3 Methods

3.1 Molecular biology methods

3.1.1 Polymerase chain reaction (PCR)

The polymerase chain reaction (PCR, [85]) is a widely used method for amplifying a few copies of fragments of double stranded DNA across several orders of magnitude. To perform the reaction, the basic requirements are: firstly a DNA template including the sequence of interest; secondly a pair of short oligonucleotides that are complementary to a 20-30 bp long sequence at either the 5' or 3' end of the target fragments, and hence named as forward and reverse primers, respectively; and finally a heat-stable DNA polymerase and mixture of four deoxyribonucleotide triphosphates (dNTPs) for DNA strand synthesis. The reaction mix is subjected to repetitive cycles of DNA double strand melting, primer annealing and strand synthesis with the DNA polymerase, and produces the template fragment flanked by the primer sequences in large quantities.

For the purpose of restriction endonuclease-based cloning, gene-specific PCR primers were designed to add recognition sites for restriction endonucleases to the 5' and 3' end of the amplified DNA sequence. To allow directional cloning, recognition sites for different restriction endonucleases were added with either primer. In order to prevent errors in the amplified DNA sequence, the *Pfu* polymerase that possesses proofreading activity was used (protocol listed in Table 3.1.1 & 3.1.2). Full-length and truncated forms of human NIBP and Ehoc-1 gene were amplified using cDNA clones ordered from RZPD. For yeast genes, extracted wild-type yeast genomic DNA was used as the template (provided by Franziska Zimmermann, MDC).

Another important PCR experiment is colony PCR to confirm the correct colonies after transformation into *E. coli* (a general protocol is shown in Table 3.1.3 & 3.1.4.). *Taq* polymerase is used to ensure higher efficiency at lower costs. Gene-specific or vector-specific primers can be used in colony PCR, however the annealing temperature in the PCR program must be adjusted according to the primers.

Compound	Concentration
Template DNA	<10 ng
Forward primer	0.3 µM
Reverse primer	0.3 µM
dNTPs	600 µM
Pfu reaction buffer	1x
DMSO	1 µl
Pfu polymerase	2.5 U
H_2O	up to 50 µl

Table 3.1.1 Preparative PCR setup

Steps		Temperature	Time
1. Initialisation		95 °C	2 min
2. Thermal cycles	DNA melting	95 °C	30 s
(x 30)	Primer annealing	55 °C	30 s
	Strand extension	68 °C	~1 min for 1 kb
3. Final extension		68 °C	7 min
4. Hold		10 °C	For ever

Table 3.1.2 Preparative PCR program.

Compound	Concentration
Single colony	x1
Forward primer	0.3 µM
Reverse primer	0.3 µM
dNTPs	200 µM
Taq reaction buffer	1x
Taq polymerase	3 U / 100 µl
H_2O	up to 20 µl

Table 3.1.3 Colony PCR setup

Steps		Temperature	Time
1. Initialisation		95 °C	5 min
2. Thermal cycles	DNA melting	95 °C	30 s
(x 25)	Primer annealing	55 °C	30 s
	Strand extension	68 °C	~1 min for 1 kb
3. Final extension		68 °C	7 min
4. Hold		10 °C	For ever

Table 3.1.4 Colony PCR program.

3.1.2 DNA gel electrophoresis

DNA gel electrophoresis was used to determine the size of PCR products after either preparative or colony PCR. To prepare a 1% agarose gel, 0.5 g agarose was mixed with 50 ml TAE buffer and heated for about 1 minute in a microwave until the agarose had completely dissolved. The dissolved gel was filled into a gel chamber and supplemented with 0.5 µg/ml ethidium bromide (EB). Samples were mixed with DNA sample buffer and loaded on the gel, as well as a DNA molecular weight marker (for instance the 100 bp DNA ladder from NEB). Electrophoresis was performed at constant current of 220 mA in TAE buffer, and the gel was analyzed under UV light. The size of each band on gel can be estimated by comparing its position with the DNA marker bands.

3.1.3 DNA purification

Plasmid preparation: Colonies picked from agar plates were grown overnight in 5 ml LB medium with appropriate antibiotics. Then the plasmids were extracted using the QIAprep Spin Miniprep Kit. The concentration and quality of mini-preparations can be determined with the Nanodrop photometer.

PCR purification: PCR products and endonuclease-treated fragments were purified using the QIAquick PCR Purification Kit to remove the remaining primers, nucleotides and enzymes.

Gel extraction: DNA samples, for instance endonuclease-treated empty vectors, can also be extracted from agarose gel using the QIAquick Gel Extraction Kit to remove the agarose and EB.

3.1.4 Restriction endonuclease-based cloning

Firstly, the PCR products and empty target vectors were digested by double digestion reaction of the restriction endonucleases *Bam*HI and *Not*I (in the presence of appropriate reaction buffer, according to manufacturer's protocols) to obtain suitable sticky ends for annealing. After purification, insert and vector were connected with T4 DNA ligase (performed according to manufacturer's protocols), and transformed into DH5α *E. coli* cells and grown overnight at 37 °C on LB agar plates with appropriate antibiotics. The positive colonies were detected from the overnight plates by looking

for visible bands with the correct size after colony PCR. Finally, the correct recombinant plasmids were extracted by mini-preparation, for subsequent transformation into suitable *E. coli* cells (for instance cells of BL21(DE3) strain) and protein synthesis. For DNA sequencing, the purified plasmids were sent to Eurofins MWG Operon.

3.1.5 Transformation of chemically competent *E. coli* cells

Transformation of recombinant vectors into *E. coli* cells was performed with chemically competent cells according to the protocol by Hanahan [86]. An overnight culture of *E. coli* BL21(DE3) or *E. coli* DH5α was diluted 1:100 in 100 ml fresh LB medium and grown at 37 °C to OD_{600nm} = 0.6. Cells were cooled down on ice and harvested at 4 °C by centrifugation for 5 min at 2,500 x *g*. The pellet was resuspended in 33 ml pre-chilled TFB 1 buffer, incubated for 30-45 min on ice and centrifuged for 5 min at 2,500 x *g*. The pelleted cells were then resuspended in 8 ml pre-chilled TFB 2 buffer, and 100 µl aliquots were flash-frozen in liquid nitrogen and stored at -70 °C.

For heat shock transformation, aliquots of competent cells were first thawed on ice. Then the ligation product or ~10 ng purified plasmid was added to each aliquot. Reactions were incubated on ice for 5 min, followed by a heat shock at 42 °C for 1 minute. The tubes were returned to ice, and let still for 10 minutes. 500 µl LB medium was added to each tube, and cells were shaken at 37 °C for 1 h. A certain amount of the cultures was streaked out on LB agar plates containing appropriate antibiotics. After overnight incubation at 37 °C, single colonies grown on the plates could be picked and used in protein expression tests.

3.2 Protein purification and characterization

3.2.1 Test expression

The recombinant plasmid carrying the gene of interest was transformed into *E. coli* cells (strains BL21(DE3) or BL21(DE3)-T1R), and plated on LB agar containing 100 μg/ml ampicillin. 2-3 single colonies were picked from each overnight plate. Each colony was inoculated in 2 ml OvernightExpress instant TB medium containing ampicillin, and shaken overnight at 37 °C, 200 rpm.

The next morning, 1 ml cell culture was collected from each tube. Cells were pelleted by centrifuge for 3 min at 10,000 rpm, then disrupted by adding 40 μl cell disruption buffer (see 2.4) and standing 2 h at room temperature. The soluble part was then separated from the whole-cell extract by centrifugation for 5 min at the top speed of a table centrifuge. Both the whole-cell extract and the soluble part were loaded on SDS-PAGE. The presence of target protein was detected by Coomassie staining, silver staining or Western blotting.

