Spring 5-8-2011

Design and Bio-production of a Nanoparticle Avian Influenza Vaccine

Kashif N. Ather
University of Connecticut - Storrs, kashif.n.ather@gmail.com

Design and Bio-production of a Nanoparticle Avian Influenza Vaccine

Honors Thesis

Kashif Ather
Molecular and Cell Biology
Research Advisor: Dr. Peter Burkhard
Honors Advisor: Dr. James Cole
Spring 2011

Table of Contents

Acknowledgements:

I would like to thank Dr. Peter Burkhard for taking me into his lab and guiding me through this project. I would also like to thank Tobias Neef for teaching me all of the procedures and providing me with advice and mentorship. Special thanks to the other members of the Burkhard lab, including Dr. Tais Pimentel, Newton Wahome, and Dr. Yongkun Yang, for all of their assistance. I would like to thank Dr. James Cole for his support in the writing of my Honors thesis. Finally, I would like to thank my family and friends for their continued encouragement and support.

List of Abbreviations

BN5C: BN5C-M2eN-mh-ch-8r

S43: S43-9C-2IAd-M2eN

DLS: Dynamic Light Scattering

E. coli: Escherichia coli

Kb: Kilo-bases

His-tag: Histidine-tag

IPTG: Isopropyl-β-D-thiogalactopyranoside

LB: Luria-Bertani broth

OD: Optical Density

pPEP-T: pPEP-T vector plasmid

SDS-PAGE: Sodium Dodecyl Sulfate Polyacrylamide gel electrophoresis

EM: Electron Microscopy

Tris: tris(hydroxymethyl)aminomethane

CAPS: *N*-cyclohexyl-3-aminopropanesulfonic acid

TCEP: *tris*(2-carboxyethyl)phosphine

Ni-NTA: nickel-nitrilotriacetic acid

kDa: kilo-Daltons

pI: Iso-electric point

List of Figures

Abstract

Influenza is one of the most common diseases in the world and the cause for numerous deaths every year. The primary method of combating the disease is the influenza vaccine, which is produced by inoculating chicken eggs with inactivated virus. Some problems with this method include allergic reactions to the chicken egg preparation, as well as the risk caused by exposure to the phages present in the vaccine. An emerging solution is to use Self-Assembling Polypeptide Nanoparticles (SAPN) to elicit an immune response in the body, rather than using inactivated viruses. Both the C-terminus and N-terminus of the polypeptide can be modified, making the construct an ideal platform for vaccine design. This project focuses on the synthesis, purification, and refolding of two peptide constructs, BN5C and S43, which are specific protein sequences that under the right conditions will refold into the 3-dimensional structures necessary for producing an immune response. The BN5C construct proved difficult to purify. Thus, changes were made for the S43 construct; notably, the addition of histidine tags to increase the efficiency of purification. The size, shape, and homogeneity of the nanoparticles were tested using dynamic light scattering and electron microscopy in order to determine the optimal refolding conditions for the nanoparticles.

Introduction

Influenza is an extremely prevalent disease and is the cause of many deaths every year. The World Health Organization estimates that there are 25-30 million cases of influenza per year in the United States alone. This number balloons up to 600 million cases per year when taking the entire world into account. Influenza is a highly contagious disease that is spread by the influenza virus, usually transferred from human to human. Influenza epidemics occur seasonally and can affect up to 15% of the population [1]. The virus can also be transferred from animal to human in some cases, as evidenced by avian influenza [2]. Viruses that are transferred from animal to human are potentially more dangerous than the human variations, as the human immune system is not familiar with these specific strains or how to combat them. Additionally, the virus can also cause a worldwide pandemic, seen as recently as 2009, where a previously unseen strain of the virus causes mass infections. The reason for influenza pandemics, which have occurred many times throughout history, are mutation of the genes of the influenza virus; specifically, the genes for the surface proteins, which makes it harder for the immune system to fight the disease [2]. Some of the most dangerous influenza pandemics have been caused by mutated avian influenza viruses that have adapted the ability to infect humans [1, 2].

The most effective way to combat the influenza virus continues to be vaccination [2, 3]. Currently, commercial influenza vaccines are made from inactivated influenza virus propagated in embryonic chicken eggs [1]. However, there are several limitations to the current vaccination options. Because vaccines are developed using chicken eggs, members of the population who have allergies to eggs cannot use the vaccine. The constant mutation of the influenza virus means that different strains circulate every year, and it becomes a challenge to manufacture vaccines

quickly enough to make them readily available to the public [3, 4]. Furthermore, there are some strains that cannot be propagated in eggs making it impossible to produce a vaccine [3].

Self-Assembling Polypeptide Nanoparticles (SAPN) offer a new method of creating vaccines for many diseases. Nanoparticles have shown to be a viable solution to combat malaria [5] and SARS [6], among others. Nanoparticles are known to be an effective adjuvant for influenza vaccines [7, 8], and can be used to load antigens for an immune response [9]. One of the distinctive features of the nanoparticles used in this study is that they form a repetitive antigen display after self-assembly, promoting a strong immune response [6, 10]. Surface proteins, or fragments of these proteins, for common pathogens, such as avian influenza, can be engineered into the nanoparticle protein sequence to produce a specific immune response [10]. Various helper T-cell epitopes can be included in the structure of the nanoparticles for repetitive display to boost immune responses.

Because the architecture of the peptide nanoparticles is similar to that of virus capsids, they are able to generate an immune response with high efficiency while avoiding the use of sometimes toxic adjuvants [10]. The peptide nanoparticles also solve many of the problems involved with currently produced vaccines. Since the nanoparticles are peptide based, no attenuated or inactive viruses are needed making them a safer and more stable option [11]. The nanoparticles used in this study can be utilized in people who are allergic to eggs and can be adapted to a wide variety of strains based on the epitopes incorporated into the protein sequence. They are also able to be produced more cheaply and faster than traditional vaccines.

The BN5C-M2eN-mh-ch-8r construct and the S43-9C-2IAd-M2eN-ch construct are two polypeptide nanoparticle constructs that were designed to elicit an immune response to avian influenza. The BN5C construct is composed of a tetramer and a trimer connected by a linker

region. Both ends of the construct contain B-cell epitopes. Near the N-terminus of the polypeptide is an avian influenza epitope called M2e. The M2e epitope is based on the ecto-domain of the influenza matrix protein 2 (M2), which is important in viral replication [12]. M2e is highly conserved across strains of influenza and has not changed very much in the last century making it an excellent epitope for an influenza vaccine [12]. The C-terminus of the construct contains a cleavage peptide for the influenza surface protein hemagglutinin. The protein comes in many different forms, but the H1 form was attached to the construct for this experiment. The B-cell epitopes are included to trigger an immune response similar to an actual response to avian influenza. The center of the construct contains an epitope for helper T-cells. The epitopes are not specific to influenza but trigger a general immune response and contribute to viral elimination and lung protection [13]. The epitope for helper T-cells is included to boost the overall immune response.

Construct	Protein Sequence
BN5C	MGSLLTEVETPTRNGWECKCSDSSGSLYRLTVIIDDRYESLKNLITLRADRLEMIINDNVSTLRALLMGGRLLARLEELERRLEELARFVAAWTLKAAAVDLELAALRRRLEELARGPGSIQSRGLFG
S43	GSLLTEVETPTRNGWESKSSDSSGSLYRLTVIIDDRYESLKNLITLRADRLEMIINDNVSTILASITGGRLLLKLAELERRLEELERRLEYTSRAIQVVRARMEELERRLEELERRGHHHHHH

Table 1. Protein sequences for the BN5C-M2eN-mh-ch-8r-H1 and S43-9C-2IAd-M2eN-ch nanoparticle constructs. There are epitopes at both ends of the sequences and the S43 sequence contains a his-tag for more efficient purification.

The S43 construct is very similar in that it is also composed of a tetramer and a trimer connected by a linker region. The tetramer and trimer have a very similar sequence of amino acids, respectively, as the BN5C sequence, but the linker region connecting them differs between the two. Again, the construct contains an M2e epitope as well as a helper T-cell epitope. Besides the linker region, the main difference between the S43 and BN5C constructs is that the S43

construct contains a histidine tag (his-tag). The his-tag was included in order to make the protein easier to purify due to the affinity of his-tags for Nickel-NTA resins.

The structure of the nanoparticle can change depending on the conditions that the protein is placed in. The structure depends on the intra-molecular interactions of the protein, which provide stability for the nanoparticle construct [14]. Both the BN5C and S43 constructs have a coiled-coil structure that is an important part of the nanoparticle design. Notably, many of the surface proteins of pathogens, including avian influenza, contain coiled-coil sequences [10], which allow the nanoparticles to be easily modified to combat disease. Coiled-coil domains are an extremely versatile structure motif and are responsible for structural stability and the refolding capabilities of the nanoparticles [15]. The stability of the coiled-coils stems from the hydrophobic interactions between the helices. However, in order to correctly form particles, the intra-molecular interactions need to be maintained while avoiding interactions between separate nanoparticles [10, 16].

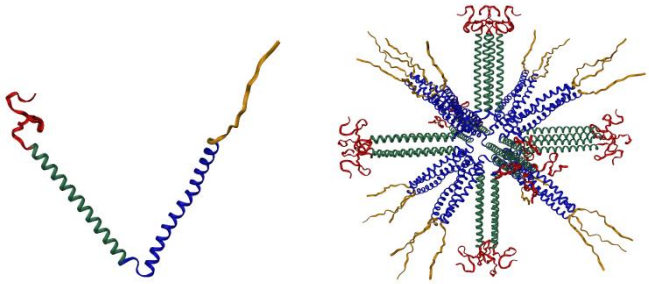


Figure 1. A computer generated model of what the basic BN5C construct monomer (left) and complete polypeptide nanoparticle (right) should look like. The tetramer and trimer, represented in green and blue, assemble into coiled-coil structures in the fully assembled nanoparticle. The B-cell epitope is represented in red.

