o}C and the pellets were frozen at -80 ^{o}C O/N.

2.4 Purification of Fab2 proteins

2.4.1 Protein A column-affinity chromatography

The frozen cell pellets were thawed and lysed with lysozyme (0.3 mg/ml), Dnase I (2.5mg/30ml) in 30 ml Lysis buffer (50 mM Tris pH 8.0, 300 mM NaCl, 0.5 mM PMSF) with vigorous stirring for 1hr at room temperature. The cell debris was spun down at 15000 rpm for 45 min at 4 ^{o}C. The supernatant was loaded onto the 1ml protein A column at 1ml/min. The column was washed with 50 column volume (CV) of 10 mM Tris pH 7.5, 500 mM NaCl at 1 ml/min. The column was eluted with 0.1 M citric acid, pH 3, and the eluate was collected in 3 ml per fraction in tubes prefilled with 1.5 ml 1 M Tris pH 8.0, (mixed immediately after elution). After all the protein was eluted out (checked the eluted fractions on UV spectrometer), the column was regenerated with 3 CV of 0.1 M H_3PO_4, and equilibrated with 10 CV of TBS at 1 ml/min (5 ml/min for 5 ml column). The eluted fractions were run on a SDS-PAGE gel with a standard protein marker and the fractions of interest were pooled together.

2.4.2 High S resin purification-Ion exchange chromatography

Fabs purified from protein A column were pooled and buffer exchanged (dialysis or Amicon concentration) into Low Salt NaOAc Buffer C (10 mM NaOAc, pH 5.0, 50 mM NaCl). High S resin suspension of 1ml was added (for 2mg < Fab <5 mg, 2 ml suspension was used, leading to 1 ml CV) and the column was washed with 20 CV of Low Salt NaOAc Buffer C. The Fab solution was loaded onto the column, and the flow through was collected. The column was eluted in a stepwise manner with an increasing salt gradient; 5 CV of Low Salt NaOAc Buffer C

(final: 50mM NaCl), 90%C/10%D (final: 100 mM NaCl), 80%C/20%D (final: 150 mM NaCl), 70%C/30%D (final: 200 mM NaCl), 60%C/40%D (final: 250 mM NaCl), 50%C/50%D (final: 300 mM NaCl), 40%C/60%D (final: 350 mM NaCl), or until no protein came out (checked the OD at 280nm). The eluate was collected in small fractions (<=4 ml and <=5 CV) and ran on a SDS-PAGE. The clean Fab fractions were buffer exchanged into Fab Storage Buffer (10 mM Tris pH 7.5, 50 mM NaCl), aliquoted and stored in -20 ^{o}C.

2.5 Large scale production of ΔC209 P4-P6 RNA

The ΔC209 P4-P6 plasmid was transformed into JM109 competent cells by heat shock transformation. Glycerol stocks were prepared with 750 µl overnight cell culture and 150 µl 80% glycerol, mixed and stored at -80 ^{o}C. Large scale DNA extraction was done using Qiagen Maxiprep kit. The concentration of the extracted DNA was measured spectrophotometrically. To linearize the extracted plasmid a large scale Ear I digestion was carried out with100 µg plasmid, 10x NEB buffer 4, and Ear I [final: 0.2 U/(µg plasmid)] at 37 ^{o}C, O/N. A 2% agarose gel was run with a 100bp ladder to check digestion. The digested DNA was PCA extracted and ethanol precipitated to remove salts, enzymes as well as RNAse A.

2.5.1 Large scale *in vitro* transcription of ΔC209 P4-P6 RNA

A typical reaction of 12 ml transcription was carried out with 10x RNA Pol Buffer (for 1x: 40 mM Tris pH 7.9, 6 mM MgCl2, 2 mM spermidine-3HCl, 10 mM DTT), 1M MgCl2, 25 mM NTPs, pH 7.7 (contains 25 mM of ATP, CTP, UTP, GTP), 1ug/ul Ear I linearized plasmid (final: 50ug plasmid/ml), autoclaved miliQ water, 4 U/µl RNase Inhibitor from Qiagen (final: 0.01U/ul), Tippase (final: 0.008 U/µl) and 1.13 µg/µl T7 RNAP (final: 45µg/ml) at 37 ^{o}C,

O/N. A 2% agarose gel was run to check the transcription. DNase I (RNase free) and CaCl2 (final 1mM), was added to final concentration 0.5 Unit per μg of Plasmid DNA and incubated at 37 ^{o}C for 30 min. Amicon filter 10 kD MWCO 50 ml falcon was used to buffer exchange the ΔC209 P4-P6 RNA into 2 mM sodium citrate, pH 6.5, and concentrated to 1/10 original transcription volume (1.2 ml). A slab gel of 10% dPAGE was prepared with, 3mm x 10cm comb. Equal volume of 2X (50 mM Tris-HCl pH 8, 1 mM EDTA pH 8, 7 M Urea, 0.005% g/ml Bromophenol blue and 0.005% Xylene cyanol) RNA loading buffer was added to the transcription mixture, loaded onto the slab gel and ran at constant power of 15 Watts for 1.5 hrs. The RNA band was detected by placing the slab gel on a TLC plate with a hand held UV lamp. The RNA band was cut and eluted via electroelution in TE buffer, 50 volts for 12hrs at 4 ^{o}C. The elutant was concentrated to 1/10 original TXN volume, PCA extracted twice and ethanol precipitated immediately. The RNA was re-dissolved in TE buffer (10 mM Tris, pH 8, 1 mM EDTA), filtered using, Microcon centrifugal filter tubes, aliquoted and stored in -80 ^{o}C**.**

2.6 ^{32}P labeling of ΔC209 P4-P6 RNA

2μl 20μM transcript RNA (40 pmol), 1 μl 10x Shrimp alkaline phosphatase buffer volume, and lμl Shrimp alkaline phosphatase (SAP) were incubated at 37 ^{o}C, 30 min and followed by 15 min at 65 ^{o}C to inactivate the SAP. 10x T4 Polynucleotide kinase buffer, 2 μl γ ^{32}P ATP and T4 PNK were added and incubated at 37 ^{o}C for 30 min. An analytical dPAGE 8% was used to purify ^{32}P-labeled RNA, using ^{32}P labeling gel box (plate 14 x 20 cm). The gel was run at constant power of 15 Watts for 1.5 hrs. The gel plate was taken out of the gel box, rinsed, blot dried and covered with saran wrap. Position markers were used on the gel outside of the

sample lanes to mark the position of the ^{32}P labeled band. In dark room, one x-ray film was held on top of the on the gel and exposed for 30 sec. The X-ray film was developed on the film developer. The developed film was aligned on the gel according to the position markers and the gel band was marked in the back of the bottom plate. The gel band was cut and eluted in TE buffer on a rotator in the cold box O/N. The ^{32}P labeled ΔC209 P4-P6 was PCA extracted twice and ethanol precipitated and saved in ^{32}P coffin at -20 ^{o}C.

2.7 Dot Blot Assay

Dilutions of 50 µl 20 µM, 10 µM, 5 µM, 2.5 µM, 1.25 µM, 0.625 µM, 0.3125 µM, 0.156 µM, 78 nM, 39 nM for each Fab (dilute in 1x PBS (137 mM NaCl, 0.3 mM KCl, 0.8 mM Na_2HPO_4, 0.15 mM KH_2P0_4, pH 7.2)) in 96 well plate were made and saved at 4 ^{o}C. A BA85 nitrocellulose (S&S) and Hybond N+ filter (Amersham Pharmacia) were pre-equilibrated with 100 µl PEM (1x PBS, 0.1 mM EDTA, 10mM $MgCl_2$), 1 mM EDTA, 10 mM $MgCl_2$ for ΔC209 P4-P6 binding Fabs and incubated for 15 min at room temperature. For VCIII (glycine riboswitch), binding Fabs, 10 mM glycine was added to the binding solution prior to the incubation. The P^{32}labeled ΔC209 P4-P6 (~650 kcpm) was folded in the PEM binding buffer by incubating at 50 ^{o}C for 10 min and another 10 min incubation at room temperature. RNase IN and 1 mg/ml Heparin were added and mixed well. The ΔC209 P4-P6/Fab2 and mutant clones were added to a new microtitre plate and their duplicates for each Fab dilution were made. The folded ΔC209 P4-P6*RNA sample at a concentration of ~0.3 nM was added to each Fab dilution ranging from 4 to 2000 nM and incubated for 30 min at room temperature. The dot-blot apparatus (Bio-Rad) was assembled as per the design protocol with nitrocellulose membrane on

top of Hybond membrane. After the vacuum was applied, wells were rinsed with 100 µl PEM binding buffer and the Fab-RNA solution (total volume of 40 µl) was transferred to the corresponding wells on the dot-blot apparatus. The wells were re-washed with cold binding buffer after the Fab-RNA solution had drained. The membranes were air dried and exposed to PhosphorImage screen cassette O/N. The PhosphorImage screen was scanned and the data was quantified using ImageQuant (Molecular Dynamics). Binding constants were calculated by fitting the data to the following equation: fraction bound = $M \times [Fab]^n/(K_d^n + [Fab]^n)$, where K_d is the binding constant; M is the maximum fraction of RNA bound at the highest Fab concentration; and n is the Hill coefficient.