The colonies in which the target proteins were correctly expressed (preferably also partly soluble) were stored as glycerol stocks (overnight culture with addition of 15% sterile glycerol), which were stored at -80 °C and would be used in up-scaled expression.

3.2.2 Expression in *E. coli*

For large-scale expression, an overnight preculture was prepared with LB medium containing ampicillin, starting from a glycerol stock. The preculture was diluted 100-500 fold in LB or TB medium and shaken at 37 °C, 100 rpm until the OD_{600nm} rose to ~0.7. For His-tagged protein, 1 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG) was added to the culture, and the cells were harvested after growing at 37 °C for 4 h. For GST-tagged protein, the culture was first cooled down to 20 °C, then 0.1 mM IPTG was added to the culture, and the cells were harvested after growing at 20 °C overnight.

The cells were harvested by centrifugation at 5,000 x *g* for 10 min. The pellets were washed once with PBS buffer, then directly disrupted or stored at -80 °C for later use.

3.2.3 Production of seleno-methionine labeled protein

For production of seleno-methionine (SeMet) – labeled Tca17, a SeMet incorporation by metabolic inhibition was performed, according to a protocol by VanDuyne *et al*. [87]. The glycerol stock of the pGex6p1-Tca17 plasmid in *E. coli* BL21(DE3)-T1R was inoculated in LB medium with ampicillin and grown overnight at 37 °C. The preculture was diluted 1:100 into M9 medium supplemented with ampicillin, thiamine and $FeSO_4$ (see 2.4), and grown at 37 °C until OD_{600nm} reached ~0.5 (since cell growth is much slower than in rich medium, good aeration is necessary to shorten the waiting time.). An appropriate amount of solid amino acid supplements was added to the cell culture to inhibit amino acid synthesis, and the culture was cooled down to 20 °C. After 15 min, gene expression was induced by 1 mM IPTG, and the cells were grown overnight at 20 °C.

Purification of SeMet-labeled Tca17 was similar as native Tca17, only the buffers were: 1) more carefully degassed, and 2) supplemented with 1 mM EDTA and 5 mM DTT to prevent oxidation of SeMet.

3.2.4 Cell lysis

During lysis cells were constantly cooled on ice to avoid heating of the samples. The frozen cell pellets were thawed on ice and resuspended thoroughly in PBS buffer. Smaller volumes (<30 ml) of suspended cells were disrupted by passing twice through a French Press at 7.6 MPa pressure. Cells grown on a larger scale were disrupted by passing twice through a Fluidizer at 9 MPa pressure. Benzonase nuclease was added to the cell lysate (1 µl Benzonase per 10 ml cell lysate) to digest DNA and RNA and thus reduce the viscosity of the lysate. A cleared lysate was obtained by centrifugation at 40,000 x *g*, 4 °C for 30 – 60 min. Small aliquots of both the raw lysate and supernatent after centrifugation were saved for SDS-PAGE later, marked as whole-cell extract and soluble part, respectively.

3.2.5 Affinity chromatography

The initial step to purify the recombinant fusion protein from cell lysate was an affinity chromatography. Either a batch purification with selective affinity matrix, or a fast protein liquid chromatography (FPLC) purification with affinity chromatography columns could be used at this step.

3.2.5.1 Purification with Ni-nitrilotriacetate (Ni-NTA) agarose:

Poly-His tagged small subunits of TRAPP were purified with Ni-NTA agarose. NTA has a tetradentate chelating group that occupies four of the six sites in the nickel coordination sphere, and binds the ions tightly to the resin. The remaining ligand binding sites of nickel ions can then be coordinated with histidine residues and thus bind His-tags specifically and efficiently. Elution is achieved by the addition of the histidine analog imidazole that, if present in excess, will occupy the binding sites for histidine.

Cell pellets from 2 l culture were resuspended in 30 ml His-lysis buffer, and disrupted by passage through the French Press or Fluidizer. After centrifuged for 30-50 min at 40,000 x *g*, 4 °C, the clear lysate was applied to 2 ml Ni-NTA agarose, and incubated at 4 °C overnight on a rotator. The matrix was transferred into a polypropylene column. The self-packed column was then washed extensively with His-wash buffer (3 x 20 ml) and eluted with His-elution buffer (3 x 2 ml). The eluted protein was dialysed against 1 l His-dialysis buffer at 4 °C overnight, and stored at -20 °C for later use.

3.2.5.2 Purification with FPLC system and GST affinity column

Glutathione S-transferase (GST) is a commonly used affinity tag in recombinant protein production. GST fusion proteins can be purified directly from bacterial lysates using Glutathione (GSH) sepharose 4B, and eluted under mild conditions with 10-20 mM reduced GSH.

A GST-tagged target protein (an Ehoc-1 fragment or full-length Tca17) was purified with an FPLC system, using a 5 ml GSTrap FF column or a 15 ml self-packed GSH sepharose 4B column. Cell pellets from 8 l culture were resuspended in 60 ml GST-lysis buffer, and disrupted by passage through the Fluidizer. After centrifuged for 30-50 min at 40,000 x *g*, 4 °C, the clear lysate was loaded onto the column (equilibrated

with PBS buffer) at a flow rate of 0.3 ml/min. Then the column was washed with PBS buffer at a flow rate of 1.5 ml/min until UV absorption returned to the base line. Purified GST-fusion protein was eluted by GST-elution buffer at a flow rate of 1 ml/min, and collected as 1 ml fractions. After assessing protein purity by SDS-PAGE and Coomassie staining, purified GST-Tca17 was dialysed against 1 l GST-dialysis buffer whereas the GST-tagged Ehoc-1 fragment was dialysed against 1 l Mono S buffer A at 4 °C overnight.

3.2.6 Determination of protein concentration

To determine the concentration of protein samples, 2 µl of the sample was loaded on the NanoDrop photometer. Choosing the program for measuring protein, and using the protein buffer as blank, the absorbance of the protein at 280 nm was recorded, and the protein concentration was calculated according to Lambert-Beer's law (Equation 3.1). The molar extinction coefficients ε_{280} for several target proteins were calculated with ProtParam [88] and listed in Table 3.2.1.

$$C = {}^{A_\lambda}/_{\varepsilon_\lambda\, l} \qquad \text{(Equation 3.1)}$$

A_λ : Absorbance at wavelength λ
ε_λ : Molar extinction coefficient (M^{-1} cm^{-1}) at wavelength λ
C : Molar concentration (M)
l : Optical path length (cm)

Protein	**MW (kDa)**	**ε_{280} (M^{-1} cm^{-1})**	**A_{280} at 1 mg/ml**
Tca17	17.4	11,460	0.660
GST-Tca17	44.2	54,320	1.230
E211	22.9	33,460	1.459
GST-E211	49.8	76,320	1.534
GST-Trs20p	46.5	61,770	1.328
GST	26.8	42,860	1.597
His-Bet3p	25.3	32,430	1.284
His-Trs33p	33.7	18,910	0.561

Table 3.2.1 Extinction coefficients of several proteins used in this work.

3.2.7 Cation exchange chromatography

In ion-exchange chromatography, proteins are separated on the basis of their surface charge. Cation-exchange matrices are derivatized with negatively charged groups for the absorption of cationic proteins.

After the first affinity chromatography, GST-tagged Ehoc-1 fragment was eluted together with a lot of GST protein, and needed to be further purified. Since the isoelectric point values (pIs) of GST and the GST-tagged Ehoc-1 fragment are quite different (5.73 and 7.09, respectively), by using a running buffer of a middle pH value, it is possible to make GST anionic and the GST-fusion protein cationic. Therefore, when passing through a cation exchange column, the target protein will be bound, and later eluted with increasing NaCl concentration, while GST passes through.