The structure of the epitopes on correctly refolded nanoparticles is similar to those of pathogen surface proteins [10]. Combined with the ability to modify the peptide sequences to include epitopes for avian influenza, the characteristics of SAPNs create an ideal formula for a

new type of vaccine. The nanoparticles could induce a strong immune response without the drawbacks of current influenza vaccines, such as the inclusion of attenuated viruses, slow production, and inability to combat multiple strains. Here, the biophysical properties of the nanoparticles were assessed to determine the optimal refolding conditions.

Methods and Materials

Generation of Recombinant Cells

An oligonucleotide sequence was designed for the BN5C-M2eN-mh-ch-8r-H1 construct that contained the nucleotides needed to code for the H1 hemagglutinin cleavage peptide, as well as nucleotides on either side of the DNA coding for the peptide that could be recognized by the Xma I and Eco RI restriction enzymes. The restriction sites recognized by these enzymes are CCCGGG and GAATTC, respectively. The oligonucleotides were ordered from Integrated DNA Technologies and arrived as separate forward and reverse strands. Both the forward and reverse strands were brought to a concentration of 0.100 μg/mL. The annealing of the oligos was carried out in a micro-centrifuge (microfuge) tube by mixing 2 μL each of the forward and reverse strands, 10 uL 10X STE and 34 uL sterile distilled water and placing in a heat block, Fisher Scientific Isotemp Model 125D, for 90 seconds at 95°C and then cooling at room temperature. A 2% agarose gel was run in order to ensure the successful annealing of the forward and reverse strands.

```
5′    CCGGGGTCCAGGTAGTATACAAAGTAGAGGATTATTTGGATAATG          3′
3′        CCAGGTCCATCATATGTTTCATCTCCTAATAAACCTATTACTTAA      5′
```

Figure 2. The annealed forward and reverse strands of the oligonucleotide. The overhangs on the 5' ends of the strands enable the insertion of the oligo into the plasmid for the BN5C construct following double digestion.

The plasmid, which codes for the BN5C construct, was double digested with Eco RI and Xma I restriction enzymes by combining 20 μL plasmid with 5 μL 10X Buffer 4, 5 μL 10X BSA, and 1 μL each of Eco RI and Xma I enzyme. The mixture was incubated at 37°C for 3 hours to carry out the reaction. The mixture was purified using a Qiagen purification kit, assuming an 80% recovery of the plasmid. The cleaved ends of the plasmid were then dephosphorylated by

combining 28 μL of the double digested plasmid with 5 μL 10X AP reaction buffer, 1.2 μL Antarctic Phosphatase and 15.8 μL distilled water and incubated at 37°C for 30 minutes. The enzyme was then deactivated by incubating at 65°C for 5 minutes. The ligation of the plasmid and insert occurs most efficiently at a plasmid to insert ratio of 1 to 3, modified by the ratio of their lengths in kb. Based on their concentrations, 4 μL plasmid and 3.56 μL insert were combined with 0.5 μL of distilled water and incubated at 45°C for 5 minutes. A microfuge tube containing 1.1 μL plasmid and 6.9 μL distilled water was used as a control. The mixtures were then cooled on ice while 1 μL reaction buffer and 1 μL ligase were added to each and the ligation mixtures were incubated at 16°C and 1000 rpm overnight. This was achieved by placing an Eppendorf Thermomixer Shaker, Model EF4283A, in the 4°C refrigerator. The ligated samples were then stored at -20°C.

Three 100 μL samples of aliquoted *Escheria Coli (E. Coli)* DH5α strain cells were taken from the -80°C freezer and thawed on ice for transformation. The DH5α strain of *E. Coli* was used because this strain has a high efficiency for transformation and is useful in maintaining plasmids [17]. While on ice, 5 μL of the ligated plasmid and insert was added to one of the samples, 5 μL of the plasmid control was added to the second, and the third was left as a negative control. All of the samples were left on ice for 30 minutes, and then heat shocked for 90 seconds at 42°C in the heat block. The samples were then put back on ice. After two minutes, 800 μL of SOC medium was added to help with growth of the cells. The three samples were then incubated at 37°C for 45 minutes in the thermomixer. After removal from the thermomixer, 100 μL of each sample was spread onto an LB-Agar plate containing 100 mg/mL ampicillin, which was inverted and incubated at 37°C overnight.

After confirming no growth on either of the controls and growth on that plate containing the ligated plasmid and insert, 4 isolated colonies were scraped from the agar plate and mixed with four separate tubes of 5 mL Luria Bertani (LB) broth to be incubated overnight at 37°C and 180 rpm in the shaker, New Brunswick Scientific C25 Floor Model. The samples were then centrifuged at 4000 rpm in the Beckman Coulter Avanti Centrifuge J-20 XP. The supernatant was removed and the pellet was resuspended in 600 μL sterile distilled water. The Promega Plasmid Miniprep System was used to extract the plasmid DNA from the samples, which was sent for DNA sequencing to ensure the ligation had taken place and the samples had the proper sequence.

The process for the S43-9C-2IAd-M2eN construct was slightly different because the construct genes arrived from GenScript in the Puc57 vector, which is not an expressing vector. Therefore it needed to be cloned into a vector that could express the protein, the pPEP-T vector. Sections from the Puc57 vector and pPEP-T vector needed to be combined to create a plasmid that would code for the S43 protein. A technique called sub-cloning was used to extract the needed sections from both plasmids. Each plasmid was double digested, separately, with Eco RI and Nco I. For the pPEP-T vector, 10.6 μL plasmid, 5 μL 10X Buffer 4, 5μL 10X BSA, and 32.4 μL distilled water were mixed in a microfuge tube. In another microfuge tube, 16.2 μL Puc57 plasmid, 5 μL 10X Buffer 4, 5 μL 10X BSA, and 21.8 μL distilled water were mixed. Each of the tubes had 1 μL each of Eco RI and Nco I restriction enzymes added to them and the mixtures were incubated at 37°C for 4 hours. Following the reaction, 30 μL of each sample was combined with 6 μL loading buffer, loaded into a 2% agarose gel, and run at 80 V for 80 minutes. The gel was visualized on a Safe Imager (Invirogen) and the sections containing the desired fragments were cut out using razor blades. The larger fragment of the pPEP-T and the smaller fragment of

the Puc57 were removed and placed into separate microfuge tubes. The Qiagen Gel Extraction Kit was used to extract the plasmid from the small cut-outs of gel and the concentrations were recorded.

The pPEP-T fragment was considered the plasmid and the Puc57 fragment was considered the insert for the ligation. The insert needed to be dephosphorylated, which was carried out in the same manner as the BN5C construct. Again, a plasmid to insert ratio of 1 to 3 is the most efficient for ligation, depending on their lengths in kb. In a microfuge tube, 3.57 μL vector plasmid and 1.86 μL insert (a 1:10 dilution of the insert) were mixed with 1 μL reaction buffer, 1 μL ligase, and 2.57 μL distilled water. A control, composed of 1.1 μL plasmid, 1 μL reaction buffer, 1 μL ligase, and 6.9 μL distilled water, was mixed in a separate microfuge tube. The transformation into *E. Coli* DH5α cells was done the same way as the BN5C sequence. A sample of the plasmid was sent for DNA sequencing to ensure the plasmid had the correct sequence.

Production of Protein

Two 100 μL samples of aliquoted *E. Coli* BL21 strains were removed from the -80°C freezer and thawed on ice. The BL21 strain of *E. Coli* was used because this strain has a high level of gene expression and is very efficient for expressing a target protein. The BL21 chromosome contains the bacteriophage T7 polymerase gene that targets the T7 promoter located in the expression vectors that contain the genes of interest [17]. One of the samples was inoculated with 5 μL of plasmid coding for the BN5C-H1 protein confirmed to have the correct sequence and the other sample was used as a control. Both were left on ice for 30 minutes and then heat shocked for 90 seconds at 42°C in the heat block. After returning to ice for 2 minutes,

800 μL SOC medium was added to each to promote recovery and growth. The samples were then incubated at 37°C and 1000 rpm in a thermomixer for 45 minutes. After incubation, 100 μL from each sample were spread on separate LB-Agar plates with 100 mg/mL ampicillin and the plates were inverted and incubated at 37°C overnight. The plates were checked the next day for growth and an isolated colony from the plate containing the plasmid sample was scraped and used to inoculate 40 mL of LB broth. The LB broth with 100 mg/mL ampicillin was incubated overnight at 37°C and 180 rpm in the shaker.

Three liters of LB broth were made, split into three separate flasks for better aeration and growth. The flasks were placed in the shaker and 1 mL of ampicillin was added to each. Each of the flasks was inoculated with 10 mL of the overnight culture created from the BL21 plate. The remaining 10 mL from the overnight culture was used to create glycerol stocks of 750 μL culture and 250 μL sterile 20% glycerol that were placed in the -80°C freezer for future use, if needed. The three flasks of LB broth were kept in the shaker at 37°C and 180 rpm and their growth was monitored using a spectrophotometer (biowave S2100 Diode Array Spectrophotometer). A random sample of LB broth was taken at the start to use as a blank and the absorbance of one of the flasks was measured every 30 minutes. When the absorbance at 600 nm (A_{600}) reached between 0.5 and 0.6, the cultures were induced by adding 1 ml of 0.8 M Isopropyl-β-D-thiogalactopyranoside (IPTG) to each flask. IPTG induces the production of the target protein by the *E. Coli* BL21 cells. The flasks were then incubated in the shaker at 37°C and 180 rpm for 4 hours.

After incubation, each of the three flasks of LB broth were transferred into 1 L centrifuge tubes and centrifuged at 4000 g for 15 minutes. The supernatant was removed and the cell pellets at the bottom of each tube were combined in a 50 mL falcon tube. The weight of the combined

cell pellets was taken and the tube was stored at -20°C. The process was the exact same for producing the protein for the S43 construct.

