JOHANNES KEPLER UNIVERSITÄT LINZ | JKU

TNF Technisch-Naturwissenschaftliche Fakultät

Initial NMR Spectroscopic Investigation of the Extrinsic Protein PsbP of Photosystem II

MASTER THESIS

to obtain the academic degree

Master of Science

in the master study of

BIOLOGICAL CHEMISTRY

Submitted by:
Bc. Adriana Walnerová, BSc.

Carried out at:
Institute of Organic Chemistry

Supervisors:
Univ. - Prof. Mag. Dr. Norbert Müller
Assoc. Prof. RNDr. Rüdiger Ettrich, PhD.

Guidance:
Dipl. - Ing.in Dr. Michaela Horničáková
Kousik Chandra, PhD.

Linz, January 2013

Ich erkläre an Eides statt, dass ich die vorliegende Diplomarbeit selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt bzw. die wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe. Die vorliegende Diplomarbeit ist mit dem elektronisch übermittelten Textdokument identisch.

Linz, January 2013

.....

Adriana Walnerová

I hereby declare that, in accordance with Article 47b of Act No. 111/1998 in the valid wording, I agree with the publication of my master thesis, in full to be kept in the Faculty of Science archive, in electronic form in publicly accessible part of the STAG database operated by the University of South Bohemia in České Budějovice accessible through its web pages. Further, I agree to the electronic publication of the comments of my supervisor and thesis opponents and the record of the proceedings and results of the thesis defense in accordance with aforementioned Act No. 111/1998. I also agree to the comparison of the text of my thesis with the Theses.cz thesis database operated by the National Registry of University Theses and a plagiarism detection system.

Linz, January 2013

.....

Adriana Walnerová

Acknowledgements

I would like to thank to prof. Norbert Müller for a great supervision of my thesis as well as for letting me work on such an interesting topic and encouraging me in further education in this field. My big thanks go also to prof. Libor Grubhoffer who is together with prof. Müller a guarantee of the study program Biological Chemistry. They both made a great job housekeeping this unique curriculum and letting us students to study smoothly at both universities and fulfill our ambitions. Herewith I would also like to thank to all my classmates who made these five years truly unforgettable.

For kind consultations concerning the PsbP and provision of the vector I want to thank prof. Rüdiger Ettrich from the Institute of Nanobiology and Structural biology of GCRC in Nové Hrady. I am very grateful to Anders Pedersen and Göran Karlsson from the Swedish NMR Centre at the University of Gothenburg for letting me carry out the cell-free expression of PsbP within the scope of a workshop held at their facility.

To all the members of the Institute of Organic Chemistry I would like to thank for creating such a nice atmosphere and making every working day a little better. Especially big thanks to Dr. Michaela Horničáková who taught me every single practical detail for the lab work and brought a huge piece of optimism even when we were running into some troubles. She made me laugh every day and forced me to eat tons of chocolate, which was keeping me up especially in long winter days! Without her help, this thesis would have been much more complicated. I can’t forget Dr. Judith Schlagnitweit who helped with the recording of NMR spectra and explained even the most complicated things in a very intelligible way. Last, but not least, I am thankful to Kousik Chandra, PhD for being patient with me spending time at the spectrometer, and for useful hints for working with the spectra.

I am very grateful to my whole family who made it possible for me to study this special program and encouraged me in every aspect. Their sympathy and a sense of humour always brought me back to the ground and helped me to overcome the demanding times. Thanks to all my friends for having a wonderful time even if it was always a little too less than we wanted. Especially thanks to Gábi, who shows me that one's dreams are never impossible!

Who deserves to gain a gold medal for being with me all the time, is Petr. He supported me in every possible way and I can't imagine bearing all those things without him. His understanding and love are incredible and I cannot tell how much it means to me that he withstood all the hard times without the slightest complaint.

"Life is what happens while you are busy making other plans."