2.8 Large scale expression of ΔC209 P4-P6/Fab2, ΔC209 P4-P6/Fab2SMA and ΔC209 P4-P6/ Fab2SMS

The Fab expression and purification were scaled up for crystallization. The Fab clones were expressed as described in the Fab expression protocol (6 L CRAP media expression) and purified by Protein A column (5 ml column for every 3 L culture) and two rounds of High S resin purification. The High S resin CV was scaled up as per Fab yield after Protein A purification. The purity of the Fabs were checked using the nuclease test assay and then buffer exchanged into Fab storage buffer and stored at -80 ^oC.

2.9 Nuclease test assay

The nuclease activity test for the three Fabs were conducted by incubating the 10 µM Fab fractions after second round of high S with a ^{32}P-labeled unrelated RNA in this case, 1 µl (~130 kcpm) *VCLD1 for 45 mins at 37 ^oC and ran on a denaturing PAGE (pre-run for 15 mins) after

mixing with equal amounts of 2x RNA loading dye (50 mM Tris-HCl pH 8,1 mM EDTA pH 8,7 M Urea, 0.005% g/ml Bromophenol blue and 0.005% Xylene cyanol). Water, TE buffer and the Fab storage buffer were used as negative controls. The gel was run at constant power of 15 Watts for 40 mins and dried in the gel drier. The dry gel was exposed to PhosphorImager screen O/N and the percentage of intact RNA was quantified using ImageQuant (Molecular Dynamics).

2.10 Mobility shift Assay

ΔC209 P4-P6 was folded as following: 1 µl 40 µM ΔC209 P4-P6, 6.5 µl 10mM Tris 7.5, 1 µl 250mM $MgCl_2$, 1.5µl 1M NaCl was incubated at 50 ^oC, for 20min and set at room temperature for 10min. To prepare the Fab/RNA complex, 1 µl 40µM ΔC209 P4-P6, 5.7 µl 10mM Tris 7.5, 1 µl 250mM $MgCl_2$, 1.5 µl 1M NaCl was incubated at 50 ^oC for 20 min and set at room temperature for 10 min. 0.8 µl 40 µM ΔC209 P4-P6 Fab2/ΔC209 P4-P6 Fab2 mutant was added and incubated at room temperature for 30 min. Subsequently, ΔC209 P4-P6 was run with ΔC209 P4-P6/Fab2, ΔC209 P4-P6/Fab2SMA and ΔC209 P4-P6/Fab2SMS complexes at different stoichiometric ratios of 1, 1.1, 1.2, 1.3 and 1.4 equivalents on a 8% native gel (8% PAGE (29:1 acrylamide:bisacrylamide), 5 mM $MgCl_2$ filtered through 0.4µ filter). The samples were run using 5x native gel loading dye, on a 8% native gel with Mg at room temperature, constant 120V for 40min. The gel was stained with ethidium bromide solution and visualized with UV illuminator.

2.11 Crystal screening of ΔC209 P4-P6/Fab2 mutant complexes using sitting drop method.

The folding conditions for 1 equivalent of ΔC209 P4-P6 complexed with 1.1 equivalent of Fab clone are as follows: ΔC209 P4-P6 (1 Eqv), 25 mM $MgCl_2$, 50 mM NaCl, 10 mM Tris

HCl 7.5, were incubated at 50 ^{o}C for 20 min and the Fab clone (1.1 Eqv) was added along with 0.5 mM Spermine·4HCl, RNAse inhibitor and incubated at room temperature for 30 min, spun down and the supernatant was collected. The drop volume used was 1µl comprising of 0.5 µl of sample and 0.5 µl of reservoir buffer. There were 3 crystal screens used, Hampton Crystal Screens I&II, Hampton Natrix Screens I&II and Hampton Index Screens with a total of over 300 conditions. Crystal screens were carried out under both 20 ^{o}C and 4 ^{o}C conditions.

The optimization on crystal trays was done by the hanging drop method with the buffer and reservoir solution in 1:1 ratio with a drop volume of 2 µl on a siliconized slide sealing a reservoir volume of 500 µl. The crystals from optimized trays were fished out into a cryoprotectant and frozen in liquid nitrogen and shipped through dry shipper to the Light synchrotron service at Brookhaven National Laboratory at New York. The composition of cryoprotectant for two optimized conditions are as follows 0.2 M DL-Malic acid pH 7.0, 22% w/v PEG 3, 350, 20% glycerol, 30mM MgCl2 and 0.5mM Spermine·4HCl (Figure 11A). 10 mM Tris pH 7.5, 2.0 M NaCl, 10% w/v PEG 6000, 20% glycerol, 30mM MgCl2 and 0.5mM Spermine·4HCl (Figure 11B).

CHAPTER 3 RESULTS AND DISCUSSION

3.1 Protein Engineering and existing crystal contacts of Fabs in the Fab-RNA complexes

Protein engineering techniques are widely used to assist in protein crytsallization (21). Understanding the physical properites of a protien is vital for engineering changes which can facilitate crystallization. For example the solubulity of proteins, an important criterion for crystallization is confered by the hydrophilic patches on surface of the protein. Systematic mutation on the surface of the catalytic domain of HIV-1 integrase dramtically improved its solubility (22,23). The other intrinsically structured elements, like the N- and C-termini which are very dynamic, unstructured and hinder crystallization were dealt by chopping off with proteolysis (24,25). The use of fusion proteins and affinity tags for obtaining crystallizable proteins also greatly facilitated protein crystallization (26-29). As a good crystallization module to begin with, Fab generally expresses well and readily crystallized alone or complexed with soluble protein antigens without significant efforts of surface engineering. Large amount of efforts on Fab crystallization were focused on the homogeneous Fab preparation and purification through enzymatic cleavage at the hinge region to separate Fab from the Fc portion of immunoglobulins as well as crystallization techniques (30,31). These Fabs or Fab antigen complexes can be crystallized with a good success rate without further surface engineering. Similarly the Fab-4D5 that our Fab framework was originated from was crystallized without optimization on the surface residues (32). Hence the optimization based on solubility issues or other purification techniques was not a major concern.

To explore the possibility of improving crystallizability through engineered crystal contacts, we analyzed the frequenlty occuring crystal contacts in the Fab/RNA complex crystals

involving Fab4D5 derivatives. The Fab mediated crystal contacts on two of the Fab chaperone RNA complexes ΔC209 P4-P6/Fab2 (PDB code: 2R8S) and class I ligase/BL3-6 (PDB code: 3IVK) are shown in Table 1. An extended anti parallel β-sheet forming a Fab homodimer was found to form a main inter-molecular Fab-Fab crystal contact in the 2R8S structure. Interestingly this crystal contact is not found to occur in 3IVK, though both Fabs, Fab2 and Fab BL3-6 possess identical sequence in this region. These two Fab-RNA crystal structures are the only two available with the Fab-4D5 framework and it has been observed that there are no crystal contacts or any frequent residues which are prone to make crystal contacts. Comparing the crystal structures of protein complexed with Fab-4D5 derivatives, full length KcSA (33) and the Fab-4D5 structure alone, there were no conserved crystal contacts or contact residues present either. Lack of conserved crystal contacts of the Fab/antigen complexes is reminiscent of the common belief that the rational engineering of crystal contacts is a rare approach and may not be ubiquitously applicable (34). In analyzing the crystal contact in the 2R8S structure, besides crystal contacts mediated with hydrogen bonding, we noticed an important crystal contact made by a hydrophobic patch recognizing the extruding base and sugar R-A125 (R indicates RNA) of another molecule. This hydrophobic patch is composed of the non-polar part of the side-chains or main chains of HC-Gly42, HC-Lys43, HC-Gly44 and LC-Gln100 (HC, heavy chain; LC, light chain), encircling two thirds of the sugar and base moiety of R-A125 from neighboring molecule without the contribution of the direct hydrogen bonding. This contributes partially to the choice of surface engineering for crystallization enhancement.

3.2 Selection of the residues for surface entropy reduction for ΔC209 P4-P6/Fab2

Random protein-protein interactions take place in protein crystallization making it a stochastic phenomenon (35). It has been reported that crystal patches favor smaller interfaces and polar residues. Consistent with this finding, surface entropy reduction method generates 'low-entropy' surface patches through site-directed mutagenesis of the surface lysines and glutamates to alanines. These low-entropy surface patches act as hot-spots and promote crystallization.

Using a model system of the globular domain of the human regulatory protein RhoGDI, the Derewanda group mutated and screened the protein constructs for crystallization with the SER approach (11,12,36,37). They found that most of the SER mutants have increased crystallizability while the double or triple cluster mutants (or more) involving lysines and glutamates in close proximity showed greatest tendency to yield new crystal forms. It has also been found that the crystal contacts for the double and triple mutants were in most cases directly mediated by the mutated epitopes (or causing steric hindrance when the wild type lysines or glutamates were engineered in), validating the beneficiary role of the SER mutagenesis to crystallization. The success in crystallization of the RhoGDI mutants has inspired the protein crystallographers. Many novel proteins have now been crystallized with the SER approach, including RGSL domain of human PDZRhoGEF (38), Y. pestis LcrV antigen (11), Tyrosine kinase domain of IGF-1 (39), EscJ protein (40), HIV CcmK4 capsid protein (41), etc. Interestingly, this approach has also enabled the crystallization of the protein complexes, such as the complex of c-Src and its inactivator Csk using a Csk mutant (42) and the complex of two pseudo-pilins EpsI and EpsJ from the type 2 secretion system of Vibrio vulnificus using a EpsI mutant (43). This application in the protein-protein complexes highlights the generality of the

SER approach and extends the method beyond individual component as the high-entropy patches occur outside complex interfaces.