Purification of the Protein

Approximately 4.5 grams of the BN5C-H1 protein was thawed on ice and resuspended in 20 mL of cold Cracking Buffer (20 mM citrate pH 6, 150 mM NaCl, 0.2% cholate, 5 mM EDTA). The falcon tube containing the protein in Cracking Buffer was placed in the sonicator, Masonix Sonicator 3000, suspended in an ice water bath to keep the sample cold during sonication. The protein/buffer was sonicated at power level 5-6 for 5 minutes, alternating 4 seconds on and 6 seconds off. The falcon tube was then centrifuged at 8000 rpm for 15 minutes. The supernatant was removed and placed in a falcon tube for analysis (Wash 1). The pellet was resuspended in 20 mL of Cracking Buffer by pipetting and vortexing. It was again centrifuged at 8000 rpm for 15 minutes and the supernatant was set aside (Wash 2). The pellet was then resuspended in 20 mL of Wash Buffer (20 mM citrate pH 6, 150 mM NaCl) and centrifuged at 8000 rpm for 15 minutes. The supernatant was removed and set aside (Wash 3). The pellet was resuspended in 25 mL of Wash Buffer and the process was repeated (Wash 4). Then, the pellet was resuspended in 20 mL of Inclusion Body Dissolving Buffer (20 mM citrate pH 3.5, 9 M urea, 0.3 mM TCEP). It was sonicated again, this time at power level 6.5 for 6 minutes, alternating between 4 seconds on and 6 seconds off. This was followed by centrifuging at 12,000 rpm for 15 minutes. The supernatant was removed and set aside (Wash 5) and the pellet was resuspended in 4 mL Inclusion Body Dissolving Buffer. All of the samples were run through Sodium Dodecyl Sulfate Polyacrylamide gel electrophoresis (SDS-PAGE) to test for the presence of protein. This was done by mixing 7 μL of each sample with 7 μL of SDS-PAGE

buffer and heating at 95°C for 5 minutes in the heat block. Then each sample was run through a polyacrylamide gel first for 15 minutes at 70 V, then at 180 V until the sample reached the bottom of the gel. The gel was submerged in Coomassie Blue (2.5 g brilliant blue, 40% distilled water, 50% ethanol, 10% glacial acetic acid) for 2 hours on a rotator (Lab-line MAXI Rotator Mixer Model 4631) and then rinsed in de-staining solution (15% methanol, 10% acetic acid). The gel was examined for the presence of protein.

The next step of purification required an ion exchange column which was created by using a syringe to insert 8 mL of SP-sepharose resin (Qiagen) into a column with a closed stopcock. Using the syringe to push the slurry into the bottom of the column minimized bubbles. The column was allowed to settle and was then equilibrated with 80 mL Inclusion Body Dissolving Buffer. The samples from the previous step that showed evidence of the presence of protein were combined and passed through the column and the flow through was collected. The column was then was washed with 80 mL Inclusion Body Dissolving Buffer and the flow through was collected. A salt gradient was used to elute the protein from the column. Each gradient step was 40 mL and the first 20 mL were collected in 2 mL fractions. The steps in the gradient were as follows: 0 M NaCl, 50 mM NaCl, 100 mM NaCl, 250 mM NaCl, 500 mM NaCl, and 1 M NaCl (20 mM citrate pH 3.5, 9 M urea, 0.3 mM TCEP). Finally, an elution buffer with pH 8.5 was passed through the column and ten 2 mL fractions were collected. The fractions were checked for the presence of the protein of interest by recording their optical density at 280 nm (OD_{280}) in the spectrophotometer. Those fractions with high values for OD_{280} were analyzed by SDS-PAGE. The fractions that tested positive for the protein were combined and dialyzed into 20 mM Tris pH 7, 5 % glycerol because of the corrosive effects of remaining in urea solution for too long. A strip of dialysis tubing was cut and equilibrated first in heavy metals

cleaning solution, and then the Tris pH 7.5 buffer. The dialysis tubing was filled with the protein in urea solution, clipped on both ends, and submerged in the urea-free buffer overnight.

The protein then underwent purification through a Q-sepharose column for the removal of endotoxin. The protein was dialyzed into a urea buffer (20 mM CAPS pH 10.5, 9 M urea, 0.3 mM TCEP) for denaturation. This time, 8 mL of Q sepharose resin (Qiagen) was pumped into the bottom of the column via syringe and allowed to settle. The column was equilibrated with the dialysis buffer and then the sample containing the protein was passed through the column. Fractions were collected as above, first of the urea buffer mentioned above (20 mM CAPS pH 10.5, 9 M urea, 0.3 mM TCEP), followed by a gradient of a lower pH buffer (20 mM Tris pH 7.5, 9 M urea, 0.3 mM TCEP) with the following steps: 0 M NaCl, 50 mM NaCl, 100 mM NaCl, 500 mM NaCl, 1 M NaCl. The OD_{280} of each fraction was measured and those with high values were tested using SDS-PAGE. The fractions that tested positive for the presence of the protein were collected and dialyzed into the same urea-free buffer as above (20 mM Tris pH 7, 5 % glycerol). Another Q-sepharose column was run with a smaller gradient to further purify the protein. All of the steps and buffers were the same, except the gradient was as follows: pH 10.5, pH 7.5 0 M NaCl, 100 mM NaCl, 200 mM NaCl, 300 mM NaCl, 400 mM NaCl, 500 mM NaCl, 1 M NaCl. They were again collected in 2 mL fractions, with the exception of the 1 M NaCl, which was collected as a whole in a 50 mL falcon tube. Again, the OD_{280} of the samples were tested, and those with high values were run through a SDS-PAGE.

The purification for the S43 construct was much different, due to the presence of a histidine tag (his-tag) within the protein. The his-tag meant that the protein could be purified more efficiently via a Nickel affinity column. The cell pellet obtained via centrifugation was resuspended in three times its weight of Lysis Buffer (9 M urea, 10 mM Tris pH 8, 100 mM

NaH_2PO_4, 20 mM imidazole). The falcon tube was suspended in a bucket of ice water and sonicated at power level 5, for 5 minutes alternating 4 seconds on, 6 seconds off. The sonicated cells were then centrifuged at 30,600 g for 45 minutes. The supernatant, referred to as the lysate, was removed and the debris pellet was discarded. The Nickel column was created by pipetting 4 mL of Ni-NTA slurry from Qiagen into a 15 mL falcon tube. The slurry was centrifuged at 4000 rpm for 3 minutes. The supernatant (ethanol) was removed and 3 mL of lysis buffer was added to the falcon tube to wash the beads. The mixture was then centrifuged at 4000 rpm for 3 minutes. The lysis buffer was removed by pipette and the Ni-NTA slurry and cell lysate were combined in a 50 mL falcon tube. The tube was mixed gently via a rotator mixer, Lab-line MAXI Rotator Mixer Model 4631, at power level 4 overnight.

The Ni-NTA/lysate mixture was poured into a closed column and allowed to settle for several minutes. The stopcock was opened and the flow through was collected for analysis. The column was then washed with 80 mL of Lysis Buffer, which was also collected. Then, 50 mL of Phosphate Buffer (8 M urea, 10 mM Tris pH 8, 500 mM H_2PO_4, 20 mM imidazole) was run through the column and collected in a falcon tube. The next step was to run a pH gradient through the column. The pH Buffer (9 M urea, 20 mM citrate, 100 NaH_2PO_4, 20 mM imidazole) was split into three 50 mL fractions that were equilibrated to pH 6.3, pH 5.9, and pH 4.3. The buffers were run through the column in that order and collected for analysis. An elution buffer (8 M urea, 10 mm Tris pH 8, 100 NaH_2PO_4) was used to create an imidazole gradient to run through the column. The gradient was as follows: 100 mM imidazole, 250 mM imidazole, 500 mM imidazole, 1 M imidazole. For the gradient, 30 mL of each imidazole concentration was run through the column and collected in 2 mL fractions. The OD_{280} for the fractions from the imidazole gradient were read and the peaks from each step in the gradient were analyzed by

SDS-PAGE along with samples from each of the buffers passed through the column and the flow through. The samples that appeared to contain the protein of interest were combined and dialyzed into Imidazole Free Buffer (8 M urea, 20 mM Tris pH 7.5, 150 mM NaCl, 2 mM EDTA) overnight, using the same process as above, then into 0 M urea to protect against the corrosive effect of urea on proteins. The sample was divided into four falcon tubes and frozen at -80°C.