A Mono S cation exchange column was first equilibrated with 1 ml 1 M NaCl and washed with 10 ml Mono S buffer A. After loaded with the dialysed protein solution, the Mono S column was washed with Mono S buffer A until the UV absorption returned to the base line. Finally, the target GST-fusion protein was eluted with a gradient of Mono S buffer B from 0% to 60%. The eluted protein was dialysed against 1 l GST-dialysis buffer at 4 °C.

3.2.8 Tag removal

To cleave the affinity tag from the recombinant protein, specific protease cutting sites were introduced in the expression constructs, directly following the affinity tag sequence. If the protein was expressed from the pQLink-H vector, the His-tag was removed with TEV (tobacco etch virus) protease. From proteins expressed from the pGex-6p-1 vector, the GST-tag was removed with PreScission protease.

To remove the GST tag, the fusion protein was mixed with PreScission protease at a ratio of 30:1 (fusion protein:protease, w:w), and dialysed against 1 l GST-dialysis buffer at 4 °C overnight. Since the PreScission protease was also fused with a GST-tag, the protein mixture could be further purified by a second affinity chromatography. The untagged target protein would stay in the flow-through, while the GST-tag, uncut fusion and PreScission protease would remain bound to the affinity column.

A 5 ml GSTrap column or a 15 ml GSH sepharose 4B column was equilibrated with GST dialysis buffer, then the protease-treated protein mixture was loaded at a flow rate of 0.3 ml/min. After washing the column with GST dialysis buffer at a flow rate of

1.5 ml/min until UV absorption went back to base line, the column was cleaned with GST elution buffer at a flow rate of 1.0 ml/min. The flow-through fractions were collected and further purified.

3.2.9 Gel filtration chromatography

Gel-filtration, also known as size-exclusion chromatography, can separate proteins on the basis of differences in size as they pass through a gel filtration medium packed in a column. The medium is composed of inert polymer beads with a defined pore diameter. Before sample loading, the column is first equilibrated with buffer, which fills the pores of the matrix and the space in between the particles. When applied to such a column, proteins will diffuse in and out of the pores of the matrix. Smaller molecules move further into the matrix and so stay longer on the column. As a result, the proteins from a mixture can be separated according to their size and shape, each protein eluting at a distinct volume.

For better separation, the target protein after second affinity chromatography was concentrated to 1-2 ml (for 16/60 column) or 3-5 ml (for 26/60 column), and loaded on to a Superdex 75 Hiload 16/60 or 26/60 column equilibrated with the appropriate protein buffer (see 2.4). The elution profile was monitored by measuring the absorption at 280 nm, and elution peaks were collected at 1 ml fractions and analyzed with SDS-PAGE. SeMet-labeled Tca17 was purified according to the same protocol, but SeMet Tca17 protein buffer was used instead.

3.2.10 Protein SDS-PAGE (sodium dodecylsulfate polyacrylamide gel electrophoresis)

Discontinuous SDS-PAGE was used to determine the composition of protein samples. The polyacrylamide gels were produced by co-polymerization of acrylamide and bisacrylamide in the presence of the radical starter ammonium peroxodisulfate (APS) and the catalyst tetramethylethylenediamine (TEMED). By varying the acrylamide concentration, the pore size of the gel was changed, which made it suitable for different protein molecular weight ranges. A 15% gel is most suitable for linear separation of proteins between 12 to 43 kDa.

SDS is an anionic detergent that denatures proteins and binds to proteins at a constant ratio. In the presence of SDS, proteins migrate in an electric field at a

velocity that is approximately proportional to the logarithm of their molecular mass. Comparing the position of protein bands to a marker lane with proteins of known size allows an estimation of the size of the protein samples.

Besides SDS, protein samples are normally heated in the presence of a reducing agent, such as DTT or βME, which further denatures the proteins by breaking disulfide bonds, therefore breaks up some higher order structures. This is known as reducing SDS-PAGE.

If protein samples need to be analyzed in the presence of disulfide bonds, non-reducing SDS-PAGE may be used instead, for which the protein samples are prepared with addition of SDS, but without boiling or reducing agent. The poly-acrylamide gels and electrophoresis buffer are prepared with the same protocol as reducing SDS-PAGE.

For reducing SDS-PAGE, the protein samples were mixed with 2 x SDS sample buffer at 1:1 ratio and boiled for 5 min at 95 °C. For non-reducing SDS-PAGE, the protein samples were mixed with 2 x non-reducing SDS sample buffer at 1:1 ratio, and then directly loaded without boiling. After loading the samples, electrophoresis was carried out in a chamber filled with electrophoresis buffer, at 120 V for 15 min (through the stacking gel) and then 60-90 min at 180 V (through the resolving gel).

3.2.11 Staining of SDS-PAGE

For Coomassie staining, the gel was first incubated for several minutes in Coomassie staining solution. After washing briefly with water to remove the residual staining solution, the gel was incubated in destaining solution, and shaken at room temperature until it was completely destained.

Silver staining was done according to the 30 min stain protocol by Nesterenko *et al*. [89]. The gel was first fixed for 5 min in fixation solution, rinsed twice, washed 5 min and rinsed again three times with water. After pretreatment for 5 min in 50% acetone and 1 min in reducing solution, the gel was rinsed with water and impregnated with silver stain solution for 8 min. After another rinsing with water, the gel was developed with development solution until bands started to appear (usually after 10 to 30 s) and stopped immediately with 1% acetic acid solution.

3.2.12 Concentration of protein

Purified target protein was concentrated with ultrafiltration devices of suitable NMWL (nominal molecular weight limit). While water and small molecules (molecular weight lower than the NMWL) can pass through the membrane during centrifugation, the protein is retained and thus concentrated. The protein samples were loaded onto the ultra-filtration membrane and centrifuged at 4 °C, 3,500 x *g*, until the desired concentration was reached. To prevent local overconcentration, which might lead to protein aggregation, the protein samples were remixed by gentle pipetting several times during the concentration process.

3.2.13 Circular dichroism (CD) spectrum

Circular dichroism (CD) refers to the phenomenon that left and right circularly polarized light is absorbed differently by optically active chiral molecules, for instance proteins. Besides the intrinsic optical activity of the individual amino acids, proteins give further CD-signals that arise from the coupling of chromophores in secondary structure elements such as α-helix and β-sheet, as well as random coils [90]. Each of these elements gives rise to a characteristic spectrum (see Fig 4.2.1 b), so that the CD spectrum of a protein allows to estimate the relative presence of the different secondary structure elements.

For CD spectrum recording, purified Ehoc-1 fragment at a concentration of 0.8 mg/ml was dialysed against low-salt buffer (see 2.4). The spectrum was collected at 20 °C from 260 nm to 190 nm with a 0.1 cm pathlength cuvette.

3.2.14 Protein buffer screen

To enable crystallography experiments, the protein samples must be stable over long periods of time. In order to find a suitable protein buffer to stabilize Tca17, a buffer screen was performed with the fluorescence-based thermal shift assay (TSA) method, also referred as differential scanning fluorimetry (DSF) [91].

TSA is based on temperature-induced protein denaturation, monitored using the environment-sensitive fluorescent dye Sypro Orange. In the presence of a native protein, the dye is naturally quenched. When the protein starts to denature in response to increasing temperature, thus exposing its hydrophobic core, the dye will

react and start to fluoresce. With the fluorescence data collected by the real-time PCR detection system, a melting curve is generated. The peak value of its first derivative curve is recorded as the melting temperature T_m, which is the temperature midpoint for the protein unfolding transition.