The crystallization of the protein-protein complexes using SER approach is a great inspiration for improving the crystallizability of Fab/RNA complexes. It suggests the possibility of crystallizing the macromolecular complexes by surface engineering of one single component, greatly increasing the generality of the Fab crystal chaperone. In addition, the SER approach creates local patches with reduced hydrophilicity (lysines and glutamates to alanines or serines mutations), consistent with our observation of the crystal contact formation (previous section).

The probable entropy causing residues on ΔC209 P4-P6 binding Fab2 were identified by visual inspection of the crystal structure of ΔC209 P4-P6/Fab2 (PDB code: 2R8S) using 3D visualization software pyMOL. There are a total of 25 lysine residues (5.6% of total amino acids) and 16 Glutamate (3.5% of total amino acids) residues. Though the total lysine and glutamate content (10%) in Fab2 is lower than that of RhoGDI (a total of 20%), most of the lysine and glutamate residues are surface residues and solvent exposed.

Based on the visual inspection of ΔC209 P4-P6/Fab2 crystal complex (2R8S) the lysine and glutamate residues on the surface were identified and classified into two groups. Group A corresponds to the most probable entropic residues HC-K76, LC-K169, LC-K190, HC-K217 and LC-E123. The Group B corresponds to the probable (less protruding) entropic residues LC-K145, LC-K107, HC-K65, HC-K218, HC-K214, and HC-E220 (Figure 1).

On studying the protruding surface residues the physical interaction of the chosen residues with the other residues was also taken into consideration, such as proximal residues that may be involved in salt or electrostatic interactions. These counter residues were also identified

as mutation candidates paired with their interaction partners. In the 2R8S structure, HC-E220, HC-K217, HC-K218, LC-E123 are clustered together and LC-E123 was found to be at an interacting distance (4.10 Å) with HC-K217. HC-Lys 214 and HC-Lys 218 were found to participate in crystal contacts using main-chain atoms while three residues (HC-Lys 217, HC-Glu220 and LC-Glu123) form crystal contacts through side chain atoms.

The surface-entropy reduction server (http://services.mbi.ucla.edu/SER/), recently developed by Eisenberg's and Derewenda's group, was used to strengthen the choice of the surface mutations (44).

This program analyses the primary sequence of the protein entity and predicts the amino acids for surface entropy reduction. It takes into account the secondary structure prediction to which a lower score is assigned, residues predicted to be in the coil region to which higher scores are assigned, and the entropy profiles which are generated using side chain entropy profile tables (Sternberg table) whose results are normalized against lysine the highest entropic residue. At the last step the sequence query is subjected to PSI-BLAST to check for conserved residues and these residues are given a lower score (but not necessarily excluded from the mutation candidates). The residues with overall high scores are predicted as the top candidates for SER mutation. As the program does not have provision for accepting multiple peptide chains, the heavy chain and light chain sequences of the Fab2 were fed to the program in tandem as a single peptide. Table 2 summarizes the results generated by the program. Of the three clusters, Cluster 2 sits in the cleft between constant domain and variable domain of Fab2, although it is minimally solvent exposed, given the location, it is very unlikely to be accessible to form intermolecular

contact. The residues predicted in both cluster 1 and cluster 3 overlap with the ones identified by visual inspection of the ΔC209 P4-P6/Fab2 crystal complex (2R8S).

Combining both of these results, a total of five mutations were identified to begin with, namely HC-K217, HC-K218, HC-E220, LC-K190 and LC-E123. Construction of two super mutants, Supermutant Alanine (Fab2 SMA) and Supermutant Serine (Fab2SMS) were decided upon as to exploit the benefits of both alanine and serine residues as suggested in the study model of crystal structure of outer surface protein A of *Borrelia burgdorferi (45)*.

3.3 Construction of the ΔC209 P4-P6/Fab2 mutants

Kunkel mutagenesis was used to construct the expression strains for Fab 2 Supermutant Alanine and Supermutant Serine (46). Single strand DNA synthesis of the phagemid encoding the Fab2 gene was performed. It is important to note that the phagemid encodes the Fab2 and phage surface protein gene III in tandem. Hence a stop codon had to be incorporated right after the heavy chain gene sequence to express the Fab2 alone terminating the expression of phage surface protein. The single strand of DNA encoding Fab2 was generated through electroporation into CJ236 (uracil deglycosidase deficient electrocompetent cells) and amplification with M13 helper phage co-infection. After phage particle lysis and purification uracil rich ssDNA was isolated (Figure 2). The single strand DNA was used as Kunkel template to incorporate the respective alanine and serine mutations with specific Kunkel primers. The light chain alanine mutations LC-E123A and LC-K190A were introduced by seperate Kunkel primers as they lie far apart in the light chain. A single primer was sufficient in case of the heavy chain alanine

mutations including the Amber stop codon as the mutations lie in close proximity (as discussed above).

All the three Kunkel primers were annealed simultaneously with the Fab2 ssDNA. The alanine mutations were more efficient than the serine mutations as 3 out of 4 clones had all the mutations incorporated. The serine mutations did not incorporate all the mutations in one step. As a result, Fab2SMS were constructed with the two light chain mutations and heavy chain mutation introduced sequentially. Both Fab2SMA and Fab2SMS plasmids were transformed into the E coli 34B8 strain (a derivative of E. coli strain W3110 with the double protease deletion ompT/degP) (6) for expression.

3.4 Optimization of the expression and purification of Fabs

The optimization of Fab2 expression is vital for obtaining large amounts of Fabs enough for crystal screening. According to the previous protocol, Fabs were typically expressed in two steps. The first step involved inoculation of *Escherichia coli* 34B8 containing Fab2 gene in 25 ml 2YT with ampicillin and shaken at 250 rpm overnight at 37 ^{o}C. In the second step, 10 ml overnight cell culture was subcultured (1/100 dilution) into 1L CRAP media in 2.8 L baffled flask and shaken at 250 rpm and 30 ^{o}C for 24 h. The Fab expression underwent some serious issues on following the previous protocol in the PI's lab. Although Fab-4D5 derivatives can be produced in large quantity with fermentation method (47), for easy in-house replication, we decided to optimize Fab expression in shake flask culture. We first tried to express Fab with fresh colony or fresh transformation, which didn't lead to consistent expression. Addition of

extra ampicillin was carried out to supplement the 24 h shaking at 30 ^{o}C, but the expression could not be revived.

Based on the study on aeration in shake flask by McDaniel and Bailey (48) Fabs were expressed in aluminum-foil-capped 2.8 L baffled flask at 300 rpm with various culture volumes. Table 3 shows that, with starting overnight culture growing at 37 ^{o}C, the 1 L and 1.25 L final culture didn't produce protein Fab while the 500 ml culture consistently produced Fab at ~ 1 mg/L. This is contrary to previous observation using the same parameters where 1.25 L culture produced high yield (per liter) than the 1 L and 500 ml culture (1). Subtle differences in ambient temperature and oxygen fluctuation, oxygen rate transfer in the particular shaker may contribute to the difference in Fab expression in specific lab environment. Further optimization of the expression was carried out by lowering the starting culture temperature to 30 ^{o}C (Table 4), which increased the yield to ~3 mg/L. The higher yield with the starting culture temperature at 30 ^{o}C may be attributed to healthier cells judging from the density of the cells or reduced inorganic phosphates level which leads to earlier induction of the phoA promoter (49).

Subsequently the Fabs were affinity purified on protein A column and then subjected to ion-exchange column purification using High S resin. One round of High S resin purification removes short fragments visible on the SDS-PAGE and most of the nucleases while a second round of High S resin purification will remove residual nuclease activity as examined by nuclease test, resulting clean Fab for crystal screening. The Figures 3, 4 and 5 show the SDS-PAGE fractions of several rounds of purification of ΔC209 P4-P6/Fab2SMA, ΔC209 P4-P6/Fab2SMS and ΔC209 P4-P6/Fab2 respectively. The nuclease test activity showed that the

Fabs retained minimal nuclease activity and can be used for crystallization in complex with ΔC209 P4-P6 RNA (Figure 6).

3.5 Characterization of Fab2SMA and Fab2SMS

Filter binding assay was used to characterize the binding between mutant Fabs with ^{32}P labeled ΔC209 P4-P6. Results (Figure 7) showed that the binding affinities of Fab2SMA and Fab2SMS binding to ΔC209 P4-P6 were similar to that of the parent Fab2. This is expected as the mutation sites were far away from the antigen binding regions. In order to set up the crystallization trays, it is mandatory to confirm the stoichiometric equation for complex formation between mutant Fabs and ΔC209 P4-P6. A native gel mobility shift assay was carried out (Figure 8) and the results showed that both Fab2SMA and Fab2SMS form a 1:1 complex with ΔC209 P4-P6, similar to Fab2. The faster mobilities of the complexes formed by either Fab2SMA or Fab2SMS reflect the positive charge shaved off on Fab2SMA and Fab2SMS when compared to Fab2. The working stoichiometric ratio (RNA : Fab = 1 : 1.1) for complete complexing of the RNA in the crystal screening was determined.

