Part III
Materials and Methods
Chapter 7
Purification, crystallization, data collection, processing and refinement

7.1 Cloning and expression of the NusG full length, NTD and CTD domains:

NusG full length (1-182) from *Vibrio cholerae* O395 was cloned into pET-28 (a+) vector under the control of T7 promoter. The complete sequences of this construct were confirmed by DNA sequencing. The 6xHis-tagged NusG full length (NusG-FL) was expressed into BL21 (DE3) strain of *E. coli*. The NTD and CTD domains of NusG were cloned separately in pET28a+ vector (Figure 7 a, b) and expressed in BL21 (DE3) *E. coli* cells.

The sequences of the forward and reverse oligos along with the recognition sequences of the restriction enzymes (bold) are listed in Table 1.

Table 1: Primer Sequences

<table>
<thead>
<tr>
<th>Clone name</th>
<th>Forward Primer- with NdeI site</th>
<th>Reverse Primer – with BamHI site</th>
</tr>
</thead>
<tbody>
<tr>
<td>NusG-FL</td>
<td>5’GGAATTCCATATGATGAGTTGAAGCTCCCAAAAAACGC 3’</td>
<td>5’GGGGATCCTTAAATCAAGCTTTCCACTTGACCAAATTCG 3’</td>
</tr>
<tr>
<td>NusG-NTD</td>
<td>5’GGAATTCCATATGATGAGTGAAGCTCCCAAAAAACGC 3’</td>
<td>5’GGGGATCCTTATGGTGCTTCGCTGGCTTTCTCAAGACG 3’</td>
</tr>
</tbody>
</table>
Clone name | Forward Primer - with Ndel site | Reverse Primer – with BamHI site
--- | --- | ---
NusG-CTD | 5'GGAATTCCATATGCCTCAAACCATGTTTTGAAGCAGGTGAAG 3' | 5'CGGGATCCCTTAATCAAGCTTTCCACTTGACCAAATTCG 3' 
Rho | 5'GGAATTCCATATGATGAATCTAACAGAACTGAAGAAC 3' | 5'CGGGATCCCTTCATAAATCATTCATCCATCTTAGAA 3' 

**PCR:** The annealing temperature $T_m$ was chosen to be approximately 5K below the lowest (theoretical) melting temperature of either oligos and 31 PCR-cycles were carried out. The final volume of the reaction mixtures were made up to 100 μl and the program used for the second step PCR was as following.

(i) Initial denaturation at 94°C for 5 minutes.
(ii) Denaturation at 94°C for 30 seconds
(iii) Annealing at 65°C for 30 seconds
(iv) Polymerization at 72°C for 45 seconds
<table>
<thead>
<tr>
<th>Components</th>
<th>Final concentrations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Template</td>
<td>20 ng</td>
</tr>
<tr>
<td>Forward primer</td>
<td>0.8 pmol</td>
</tr>
<tr>
<td>Reverse primer</td>
<td>0.8 pmol</td>
</tr>
<tr>
<td>dNTP mix (dATP, dGTP, dTTP and dCTP)</td>
<td>250 pmol</td>
</tr>
<tr>
<td>10X Buffer</td>
<td>1X</td>
</tr>
<tr>
<td>Taq Polymerase</td>
<td>2 unit</td>
</tr>
<tr>
<td>H₂O to made up the</td>
<td>100 µl volume</td>
</tr>
</tbody>
</table>

**Ligation:** The PCR products were purified with the Qiagen “PCR purification kit”, digested with restriction enzymes (NdeI and BamHI) in one reaction, purified and finally ligated into the target vector. The vector was previously digested with NdeI and BamHI (NEB) enzymes.