John Lennon

Abstract

The extrinsic protein PsbP from Photosystem II of *Spinacia oleracea* was prepared by expression as a recombinant fusion protein in genetically engineered *E. coli* cells. In addition the PsbP was also synthesized using a cell-free expression method with a ribosomal extract. To improve upon the initially poor stability of PsbP, isolation and purification conditions and procedures were optimized to preserve PsbP in its folded monomeric state. The protein was enriched with stable C-13 and N-15 isotopes in order to enable investigation by means of multi-dimensional nuclear magnetic resonance. The PsbP gene construct included a 6x histidine tag at the N-terminus for selective purification. His-PsbP was isolated from bacterial *E. coli* cells and purified first with Ni^{2+} affinity chromatography and ion-exchange chromatography. At that stage the His-tag was cleaved off by a thrombin protease and the reaction mixture was purified via gel filtration. The final pure protein was obtained in sufficient amounts from 1 L of minimal expression medium (M9) – 15 mg of (N-15) PsbP and 21 mg of (N-15, C-13) PsbP. The labelled proteins were used for recording of 2D and 3D NMR spectra. These were used to start the initial backbone resonance assignment of PsbP. A total of 42 amino acids could be assigned sequence specifically so far. Interestingly, these amino acids were found to be located mostly in two distinct regions of the PsbP sequence, where they are missing in the previously determined crystallographic structure. Preliminary experiments under optimized conditions show the way, how to proceed towards full NMR assignment of PsbP.

Table of Contents

1. List of Abbreviations

APS = ammonium persulfate

DEAE = diethylaminoethanol

DSS = 3 – (trimethylsilyl) – 1 – propanesulfonic acid-d_6 sodium salt (98% D)

DTT = dithiothreitol ((2S,3S)-1,4-bis(sulfanyl)butane-2,3-diol)

EDTA = ethylenediaminetetraacetic acid

HSQC = heteronuclear single quantum coherence spectroscopy

LB = lysogeny broth medium

M9 = minimal medium

NMR = nuclear magnetic resonance

NOESY = Nuclear Overhauser effect spectroscopy

PAGE = polyacrylamide gel electrophoresis

PS = photosystem

QAE = Diethyl(2-hydroxypropyl)ammonioethyl

SDS = sodium dodecyl sulfate

SP = sulfopropyl

TEAE = Triethylaminoethyl

TEMED = tetramethylethylenediamine

2. Introduction

2. 1. Evolution of oxygenic photosynthesis

Photosynthesis takes place in all kinds of higher plants, algae and cyanobacteria. Oxygenic photosynthesis provides the Earth with stable oxygen supply and ensures all life forms with the necessary fuel. The rise of photosynthetic organisms started to shape the biosphere 2.3 billion years ago when the green algae started to produce at that time toxic molecular oxygen which now makes 21% of the atmosphere [1, 2]. Recent research indicates that these blue-green algae, Cyanobacterium, might be even older [3].

Photosynthetic reactions are carried out with the help of roughly 100 proteins spread throughout the membrane of photosynthetic organelles. These form four major multi-protein membrane complexes: photosystem I and II, cytochrome b_6f and F-ATPase. The outcome of the series of reactions going on in these complexes is conversion of sunlight into chemical energy with synchronous water reduction yielding molecular oxygen and reducing compounds like NAD(P)H [4].

When thinking about evolution of oxygenic photosynthesis, one might turn the attention to the anoxygenic one, its probable precursor. This process is carried out nowadays in various kinds of bacteria which are able to use protein complexes similar to photosystem I or photosystem II, respectively. Such a nice example is a purple bacterium *Rhodopseudomonas palustris* having reaction center which mimics photosystem II and ensuring cyclic electron flow when grown in light. If the bacterium does not have an access to the light energy, it can easily switch off its photosynthetic machinery and use only cellular respiration as energy source. On the opposite, the anaerobic green bacterium *Chlorobium tepidum*, extracts the energy from hydrogen sulfide as substrate, which in combination with light ensures only linear electron flow [5, 6]. How is it possible that such traits of genes encoding single photosystems are found among these bacteria? How did these genes happen to be contained in higher plants yielding both complexes? Answer might be a missing link in the evolutionary tree of bacteria, which was proposed to be a „protocyanobacterium“ – an organism which is capable of expressing both anoxygenic photosystems [7].