3.6 Crystallization

Large scale Fab expression using the protocol established above was performed. The typical yields of Fabs after large scale expression were around 3.8 mg/L for Fab2SMA, 2.8 mg/L for Fab2SMS and 2.6 mg/L for Fab2 after protein A column purification. After two rounds of High S resin purification, half of these yields were retained. Thus, 6-10 L cell culture provided sufficient Fab proteins for crystal screening. To obtain RNA with crystallizable quality, ΔC209 P4-P6 plasmid was linearised using Ear I digestion (Figure 9). Large scale transcription (~12 ml)

of ΔC209 P4P6 was performed (Figure 10) followed by denaturing PAGE purification. Once the Fab-RNA complexes were prepared according to the determined stoichiometric ratio, the crystal conditions of the three complexes were screened using Hampton Crystal Screen I and II, Hampton Index Screen, and Hampton Natrix Screen I and II at both 20 ^{o}C and 4 ^{o}C with sitting-drop method in the 96-3 well Intelli plates side-by-side. Considering the well solution and temperature, a total of ~600 conditions were screened for each Fab/RNA complex. Crystal trays were monitored over a period of two months. At 20 ^{o}C, Fab2-RNA complex did not yield any crystals, which is similar to the screening observed previously (1). However, three conditions for Fab2SMA-RNA complex and one condition for Fab2SMS-RNA complex were obtained. Figure 11 shows the crystals and their respective crystal conditions. At 4 ^{o}C, crystals for all three complexes were obtained in Crystal Screen I 26 (Figure 11 and Table 5), from which the crystal structure of ΔC209 P4-P6/Fab2 complex was originally solved (1). In addition, we have also obtained crystals for all three complexes in a similar condition, Crystal Screen I 5 (Table 5). Combining the results at both 20 ^{o}C and 4 ^{o}C, crystal screens results from both 20 ^{o}C and 4 ^{o}C show that crystal forming rate is greatly increased with the Fab2SMA and Fab2SMS mutant forms. The crystal conditions obtained were then reproduced and marginally optimized by varying the pH, salt and precipitant concentration (depending on specific conditions) to obtain the range for crystal formation for different conditions (Table 6).

The X-ray diffraction of two preliminary crystal forms were analyzed at beam line X25 and processed with HKL2000. The first crystal belongs to space group P1 and diffracts to 8 Å (condition shown in Figure 11A) while the second one belongs to space group C2 and diffracts to 6 Å (condition shown in Figure 11B).

As P4-P6 was crystallized several times before (1,3,7), optimization of the crystal conditions to obtain bigger and better crystals were not pursued.

The fact that Fab2SMA and Fab2SMS promote better crystal forming rates when complexed with ΔC209 P4-P6 proves that the surface engineering is a valuable tool for improving crystallizability.

3.7 Generality of the SER mutations

In order to test the generality of the SER mutations (mutations are far away from the CDR region of the Fabs) to Fabs binding to different antigens, all five alanine mutations were incorporated into two glycine riboswitch binding Fabs, VCIIIFab18 and VCIIIFab20, to construct VCIIIFab18SMA and VCIIIFab20SMA respectively. Expression of both Fabs gave reasonably good yield, ~3 mg/L. Filter binding assays show that VCIIIFab18SMA and VCIIIFab20SMA bind VCIII RNA with affinities similar to their corresponding parent Fabs (Figure 12). This result indicates that it is possible to apply the identified SER mutations to the Fabs with the same framework (Fab-4D5) to create a better RNA crystal chaperone.

Table 1: Crystal contacts residues in several Fab-4D5 derivative containing structures.

Crystal structure	Fab-Fab Interaction	Fab-antigen Interaction
ΔC209P4-P6/Fab2 (2R8S)[a]	H-Gly126N L-Glu123OE2 H-Asn212ND2 H-Ser211O H-Thr213OG1 H-Glu220OE1,OE2 H-Lys217NZ *H-Asn212O* *H-Lys214N* *H-Lys214O* *H-Asp216N* *H-Asp216O* *H-Lys218N*	L-Ser202OG L-Gln199NE2 L-Val110N L-Ser14OG L-Arg18N H-Glu89OE2 H-Arg87NH1, NE
BL3-6/BL ligase I (3IVK)[a]	H-Ser88OG L-Ser204OG	H-Lys191NZ L-Asn211ND2 L-Lys141NZ L-Thr207N L-Lys208NZ H-Gly145O
FL-KcSA/Fab (3EFF)[a]	H-Thr216OG1 H-Asn215ND2 H-Ser214O H-Thr127OG1 H-Ser126OG	

[a] Residue number is kept the same as from PDB file and the italicized residues take part in the anti-parallel β sheet formation.

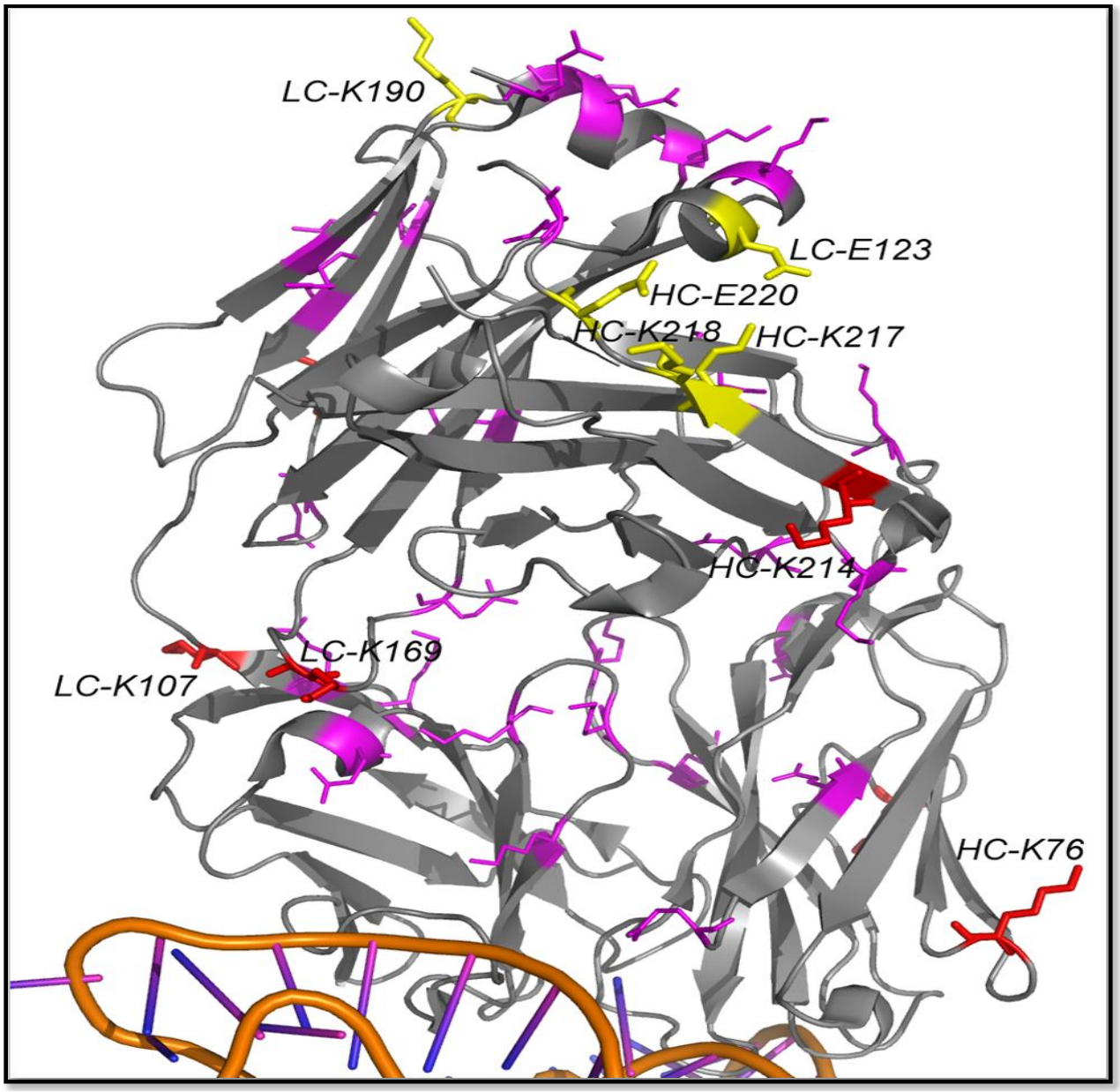

Figure 1: Crystal structure of ΔC209 P4-P6/Fab2 complex (2R8S).

RNA is only partially shown. Five SER mutation sites are shown in yellow. Other residues picked by visual inspection are shown in red. The rest of the lysines and glutamates are shown in magenta.