<table>
<thead>
<tr>
<th></th>
<th>Standard reaction</th>
<th>Positive control</th>
<th>Negative control</th>
</tr>
</thead>
<tbody>
<tr>
<td>10X Ligase Buffer</td>
<td>1 µl</td>
<td>1 µl</td>
<td>1 µl</td>
</tr>
<tr>
<td>(Promega)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pET 28a+ Vector</td>
<td>1 µl</td>
<td>1 µl</td>
<td>1 µl</td>
</tr>
<tr>
<td>(Promega)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCR product</td>
<td>1 µl</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>T4 DNA Ligase</td>
<td>1 µl</td>
<td>1 µl</td>
<td>1 µl</td>
</tr>
<tr>
<td>(Promega)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H₂O</td>
<td>6 µl</td>
<td>6 µl</td>
<td>7 µl</td>
</tr>
<tr>
<td>Total</td>
<td>10 µl</td>
<td>10 µl</td>
<td>10 µl</td>
</tr>
</tbody>
</table>
**Transformation:** The ligation product was directly transformed into electrocompetent cells (XL1-blue, Stratagene) according to the manufacturer’s instructions and plated on agarose plates with 30μg/ml kanamycin.

**Analysis:** Plasmids were purified from colonies by mini-preps (Promega or Qiagen) and analysed on agarose gels after digestion with the corresponding enzymes. The sequence of the clones has been confirmed by sequencing.

7.2 Expression and purification of recombinant proteins in E. coli

7.2.1 Protein expression

Overexpression of the 6XHis mutant proteins was carried out in *Escherichia coli* strain BL21 (DE3) (Novagen) in LB medium at 37°C. The BL21 (DE3) competent cells were transformed with the recombinant pET28a (+) plasmids containing the insert. From the transformed colonies, one single colony was inoculated in 2ml of fresh LB medium containing kanamycin up to a final concentration of 30 μg/ml and grown overnight at 37°C. 50 μl of this overnight culture was inoculated to 5 ml of fresh LB medium and allowed to grow till OD₆₀₀ reached to ~0.4-0.5. 1 ml aliquot was taken out at that point to preserve uninduced cell and rest of the culture was induced by IPTG with a final concentration of 0.2mM followed by growth for another 4 hours at the same temperature. Then 500μl aliquots from both uninduced and induced cells were harvested in a 1.5 ml microcentrifuge tube by centrifugation at 12000 g for 2 min and the resulting pellet was suspended in 30 μl of H₂O, vortexed and mixed with 10μl of 4x SDS-PAGE sample buffer (250 mM Tris.HCl of pH 6.8,
10% SDS, 40% glycerol, 5% Bromophenol blue and 20% β-mercaptoethanol). The cells were lysed by boiling in water bath for 5 min and after centrifuging at 12000 g for 5 min, 10µl of the supernatant was loaded in 12.5% SDS-polyacrylamide gel (62) run at 200V and 20 mAmp. Overexpression of the recombinant mutant protein was checked in the gel stained with Coomassie R-250 (Sigma) and appearance of an approximately 22 kDa protein in the induced cells compared to the uninduced cells confirmed the overexpression of the recombinant mutant proteins.