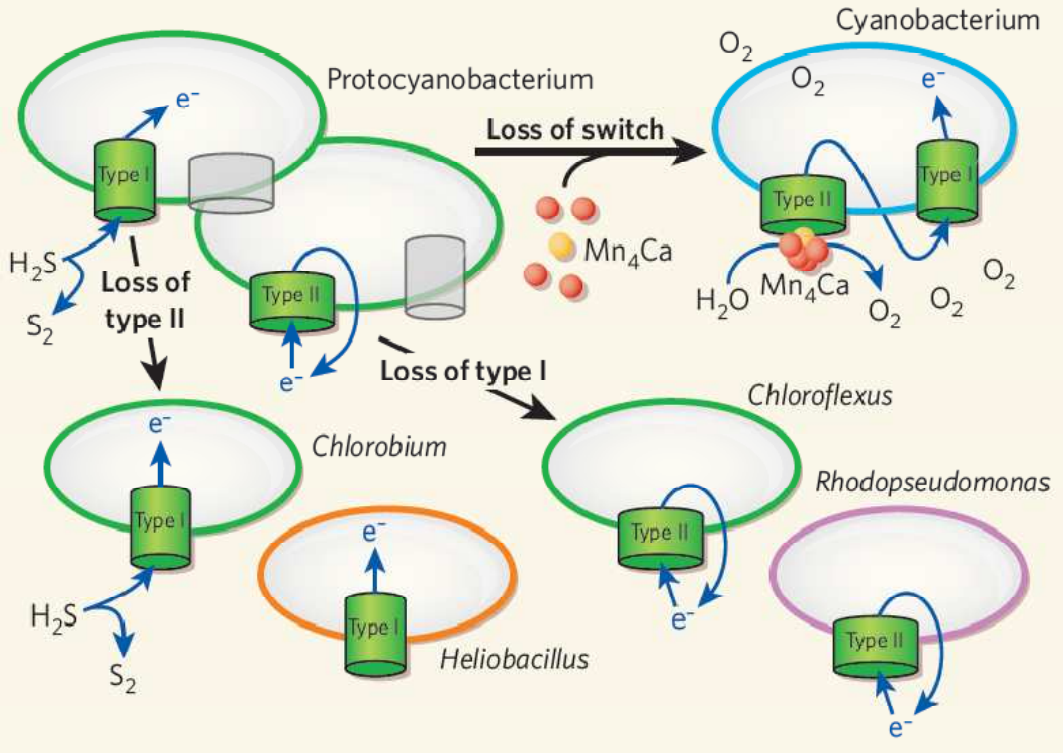


Figure 1: *Evolution of photosynthesis in cyanobacteria* [4].

Comparison of genomes of several cyanobacteria and anoxygenic photosynthetic bacteria revealed a core gene set standing behind the photosynthesis process. It is assumed that PS I is the older prototype from which PS II evolved eventually. The protocyanobacterium probably alternated the expression of PS I and PS II by presence of regulatory switch. A possible mutation under feasible circumstances lead to a loss of this regulatory switch and eventually to an incorporation of the water splitting center composed of four manganese and one calcium atoms. With this novel catalyst, a linear flow of electrons is employed which is used for water oxidation in true cyanobacteria. Without the manganese core, the two photosystems would not be able to coexist with each other since they share parts of electron transport chain being cyclic and linear at one time thus leading to a self-destruction [4].

Nevertheless, evidence about the existence of this protocyanobacterium has not yet been found and therefore the hypothesis is still waiting to be supported by solid evidence.

2.2. Photosystem II – a multiprotein complex

PS II is capable of generating an electrochemical potential of +1.1 V, which is sufficient to remove 2 e^- from each of two water molecules and at the same time consuming 4 photons. These four electrons are combined with 4 more photons in PS I, which pushes them on the electron transporter $NADP^+$. Electrons are finally transferred to a carbon dioxide molecule, which serves as basis for the production of sugars and higher organic molecules [4].

The basis for the water splitting to occur is the so-called oxygen evolving center, represented by the Mn_4Ca cluster. The complex sits at the outer surface of the photosynthetic membrane. Two water molecules enter in at one side and the molecular oxygen leaves it at the opposite side. The process is achieved by removal of four e^-, each from one manganese atom, which are replaced by electrons from water. Electrons are afterwards transferred to tyrosine residues from which they travel to chlorophyll oxidized by light absorption [8].