Self-Assembly of the Nanoparticles

After collection of reasonably pure protein, the refolding and self-assembly of the BN5C-M2eN-mh-ch-8r-H1 nanoparticles was attempted by dialyzing the protein into different buffer conditions. The concentration of the protein in solution was estimated, and it was concluded that refolding should be done at the concentration at which the protein was purified. The dialysis was done as above, with the dialysis tubing equilibrated first in heavy metals cleaning solution, then a small quantity of the target buffer before adding the protein and dialyzing it overnight in said buffer. The refolding process that was used is called "direct refolding" which requires that the protein first undergo dialysis into an 8 M urea variation of each buffer, followed by a 0 M urea concentration. The 8 M urea buffers were kept at room temperature, while the 0 M urea buffers were kept at 4°C. The protein was filtered to remove impurities during the switch from 8 M urea buffer to 0 M urea buffer. The first attempt at refolding was done with a buffer of 20 mM Tris pH 8.5, 150 mM NaCl, 5 % glycerol (first dialyzed into 8 M urea, then 0 M urea) to get an idea of the general characteristics of the construct under these conditions. A sample of protein was taken for analysis with Dynamic Light Scattering (DLS) and electron microscopy. Then, the

protein was split into six equal samples and refolding was attempted under the following conditions:

A) 20 mM Citrate pH 6, 150 mM NaCl, 5% glycerol

B) 20 mM Tris pH 8.5, 150 mM NaCl, 5% glycerol, 0.1 mM TCEP

C) 20 mM CAPS pH 10.5, 150 mM NaCl, 5% glycerol

D) 20 mM Tris pH 8.5, 300 mM NaCl, 5% glycerol

E) 20 mM Tris pH 8.5, 0 M NaCl, 5% glycerol

F) 20 mM Tris pH 8.5, 150 mM NaCl, 5% glycerol, 0.1% sodium deoxycholate

As above, the refolding process required that for each condition the protein was first dialyzed into 8 M urea buffer at room temperature overnight, then transferred into a urea-free buffer for that condition, and dialyzed overnight at 4°C. A sample of protein was taken from each condition and analyzed using Dynamic Light Scattering (DLS), Malvern Zetasizer S Model, to examine the size, aggregation, and dispersion of the particles that had formed. Samples were then taken from each condition and further examined using Transmission Electron Microscopy (TEM) to visualize the nanoparticles. After analyzing the results from the DLS machine and TEM, a new set of conditions in which to test refolding of the nanoparticles was created. The second set of conditions was as follows:

A) 20 mM Tris pH 8.5, 0 M NaCl, 5% glycerol

B) 20 mM Tris pH 8.5, 150 mM NaCl, 5% glycerol, 0.1% sodium deoxycholate

C) 20 mM Tris pH 8.5, 150 mM NaCl, 5% glycerol, 0.2% sodium cholate

D) 20 mM Tris pH 8.5, 0 M NaCl, 5 % glycerol, 0.1% sodium deoxycholate

E) 20 mM Tris pH 8.5, 0 M NaCl, 5% glycerol, 0.2% sodium cholate

F) 20 mM Tris pH 8.5, 50 mM NaCl, 5% glycerol

Again, the refolding process required that for each condition the protein was first dialyzed into 8 M urea buffer at room temperature overnight, then transferred into a urea-free buffer for that condition, and dialyzed overnight at 4°C. Samples were taken from each condition and analyzed using the DLS machine and were also examined using TEM to visualize the nanoparticles.

A different method called "quick refolding" was implemented to test the refolding conditions for the S43-9C-2IAd-M2eN construct. A small amount of the protein was thawed at 4°C and denatured by dialyzing it into urea buffer (9 M urea, 20 mM Tris pH 7.5, 150 mM NaCl, 5 % glycerol). The protein was then filtered to remove impurities and brought to a concentration of 1.0 mg/mL. 150 μL of the protein was added drop-wise to 2.85 mL of buffer for each refolding condition to obtain a final concentration of 0.05 mg/mL. The protein was added while the buffer was spinning extremely fast to ensure proper mixing of the protein in the buffer. The 3 mL total of protein and buffer was pipetted into dialysis tubing and placed in a beaker containing that buffer overnight at 4°C. The first type of condition to be tested was an ammonium sulfate buffer (20 mM Tris pH 7.5, 20 mM $(NH_4)_2SO_4$, 5% glycerol and 20 mM Tris pH 7.5, 50 mM $(NH_4)_2SO_4$, 5% glycerol). The results from the quick refolding were analyzed using DLS and electron microscopy. A new set of conditions were designed to further test the refolding of the S43 construct, as follows:

A) 20 mM Tris pH 7.5, 50 mM NaCl, 5 % glycerol

B) 20 mM Tris pH 7.5, 150 mM NaCl, 5% glycerol

C) 20 mM Tris pH 8.5, 150 mM NaCl, 5% glycerol

The samples were removed from dialysis tubing and pipetted into falcon tubes. Samples were taken for each condition to analyze using DLS and electron microscopy. Based on the results from the first set, a second set of refolding conditions was designed. Quick refolding was used again for the second set, however this time 250 μL of protein was refolded in 4.75 mL of filtered buffer, for a final volume of 5 mL. Again, the starting and final concentration of the protein was 1.0 mg/mL and 0.05 mg/mL respectively. The second set of conditions was as follows:

A) 20 mM Tris pH 7.5, 5 % glycerol
B) 20 mM Tris pH 7.5, 100 mM NaCl, 5% glycerol
C) 20 mM Tris pH 7.5, 50 mM NaCl, 5% glycerol, 0.2% sodium cholate
D) 20 mM Tris pH 7.5, 50 mM NaCl, 5% glycerol, 0.1% sodium deoxycholate

Following the refolding, samples were taken from each condition and analyzed using DLS, as well as TEM to visualize the nanoparticles.

Results

Protein Purification

SDS-PAGE was used to test for the presence of protein throughout the process of protein purification. The BN5C-M2eN-mh-ch-8r monomer has a molecular weight of 14.39 kDa but there could be cases where the protein forms a dimer or trimer, in which case the molecular weight would be double or triple that value, respectively. Normally, these bonds would be disassociated by SDS, but the coiled-coil interactions and di-sulfide bonds are very strong and SDS does not fully disrupt them. Therefore, when examining the gels, a band at the level corresponding to those values on the molecular weight marker could still indicate the protein of interest. Since the buffers that were used in the purification contain urea, which denatures proteins, it is more likely that we will find evidence of the monomer via a band at 14.39 kDa.

The first gel examined the results of the inclusion body preparation that immediately followed the lysing of the cells via sonication. The successive supernatants of the centrifugation (Washes 1-5) were tested along with a sample from the pellet, which had been dissolved in Inclusion Body Dissolving Buffer. It was anticipated that the protein would most likely be present in the fifth wash (Supernatant 5), and possibly the pellet as well. The fifth supernatant was a result of sonicating the pellet in Inclusion Body Dissolving Buffer and centrifuging at 12,000 rpm for 15 minutes. The results confirmed this, however the protein was left in urea solution for too long and it was believed that it was no longer usable for refolding. The process was restarted by again growing three 1 L cultures, combining the cell pellets, and sonicating in Cracking Buffer and Inclusion Body Dissolving Buffer. This time however, after the fifth centrifugation the pellet was resuspended in 2 mL Inclusion Body Dissolving Buffer so that the protein would be more concentrated.

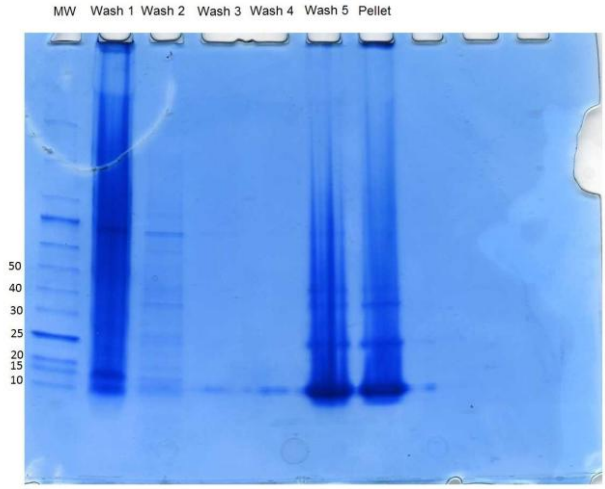


Figure 3. SDS-PAGE from inclusion body purification. The far left well is a molecular weight marker and the other wells are marked Wash 1 through 5, the resulting supernatants from centrifugation during the purification, and the resuspended cell debris pellet. The darkest bands in the Wash 5 and Pellet lanes fall between 10 and 15 kDa and confirm the presence of the protein of interest.

The gel was run once again with samples from each wash (supernatant) and the pellet. Again, it was expected that the BN5C-M2eN-mh-ch-8r-H1 protein would be present in Wash 5 as well as the pellet. The result is shown in Figure 3. The protein of interest has a molecular weight of 14.39 kDa so it would be represented by a band level with the space between the markers for 10 kDa and 15 kDa. The dark bands present at this level in the lanes for Wash 5 and the pellet show that the protein is present in both of these samples. Since the bands at this level are the darkest, the protein of interest is the most abundant protein in these washes. There are several lighter bands present in both lanes as well, indicating other contaminants. The protein is not very pure, but since this is only the first step of purification that is to be expected. Washes 3 and 4 also show faint bands at the correct weight, but they are very light, which indicates that there is very little of the protein contained in these samples. Thus, it would not be worthwhile to dilute the collection of protein at this level by adding the volumes of these washes if there is only

going to be a small gain in the amount of protein. Wash 5 and the pellet suspended in I.B. Dissolving Buffer were combined into a 50 mL falcon tube for the next step in the purification.

The SP-sepharose gel was set up and infused with the protein, then extracted with a NaCl gradient (50 mM, 100 mM, 250 mM, 500 mM, and 1 M) followed by a pH 8.5 buffer. The samples were collected in 2 mL fractions, ten for each buffer. The optical density of the samples at 280 nm (OD_{280}) was read using a spectrophotometer and used to determine which fractions should undergo further analysis using gel electrophoresis. The lower concentrations of salt, 50 mM and 100 mM, all had OD_{280} readings under 0.1, indicating that there was not much protein in the sample. Starting with the 250 mM NaCl gradient, the OD_{280} readings began to increase in value. Generally, values that had an OD_{280} greater than 1.0 were believed to contain a significant amount of protein and samples from these sets of fractions were analyzed via SDS-PAGE. There was at least one sample from each of the 250 mM, 500 mM, 1 M, and pH 10.5 gradients used in the gel electrophoresis.