### 7.2.2 Purification of 6xHis-tagged Proteins

Purification of the His-tagged protein by Ni²⁺-NTA is based on the remarkable affinity between the six consecutive histidine residues — ‘the 6×His tag’, either at the N-terminal or at the C-terminal of the proteins, and Ni²⁺ ion immobilized on Nitrilotriacetic acid (NTA) resin. This affinity is even greater than that of the antigen-antibody or the enzyme-substrate interactions. Untagged proteins present in the cell lysate having histidines on the surface/close proximity to the surface may also bind to the Ni²⁺-NTA along with the desired ‘His tagged’ protein but their interactions with Ni²⁺-NTA, in practice, will be much weaker compared to the specific interactions of the 6×His tagged protein with Ni²⁺-NTA. So, the unbound and/or nonspecifically bound proteins to the matrix can easily be washed away under relatively stringent conditions without affecting the binding of the target protein. The target protein can be elution from the column either with higher concentration of imidazole or slight reduction in pH.
For the purification of the full length NusG (NusG-FL), overnight cultures of BL21 (DE3) cells harboring the recombinant pET28a (+) plasmid that carries the corresponding inserts were inoculated (1:100 dilution) to fresh LB medium supplemented with Kanamycin (final concentration of 30 µg/ml). The cultures were then grown at 37°C with vigorous shaking until the OD₆₀₀ reached to approximately 0.4-0.5. Expression of the desired proteins was induced at that point by addition of IPTG to a final concentration of 0.2mM followed by an incubation of another 4 hours at 37°C with vigorous shaking. The bacterial cells, harvested from the liquid culture by centrifugation at 4500 rpm for 15 min at 4°C, were resuspended in ice-cold lysis buffer having 20 mM Tris (pH 7.5), 100 mM NaCl, 10mM EDTA, 1% Triton X-100. Lysozyme (final concentration 1mg/ml) and PMSF (final concentration 1 mM) were added in the suspended solution right then and incubated on ice for 30 minutes. The suspension was sonicated using 15 bursts of 10 secs at 200 W with an interval of 10 secs. The supernatant was separated from the insoluble cell component that comes into pellet by centrifugation of the crude lysate at 12000g for 45 min at 4°C. The pellet was resuspended in 0.1 culture volume of the lysis buffer and again centrifuged at 12000g for 45 min at 4°C. After separating the supernatant, the weight of the pellet was measured and resuspended in the solubilization buffer containing 50mM CAPS (pH 11.0), 0.3% N-lauryl sarcosine, 2mM DTT at the final concentration of 20mg/ml. It was incubated at room temperature (RT) for fifteen minutes followed by a centrifugation step at 12000g for 15 min at room temperature. The supernatant was collected carefully without disturbing the pelleted debris and dialysed against the lysis buffer. To ensure complete removal of N-lauryl sarcosine, turbidometric assay was performed. Briefly, a standard curve of serially diluted N-lauryl sarcosine was prepared at OD 405nm and then the supernatant sample was analysed to
determine residual amount of the detergent. The sample was then taken for affinity chromatography.

Procedure for the purification of Nus-NTD or Nus-CTD is almost same. Briefly, for the purification of the Nus-NTD (or Nus-CTD), overnight cultures of BL21 (DE3) cells harboring the recombinant pET28a (+) plasmid that carries the corresponding inserts were inoculated (1:100 dilution) to fresh LB medium supplemented with Kanamycin (final concentration of 30 μg/ml). The cultures were then grown at 37°C with vigorous shaking until the OD$_{600}$ reached to approximately 0.4-0.5. Expression of the desired proteins was induced at that point by addition of IPTG to a final concentration of 0.2mM followed by an incubation of another 4 hours at 37°C with vigorous shaking. The bacterial cells, harvested from the liquid culture by centrifugation at 4500 rpm for 15 min at 4°C, were resuspended in ice-cold lysis buffer having 50 mM NaH$_2$PO$_4$ (pH 8.0), 500 mM NaCl. Lysozyme (final concentration 1 mg/ml) and PMSF (final concentration 1 mM) were added in the suspended solution right then and incubated on ice for 30 minutes. The suspension was sonicated using 15 bursts of 10 secs at 200 W with an interval of 10 seconds. The supernatant was separated from the insoluble cell component that comes into pellet by centrifugation of the crude lysate at 12000g for 45 min at 4°C. The cleared lysate containing the soluble 6×His tag protein was applied onto a column packed with 1.5 ml slurry of Ni$^{2+}$-NTA resin, pre-equilibrated with the lysis buffer. Weakly/non-specifically bound proteins were washed from the resin with the buffers having gradually increased imidazole concentrations (5mM, 10mM and 15 mM). The recombinant His-tagged proteins were stripped from the Ni$^{2+}$-NTA columns with 150 mM imidazole in the same buffer and the elution was monitored by measuring the absorbance at [47]
280 nm. The purity of the eluted protein was checked in 15% SDS-PAGE. The eluted fractions contain the target proteins with few minor contaminants. These fractions containing NusG-NTD (or NusG-CTD) were separately pooled and concentrated using 3 kDa cut-off centrifugal unit of Millipore at 4500g at 4°C.