2.3. Psb family of proteins

Psb proteins are located in PS II, a multi-protein and pigment complex located in the first place of light-induced photosynthetic reactions. In cyanobacteria the extrinsic part of PS II is formed by three proteins – PsbO, PsbU and PsbV [9]. Together with 17 transmembrane proteins of various sizes, they form the PS II. While the transmembrane part is conserved from bacteria to higher plants, individual species show high variation in the extrinsic moieties.

The PsbO (33 kDa) is present in both, eukaryotic and prokaryotic organisms in which it plays a major role in stabilizing the manganese center and maintaining its function. The other, smaller proteins differ in between the species. In higher plants, one can find three other subunits – PsbQ, PsbP and PsbR whereas in cyanobacteria these are replaced by PsbU and PsbV. Recent studies proposed existence of additional 20 kDa PsbQ-like protein called PsbQ´ in red algae and fifth extrinsic 13.5 kDa subunit in diatoms [10].

PsbO was proved to be a requisite in a proper functionality of PS II and OEC, respectively, in various experiments. In one of the studies higher plants treated with $CaCl_2$ lacked PsbO protein as well as adjacent PsbP and PsbQ but the manganese cluster was kept intact [11, 12].

Nevertheless, the rate of oxygen evolution dropped down rapidly. At the same time it was proved that PS II void of PsbO showed slower electron flow from the Mn_4CaO_5 cluster to tyrosine residues [13].

PsbR is a 10 kDa protein found only in higher plants and green algae. PsbR seems to bind to the lumenal side of thylakoids via its highly hydrophobic C – terminus. This part is also responsible for lowered solubility of PsbR and its tendency to precipitate [14]. Although very small, the PsbR protein seems to be necessary for PsbP incorporation into PS II [15].

PsbQ is a 16.5 kDa protein sitting right next to PsbP. PsbQ together with PsbO and PsbP are essential for stability and proper functionality of oxygen evolving center in PS II. Function and mode of action of PsbP and PsbQ is still not fully understood, but it has been suggested that these proteins are responsible for the modulation of physiological calcium and chloride concentration [16]. If these conditions are disturbed, the rate of oxygen evolution is decreasing significantly. PsbP and PsbQ also serve to stabilize the OEC, though none of them seems to directly bind one of the manganese atoms [17]. These cooperating proteins tightly control access of substrates to the membrane and protect it from exogenic reductants except water [18].

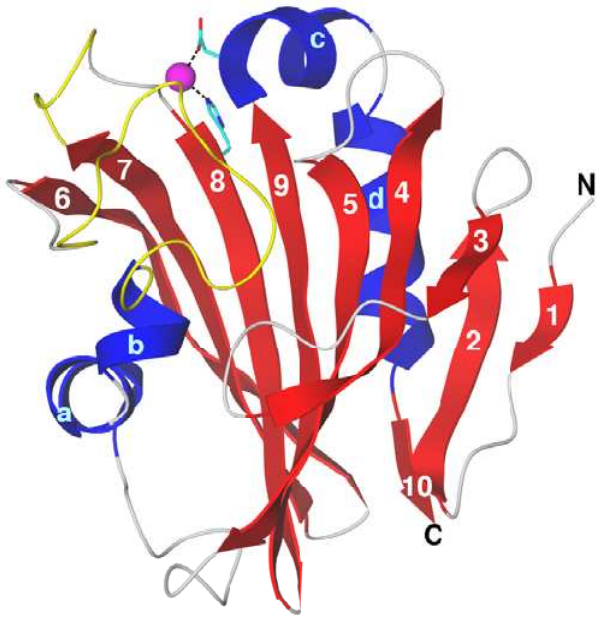


Figure 2: *X-ray crystallographic structure of PsbP from Spinacia oleracea with three modeled parts – N-terminus and two internal loops (in yellow)*[19].