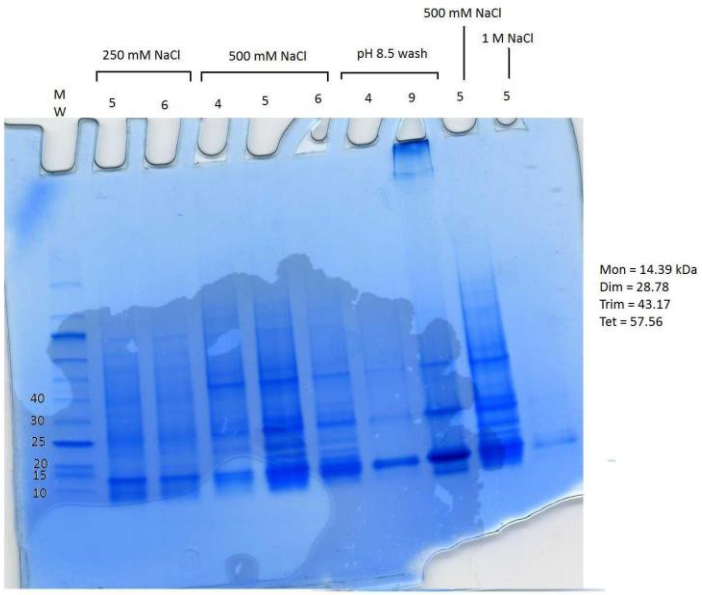


Figure 4. SDS-PAGE of fractions from SP-sepharose column purification. The far left column contains a molecular weight marker. The numbers below the buffer labels indicate the fraction the sample was taken from. The dark bands in the two fractions from the pH 8.5 represent the

presence of the protein of interest. The lighter bands at higher molecular weights may be the dimer and trimer of the protein.

Samples from fractions with high OD_{280} readings were run in the gel. The fifth fraction from the 500 mM NaCl buffer had an extremely high OD_{280} that was greater than 2.0 and thus unable to be read accurately by the spectrophotometer. The result is shown in Figure 4. The majority of the lanes show many bands at different heights, and none of the bands are very dark, indicating that there was not a large amount of protein. However, the two lanes containing samples from the pH 8.5 wash had dark bands slightly below the 15 kDa band in the molecular weight marker, which matches with the known size of the BN5C construct, 14.39 kDa. The bands are darker than any of the other bands on the gel and are by far the darkest in their respective lanes, showing that the protein of interest is more pure in these samples. There are lighter bands at higher molecular weights in these lanes but they match up with the molecular weights of the dimer and trimer of the BN5C construct, which were calculated to be 28.78 kDa and 43.17 kDa, respectively. These bands could be due to coiled-coil trimer formation or disulfide bond formation, which are strong enough to not be fully disrupted by SDS. There is a band just underneath the band representing the monomer of the protein in the sample for the ninth fraction of the pH 8.5 wash, but these are still the purest samples and the SP-sepharose column is only the second of three purifications. The fourth through tenth fractions from this wash had the highest OD_{280} readings and these seven fractions were combined in a 50 mL falcon tube in preparation for the next step of purification.

The next step in the purification for the BN5C construct was a Q-sepharose column. The OD_{280} for each of the fractions were read to estimate the amount of protein present. The values for the pH 10.5, 0 M NaCl, 100 mM NaCl, and 1 M NaCl fractions were essentially 0, but some

of the 500 mM fractions did have relatively high readings. However, the values were all under 1.0, much lower than the OD_{280} readings taken from the SP-sepharose column.

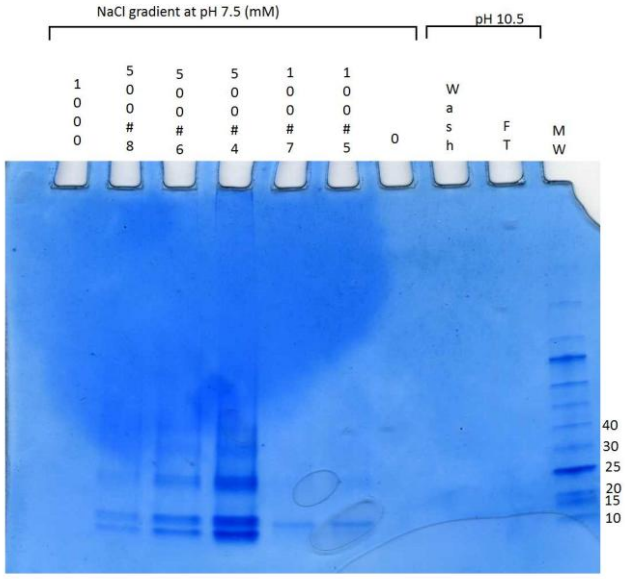


Figure 5. SDS-PAGE from Q-sepharose column purification. The far right lane contains a molecular weight marker. The only lanes that have visible bands are those with samples from the 500 mM NaCl wash. The fractions appear to co-elute. There also appears to be a band below the band associated with the monomer of the protein, which could represent impurities.

Samples from the fractions with the highest OD_{280} values from each wash were taken for further examination using SDS-PAGE. The results are shown in Figure 5. The only visible bands are in lanes that correspond to fractions from the 500 mM NaCl wash. The samples appear to completely co-elute, and while it is possible that all of the bands represent the protein of interest, it is not likely. Furthermore, in all three lanes of the 500 mM NaCl wash there is a band beneath the band which corresponds to the protein of interest. This appears to be some sort of contamination which means that the protein is not pure and will make the refolding step more difficult. There was a large step in the gradient from 100 mM NaCl to 500 mM NaCl and smaller gradient steps could yield a purer protein. Creating more steps in the gradient would at least create a clearer image of how the protein was purifying.

The Q-sepharose column was repeated with a smaller gradient. The OD_{280} readings for all of the fractions from the purification were extremely low. With the exception of a few fractions from the 200 mM NaCl and 300 mM NaCl gradients, the values were all essentially 0.

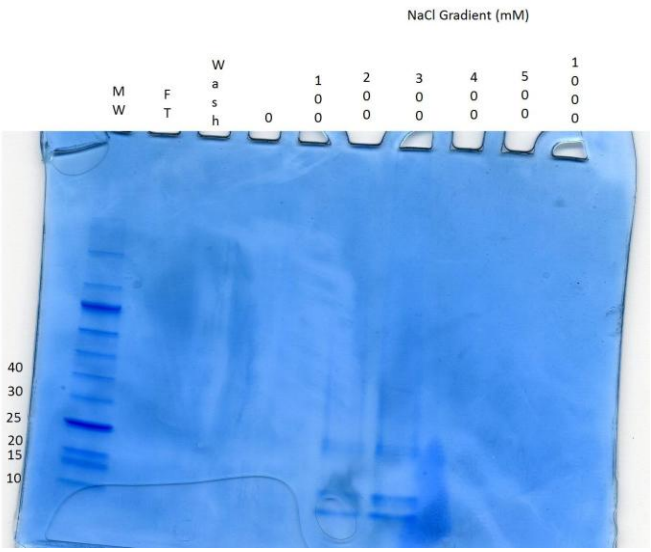


Figure 6. SDS-PAGE from second Q-sepharose purification. The far left lane contains a molecular weight marker. Bands appear in the 200 mM and 300 mM lanes, but are very faint. Again the samples seem to co-elute and there are bands visible beneath those that correspond to the protein of interest.

Samples from each of the washes as well as the flow through were used for SDS-PAGE analysis. The result is shown in Figure 6. The bands are much fainter than they were in the previous Q-column gel. The bands are in similar order to the previous gel, and it appears that the second column purification did not result in purer protein. In addition, the fractions still appear to co-elute, which is a positive sign in terms of the bands representing the protein that we are looking for. While the protein does not seem to be very pure and much of it was lost during the many rounds of purification, the only step to take at this point is to attempt to refold the construct. The fractions from the 200 mM NaCl and 300 mM NaCl washes were combined and filtered. Using the absorbance at 280 nm (A_{280}) it was estimated that the concentration of the protein was about 0.05 mg/mL. The refolding of the protein was attempted at this concentration.

The purification for the S43-9C-2IAd-M2eN construct was much simpler due to the his-tag contained in the protein. Thus, only one gel was run.. The S43 construct has a molecular weight of 14.19 kDa, so bands in the gel that represent the monomer of this protein will still be slightly below the 15 kDa molecular weight marker. Again, it is possible that the dimer or trimer of the protein could appear in the gels, but since the purification was done in urea buffer this is less likely and the monomer will appear more readily.

The his-tag in the construct enables the construct to bind very efficiently to the nickel beads in the Ni-NTA column. It was assumed that the protein would elute with the pH 5.9 and pH 4.3 washes. The OD_{280} of the imidazole fractions were measured using the spectrophotometer to find the peaks from each wash. The peak of each wash in the imidazole gradient represents the fraction most likely to contain the protein of interest.

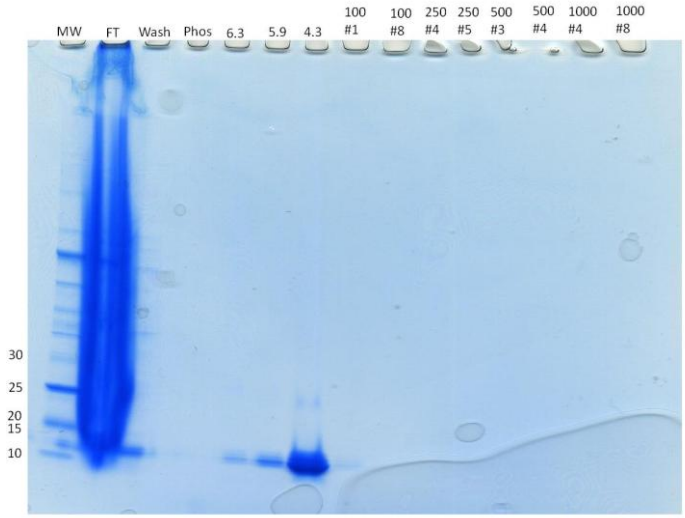


Figure 7. SDS-PAGE from Ni-NTA column purification. The far left lane contains a molecular weight marker. The extremely bright band in the pH 4.3 wash indicates a large amount of protein present. The band matches the molecular weight of the desired protein and there are no other visible bands. There is also a faint band at the same height in the pH 5.9 wash.