Table 2: Query results from the web-based SERp server

Cluster 1	Cluster 2	Cluster 3
KADY**EK**H**K**VYACE	**QQK**PGKAPK	KK**V**E**PK**
SERp score : 5.54	SERp score : 4.21	SERp score : 3.64
LC-E187, K188, K190	LC-Q37, Q38, K39	HC-K217, K218, E220

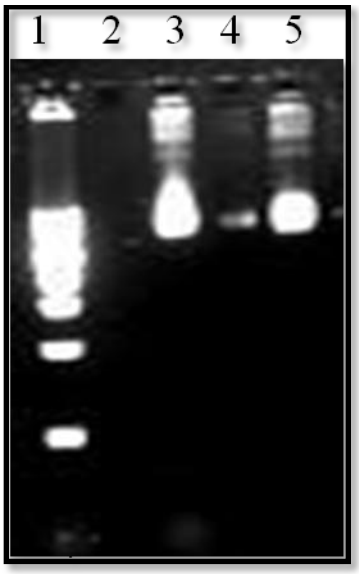

Figure 2: Single strand DNA extraction ΔC209 P4-P6/Fab2 without stop codon.

Lane 1: 1kb Marker, Lane 3: ss DNA ΔC209P4P6Fab2, Lane 5: ss DNA VCIIIFab18, Lanes 2,4: empty

Table 3: Trials for media optimization for Fab expression - variation in media volume

Fab2 and mutants	1L CRAP	1.25L CRAP	500ml CRAP
OD_{600nm} O/N	7-9	7-9	7-9
$OD_{600nm\ 24\ hrs}$	4-5	3-4	6<
Protein concentration	Nil	Nil	1.14mg/L

Table 4: Variation in temperature of starting culture

ΔC209P4P6Fab2SMA	30 ^{o}C Starting culture	37 ^{o}C starting culture
O/N OD_{600nm}	7.117	8.6276
After 24 hr OD_{600nm}	10.485	5.989
Protein expression	2.12mg/L	1.14mg/L

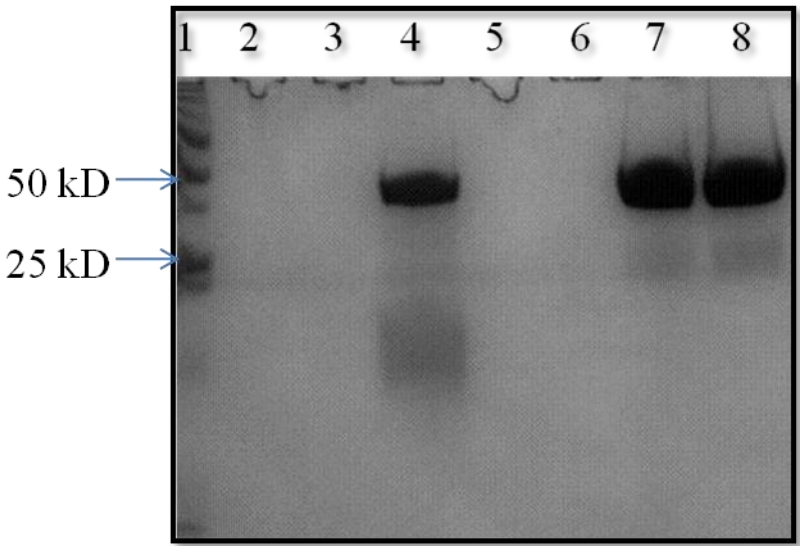

Figure 3: Large scale expression and purification of ΔC209 P4-P6/Fab2SMA.

Lane 1: Protein marker, Lane 4: Protein A column purification fraction, Lane 7:1st High S resin purification fraction, Lane 8: 2nd High S resin purification fraction

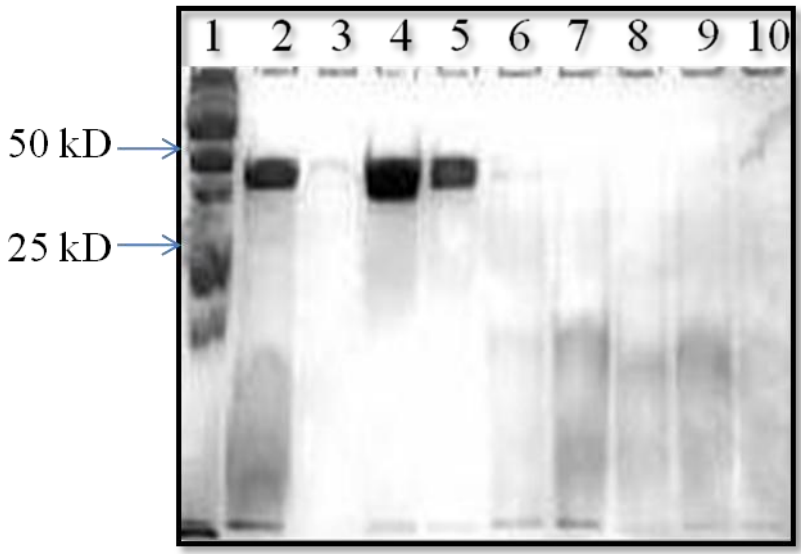

Figure 4: Large scale expression and purification of ΔC209 P4-P6/Fab2SMS.

Lane1: Protein marker, Lane 2: Protein A purification fraction, Lane 3: flow through from 1st high S resin purification, Lane 4: 1st high S purification fraction, Lane 5-10: 2nd high S purification fractions

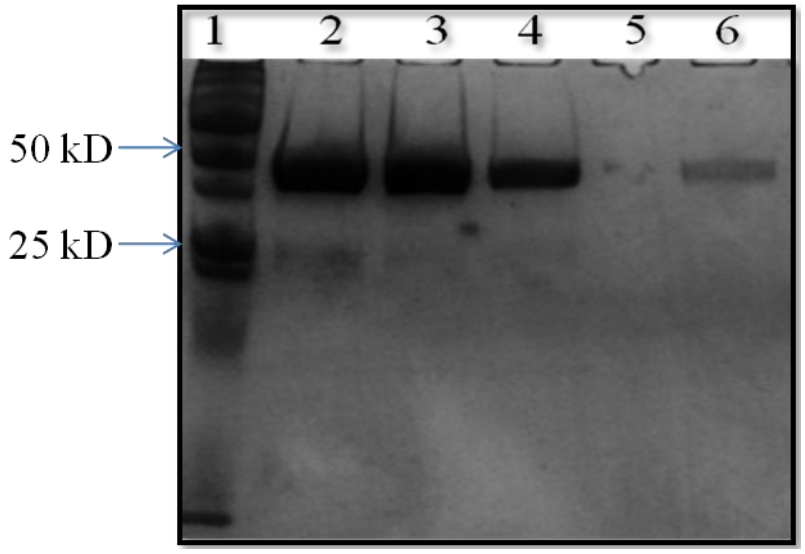

Figure 5: Large scale expression and purification of ΔC209 P4-P6/Fab2.

Lane 1: Protein marker, Lane 2: Protein A purification fraction, Lane 3: 1st high S resin purification fraction, Lane 4-6 2nd High S purification fractions

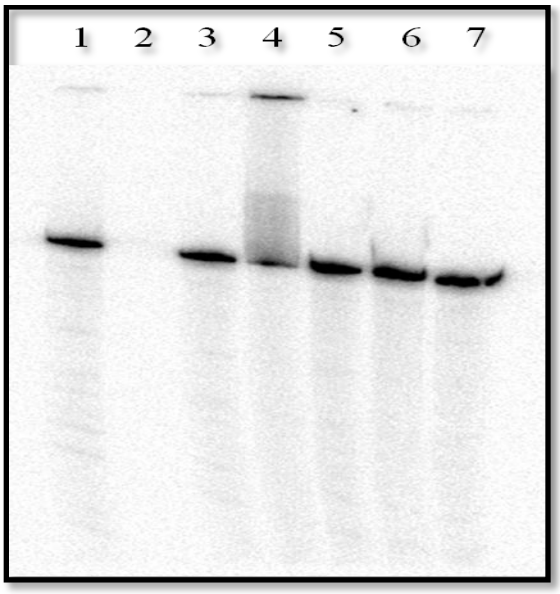

Figure 6: The nuclease activity gel analysis for all the Fabs.

Lane 1: Fab2, 87% intact *VCLD1 RNA, Lane 2: empty, Lane 3: Fab2SMA, 91% intact *VCLD1 RNA, Lane 4: Fab2SMS, 89% intact *VCLD1 RNA, Lane 5: Fab storage buffer, 90% *VCLD1 RNA, Lane 6: TE buffer, 92% *VCLD1 RNA and Lane 7: double distilled water, 92% *VCLD1 RNA.