In the buffer screen for purified Tca17, a 96-well PCR plate was used. In each well, 20 µg Tca17 protein was dissolved in 30 µl screen buffer, and then mixed with 20 µl 2.5 x Sypro Orange water solution. The sample was heated from 20 °C to 100 °C with a rate of 1 °C per minute. Changes in fluorescence intensity were monitored, and the wavelengths for excitation and emission were 490 and 530 nm, respectively. The melting temperature in each buffer condition was compared with a reference buffer, and finally a stabilizing buffer condition was designed for Tca17, based on the conditions where T_m was significantly higher than in the reference buffer (T_m - $T_{mr} \geq 5$ °C).

3.3 Protein crystallography

X-ray diffraction analysis of crystals is the most powerful methods to determine the three-dimensional structure of biological macromolecules, most importantly nucleic acids and proteins. In comparison to other methods of structural biology, for instance nuclear magnetic resonance (NMR) spectroscopy or electron microscopy, protein crystallography has the advantage of its ability to obtain atomic resolution structures of proteins of various types and sizes, as long as crystals can be obtained.

Because X-rays have a wavelength (around 1 Å) close to the distance separating covalently bound atoms, when a beam of X-rays strikes a crystal, scattering from atoms in the unit cell will result in a distinct diffraction pattern. From the collected diffraction pattern and the related phase information, a three-dimensional image of the electron density can be derived. If the quality of this electron density map is high enough, the mean positions of the atoms in the crystal, and therefore the three-dimensional structure of the whole macromolecule (or macromolecular complex) can be determined accordingly.

Most figures and equations used in Chapter 3.3 are adapted from the crystallography textbook by Drenth [92].

3.3.1 Crystallization

Growing the protein crystals required for X-ray structure determination is still a bottleneck step in structural biology. Crystallization of macromolecules may be separated into three main steps: nucleation, growth and cessation of growth. Firstly the protein solution is brought to a super-saturation state, where a few crystallization nuclei will form. Such a “nucleus” can be a dust particle, a microcrystal as seed, or a small protein aggregate. Nuclei may grow into crystals by recruiting protein molecules from the solution. Protein molecules associate to crystal lattice in an orientation dictated by crystal symmetry, until a point when further crystal growth is terminated either by a subsaturation free protein concentration, or by lattice defects accumulated at the crystal surface (Fig 3.3.1).

Theoreticaly, the best crystals should grow in a low level of supersaturation. Without supersaturation, nuclei will not form, and the protein solution will stay clear. On the other side, maintaining a high supersaturation would lead to formation of too many

nuclei, and therefore result in amorphous precipitate or many microcrystals. Also, crystals should grow slowly to reach a maximal degree of order in their structure.

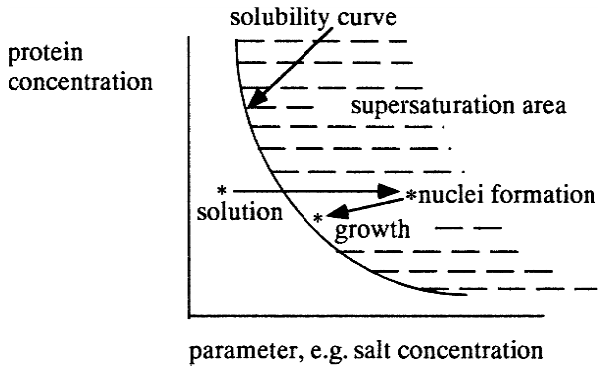

Fig 3.3.1 A typical solubility curve for a protein, as a function of salt concentration or other parameters, taken from [92].

The most widely used macromolecular crystallization technique is the vapor-diffusion method. A small drop containing protein, buffer and precipitant is equilibrated against a reservoir solution containing similar buffers and a higher concentration of precipitant, in a sealed chamber under constant temperature. Upon vapor diffusion, the precipitant and protein concentration in the protein droplet is usally increased, until protein solution reaches supersaturation, and crystal growth is initiated. Two types of vapor-diffusion methods were applied in this work: the hanging drop method and sitting drop method (Fig 3.3.2).

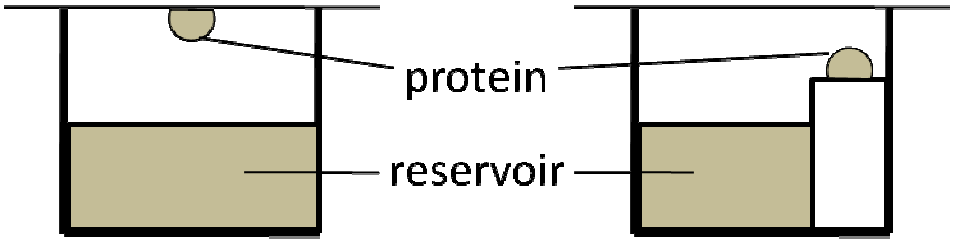

Fig 3.3.2 Diagrams of hanging drop method (left) and sitting drop method (right).

Besides the requirement for a highly stable and homogenous protein sample (>95% purity, single oligomerization state), the success of crystallization relies also on circumstances of the crystallization experiment. Crystallization depends on numerous parameters that can be optimised, such as protein concentration, ionic strength, pH value, temperature, and probably most important, precipitants, which define a multi-dimensional parameter space. Exhaustive searching of this parameter space in crystallization experiments varying all possible parameters in all possible combina-

tions is usually impossible. The best way to screen the multi-dimensional crystallization space is with sparse matrix screens. These screens try to cover the whole range of crystallization conditions and are formulated on the basis of precipitant solutions that have been successfully used in crystallization trials before. A variety of such screens, designed to meet different requirements, are available commercially. Although more difficult to implement than buffer or precipitant screens, a few different protein start concentrations and crystallization temperatures should also be tested. Once an initial condition is found, where microcrystals or spheroids grow, further optimisation is performed with systematic fine grid screens around the initial condition, until the best crystals possible can be grown.

Among all types of precipitants, polyethylene glycols (PEGs) of different length are found to be the most effective. Like salts, PEGs compete with protein for water and exert excluded volume effects. But unlike salts, PEGs can also decrease the effective dielectric constant of the solution, which increases the effective distance over which protein electrostatic effects occur.

3.3.2 Crystal mounting and cryo-protection

After inspected by microscope, the crystal selected for a diffraction experiment is captured by a fiber loop of suitable size, supported by surface tension of a drop of cryo-protected liquid. Then the whole loop is flash frozen in liquid nitrogen, and mounted for data collection at 100 K. Comparing with data collection at room temperature, this technique greatly reduces radiation damage caused by the passage of X-ray beam. This is especially necessary when a high-intensity synchrotron beam is used as the X-ray source.

For a successful measurement, the protein crystal must be soaked in a suitable cryo-protection solution. When frozen, the cryo-protectant solution should form vitrified water (amorphous glass state) rather than crystalline ice, which would also diffract X-rays, interfere with the diffraction pattern of the protein crystal which it might physically destroy. On the other hand, soaking in this solution should not damage the crystal. A suitable cryo-protection solution can normally be prepared by adding 10-30% (v/v or w/v) cryo-protectant, for instance glycerol, small PEGs or sucrose, to the crystallization solution from which the crystal grew. However, when a higher concentration of salt is present in the initial protein buffer, the protein drop will not decrease much in size during crystallization, and the effective precipitant concentra-

tion in protein drop is lower than reservoir solution. In this case, a dilution of the reservoir solution will be necessary to prepare the cryo-protection solution.