A sample of the fraction from each step in the imidazole gradient with the highest OD_{280} was analyzed via SDS-PAGE along with samples from the flow through, lysis buffer wash,

phosphate buffer wash, and each pH wash. The results are shown in Figure 7. The large dark band in the lane for the pH 4.3 sample shows the presence of the protein of interest. The height of the band matches the molecular weight for the S43 construct, 14.19 kDa. There is a very faint band below this, which may indicate a small amount of impurity. The faint bands in the pH 6.3 and 5.9 washes are at the same level and are likely also evidence of the protein of interest. These are the only three clear discernable bands on the entire gel. The band in the pH 4.3 wash is very dark, indicating that there is a lot of protein present, and it appears to be quite pure as well. The lighter bands from the samples of the other two pH washes are very light, and although they also seem to be pure, it is not worth diluting the concentration of the protein by combining these washes with the pH 4.3 wash.

Dynamic Light Scattering Reports

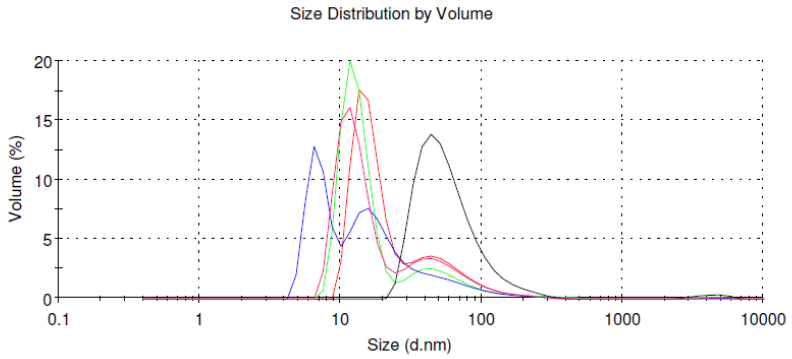
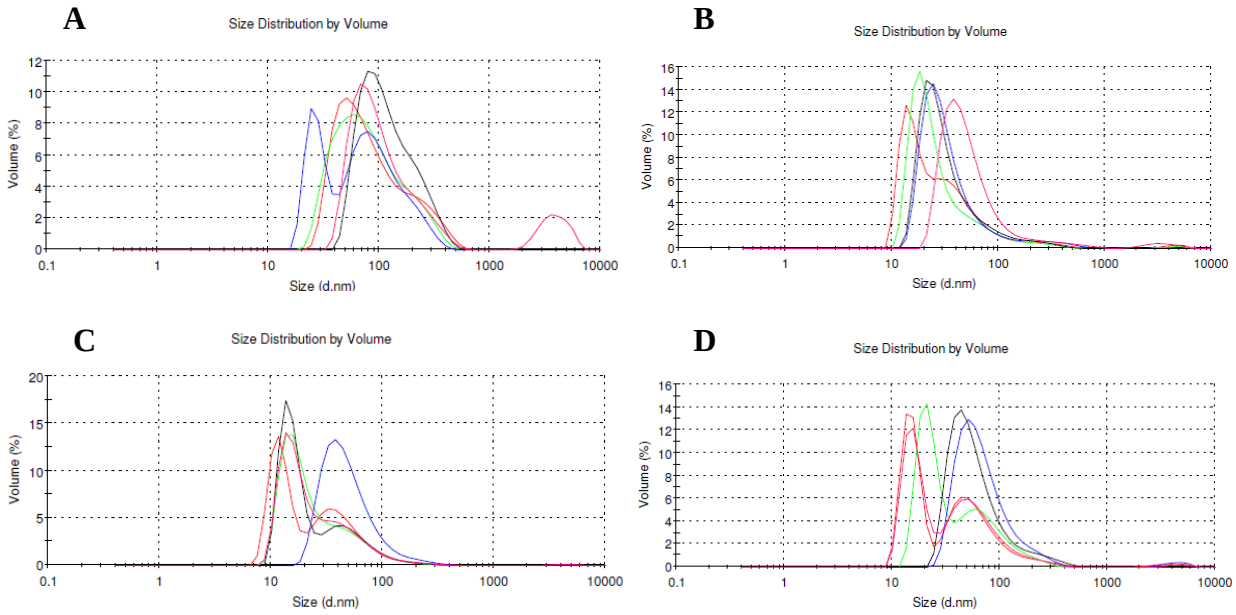


Figure 8. DLS report for first refolding attempt of BN5C-M2eN-mh-ch-8r-H1 construct. The first refolding attempt was done using a buffer containing 20 mM Tris at pH 8.5, 150 mM NaCl and 5% glycerol. There is not much agreement in size distribution.

Dynamic Light Scattering (DLS) reports are used to analyze the assembled nanoparticles with respect to size, aggregation, and dispersion of particles. The goal of refolding is to achieve nanoparticles that are uniform in size and shape. Ideally, the graph would be very narrow and tall with the peak at the correct size in diameter for the construct. It is expected that the correctly

refolded nanoparticles will be about 25-35 nanometers in diameter. Throughout the refolding process of the BN5C-H1 construct, the protein being refolded had a very low concentration. A baseline condition for BN5C-M2eN-mh-ch-8r-H1 was analyzed using DLS. The results are shown in Figure 8. The DLS report indicates the sample is poly-dispersed, as there is no agreement in the size of the particles. There is no evidence of uniformity among the particles as none of the sample runs match with each other, and there are several peaks visible within each run. The report indicates that the particles are not the correct size and the sample was not homogeneous. The peaks at higher molecular weight could be due to aggregation of the protein or incorrect formation of the nanoparticles. Despite the poor results, TEM images of particles from this refolding condition were still taken. Since these results did not give us positive refolding results, the next refolding attempt had a very wide range of conditions.



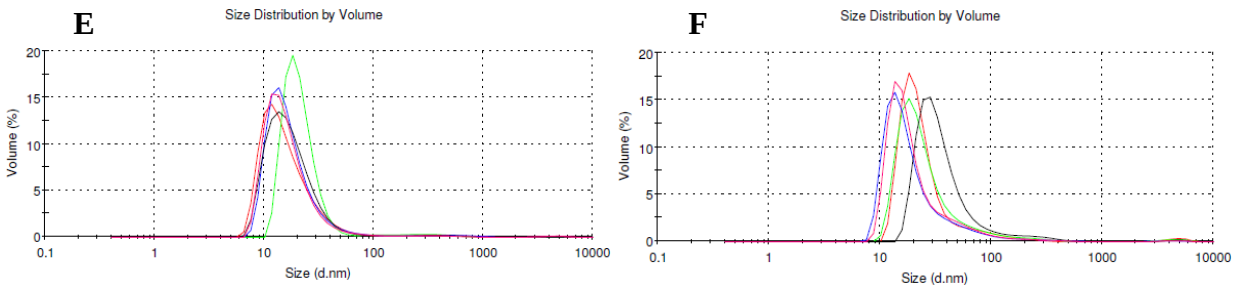
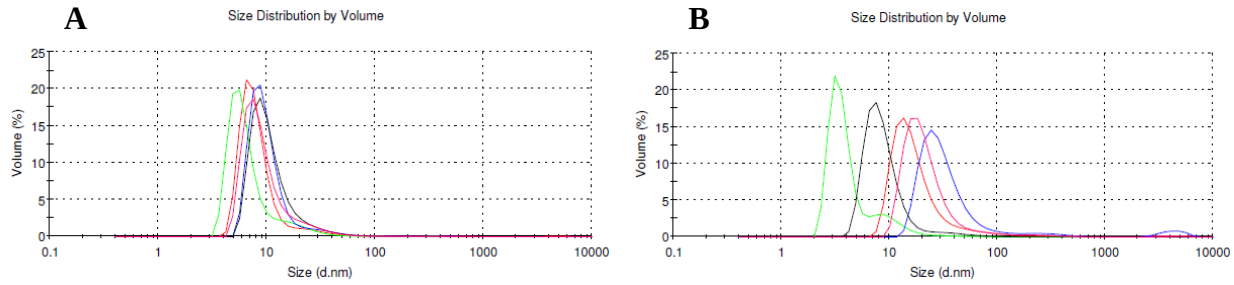


Figure 9. (A) DLS report of BN5C construct refolded in 20 mM Citrate pH 6, 150 mM NaCl, 5% glycerol. (B) DLS report of BN5C construct refolded in 20 mM Tris pH 8.5, 150 mM NaCl, 5% glycerol, 0.1 mM TCEP. (C) DLS report of BN5C construct refolded in 20 mM CAPS pH 10.5, 150 mM NaCl, 5% glycerol. (D) DLS report of BN5C construct refolded in 20 mM Tris pH 8.5, 300 mM NaCl, 5% glycerol. (E) DLS report of BN5C construct refolded in 20 mM Tris pH 8.5, 0 M NaCl, 5% glycerol. (F) DLS report of BN5C construct refolded in 20 mM Tris pH 8.5, 150 mM NaCl, 5% glycerol, 0.1% sodium deoxycholate.

Again, the nanoparticles should be about 25-35 nanometers in diameter if they are refolded correctly. The first set of conditions did not yield very positive results from the DLS, as shown in Figure 9. Condition A showed evidence of aggregation and appears to very poly-disperse, as there is not much agreement in size. Condition B is still poly-disperse but there is less aggregation of the protein. The same is true for Condition C while Condition D exhibits both high aggregation and poly-dispersity. The particles present in Condition E appear to be relatively uniform in size, but the diameter is slightly smaller than what was expected for this construct. However, it still shows the best results out of all of the conditions tested in this screen. Condition F is still poly-disperse and while there is no agreement in size, there doesn't appear to be much aggregation. The results from the first screen were taken into account when developing conditions for the second screen.