The mature part of the PsbP protein consists of 190 amino acids forming a 20.2 kDa polypeptide. X-ray crystallographic structures of PsbP from higher plants have been determined over past years. The newest 3D structure is of the PsbP from *Spinacia oleracea*[19]. It was obtained with high 1.98 Å resolution. In crystals PsbP consists of a central anti-parallel β-sheets surrounded by two α-helical domains. The crystalline protein contained all parts of folded PsbP, but still two internal regions were left unresolved in the electron density map together with the N-terminal region (Figure 2). Parts located in the middle region of PsbP were modeled as loops and the terminal segment as β-sheet[19].

The N-terminus of PsbP is most probably crucial for the functionality of OEC and of course the oxygen evolution. PsbP truncated by 15 N-terminal amino acids was still able to bind to PS II membrane, but it was not able to switch on the oxygen production. This flexible region is also necessary for Ca^{2+} and Cl^- retention [20].

The association of PsbP to PsbQ and their assembly to the PS II membrane still leave many questions open. The most accepted model assumes that PsbO binds to the PS II membrane followed by PsbP to which PsbQ is attached in the end. Recently, contradicting results were obtained concerning the PS II assembly in green algae. PsbP and PsbQ were able to attach themselves to photosynthetic membrane without the help of PsbO [21].

3. Materials and Methods

3.1. Analysis

3. 1. 1. SDS-PAGE

For the routine analysis of protein samples, sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was employed. This technique is used for the separation of proteins according to their size only. Acrylamide and bisacrylamide form a matrix by polymerization through which the proteins migrate from cathode to the anode in a proper buffer. This is facilitated by addition of a strong detergent SDS, which destroys the native structure of the protein and gives it a net negative charge. Another agent, DTT, is added to the electrophoretic sample to prevent oligomerization rising from the formation of disulfide bridges between cysteines. This method is widely used for the check on protein purity, size and roughly concentration. Proteins can be isolated from the gel and refolded as well [22].

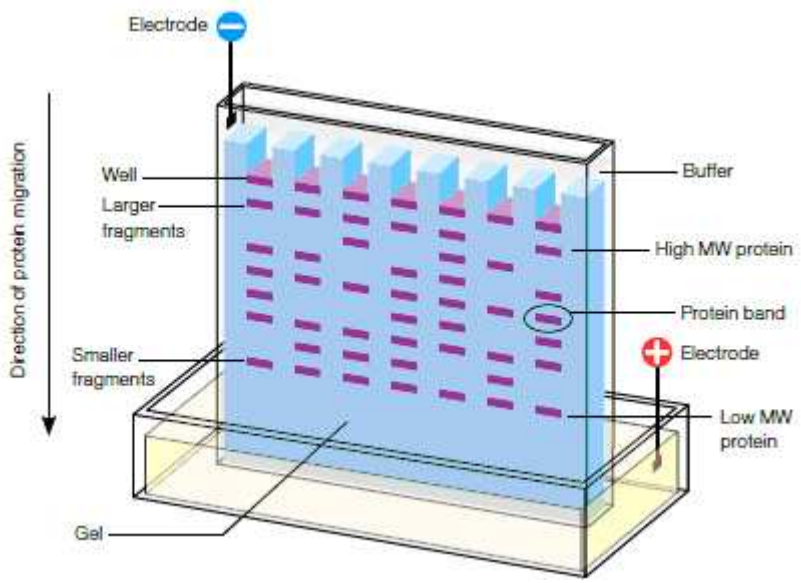


Figure 3: *Principle of SDS-PAGE* [23].

3. 1. 2. Native PAGE

Native polyacrylamide gel electrophoresis is used for the analysis of proteins under non-denaturing conditions. In that case neither reducing agents such as DTT or 2-mercaptoethanol nor detergents like SDS are added. Overall electrophoretic mobility of the protein is determined by its internal charge at given pH and hydrodynamic size. This method is therefore very useful for the study of protein complexes, oligomers and binding events as well as protein susceptibility to temperature or different buffers.

3. 1. 3. Concentration determination

In order to calculate protein concentration, its UV absorbance is measured. Absorbance in near UV depends mostly on aromatic amino acids tyrosine, tryptophan and in lesser extent on phenylalanine and disulfide bonds between two cysteine residues.