3.3.3 Diffraction data collection

The internal structure of a crystal can be understood as a three-dimensional lattice made up of identical small building blocks, called asymmetric units. The asymmetric units are arranged into unit cells by crystallographic symmetry operations. A unit cell defines the crystal coordinate system by three edges (*a*, *b*, *c*) and three angles (α, β, γ). Depending on internal symmetry, crystals can be grouped into different space groups.

According to W. L. Bragg, diffraction of X-rays by a crystal lattice can be described as reflection on discrete sets of lattice planes. Each set of these equivalent, parallel planes can be describes with Miller indices *h*, *k* and *l*, which define how often a set of parallel lattice planes intersects with the unit cell edges *a*, *b*, *c*. Diffracted beams from equivalent lattice planes will result in constructive interference and, hence, an observable reflection only if Bragg's law is fulfilled:

$$2d_{hkl}sin\theta = n\lambda \qquad \text{(Equation 3.2)}$$

d_{hkl} : distance between lattice planes
θ : Bragg angle
λ : wavelength of the X-rays

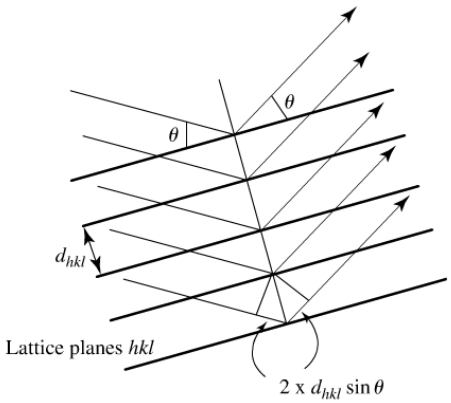

Fig 3.3.3 Diffraction from lattice planes, taken from [93]

Equation 3.2 can be explained by Fig 3.3.3. X-rays diffracted on parallel planes will travel along different lengths before adding up to each other. In order to obtain constructive interference after diffraction, the path difference $2d_{hkl}sin\theta$ should equal an integer multiple of the wavelength λ. From the equation, it can also be derived

that, theoretically, the minimal distance between lattice planes would be *λ/2*, which defines the maximal resolution of the diffraction experiment at a given wavelength.

The diffraction pattern of the crystal lattice is also a lattice, known as the reciprocal lattice. Similar to the crystal lattice, the reciprocal lattice can be represented by lattice parameters *a**, *b**, *c**, α*, β*, and γ*. The direction of *a** is perpendicular to the plane defined by direct-space axes *b* and *c*, and its magnitude is reciprocal to the magnitude of *a*. Each point (*h k l*) in the reciprocal lattice corresponds to a set of parallel lattice planes (*h k l*) in the real space lattice.

Another important diffraction pattern is the diffraction from a single unit cell. X-rays scattered from each atom in the unit cell will infere with each other, depending on the diffraction direction and atomic positions. This will result in a complicated diffraction pattern where no separated reflection points can be detected, and the reflection amplitude and phase vary continuously.

Since a crystal is a lattice built up by single unit cells, the X-ray diffraction pattern of a crystal is the convolution of the scattering from the atoms in one unit cell and the lattice. In other words, the observed diffraction pattern is the scattering pattern of a single unit cell observed only at reciprocal lattice points.

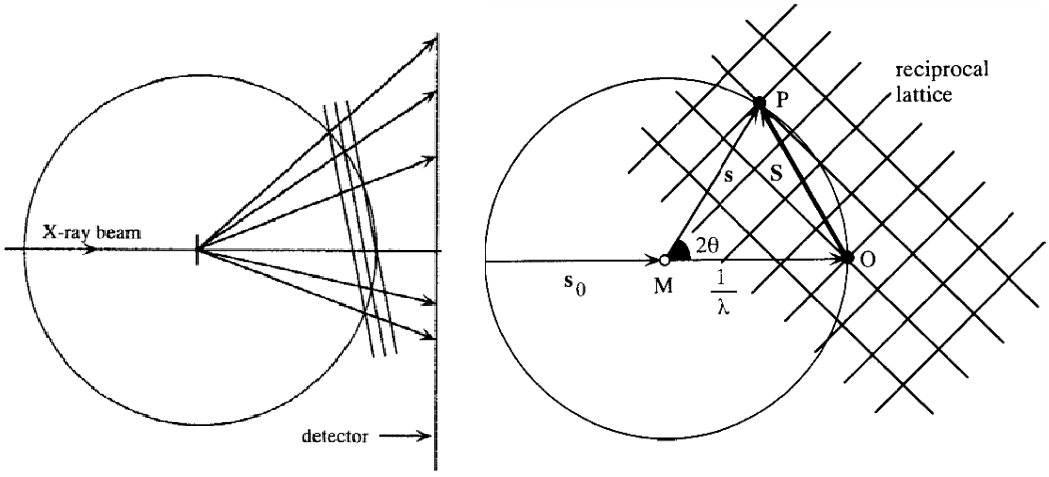

Fig 3.3.4 Ewald sphere as a tool to understand diffraction from a crystal, taken from [92].
s_0: incident X-ray beam of wavelength λ; s: diffracted beam; S: scattering vector;
M: center of the Ewald sphere and crystal position; θ: scattering (Bragg) angle;
O: origin of reciprocal lattice; P: reciprocal lattice point located on the Ewald sphere, giving rise to an observable reflection.

X-ray diffraction from a crystal can be best explained by the reciprocal lattice and the Ewald consruction. Here, Ewald sphere is an imaginary sphere in the reciprocal

space, which is centered on the crystal, has a radius of *1/λ*, and passes through the origin of the reciprocal space. An incident X-ray wave through the crystal at the centre (M) of the sphere and the origin of reciprocal space (O) will give rise to an observable reflection if, and only if, the scattered wave s leaves the Ewald sphere at a point that coincides with a reciprocal lattice point (P) (Fig 3.3.4 right). Braggs law will be fulfilled for the given wavelength λ, scattering angle θ and orientation of crystal and reciprocal lattice. Such a scattered beam will project on the detector which is set perpendicular to the incident X-ray wave, at a certain distance to the crystal, and be recorded as a diffraction spot (Fig 3.3.4 left).

As the crystal is rotated during a diffraction experiment, the reciprocal lattice is also rotated, bringing different reciprocal lattice points onto the surface of the Ewald sphere, where the associated reflections are recorded. In order to use the collected data later for structure determination, the diffraction experiment must be designed so as to collect as many unique spots as possible, with adequate experiment time and X-ray dose.

3.3.4 Data processing

Before collection of full datasets, it is normally necessary to take a few diffraction images at different angles, and process the preliminary data with indexing programs like iMosflm [94]. The first step is to determine the correct space group. This can be done by generating a list of vectors between the collected reflections, among which the three shortest non-coplanar vectors are selected. These will be used for defining a trial unit cell. Based on the trial space group, each collected reflections is indexed. A correct space group will produce integer values of *h, k,* and *l*, for all reflections. Based on the preliminary data and a correctly assigned space group, the orientation of the crystal can be determined, and an optimized data collection strategy will be suggested by the indexing program. After data collection of the full datasets, the indexing process is also performed for all reflections.

Besides indexing, each reflection also has an important value, the intensity of that reflection, which needs to be extracted from area detector (eg. CCD detector) data via ‘integration’. This is related to the fact that each reflection is three-dimensional, and might be only partially recorded in some images. In the integration process, the computer scans through all fully and partially recorded spots of the same *h, k, l*

values, add up the intensity for each reflection and discard the redundant partial reflections.