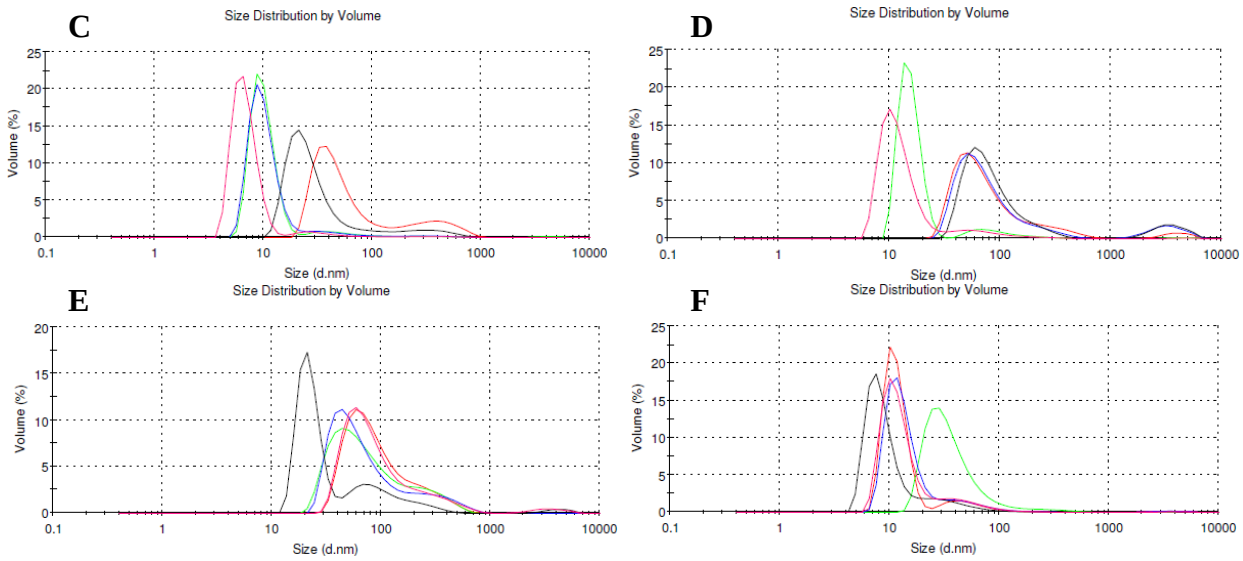


Figure 10. (A) DLS report of BN5C construct refolded in 20 mM Tris pH 8.5, 0 M NaCl, 5% glycerol. (B) DLS report of BN5C construct refolded in 20 mM Tris pH 8.5, 0 M NaCl, 5% glycerol, 0.1% sodium deoxycholate. (C) DLS report of BN5C construct refolded in 20 mM Tris pH 8.5, 0 M NaCl, 5% glycerol, 0.2% sodium cholate. (D) DLS report of BN5C construct refolded in 20 mM Tris pH 8.5, 0 M NaCl, 5% glycerol, 0.1% sodium deoxycholate. (E) DLS report of BN5C construct refolded in 20 mM Tris pH 8.5, 0 M NaCl, 5% glycerol, 0.2% sodium cholate. (F) DLS report of BN5C construct refolded in 20 mM Tris pH 8.5, 50 mM NaCl, 5% glycerol.

The results from the first screening showed that the condition with no salt and the condition with sodium deoxycholate yielded the best results. Thus, the conditions tested for the second screening all contained low or no salt and/or a detergent such as cholate or deoxycholate. The results are shown in Figure 10. All of the conditions exhibited a lot of poly-dispersity, with the exception of Condition A. Conditions A and F were the only samples to not show evidence of aggregation and both of these conditions contained peaks at a diameter that was smaller than expected for the BN5C-H1 construct. This could indicate protein that did not fold correctly or incompletely folded. In general, the results for the second screening were not any more positive than those of the first screening, despite developing the screen based on the best results from the previous one. Fifteen different conditions were tested using the BN5C-H1 construct, and none displayed good refolding according to the DLS results.

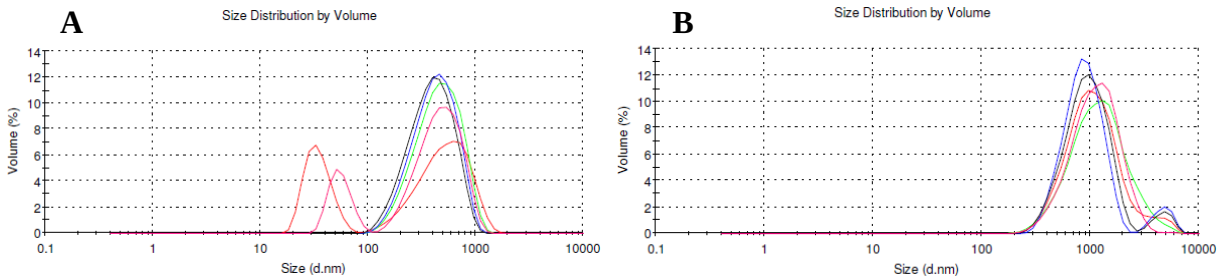


Figure 11. (A) DLS report of S43 construct refolded in 20 mM Tris pH 7.5, 20 mM $(NH_4)_2HPO_4$, 5% glycerol. (B) DLS report of S43 construct refolded in 20 mM Tris pH 7.5, 50 mM $(NH_4)_2HPO_4$, 5% glycerol.

The first test conditions for the S43-9C-2IAd-M2eN construct were buffers that contained ammonium sulfate. While previous constructs similar to S43 had shown positive results refolding in ammonium sulfate, the DLS reports generated for these conditions were not encouraging, as shown in Figure 11. While there is not as much poly-dispersity as was witnessed with the refolding of the BN5C construct, there seemed to be a high level of aggregation. Similar to the previous construct, the S43 nanoparticle is expected to have a diameter of around 25-30 nanometers. The peaks shown in Figure 11 are over 50 nanometers for both refolding conditions. The next step involved screening the refolding of the S43 protein with more conventional conditions to get a general idea of the refolding properties of the particular construct.

The results of the first screen for the S43 construct (not shown) were not encouraging. Condition A is both poly-disperse and appears to be highly aggregated. Condition B shows only the solvent peak, implying that there was no protein present in the sample. This needs to be further examined using electron microscopy to determine the cause of the lack of particles. Condition C contains much less aggregation and the peaks are around the diameter that the S43 construct is expected to be, but there is a lot of poly-dispersity, meaning that the particles are not uniform in size. These results were taken into account when creating the next set of screening conditions.

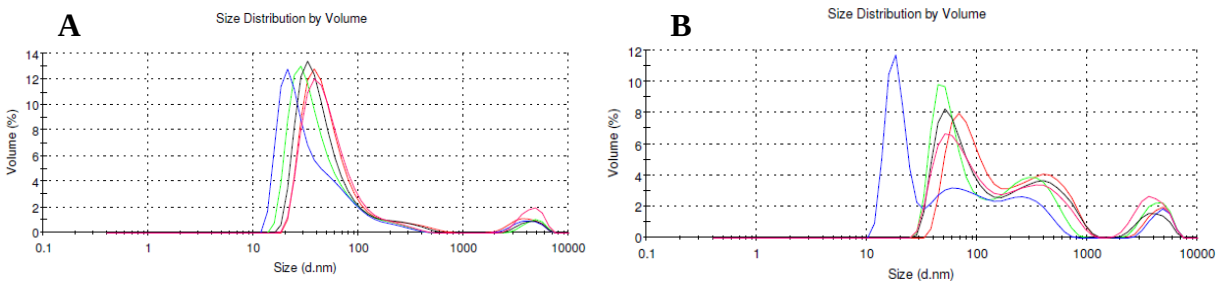


Figure 12. (A) DLS report of S43 construct refolded in 20 mM Tris pH 7.5, 50 mM NaCl, 5% glycerol, 0.2% sodium cholate. (B) DLS report of S43 construct refolded in 20 mM Tris pH 7.5, 50 mM NaCl, 5% glycerol, 0.1% sodium deoxycholate.

The second screen of the S43 construct contained four different conditions, two of which are shown in Figure 12. Two of the conditions, which did not contain any detergent, only showed solvent peaks on the DLS reports and are not shown. It is believed that this may be due to massive aggregation. The condition which contained sodium cholate, was only slightly poly-disperse and peaked around 50 nm, which is slightly higher than the expected size of the nanoparticles. The condition which contained sodium deoxycholate, was poly-dispersed and contained peaks in the expected range of the nanoparticle diameter, but also showed signs of aggregation.

Electron Microscopy

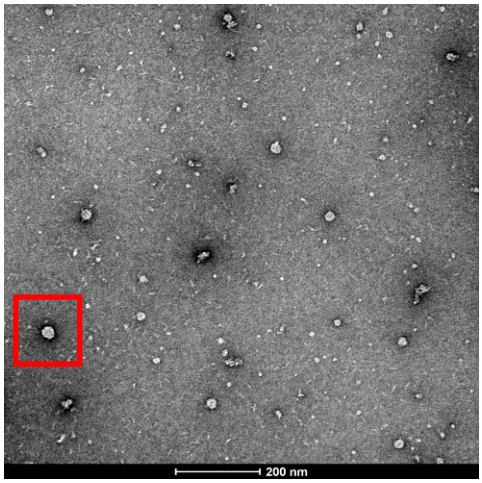


Figure 13. EM image of the BN5C-M2eN-mh-ch-8r-H1 protein after refolding in 20 mM Tris at pH 8.5, 150 mM NaCl and 5% glycerol. The particle within the red square is an example of what a correctly formed nanoparticle might look like under electron microscopy.