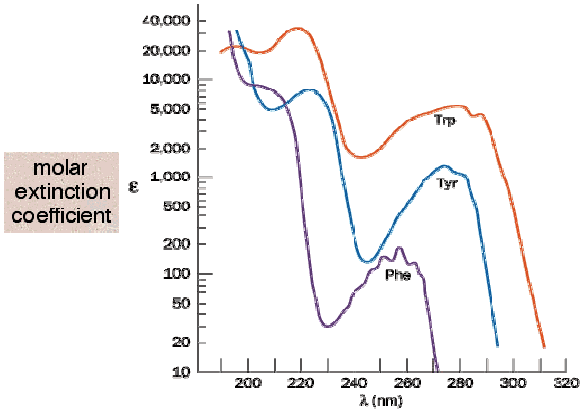


Figure 4: *UV absorbance profile of aromatic amino acids* [24].

The Beer-Lambert law is employed to calculate exact concentration of the protein in a solution. This value may be however disturbed by presence of nucleic acids whose extinction at 280 nm can be 10x higher than the one of the protein [25].

The formula of the Lambert-Beer law is

$$A_{280} = \varepsilon \cdot c \cdot l$$

where A_{280} is absorption at 280 nm, ε is extinction coefficient, c is the concentration in mol/l and l is optical path length in cm. With given protein-specific extinction coefficient, the concentration calculation is a straight-forward process [26].

For the determination of DNA concentration, absorbance is measured at 260 nm. Since the concentration of double stranded DNA with $Abs_{260} = 1$ is 50 µg/ml, we can use following formula for the calculation of plasmid DNA concentration:

$$\text{Concentration } (\mu g / ml) = \text{Abs}_{260} * 50 * \text{dilution factor}$$

3. 2. Protein Expression

3. 2. 1. In vivo expression

Preparation of a recombinant protein nowadays is a routine work in most of the biotechnology laboratories. One can choose from a great variety of eukarotic and prokaryotic host organisms, which are able to synthesize the desired protein. The most extensively studied organism in this sense is a gram-negative bacterium *Escherichia coli*. These genetically engineered cells are able to produce vast amounts of recombinant protein in a reasonable time at low cost. However *E. coli* machinery is not able to produce additional post-translational modifications like glycosylation, phosphorylation or acetylation. In that case, yeasts, insect or mammalian cells are of better choice. Biggest challenge lies in the preparation of membrane proteins. These are frequently toxic to the cells or might not be folded properly. This is one of the reasons why many researchers try to synthesize such proteins *in vitro*, i.e. in a cell-free environment.

The procedure starts with the preparation of a plasmid – a circular double stranded DNA, which is often found in bacteria. Such a commercially available plasmid is adapted by an

insertion of desired sequence, which codes for the protein of interest. A novel plasmid bearing the necessary genetic information is afterwards transformed into bacterial cells, which are able to read this novel genetic information. These special bacterial strains are programmed to overexpress the protein and fairly high amounts of recombinant protein can be produced by this method. Two kinds of media are generally used for a bacterial growth – rich and minimal medium. The rich medium, for instance LB, contains yeast extract and casein hydrolysate as a source of amino acids and nitrogen. This kind of media is used for routine and easy cultivation of bacteria. The minimal medium, here M9, consists of minimal possible amount of nutrients necessary for a bacterial growth. Amino acids or source of nitrogen and carbon must be added to the medium together with metal ions. In the case of isotopicaly labeled protein, these amino acids (or cheaper precursors glucose and ammonium sulphate) are ^{15}N, ^{13}C and sometimes also ^{2}H labeled.

Antibiotics are used to prevent unwanted growth of different bacterial strains. Since the plasmid codes for an antibiotic resistance, the transformed cells are growing also in the presence of the given antibiotic. When the cells reach log phase, they are induced with IPTG which mimics allolactose and switches on the transcription of lac operon. The Lac operon is a regulatory part of DNA which is located upstream the inserted sequence coding for the target protein. After the induction, bacteria start to express the wanted protein together with their own. If grown at 37°C at that stage, the cells might get overwhelmed by the newly formed recombinant protein and start to deposit it in so called inclusion bodies. These formations are basically clumps of the aggregates from which it is sometimes challenging to recover the protein. This problem can be circumvented by lowering the incubation temperature to 28°C and decrease the rate of protein formation.