7.2.3 Removal of His-tag with thrombin

Prior to the cleavage of the His-tag, the proteins were transferred to a buffer containing 50 mM Tris.HCl (pH 8.5) containing 300 mM NaCl which is the buffer for the optimum pH (8.5) and activity of thrombin. NaCl was added to avoid any sort of aggregation that may form in the highly concentrated protein and its concentration was standardized accordingly. This buffer exchange was necessary as the high salt and imidazole present in the elution buffer may interfere with the cleavage by thrombin. Moreover, it is reported that imidazole can hamper the success rate of crystallization as the presence of imidazole often results in protein aggregation (63). For the cleavage reaction, the amount of thrombin, temperature and time of incubation were calibrated previously. However, for each individual protein, certain modifications and optimization were found to be necessary. In general, about 6-10 mg of the fusion protein was digested with 2 units of thrombin (Novagen) at 4°C for 38-40 hr. where final reaction volume was restricted to 50μl.

7.2.4 Size exclusion chromatography using Sephacryl S-100

After the thrombin cleavage, the proteins were further purified from the cleaved His tag, other few minor contaminants eluted from the Ni²⁺-NTA column and thrombin by size exclusion chromatography using Sephacryl S-100 (GE Healthcare Biosciences). The reaction
mixture set for the thrombin cleavage was loaded directly on a Sephacryl S-100 column of dimension 85×1.5 cm, pre-equilibrated with 50 mM Tris.HCl (pH 8.0) containing 300 mM NaCl. 0.02 % Na-Azide, pH 8.0 was added to the buffer to avoid any sort of microbial contamination.

7.2.5 Preparation of Psu WT and T123C mutant proteins

The wild type Psu protein was expressed and purified by a three step process involving a 25% ammonium sulphate precipitation in buffer containing 10 mM Tris–HCl pH 7.9, 0.1 mM EDTA, 0.1 mM DTT and 5% (v/v) glycerol followed by purification through Q-sepharose (GE Healthcare) and CM sepharose columns (GE Healthcare). Psu was eluted between 50 and 150 mM NaCl and the eluted fractions were stored in 20 mM Tris–HCl pH 7.9, 0.1 mM EDTA, 0.1 mM dithiothreitol and 100 mM NaCl. This procedure yielded about 95% pure protein. The T123C change was made by site-directed mutagenesis on the zero-cys Psu and was purified in a similar way as WT.

7.3 Crystallization of the target proteins

7.3.1 Crystallization of NusG-NTD and NusG-CTD

The proteins were concentrated to ~10–12 mg/mL in a buffer containing 50 mM Tris (pH 7.0) and 300 mM NaCl. Crystal growth was accomplished using hanging drop vapour diffusion method against 700 μL of reservoir and incubated at either 4°C or 20°C. PEG grid screen, Ammonium Sulphate Grid Screen, Crystal Screen and Crystal Screen 2 from Hampton and Qiagen Nextal Suite were used as precipitants.

[49]
7.3.2 Data Collection and Processing

Crystals of NusG CTD were fished out from the crystallization drops using a 10 mm nylon loop (Hampton Research, Laguna Niguel, California, USA), briefly soaked in a cryoprotectant solution consisting of 7% (v/v) glycerol, 300 mM NaCl in 50 mM Tris, pH 7.0 and flash-frozen in a stream of nitrogen (Oxford Cryosystems) at 100 K. A native diffraction data set was collected to 2.8 Å resolution (Figure 6.5) using an in-house MAR Research image plate detector of diameter 345 mm and Cu Kα radiation generated by a Bruker-Nonius FR591 rotating-anode generator equipped with Osmic MaxFlux confocal optics and running at 50 kV and 90 mA. A total of 131 frames were collected with a crystal-to-detector distance of 250 mm. The exposure time for each image was 4 min and the oscillation range was maintained at 1°. Data were processed and scaled using imosflm (64). Data-collection and processing statistics are given in Table 2.