Regardless of the results from the DLS report, samples from each of the refolding conditions performed for the BN5C construct were analyzed with electron microscopy. The EM images show the nanoparticles, or lack thereof, in the sample and are a more definitive way to determine if the refolding was successful. An EM image for the first refolding attempt, in 20 mM Tris pH 8.5, 150 mM NaCl, 5% glycerol, is shown in Figure 13. The red box encloses an example of a well shaped nanoparticle. The particle is spherical and the black ring around the particle indicates that the particle has thickness, meaning that the spherical shape exists in three dimensions. The size of the particle appears to be within the range of 25-35 nanometers that is expected for the BN5C construct. But although there is evidence of a correctly formed particle, this refolding event was still not successful. There were very few particles of this nature present in the EM images for this refolding condition and much of the protein formed small misshaped particles or strings. Some of the images from this refolding attempt show massive aggregation. A good refolding condition would contain many correctly formed particles with very little unfolded protein or aggregates.

The EM images for the first refolding screen, for which the DLS results are shown in Figure 9, were not any more positive. The images for Condition A show some spherical particles, but none of them display any thickness. There is also a lot of aggregation visible. The images for Condition B display mostly misshaped or small particles, again with no thickness, or aggregated protein. There are a small number of particles that look like they may have formed correctly, but not enough to consider it a successful refolding. Almost all of the images for Condition C contain massive aggregation. Again, the images for Condition D contain a lot of aggregation.

There are a few examples of spherical particles, but they do not appear to have any thickness. Condition E contains a lot of misshaped particles and protein strands, while Condition F only contains aggregated protein.

The second set of screening conditions was not any better. Condition A only showed misshaped strands and particles that did not form correctly. The images for Condition B and Condition C contained mostly aggregated protein. Images for Condition D did not show any protein but contained a lot of crystal-like fragments, most likely due to an interaction between the deoxycholate and the stain used for the electron microscope. The same goes for Condition E and the cholate. The images for Condition F did not show any nanoparticles, only small misshaped strands of protein.

The first refolding attempt for the S43-9C-2IAd-M2eN-ch construct was done in a buffer containing ammonium sulfate. The EM images for the 20 mM condition contain small misshaped particles that do show evidence of having thickness, but are not spherical nor the size that we would expect. The 50 mM condition shows more of the same, with very small non-spherical particles. The first screen of the S43 construct did not show very many particles either. The images from each of the three conditions all exhibited a lot of protein aggregation, with very few particles around the expected 30-50 nm size.

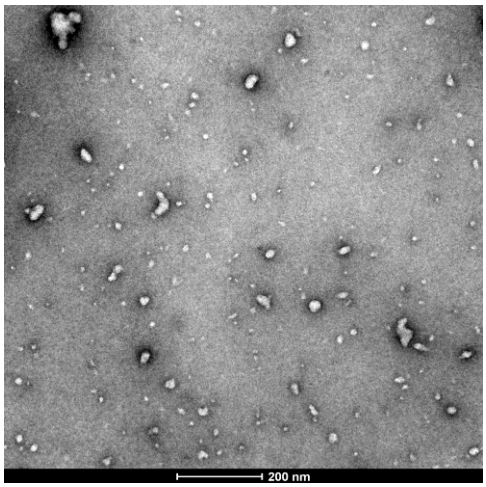


Figure 14. EM image of the S43-9C-2IAd-M2eN construct refolded in 20 mM Tris pH 7.5, 50 mM NaCl, 5% Glycerol, 0.1% Sodium deoxycholate. Several nanoparticles of the correct size are visible and there is little aggregation or unfolded protein.

The second screen had mixed results. The two conditions that did not contain detergents displayed only aggregated protein. The third condition, with sodium cholate, had a few nanoparticles of the correct size but was composed mostly of misshaped protein and aggregation. The fourth condition from the second screen, containing sodium deoxycholate, is the best in terms of nanoparticle formation. The EM image for this condition is shown in Figure 14. There are several examples of nanoparticles with the expected diameter for the S43 construct. There is still some aggregation and unfolded protein present, but not as much as the other conditions. The relative success of this condition indicates that deoxycholate could be an important factor in the refolding of this construct.

Discussion

The results of the biophysical analysis of the BN5C-H1 construct did not reveal any promising refolding conditions, despite the number of screens performed. The S43 construct, on the other hand, showed some success using sodium deoxycholate, though not with any other type of condition. In general, the refolding of the constructs whose structure contain a tetramer and a trimer, such as the BN5C and S43 constructs, have not been as successful as structures containing a pentamer (Neef and Burkhard, unpublished data). It is unclear whether this is more due to the structure itself, or the relative unfamiliarity with the refolding properties of the tetramer/trimer combination.

One reason for the lack of nanoparticles observed with the BN5C-H1 construct may have been the low concentration of protein used for refolding. The bands present on the SDS-PAGE gels during the latter stages of purification were very faint. After the various rounds of purification, very little protein was actually obtained, and it had a very low concentration, about 0.05 mg/mL. The fact that the protein was already at such a low concentration made it more difficult to work with and may have contributed to the poor results seen in the DLS graphs and EM images. There was little protein present to assemble into nanoparticles, which correlates to the lack of spherical nanoparticles observed in the EM images of the refolding screens. Therefore, the lack of nanoparticles could be due to a lack of protein and especially the difficulty of removing impurities. The BN5C construct was never adequately purified.

This highlights the importance of the his-tag present in the S43 construct. The six histidines at the end of the sequence made the protein much easier to purify. The binding and unbinding of the protein to the Nickel column was facilitated by a change in pH and the protonation of the histidines. This resulted in a more highly concentrated protein following

purification, as indicated by the band in the SDS-PAGE gel. Since the protein was present at a high concentration, we were able to use a different refolding technique. The difference was evident in the amount of protein present in the EM images, whether or not it had correctly refolded. The ideal result of self-assembly would be many self-assembled nanoparticles with very little unfolded protein present. Therefore, having a relatively high concentration of protein is a very important element of refolding. Going forward, the his-tag appears to be an essential component of creating a nanoparticle influenza vaccine.

The best looking nanoparticles in the EM images are from the S43 construct refolded in 0.1% deoxycholate. Many distinct nanoparticles are present, but there were also signs of aggregation. The DLS report also showed aggregation and was still fairly poly-disperse. The purpose of the screen was to identify a condition in which the particles seemed to form best, and use this data to create a more specific refolding condition. This goal was achieved in that deoxycholate has been identified as an important component for the refolding of the S43 construct. The assembly of the nanoparticles depends on sustaining the intra-molecular bonds of the nanoparticle while avoiding inter-nanoparticle bonds, so it is not surprising that a detergent appears to be an important factor in the refolding of this construct. Using this data, a new set of conditions can be created based on the inclusion of deoxycholate. There is still a lot of work to be done to find a specific condition that will allow for the optimal refolding of the S43 construct. In order to proceed to testing the nanoparticles as a vaccine, the protein must be able to consistently refold with uniform results.

The S43-9C-2IAd-M2eN construct shows promising results for purifying and refolding and following further work to find the best possible refolding condition, it could move to the next stage of developing an avian influenza vaccine. Although the S43 construct has not yet

shown evidence of efficient and complete refolding, progress has been made to finding the optimal condition. Depending on the success of subsequent refolding attempts, the next step is to test the ability of the nanoparticles to elicit an immune response in animal models to determine the viability of an avian influenza vaccine composed of S43 peptide nanoparticles.

References

1. World Health Organization. Acute Respiratory Infections. Initiative for Vaccine Research, 2009.
2. Nicholls, H. Pandemic Influenza: The Inside Story. Public Library of Science Biology, 2009. 4(2): e50.
3. Wareing, M.D. and Tannock, G.A. Influenza update: vaccine development and clinical trials. Current Opinion in Pulmonary Medicine, 2002. 8(3): 209-213.
4. Fichero, E. et. al. New Strategies to Overcome the Drawbacks of Currently Available Flu Vaccines. Advances in Experimental Medicine and Biology, 2009. 655:243-252.
5. Kaba, S. A., et al. Nonadjuvanted Polypeptide Nanoparticle Vaccine Confers Long-Lasting Protection Against Rodent Malaria. The Journal of Immunology, 2009. 183(11): 7268-7267.
6. Pimentel, T. A. P. F. Peptide Nanoparticles as Novel Immunogens: Design and Analysis of a Prototypic Severe Acute Respiratory Syndrome Vaccine. Chemistry and Biology of Drug Design, 2009. 73: 53-61.
7. Okamoto, S. et. al. Influenza hemagglutinin vaccine with poly(γ-glutamic acid) nanoparticles enhances the protection against influenza virus infection through both humoral and cell-mediated immunity. Vaccine, 2007. 25(49): 8270-8278.
8. Kreuter, J. Nanoparticles as Adjuvants for Vaccines. Pharm Biotechnology, 1995. 6: 463-472.
9. Adair, B. M. Nanoparticle vaccines against respiratory viruses. Nanomedicine and Nanobiotechnology, 2009. 1(4): 405-414.
10. Senthilkumar, R. et. al. Structure based design of peptides that self-assemble into regular polyhedral nanoparticles. Nanomedicine: Nanotechnology, Biology and Medicine, 2006. 2: 95-102.
11. Zonik, B.S. et. al. Nanoparticles and the Immune System. Endocrinology, 2010. 151(2): 458–465.
12. Kowalczyk, W. et. al. Strategies and limitations in dendrimeric immunogen synthesis. The influenza virus M2e epitope as a case study. Bioconjugated Chemistry, 2010. 21(1): 102-110.
13. Kim, T.S. et. al. T cell responses during influenza infection: getting and keeping control. Trends in Immunology, 2011. [E-pub ahead of print].
14. Burkhard, P. et. al. Design of a minimal protein oligomerization domain by a structural approach. Protein Science, 2000. 9(12): 2294-2301.
15. Burkhard, P. et. al. Coiled coils: a highly versatile protein folding motif. Trends in Cell Biology, 2001. 11(2): 82-88.
16. Burkhard, P. et. al. Improving Coiled-coil Stability by Optimizing Ionic Interactions. Journal of Molecular Biology, 2002. 318(3): 901-910.
17. Sambrook, Joseph and Russel, David W. Molecular Cloning: A Laboratory Manual Third Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. 2001. A3.6, 1.118.