Lysis of the cells to release the protein can be achieved either mechanically, e.g. by sonication or using a French press, or enzymatically with the help of lysozyme.

3. 2. 2. In vitro expression

Expression of a protein without the use of living cells has been introduced already in 1961 to produce polypeptide from poly-uracil RNA sequence by M. Nirenberg and H. Matthaei. The outcome was a determination of codons responsible for the formation of individual amino acids. However since that time, cell-free protein synthesis has been established as

an alternative to a conventional *in vivo* approach. At the beginning of 1990s, wheat-germ and E. coli lysates preparation was optimized in order to produce higher yields of the target protein which was seen as the biggest drawback of this method until then. Unlike *in vivo* expression cell-free can be used for an efficient synthesis of proteins toxic to cells. This process is sometimes the only way how to produce big membrane proteins in a combination with proper detergent[43].

Cell-free expression is often used to achieve stable and uniform isotope labeling in proteins as well as for the introduction of non-natural amino acids. Labeling of proteins can be done very specifically, which gives this method great potential for application in NMR studies of proteins.

Cell-free expression can be carried out in a two different ways – the batch and the dialysis mode. A batch mode is based upon the mixing of all ingredients such as ribosome lysate, amino acids, DNA Polymerase, target DNA, low molecular weight mix and energy-regeneration system in a closed vessel, e.g. a test tube. Batch mode suffers from a nucleotide degradation and energy depletion after short time reaching a plateau within 1-2 hours of reaction. Nevertheless, upon optimization it is possible to produce up to 1 mg/ml of the target protein [27].

Dialysis mode circumvents these problems by stable energy, nucleotide and amino acids supply through the membrane whereas the low-molecular weight by-products are dialyzed from the reaction chamber into a bigger reservoir. Using this setup, the reaction can be carried out for 24 hours.

Both methods guarantee synthesis of sufficient protein amounts which are achieved within a fraction of time needed for ordinary in vivo expression. Cell-free expression offers a powerful toll to quickly and economically prepare labeled proteins [28, 29].

3. 3. Protein Purification

3. 3. 1. Affinity chromatography

This type of chromatography is based on the selective interactions of a protein and its ligand which is bound to a solid phase, the purification resin. Such interactions are usually hydrophobic, electrostatic or van der Waals forces. The protein of interest is captured in the chromatographic column via these forces whereas the impurities are flushed away. Afterwards the protein can be eluted from the column either by a competitive ligand or by change of pH, ionic strength or polarity. In this particular case, the purification makes use of high affinity of His-tag to Ni^{2+} ions. Ni-Sepharose resin is therefore employed as a stationary phase. His-PsbP is eluted using gradient of increasing imidazole concentration which binds to Ni^{2+} ions instead of the histidine groups [22, 30].