7.3.3 Crystallization, data collection and structure determination of the wild type Psu protein

The crystallization trials of Psu was carried out at 277 K and 293 K. Typically, 3 µl protein solution (5 mg/ml) was mixed with 2 µl reservoir solution containing 7.5% (w/v) PEG 6000, 5% (v/v) glycerol, 0.5 mM DTT and 300 mM NaCl in 0.1 M MES pH 6.0 at 277 K. Cube shaped crystals appeared after 7 days (Figure 6.6 a, b). For phasing, the crystals of Psu were soaked in 2mM HgI₂ in the reservoir solution for 2 hours. X-ray data were collected in-house at 100 K using a MAR Research image-plate detector of diameter 345 mm and Cu Kα radiation (1.54 Å) generated by a Bruker-Nonius FR591 rotating-anode generator.
equipped with Osmic MaxFlux confocal optics and running at 50 kV and 90 mA. A total of 129 frames were collected with a crystal-to-detector distance of 210 mm. The exposure time for each image was 4 min and the oscillation range was maintained at 1 degree (Figure 6.6 c). The data was integrated with imosflm (64) and scaled with scala in the imosflm GUI (64). Phases from the single wavelength anomalous dispersion (SAD) data were calculated in Phenix Autosol (65) with a figure-of-merit of 0.48 and 1 Hg site refined. The density was modified with mask type wang with Resolve in Phenix Autosol which produced a readily interpretable map. The initial model was built with Phenix Autobuild (33), building 142 out of 190 residues and placing 91 of the side chains. The rest of the model was built manually in Coot (66) and O (67) and iterative building cycles in Phenix Autobuild. The model was refined with Phenix Refine (65). The final R factor was 19.21 and R_{free} 21.87 including TLS (68) with excellent stereochemistry. Density was absent for the first 3 residues giving a final model of 187 residues out of 190. The Ramachandran plot in Molprobity (69) shows that more than 98% of the residues are in the most favoured region with none in the disallowed region. The data processing and refinement statistics are summarized in Table 3 and Table 4 respectively.

7.3.4 Crystallization, data collection and structure determination of Psu T123C mutant protein

The T123C change was made by site-directed mutagenesis on the zero-cys Psu and was purified in a similar way as WT. For crystallization, the protein was dialysed against buffer containing 50mm Tris (pH 8.0), 300mM NaCl. The crystallization trials of Psu was carried out at 277 and 293 K. Typically, 3 ml protein solution (5 mg/ml) was mixed with 2
ml reservoir solution containing 7.5\%(w/v) PEG 6000, 5\%(v/v) glycerol in 0.1 M HEPES pH 7.0 at 277 K. Cube shaped crystals appeared within a couple of days. Data was collected up to 3 Å in conditions similar to the WT Psu. For phasing, the coordinates of WT Psu was used for molecular replacement with Phaser (70) in CCP4 (71). Model building was done with Coot (66) and refinement was carried out with Phenix refine (65). The data processing and refinement statistics are summarized in Table 3 and Table 4 respectively.

7.3.5 Circular dichroism measurement

Circular dichroism (CD) is being increasingly recognized as a valuable technique for examining the structure of proteins in solution. CD signals only arise where absorption of radiation occurs. In proteins, the chromophores of interest include the peptide bond (absorption below 240 nm), aromatic amino acid side chains (absorption in the range 260 to 320 nm) and disulphide bonds (weak broad absorption bands centered around 260 nm). Absorption in this region (240 nm and below) is due principally to the peptide bond; there is a weak but broad n-\(\pi^*\) transition centered around 220 nm and a more intense \(\pi-\pi^*\) transition around 190 nm.

Far-UV CD spectra were acquired for protein concentrations of 1mg/ml (0.047mM) in 10 mM sodium phosphate (pH 7.4) at 25°C using a BioLogic Science Instruments (France) spectropolarimeter and cuvettes of 1 mm path length. The percentage of helicity was estimated following standard protocol.

\[
\text{Percentage of helicity} = 100 \times \frac{c_{\text{hel}}[\theta]_{222}}{\max[\theta]_{222}}
\]

(1)

Where, \(\max[\theta]_{222}\) stands for 100\% helicity and was calculated using the formula
Where, \( n \) was the total number of amino acid residues in the protein and 
\([\theta]_{222}\) was the mean residue molar ellipticity in degrees cm\(^2\) dmol\(^{-1}\).

\([	ext{eq] max[\theta]_{222} = -40,000 \times \left[1 - \left(\frac{2.5}{n}\right)\right]}

\) was obtained with the help of the following equation

\[
[\text{eq} \] \text{max[\theta]_{222} = (100M_r\theta)/(c\ell N_A)}
\]

\( \theta \) was the experimental ellipticities in degrees, \( M_r \) was the molecular mass of the protein in Da, \( c \) was the concentration of the protein in milligram per milliliter, \( \ell \) was the pathlength in centimeters, and \( N_A \) was the number of residues in protein.