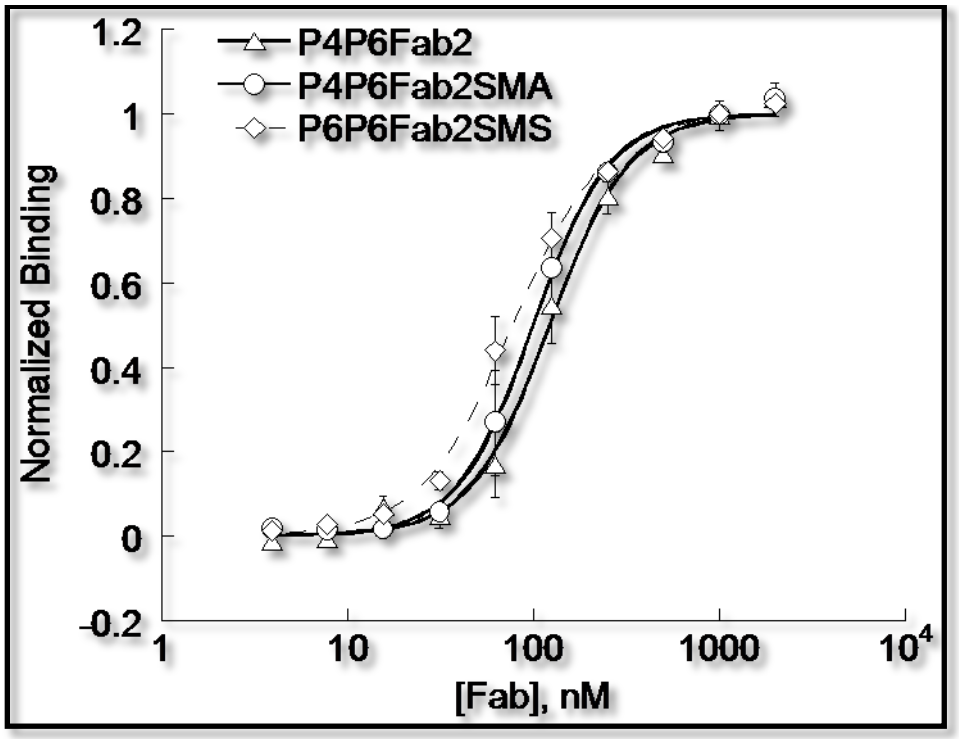

Figure 7: Binding curves for the ΔC209 P4P6 RNA binding to Fab2, Fab2SMA and Fab2SMS.

The binding affinities of ΔC209 P4-P6 with Fab2, Fab2SMA, and Fab2SMS were 122 ± 6, 100 ± 3, and 77 ± 3 nM respectively. The binding affinity of Fab2 binding to ΔC209 P4-P6 was previously determined to be 51 nM (1).

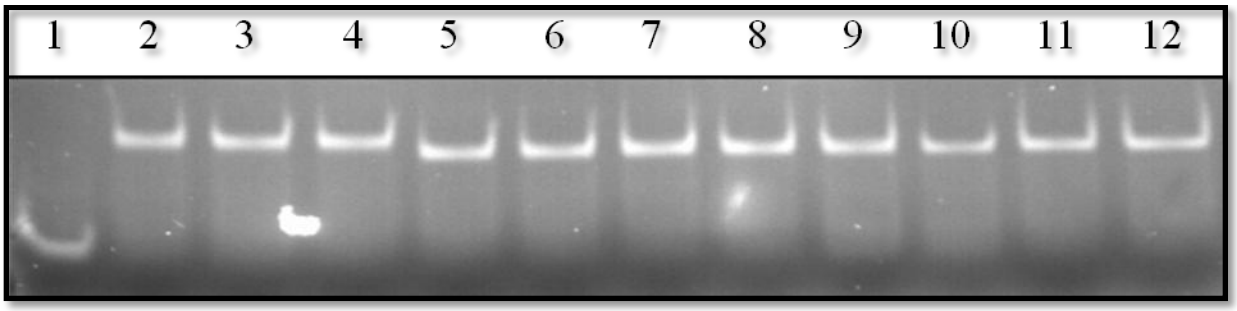

Figure 8: Native gel mobility shift assay of ΔC209 P4-P6 binding to Fab2, Fab2SMA and Fab2SMS.

Lane 1: ΔC209 P4-P6 RNA alone, Lane 2-4: ΔC209 P4-P6 complexes with Fab2 in molar ratio of 1 : 1.1, 1 : 1.2, 1 : 1.3, Lane 5-8: ΔC209 P4-P6 complexes with Fab2SMA in the molar ratio of 1 : 1, 1 : 1.2, 1 : 1.3, 1 : 1.4, and Lane 9-12: ΔC209 P4-P6 complexes with Fab2SMS in molar ratio of 1 : 1, 1 : 1.2, 1 : 1.3, 1 : 1.4.

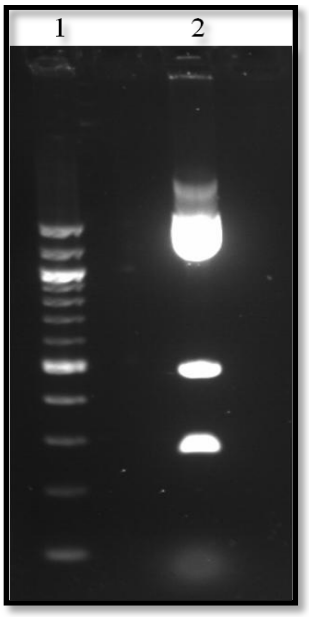

Figure 9: Large scale Ear I digestion of ΔC209 P4-P6 DNA.

Lane 1:100bp marker, Lane 2: Large scale ΔC209 P4-P6 DNA Ear I digest

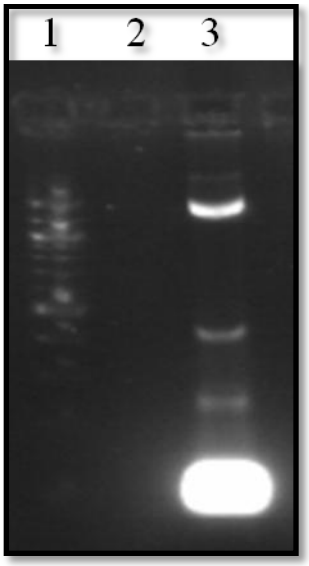

Figure 10: Large scale *in vitro* transcription of ΔC209 P4-P6.

Lane 1: 100bp marker, Lane 3: ΔC209 P4-P6 large scale transcription product

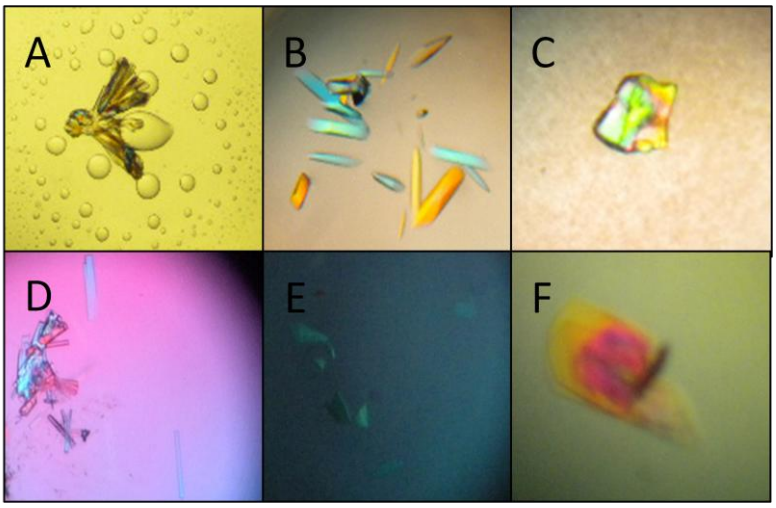

Figure 11: Crystal pictures of ΔC209 P4-P6 complexed with Fab2 mutants from initial screening.

Crystal screening was done with sitting drop method combining 0.5 µl of sample and 0.5 µl well solution. Sample buffer contains 10Mm Tris 7.5, 25 mM $MgCl_2$, 50 mM NaCl, and 0.5 mM Spermine·4HCl. (A) Rosette crystals grew from ΔC209 P4-P6/Fab2SMA in Hampton Index Screen 92 (0.15M DL-Malic acid pH 7.0, 20% w/v PEG 3,350) at 20 ^{o}C. Initial crystals diffracted to 8 Å. (B) Rod-shaped crystals grew from ΔC209 P4-P6/Fab2SMS in Hampton Crystal Screen II 1 (2.0 M NaCl, 10% w/v PEG 6000) at 20 ^{o}C. Initial crystals diffracted to 6 Å. (C) Spherulite crystals grew from ΔC209 P4-P6/Fab2SMA in Hampton Index Screen 87 (0.2 M Magnesium chloride hexahydrate, 0.1 M Tris pH 8.5, 3.4M 1, 6-Hexanediol) at 20 ^{o}C. (D) Needle crystals grew from ΔC209 P4-P6/Fab2SMA in Hampton Natrix Screen 1 (0.01 M Magnesium chloride hexahydrate, 0.05 M MES monohydrate pH 5.6, 1.8 M Lithium sulfate monohydrate) at 20 ^{o}C. (E) Single plate crystals grew from ΔC209 P4-P6/Fab2SMA crystals in Hampton Crystal Screen I 26 (0.2 M Ammonium acetate, 0.1 M Sodium citrate tribasic dihydrate pH 5.6, 30% v/v(+/-)-2-Methyl-2, 4-pentanediol) at 4 ^{o}C. (F) Plate crystals grew from ΔC209 P4-P6/Fab2SMS in Hampton Crystal Screen I 26 (0.2 M Ammonium acetate, 0.1M Sodium citrate tribasic dihydrate pH 5.6, 30% v/v(+/-)-2-Methyl-2, 4-pentanediol) at 4 ^{o}C.