Another important part of data processing is scaling of the measured intensities for each reflection. Firstly, depending on data collection strategy, the same reflection might be recorded several times on different images. This can be used to scale the individual images. Also, reflections related by crystallographic symmetry, which should have identical intensities, can be averaged to reduce the statistical error in the measurement. A correct scaling will be able to largely compensate for the effects of crystal radiation damage, X-ray absorption and some other problems in data collection.

In this study, the collected diffraction intensities were scaled by the program XSCALE of the XDS suite [95], and then converted to CCP4 formats with the program XDSCONV [95] for further processing with CCP4 programs [96]. The quality of the data is indicted with the merging *R*-factor (R_{merge}), which is computed by summation of the absolute differences of all measured intensities from their average value, divided by the sum of all intensities (see Equation 3.3).

$$R_{merge} = \frac{\sum_{hkl} \sum_i \left| \overline{I(hkl)} - I_i(hkl) \right|}{\sum_{hkl} \sum_i I_i(hkl)} \qquad \text{(Equation 3.3)}$$

3.3.5 Calculation of the electron density

The intensity of the diffracted beam *h k l* is proportional to the square of the amplitude of the structure factor **F**(*h k l*). A reciprocal lattice point P(*h k l*) can also be presented by the Scattering vector **S**, pointing from origin O to P (see Fig. 3.3.4). Therefore, the structure factor **F**(*h k l*) can also be written as **F**(**S**), which is a function of the electron density distribution in the unit cell, and can be calculated by summing over all atoms *j* in the unit cell:

$$\mathbf{F}(\mathbf{S}) = \sum_j f_j \, \exp\left(2\pi i \mathbf{r}_j \cdot \mathbf{S}\right) \qquad \text{(Equation 3.4)}$$

S : Scattering vector in reciprocal space (see Fig 3.3.4)
f_j : Atomic scattering factor of atom *j*
$\mathbf{r}_j$: Vector defining the position of atom *j* with respect to origin of the unit cell

F(S) can also be calculated by integrating over all electrons in the unit cell:

$$\mathbf{F}(\mathbf{S}) = \int_{cell} \rho(\mathbf{r}) \exp\left(2\pi i \mathbf{r}_j \cdot \mathbf{S}\right) dv \qquad \text{(Equation 3.5)}$$

ρ(**r**) : Electron density at position **r** in the unit cell

From Equation 3.5 we can derive a relationship between the electron density ρ(*x y z*) and the structure factor **F**(*h k l*):

$$\mathbf{F}(h\ k\ l) = V \int_{x=0}^{1} \int_{y=0}^{1} \int_{z=0}^{1} \rho(x\ y\ z) \exp\{2\pi i(hx + ky + lz)\}\, dx\, dy\, dz$$

(Equation 3.6)

ρ(*x y z*) : Electron density at position *x, y, z* in the unit cell

From Equation 3.6, **F**(*h k l*) is the Fourier transform of ρ(*x y z*). Therefore, ρ(*x y z*) is also the Fourier transform of **F**(*h k l*):

$$\rho(x\ y\ z) = \frac{1}{V} \sum_h \sum_k \sum_l \mathbf{F}(h\ k\ l)\ \exp\{-2\pi i(hx + ky + lz)\}$$

$$= \frac{1}{V} \sum_h \sum_k \sum_l |\mathbf{F}(h\ k\ l)|\ \exp\{-2\pi i(hx + ky + lz) + i\alpha(h\ k\ l)\}$$

(Equation 3.7)

From Equation 3.7, two factors are required to calculate the electron density in the unit cell: Whereas the structure factor amplitudes $|\mathbf{F}(h\ k\ l)|$ can be derived from the intensities *I*(*h k l*), the phase angles *α*(*h k l*) cannot be obtained directly from the diffraction pattern, and need to determined with other methods, for instance through anomalous scattering or molecular replacement.

3.3.5.1 Multiple-wavelength anomalous diffraction (MAD) phasing

In order to solve the phasing problem, crystals grown from SeMet-labeled Tca17 protein were used to measure anomalous scattering.

When heavy atoms (for instance Selenium atoms) are present in a crystal, anomalous scattering can be observed, due to the fact that in heavy atoms, electrons can no longer be considered as completely free, and will absorb X-ray strongly. As a consequence, the intensities of a reflection *h k l* and its Friedel mate $\bar{h}\,\bar{k}\bar{l}$ are no longer equal. With carefully designed experiments, the anomalous scattering information can be used to determine the phase angle of the protein reflections.

In the presence of anomalous scattering, the atomic scattering factor is no longer a real number, as in its absence, but becomes a complex number as described in Fig 3.3.5 and Equation 3.8. Each term on the right side of Equation 3.8 is a function of wavelength *λ*. As a result, the effect of atomic anomalous scattering can vary greatly at different wavelengths.

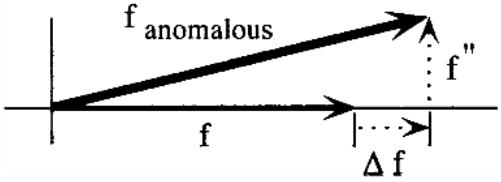

Fig 3.3.5 Argand diagram of atomic scattering factor for an anomalous scatterer, from [92].

$$f_{anom.} = f + \Delta f + if'' \qquad \text{(Equation 3.8)}$$

f : Normal atomic scattering without phase shift
Δf : Real (dispersive) correction term, often named as *f'*
f'' : Imaginary (anomalous) correction term

After collecting X-ray diffraction data from a heavy atom containing protein crystal at two or more wavelengths, the heavy atom substructure can be determined with the Patterson method. The Patterson function *P(uvw)* is a Fourier summation with intensities but no phase angles:

$$P(u\ v\ w) = \frac{1}{V}\sum_{h}\sum_{k}\sum_{l}|\mathbf{F}(h\ k\ l)|^2\ \exp\{-2\pi i(hx + ky + lz)\} \qquad \text{(Equation 3.9)}$$

P(*u v w*) : Patterson function
u,v,w : Relative coordinates in Patterson space

Equation 3.9 can be rewritten in vector form:

$$P(\mathbf{u}) = \frac{1}{V}\sum_{s}|\mathbf{F}(\mathbf{S})|^2\cos(2\pi\mathbf{u}\cdot\mathbf{S}) = \int_{r1}\rho(\mathbf{r_1})\times\rho(\mathbf{r_1}+\mathbf{u})dv \qquad \text{(Equation 3.10)}$$

P(**u**) : Patterson function at vector **u**
u : The vector relate two positions $\mathbf{r}_1$ and $\mathbf{r}_1$+**u** in real space

A special map can be calculated from Patterson function, the Patterson map. From equation 3.10, a peak in the Patterson map represents a vector between a pair of atoms or an atom with itself or a symmetry-related atom (self-vectors). Apart from the peak at origin, which is caused by all self-vectors, there are $N^2 - N$ peaks in a Patterson map from N atoms. This makes the Patterson map derived from a native protein crystal impossible to interpret. However, since the height of a Patterson peak is proportional to the product of the atomic numbers of its corresponding atom pair, the vectors between heavy atoms can be identified. By some more calculation, the absolute heavy atom positions can be derived from these Patterson peaks.

After the substructure of heavy atoms (Selenium atoms) was determined, the structure factor phase was determined using multiple-wavelength anomalous diffraction (MAD) method. MAD method makes use of the anomalous difference between **F**(**h**) and its Friedel's mate **F**(-**h**). The principal of MAD can be described with the Harker diagram (Fig 3.3.6), which is based on the following equation:

$$\mathbf{F}_N = \mathbf{F} - \mathbf{F}_A \qquad \text{(Equation 3.11)}$$

F, $\mathbf{F}_A$, $\mathbf{F}_N$: Structure factor in total, from anomalous scattering and normal scattering, respectively.