For denaturation studies, the protein was incubated in varying concentrations of GdmCl (1, 1.5, 2, 2.5 and 3M) at 25\(^\circ\)C. The protein was initially standardized for the minimum GdmCl concentration at which denaturation reaches saturation. Initial denaturation with 6M, 5M, 4M and 3M gave the same CD profiles. So, maximum of 3M GdmCl was used for performing the denaturation experiments to minimize the noise in the 200-210nm range. The far-UV CD spectrum of each sample was acquired after 20 minutes of incubation. Each spectrum was recorded as an average of 8 scans. In all experiments, contributions of the buffer to the spectra were subtracted, and mean residue ellipticities were determined before plotting the spectra. For renaturation studies of Psu WT and Psu T123C, the proteins at a concentration of 2mg/ml (0.095mM) and 1.7mg/ml (0.08) respectively were first denatured in a reaction mixture containing 3M GdmCl and then subsequently dialysed against buffer without GdmCl.
7.3.6 Fluorescence Spectroscopy

Steady state fluorescence studies were performed using a Varian Cary Eclipse fluorescence spectrophotometer for protein concentrations of 1 mg/ml (0.047) in 10 mM sodium phosphate (pH 7.4) at 25°C and cuvettes of 4 mm pathlength. For denaturation studies, the protein was incubated in varying concentrations of GdmCl (1, 1.5, 2, 2.5 and 3 M) at 25°C for 20 minutes each. Fluorescence emission from Psu tryptophans were measured using an excitation at 295 nm, and the emitted intensity was recorded at 340 nm. In all experiments, contributions of the buffer to the spectra were subtracted before plotting the spectra.

7.3.7 Cross Linking and SDS-PAGE

For the cross-linking experiments, T123C-ΔCys-Psu was incubated with the Cu-P (72) before crystallization to induce disulphide bond between the two cysteines. The disulphide bridge can be formed in the presence of Cu-P only if they are quite close to each other (~6 Å) in the tertiary/quaternary structure and are correctly oriented. The concentrations of the cross-linker and T123C-ΔCys-Psu were 5 mM and 80 μM respectively. T123C-ΔCys-Psu and Cu-P were mixed in phosphate buffer (10 mM NaH₂PO₄, pH 7.4) devoid of any reducing agents and incubation was continued for 30 min at 25°C. Non-reducing SDS-sample buffer was added to the cross-linking reactions, which were subjected to SDS-PAGE (15% polyacrylamide gel).
7.3.8 Structural analysis

B factors were calculated using Baverage in CCP4 (71). The oligomeric state of the protein was analysed using the PISA webserver (73). Figures were prepared using Pymol (74).

7.3.9 Docking of Psu dimer on the Rho hexamer

For the Rho-Psu docking, the crystal structures of the Psu dimer (PDB id. 3RX6) and the Rho hexamer (PDB id. 3ICE) were first modified by removing the ligands and water. The docking of the Psu dimer on the Rho hexamer was performed using ‘O’ (67) based on the constraints from previous experimental observations, results of the biochemical experiments in this study and maintaining a symmetric interaction at both the binding sites. The docked model so obtained was refined using the docking refinement module in Nomad-Ref server (75) with the default parameters. The binding free energy of the refined docked model was calculated using PISA (73).

7.3.10 Transmission Electron Microscopy (TEM)

Non-his tagged P67L Rho was purified in a buffer solution containing 100 mM NaH$_2$PO$_4$, 100mM NaCl pH-7.8 and 1mM ATP to induce hexamerization of Rho. N-terminal HMK his tag WT Psu was added at a volume ratio of 1:3 and incubated at room temperature (25°C) to facilitate the complex formation. The complex so formed were applied to a carbon coated TEM grid and stained with 0.5% uranyl acetate. After negative staining, the grids
were dried slowly before observation under transmission electron microscopy (FEI, Tecnai S-twin).