Table 5: List of other crystal hits for the ΔC209 P4P6 RNA binding to Fab2 and its mutants

RNA-Fab Complex	Reservoir condition	Temperature	Crystal form
ΔC209 P4-P6/Fab2SMA	0.2 M Sodium citrate tribasic dihydrate,0.1M HEPES sodium pH 7.5,30% v/v(+/-)-2-Methyl-2,4-pentanediol (Crystal Screen I 5)	4 ^{o}C	Plates
ΔC209 P4-P6/Fab2SMS	0.2 M Sodium citrate tribasic dihydrate,0.1M HEPES sodium pH 7.5,30% v/v(+/-)-2-Methyl-2,4-pentanediol (Crystal Screen I 5)	4 ^{o}C	Plates
ΔC209 P4-P6/Fab2	0.2 M Sodium citrate tribasic dihydrate,0.1M HEPES sodium pH 7.5,30% v/v(+/-)-2-Methyl-2,4-pentanediol (Crystal Screen I 5)	4 ^{o}C	Rosettes
ΔC209 P4-P6/Fab2	0.2 M Ammonium acetate,0.1M Sodium citrate tribasic dihydrate pH 5.6,30%v/v(+/-)-2-Methyl-2,4-pentanediol (Crystal Screen I 26)	4 ^{o}C	Plates

Drop volume: Sample : Reservoir 1:1. Sample buffer contained 10Mm Tris 7.5, 25Mm $MgCl_2$, 50 Mm NaCl, 0.5Mm Spermine.4HCl

Table 6: Optimization of obtained crystal conditions.

RNA-Fab complex	Reservoir condition	Optimisation range
ΔC209 P4-P6/Fab2SMS	2 M NaCl,10%w/v PEG 6000 at 20 ^{o}C	1.7-2M NaCl and 10%-13% w/v PEG 6000
ΔC209 P4-P6/Fab2SMA	0.15M DL-Malic acid pH 7.0,20%w/v PEG 3,350 at 20 ^{o}C	0.15M-0.2M DL-Malic acid pH 7.0 and 18%-22% w/v PEG 3,350
ΔC209P4-P6/Fab2SMA	0.2 M Ammonium acetate,0.1M Sodium citrate tribasic dihydrate pH 5.6,30%v/v(+/-)-2-Methyl-2,4-pentanediol at 4 ^{o}C	pH 6.0-6.2 and 39%-45% v/v(+/-)-2-Methyl-2,4-pentanediol
ΔC209 P4-P6/Fab2SMS	0.2 M Ammonium acetate,0.1M Sodium citrate tribasic dihydrate pH 5.6,30%v/v(+/-)-2-Methyl-2,4-pentanediol at 4 ^{o}C	pH 5.9-6.1 and 40%-42% v/v(+/-)-2-Methyl-2,4-pentanediol

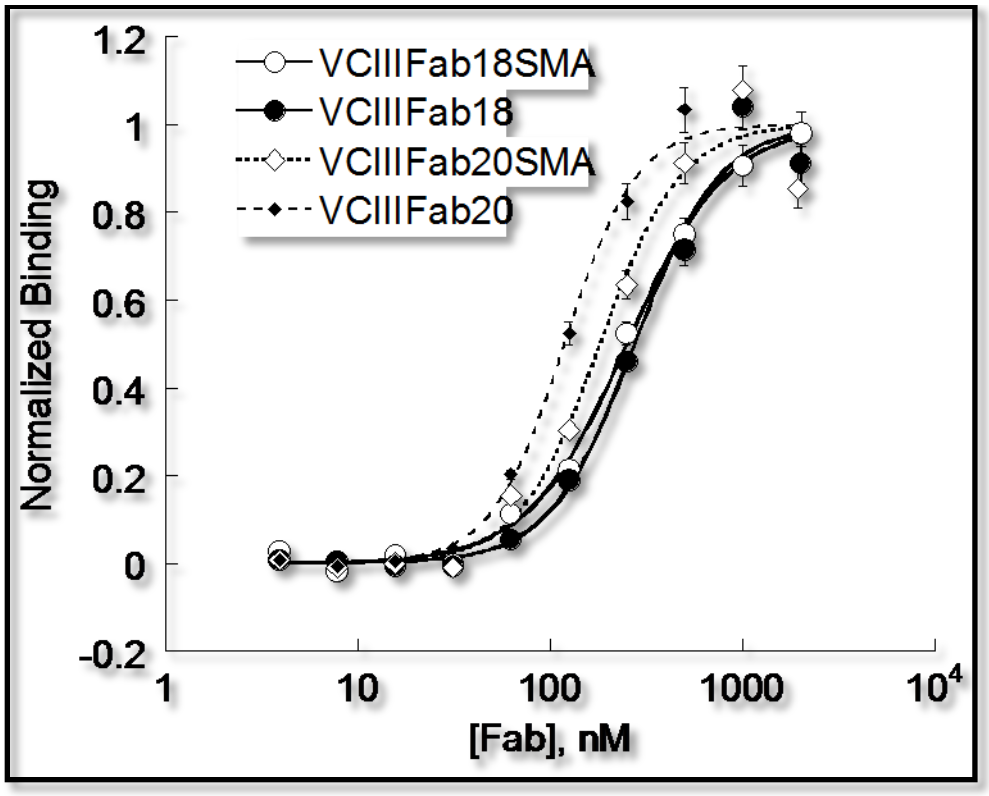

Figure 12: Surface engineered Fabs bind to VCIII glycine riboswitch.

The binding affinities of VCIII with Fab18, Fab18SMA, Fab20 and Fab20SMA were 274 ± 20, 249 ± 9, 118 ± 5, and 182 ± 17 nM respectively.

CHAPTER 4 CONCLUSION

The Surface Entropy Reduction method was successfully employed to improve the crystallizability of ΔC209 P4-P6/Fab2 complex. Interestingly, this improved crystallizability is based on mutations regardless of the preexisting crystal contacts in the parent RNA Fab complex. Our SER mutation sites were deliberately chosen in the constant domain of the Fab so that there will be minimal effect on the binding between Fab and RNA. The result of VCIII RNA binding Fabs showed that it's indeed possible to apply these mutations in a general manner. It may also be worthwhile to incorporate these mutations in the Fab framework to create a new phage displayed Fab library. This will eliminate the extra mutation steps involved after phage displayed selection if the goal is to crystallize the cognate RNA with a Fab chaperone. In addition, we have optimized the shake flask expression of the Fabs, which made the large amount of Fabs readily available for crystal screening. Combining these results, we are one step further in making a robust RNA crystallization chaperone.

REFERENCES

1. Ye, J.D., Tereshko, V., Frederiksen, J.K., Koide, A., Fellouse, F.A., Sidhu, S.S., Koide, S., Kossiakoff, A.A. and Piccirilli, J.A. (2008) Synthetic antibodies for specific recognition and crystallization of structured RNA. *Proc Natl Acad Sci U S A*, **105**, 82-87.
2. Koldobskaya, Y., Duguid, E.M., Shechner, D.M., Suslov, N.B., Ye, J., Sidhu, S.S., Bartel, D.P., Koide, S., Kossiakoff, A.A. and Piccirilli, J.A. (2011) A portable RNA sequence whose recognition by a synthetic antibody facilitates structural determination. *Nat Struct Mol Biol*, **18**, 100-106.
3. Ke, A. and Doudna, J.A. (2004) Crystallization of RNA and RNA-protein complexes. *Methods*, **34**, 408-414.
4. Ferré-D'Amaré, A.R. (2010) Use of the spliceosomal protein U1A to facilitate crystallization and structure determination of complex RNAs. *Methods*, **52**, 159-167.
5. Wu, S., Ke, A. and Doudna, J.A. (2007) A fast and efficient procedure to produce scFvs specific for large macromolecular complexes. *J Immunol Methods*, **318**, 95-101.
6. Fellouse, F.A., Esaki, K., Birtalan, S., Raptis, D., Cancasci, V.J., Koide, A., Jhurani, P., Vasser, M., Wiesmann, C., Kossiakoff, A.A. *et al.* (2007) High-throughput generation of synthetic antibodies from highly functional minimalist phage-displayed libraries. *J Mol Biol*, **373**, 924-940.
7. Juneau, K., Podell, E., Harrington, D.J. and Cech, T.R. (2001) Structural basis of the enhanced stability of a mutant ribozyme domain and a detailed view of RNA--solvent interactions. *Structure*, **9**, 221-231.
8. Cate, J.H., Gooding, A.R., Podell, E., Zhou, K., Golden, B.L., Kundrot, C.E., Cech, T.R. and Doudna, J.A. (1996) Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science*, **273**, 1678-1685.
9. Derewenda, Z.S. and Vekilov, P.G. (2006) Entropy and surface engineering in protein crystallization. *Acta Crystallogr D Biol Crystallogr*, **62**, 116-124.
10. Derewenda, Z.S. (2010) Application of protein engineering to enhance crystallizability and improve crystal properties. *Acta Crystallogr D*, **66**, 604-615.
11. Derewenda, Z.S. (2004) Rational protein crystallization by mutational surface engineering. *Structure*, **12**, 529-535.
12. Longenecker, K.L., Garrard, S.M., Sheffield, P.J. and Derewenda, Z.S. (2001) Protein crystallization by rational mutagenesis of surface residues: Lys to Ala mutations promote crystallization of RhoGDI. *Acta Crystallogr D Biol Crystallogr*, **57**, 679-688.
13. Vekilov, P.G., Feeling-Taylor, A.R., Yau, S.T. and Petsev, D. (2002) Solvent entropy contribution to the free energy of protein crystallization. *Acta Crystallogr D*, **58**, 1611-1616.
14. Yau, S.T., Petsev, D.N., Thomas, B.R. and Vekilov, P.G. (2000) Molecular-level thermodynamic and kinetic parameters for the self-assembly of apoferritin molecules into crystals. *Journal of Molecular Biology*, **303**, 667-678.
15. Petsev, D.N., Thomas, B.R., Yau, S.T., Tsekova, D., Nanev, C., Wilson, W.W. and Vekilov, P.G. (2001) Temperature-independent solubility and interactions between apoferritin monomers and dimers in solution. *J Cryst Growth*, **232**, 21-29.