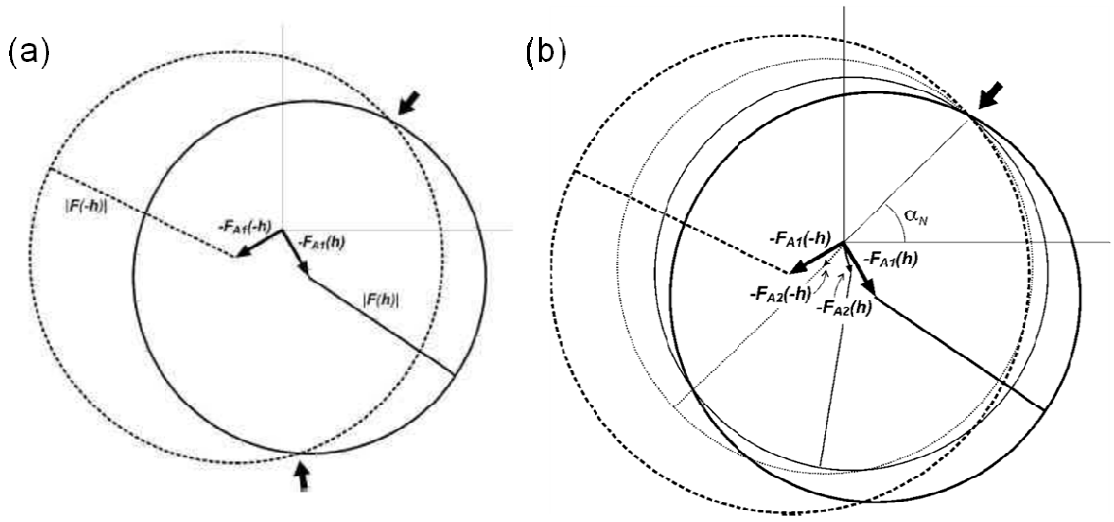

Fig 3.3.6 Harker diagram of (a) single- and (b) multiple-wavelength anomalous diffraction method, from [97].
FA1, FA2: structure factor contributed by anomalous scatterers at 2 different wavelengths; |F|: total structure factor amplitude;
α_N: phase of the structure factor F_N.

With heavy atom substructure determined by Patterson method, $\mathbf{F}_A(\mathbf{h})$ and $\mathbf{F}_A(-\mathbf{h})$ can be determined. The total structure factor amplitudes $|\mathbf{F}(\mathbf{h})|$ and $|\mathbf{F}(-\mathbf{h})|$ can be derived from diffraction pattern. Therefore a phase circle construction diagram (Harker diagram) can be drawn as Fig 3.3.6 a. The points on the two phase circles represent all the possible values for $\mathbf{F}_N(\mathbf{h})$ (solid circle) and $\mathbf{F}_N(-\mathbf{h})$ (dashed circle), respectively. Since $\mathbf{F}_N(\mathbf{h}) = \mathbf{F}_N(-\mathbf{h})$, the two points where the two phase circles meet (marked by arrows in Fig 3.3.6 a) are the possible correct solutions for $\mathbf{F}_N$.

In order to find the correct solution out of these two points, at least one more measurement at a different wavelength is performed. From the new measurement, two more phase circles can be drawn (see Fig 3.3.6 b). The point where all four circles pass will give the correct phase for $\mathbf{F}_N$, which can then be used in Equation 3.7 to calculate the electron density in unit cell. In practice, the two different wavelengths are typically chosen to maximize *f''* and minimize *f'*, referred as "peak" and "inflection", respectively. In addition, a wavelength remote from peak is used, where the anomalous scattering is very weak.

Three anomalous diffraction datasets from a crystal of SeMet-labeled Tca17 were collected at the peak, inflection and remote wavelengths, respectively. After reflection indexing and scaling, the MAD datasets were uploaded and processed with the Auto-Rickshaw server [98]. In this automatic crystal structure determination pipeline, the heavy atom substructure was determined with SHELXC and SHELXD [99], then

heavy atom positions refined and phases calculated with SHARP [100] and MLPHARE [96]. From the calculated electron density map an initial model was calculated with HELICP [101].

3.3.5.2 Molecular replacement phasing

After an incomplete atomic model of Tca17 is derived from experimental phase, it is used to calculate the initial structure factor phases of a native dataset, with molecular replacement method.

Since the two crystals used to collect the anomalous scattering datasets and the native dataset were not perfectly isomorphous, the initial task is to place the model in the correct orientation and position of the unit cell, in respect to the native dataset. The correctly oriented and positioned model will yield a Patterson map which is similar to the Patterson map computed from the experimental data. Phases can then be derived from the model, which can be used to calculate an electron density map with Equation 3.7.

3.3.6 Model building

As the last step of crystal structure determination, an atomic model of Tca17 was fitted into the electron density map. The correct fit of the protein model with the experimental data was ascertained by computing the crystallographic *R*-factor (Equation 3.14), and probing its stereochemistry (see 3.3.7).

The initial model from automated model building was visualised with COOT [102] with the experimental electron density map calculated from three-wavelength MAD data, and modified and improved manually. Then the improved initial model was used to calculate an electron density map from the native dataset with better resolution by molecular replacement phasing, using the program Molrep from the CCP4 suite [96]. An automated model building with ARP/wARP [101], using the incomplete structure as starting model was also performed to accelerate the manual work.

In order to further refine the model with respect to the high-resolution data, different types of electron density maps were calculated to fit the model to the electron density. A $2F_o$-F_c map was calculated according to Equation 3.12, which can be regarded as the sum of a simple electron density map and a difference electron density map, so as to reduce model bias. A F_o-F_c map was calculated according to Equation 3.13. This difference electron density was inspected to reveal errors and missing parts in model.

$$\rho(xyz) = \frac{1}{V}\sum_h\sum_k\sum_l \left||2\cdot F_{obs}| - |F_{calc}|\right| \exp\{-2\pi i(hx+ky+lz) + i\alpha_{calc}\}$$

(Equation 3.12)

$$\rho(xyz) = \frac{1}{V}\sum_h\sum_k\sum_l \left||F_{obs}| - |F_{calc}|\right| \exp\{-2\pi i(hx+ky+lz) + i\alpha_{calc}\}$$

(Equation 3.13)

$|F_{obs}|$: Measured structure factor amplitude
$|F_{calc}|$: Calculated structure factor amplitude
α_{calc} : Phase of structure factor, calculated from current model

Besides the atomic model of Tca17 protein, water molecules and other ligands binding to the protein were also fitted into the electron density map.

3.3.7 Model refinement

How well an atomic model agrees with the experimental data can be represented by the crystallographic *R*-factor calculated as follows:

$$R = \frac{\sum_{hkl}\left||F_{obs}| - |F_{calc}|\right|}{\sum_{hkl}|F_{obs}|} \times 100\% \qquad \text{(Equation 3.14)}$$

$|F_{obs}|$: Measured structure factor amplitude
$|F_{calc}|$: Calculated structure factor amplitude

In practice, instead of calculating an *R*-factor for all data, partial *R*-factors are often calculated. R_{free} is calculated for a test set of reflections (usually 5% of the data) that is not used in refinement, and R_{work} is calculated for the remaining structure factors that are included in model refinement. A correctly refined structure should have low values for both *R*-factors, and not more than 5% difference between them.