16. Gliko, O., Neumaier, N., Pan, W., Haase, I., Fischer, M., Bacher, A., Weinkauf, S. and Vekilov, P.G. (2005) A metastable prerequisite for the growth of lumazine synthase crystals. *J Am Chem Soc*, **127**, 3433-3438.
17. Baud, F. and Karlin, S. (1999) Measures of residue density in protein structures. *P Natl Acad Sci USA*, **96**, 12494-12499.
18. Avbelj, F. and Fele, L. (1998) Role of main-chain electrostatics, hydrophobic effect and side-chain conformational entropy in determining the secondary structure of protein. *J.Mol.Biol.*, **279**, 665-684.
19. Doig, A.J. and Sternberg, M.J.E. (1995) Side-chain conformational entropy in protein folding. *Protein Sci.*, **4**, 2247–2251.
20. Lo Conte, L., Chothia, c. and Janin, J. (1999) The atomic structure of protein-protein recognition sites. *J.Mol.Biol.*, **285**, 2177-2198.
21. Derewenda, Z.S. (2010) Application of protein engineering to enhance crystallizability and improve crystal properties. *Acta Crystallogr D Biol Crystallogr*, **66**, 604-615.
22. Dyda, F., Hickman, A.B., Jenkins, T.M., Engelman, A., Craigie, R. and Davies, D.R. (1994) Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science*, **266**, 1981-1986.
23. Jenkins, T.M., Hickman, A.B., Dyda, F., Ghirlando, R., Davies, D.R. and Craigie, R. (1995) Catalytic domain of human immunodeficiency virus type 1 integrase: identification of a soluble mutant by systematic replacement of hydrophobic residues. *Proc Natl Acad Sci U S A*, **92**, 6057-6061.
24. Dong, A., Xu, X., Edwards, A.M., Chang, C., Chruszcz, M., Cuff, M., Cymborowski, M., Di Leo, R., Egorova, O., Evdokimova, E. *et al.* (2007) In situ proteolysis for protein crystallization and structure determination. *Nat Methods*, **4**, 1019-1021.
25. Wernimont, A. and Edwards, A. (2009) In situ proteolysis to generate crystals for structure determination: an update. *PLoS One*, **4**, e5094.
26. Diguan, C., Li, P., Riggs, P.D. and Inouye, H. (1988) Vectors That Facilitate the Expression and Purification of Foreign Peptides in Escherichia-Coli by Fusion to Maltose-Binding Protein. *Gene*, **67**, 21-30.
27. Smith, D.B. and Johnson, K.S. (1988) Single-Step Purification of Polypeptides Expressed in Escherichia-Coli as Fusions with Glutathione S-Transferase. *Gene*, **67**, 31-40.
28. Smith, M.C., Furman, T.C., Ingolia, T.D. and Pidgeon, C. (1988) Chelating peptide-immobilized metal ion affinity chromatography. A new concept in affinity chromatography for recombinant proteins. *J Biol Chem*, **263**, 7211-7215.
29. Uhlen, M., Forsberg, G., Moks, T., Hartmanis, M. and Nilsson, B. (1992) Fusion proteins in biotechnology. *Curr Opin Biotechnol*, **3**, 363-369.
30. Boulot, G., Guillon, V., Mariuzza, R.A., Poljak, R.J., Riottot, M.M., Souchon, H., Spinelli, S. and Tello, D. (1988) Crystallization of Antibody Fragments and Their Complexes with Antigen. *J Cryst Growth*, **90**, 213-221.
31. Stura, E.A., Fieser, G.G. and Wilson, I.A. (1993) Crystallization of Antibodies and Antibody-Antigen Complexes. *ImmunoMethods*, **3**, 164-179.
32. Eigenbrot, C., Randal, M., Presta, L., Carter, P. and Kossiakoff, A.A. (1993) X-ray structures of the antigen-binding domains from three variants of humanized anti-

p185HER2 antibody 4D5 and comparison with molecular modeling. *J Mol Biol*, **229**, 969-995.
33. Uysal, S., Vasquez, V., Tereshko, V., Esaki, K., Fellouse, F.A., Sidhu, S.S., Koide, S., Perozo, E. and Kossiakoff, A. (2009) Crystal structure of full-length KcsA in its closed conformation. *Proc Natl Acad Sci U S A*, **106**, 6644-6649.
34. Derewenda, Z.S. (2004) The use of recombinant methods and molecular engineering in protein crystallization. *Methods*, **34**, 354-363.
35. Carugo, O. and Argos, P. (1997) Protein—protein crystal-packing contacts. *Protein Science*, **6**, 2261–2263.
36. Mateja, A., Devedjiev, Y., Krowarsch, D., Longenecker, K., Dauter, Z., Otlewski, J. and Derewenda, Z.S. (2002) The impact of Glu-->Ala and Glu-->Asp mutations on the crystallization properties of RhoGDI: the structure of RhoGDI at 1.3 A resolution. *Acta Crystallogr D Biol Crystallogr*, **58**, 1983-1991.
37. Cooper, D.R., Boczek, T., Grelewska, K., Pinkowska, M., Sikorska, M., Zawadzki, M. and Derewenda, Z. (2007) Protein crystallization by surface entropy reduction: optimization of the SER strategy. *Acta Crystallogr D Biol Crystallogr*, **63**, 636-645.
38. Longenecker, K.L., Lewis, M.E., Chikumi, H., Gutkind, J.S. and Derewenda, Z.S. (2001) Structure of the RGS-like domain from PDZ-RhoGEF: linking heterotrimeric g protein-coupled signaling to Rho GTPases. *Structure*, **9**, 559-569.
39. Munshi, S., Hall, D.L., Kornienko, M., Darke, P.L. and Kuo, L.C. (2003) Structure of apo, unactivated insulin-like growth factor-1 receptor kinase at 1.5 angstrom resolution. *Acta Crystallogr D*, **59**, 1725-1730.
40. Yip, C.K., Kimbrough, T.G., Felise, H.B., Vuckovic, M., Thomas, N.A., Pfuetzner, R.A., Frey, E.A., Finlay, B.B., Miller, S.I. and Strynadka, N.C.J. (2005) Structural characterization of the molecular platform for type III secretion system assembly. *Nature*, **435**, 702-707.
41. Pornillos, O., Ganser-Pornillos, B.K., Kelly, B.N., Hua, Y.Z., Whitby, F.G., Stout, C.D., Sundquist, W.I., Hill, C.P. and Yeager, M. (2009) X-Ray Structures of the Hexameric Building Block of the HIV Capsid. *Cell*, **137**, 1282-1292.
42. Levinson, N.M., Seeliger, M.A., Cole, P.A. and Kuriyan, J. (2008) Structural basis for the recognition of c-Src by its inactivator Csk. *Cell*, **134**, 124-134.
43. Yanez, M.E., Korotkov, K.V., Abendroth, J. and Hol, W.G.J. (2008) The crystal structure of a binary complex of two pseudopilins: EpsI and EpsJ from the type 2 secretion system of Vibrio vulnificus. *J Mol Biol*, **375**, 471-486.
44. Goldschmidt, L., Cooper, D.R., Derewenda, Z.S. and Eisenberg, D. (2007) Toward rational protein crystallization: A Web server for the design of crystallizable protein variants. *Protein Sci*, **16**, 1569-1576.
45. Makabe, K., Tereshko, V., Gawlak, G., Yan, S. and Koide, S. (2006) Atomic-resolution crystal structure of Borrelia burgdorferi outer surface protein A via surface engineering. *Protein Sci*, **15**, 1907-1914.
46. Kunkel, T.A. (1985) Rapid and Efficient Site-Specific Mutagenesis without Phenotypic Selection. *P Natl Acad Sci USA*, **82**, 488-492.
47. Muller, Y.A., Chen, Y., Christinger, H.W., Li, B., Cunningham, B.C., Lowman, H.B. and de Vos, A.M. (1998) VEGF and the Fab fragment of a humanized neutralizing antibody:

crystal structure of the complex at 2.4 A resolution and mutational analysis of the interface. *Structure*, **6**, 1153-1167.
48. McDaniel, L.E. and Bailey, E.G. (1969) Effect of shaking speed and type of closure on shake flask cultures. *Appl Microbiol*, **17**, 286-290.
49. Wang, Y., Ding, H., Du, P., Gan, R. and Ye, Q. . (2005) Production of phoA promoter-controlled human epidermal growth factor in fed-batch cultures of Escherichia coli YK537 (pAET-8). *Process Biochem*, **40**, 3068-3074.