The atomic model of Tca17 was refined over several rounds with the program REFMAC5 [103], which is based on maximum likelihood methods. The stereo-chemical quality was evaluated using the program PROCHECK [104]. The root-mean-square deviations (rmsd) from standard bond lengths and angles were kept within a reasonable range during the refinement, whereas the torsion angles between the peptide planes at the Cα atom of an amino acid were not restrained, and were used to verify the result of refinement. In a correctly refined structure, all backbone torsion angles should fall into the favored or allowed regions of the Ramachandran diagram.

3.4 Biophysical and biochemical methods

3.4.1 Analytical ultracentrifugation (AUC)

Analytical ultracentrifugation is a biophysical technique that quantitatively measures sedimentation of molecules or complexes. By analyzing the data, important characteristics can be derived about the molecules in solution, for instance their mass, the stoichiometry of a complex, or the association constants of two interacting partners. In this study, sedimentation velocity (SV) methods were used, which measure the rate at which molecules move in response to centrifugal force generated in the ultracentrifuge. From the measured data, the sedimentation coefficient (*s*) of the protein or protein complex can be calculated, which depends on molecular weight as well as molecular shape.

In practice, through some mathematical manipulation, SV data can be transformed into the sedimentation distribution *c(s)* graph, which shows how much material is sedimentated at various values of *s*. The molecular mass *M* for each peak in the *c(s)* graph can be estimated from the following equation:

$$\frac{s}{D} = \frac{M}{kT} \qquad \text{(Equation 3.15)}$$

s : Sedimentation coefficient
D : Diffusion constant, can be estimated from size and shape of the sedimented molecule
k : Boltzmann's constant
T : Absolute temperature

In order to determine the oligomerization state of Tca17 in solution, SV experiments were performed (by Marcel Jurk from the Leibniz-Institut für Molekulare Pharmakologie, Berlin) using an analytical ultracentrifuge equipped with absorption and interference optics. Protein solution and reference buffer, 400 µl each, were filled separately in the two sectors of a two-sector titan cell. SV data were recorded with absorbance scans when the cells were centrifuged at 40,000 rpm, 20 °C. Recorded SV data were evaluated with the program SEDFIT [105]. Buffer density and viscosity as well as partial specific volume of the protein were calculated using the program SEDNTERP [106].

3.4.2 Static light scattering (SLS)

Static light scattering is a technique that measures the intensity of the scattered light at multiple angles to obtain the average molecular weight of a macromolecule. When coupled with size-exclusion chromatography, SLS can be used to detect the presence of different macromolecule species in solution, and determine the mass of each species.

In this study, the oligomerization state of Tca17 in different protein buffers was determined by running the purified protein through an analytical gel filtration column (Superdex 75 10/300 GL), then through the 2-angle SLS detector. The intensity of the scattered laser light was recorded and analyzed by the program OmniSEC. For accurate results, the gel filtration column and the SLS detector had to be well equilibrated with the corresponding buffer.

3.4.3 Pull-down assays

The pull-down assay experiment is a commonly used technique to probe protein-protein interactions. The purified bait protein fused with an affinity tag is immobilized on a specific affinity matrix, then the prey proteins in solution will be incubated with the matrix. If the two protein partners really interact with each other, the prey protein will be associated with the matrix, which can be confirmed by SDS-PAGE and suitable detection methods, for instance gel staining or western blotting.

In this study, both GST- and His- pull-down assays were performed to probe physical association between Tca17 and yeast TRAPP subunits Bet3p and Trs33p. GST-tagged Tca17 was expressed in *E. coli* BL21(DE3)-T1R cells, and purified with a GSTrap column (according to the protocol in 3.2.5.2). The genes of Bet3p and Trs33p were cloned into the pQLink-H vector with an N-terminal His_7 tag. The sequence-verified expression vectors were also transformed into *E. coli* BL21(DE3)-T1R cells. Then the two proteins were purified with a HisTrap column (according to the protocol in 3.2.5.1). For GST pull-down assay, His-tag was removed with a second HisTrap column. After purification, all proteins were dialysed at 4 °C overnight against protein storage buffer, and stored at 4 °C until usage.

For GST pull-down assay, 0.25 mg of tag-free Bet3p or Trs33p was diluted into 500 µl of binding buffer 1. GST-Tca17 at 1:1 molar ratio and 40 µl GSH sepharose 4B beads (50% in binding buffer 1) were added. And then the mixture was incubated

overnight at 4 °C. After washing three times with 500 µl of binding buffer 1, the GST-Tca17 and bound prey proteins were eluted with 50 µl elution buffer 1, and examined with SDS-PAGE. As control experiments, the same pull-down assay was performed in parallel with the same molar amount of GST protein. The pull-down signals from GST-Tca17 and GST were compared to exclude false positive results.

For His pull-down assay, 0.25 mg of His_7-Bet3p was diluted into 500 µl of binding buffer 2. Tag-free Tca17 at 1:1 molar ratio and 40 µl Talon beads (50% in binding buffer 2) were added. And then the mixture was incubated overnight at 4 °C. After washing three times with 500 µl of binding buffer 2, His_7-Bet3p and the bound Tca17 protein were eluted with 50 µl elution buffer 2, and examined with SDS-PAGE. As control experiments, same amount of tag-free Tca17 alone was added directly to the Talon beads. The pull-down signal was compared with the control to exclude false positive results.

3.4.4 Co-immunoprecipitations (CoIPs)

The interactions between NIBP and Ehoc-1 were studied with co-immunoprecipitation (CoIP) experiments. NIBP fragments amplified from its cDNA clone were cloned into the pcDNA3-flag vector, while Ehoc-1 fragments amplified from its cDNA clone were cloned into pTL-HA 1 vector with restriction endonuclease-based cloning. After verification by DNA sequencing, each pair of expression vectors for flag- and HA-tagged proteins were used to transfect HEK293 cells by calcium precipitation for transient expression of the tagged proteins. (Cell transfection and CoIPs were performed in cooperation with Daniel Kümmel, a former PhD student in the same group.)

Each sample of the HEK293 cell lysate was incubated with 10 µl α-HA affinity matrix for 1 h at 4 °C. After washing 5 times with CoIP buffer, the immunoprecipitations (IPs) were resuspended in 2 x SDS sample buffer and analyzed by SDS-PAGE and Western blot using specific antibodies.

3.4.5 Western blot

The protein samples were separated on a 15% SDS-PAGE gel with an additional lane of Precision Plus protein dual color standard (from Bio-Rad). Before the immunodetection, proteins were transferred from the gel onto a 5.5 x 9 cm poly-

vinylidene fluoride (PVDF) membrane. PVDF membranes had to be activated before use by soaking briefly in methanol. Then the membrane was washed with H_2O, and incubated in semi-dry transfer buffer with 2 pieces of 5.5 x 9 cm filter paper for at least 10 minutes. The transfer was performed for 1 h with a semi-dry blotting device at 44 mA per gel with filter paper and membranes soaked in semi-dry transfer buffer.

3.4.6 Immunodetection

The proteins immobilized on a membrane were probed with immunodetection using specific antibodies. For detection of flag- and HA-tagged samples, the membrane was first blocked with 3% BSA in PBST buffer to saturate unspecific binding sites. Then the blocked membrane was incubated for 1 h at room temperature or overnight at 4 °C in primary antibody buffers, which were prepared by diluting primary antibodies in PBST buffer with 1.5% BSA (α-flag 1:6000, α-HA 1:1000). After washing 3 times with PBST buffer, the membrane was incubated for 1 h at roomtemperature in secondary antibody buffer, which was prepared by diluting the HRP conjugated secondary antibody in PBST buffer with 1.5% BSA.

Finally the membrane was washed extensively with PBST buffer to remove the weakly bound HRP conjugated antibody, and was developed using ECL Western blotting analysis reagents, and visualised by LAS-4000 gel documentation system.