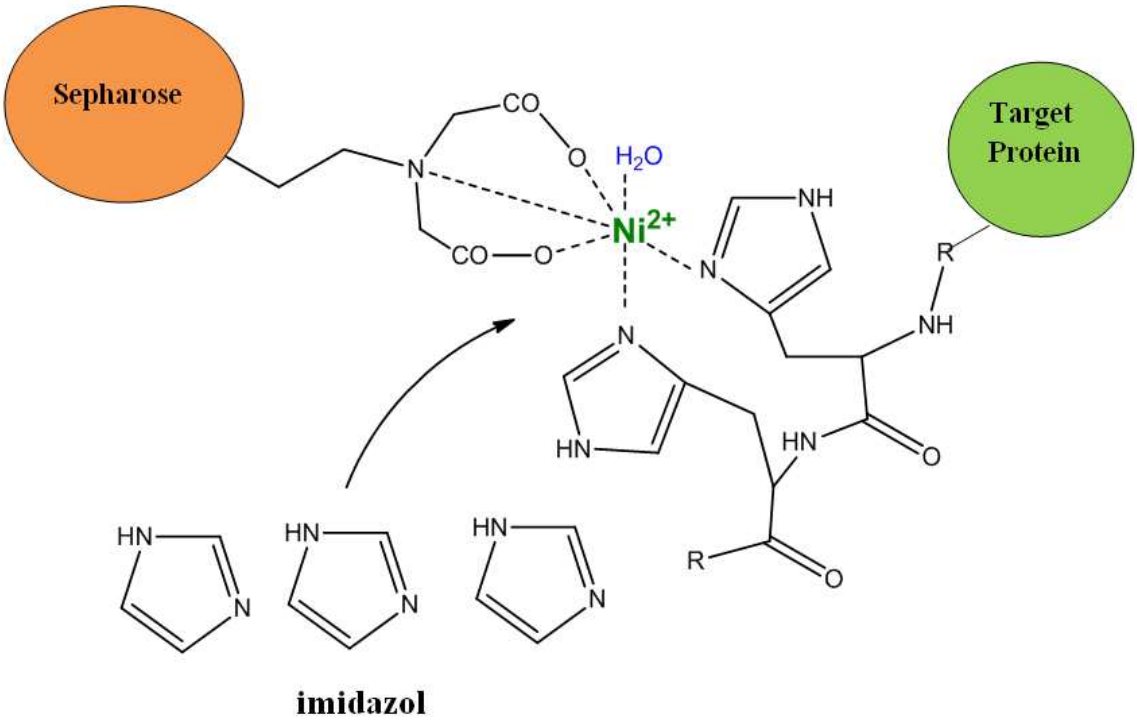


Figure 5: *Histidine groups bound to Ni^{2+} and replaced by imidazole washing.*

3. 3. 2. Ion-exchange chromatography

This method exploits the electrostatic interactions of protein charged groups and oppositely charged functional groups immobilized in the purification resin. Conditions employed in the chromatographic separation are dictated by the pI value of the respective protein. The pI is a pH value at which the protein has a zero net charge. This value depends on the amino acid composition of a protein. The pH value of running buffers and charge of the stationary phase are of a choice. For proteins negatively charged at the working pH, an anionic exchanger column is suitable.

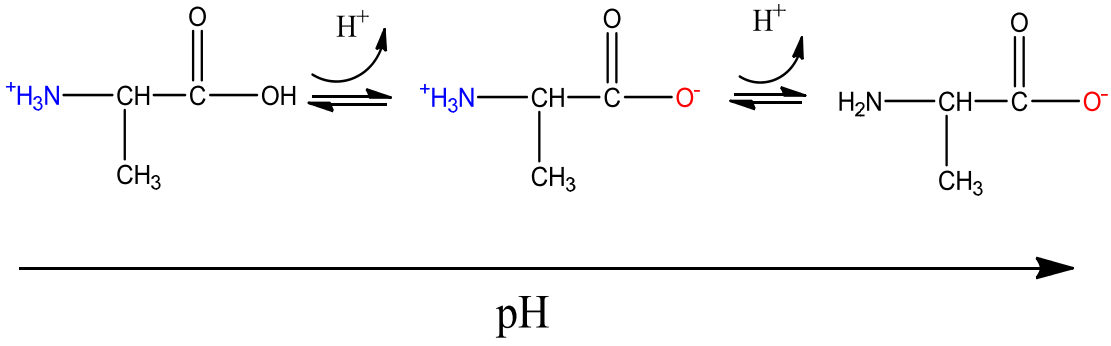


Figure 6: *Equilibrium between the differently charged states of alanine.*

Such ion exchangers containing positively charged groups are DEAE, TEAE or QA sepharoses. Cationic exchangers like SP, SE or CM Sepharose consists of positively charged groups. If the standard mode of the purification is chosen, the protein of interest will bind to the column until it is eluted by counter-ion presence. An opposite approach called „flow through mode“ ensures that the protein is not binding to the column whereas the contaminants are. His-PsbP has its pI at 7.1 which makes it suitable for the purification on a cationic exchanger, in this case a SP-Sepharose. Subsequent elution of the protein is obtained by the concentration gradient of NaCl.

Ion-exchange chromatography is one of the most versatile purification techniques used in the normal laboratory practice with various modifications possible. On top of that, it comprises of high resolving power and high binding capacity [22, 31, 32].

3. 3. 3. Cleavage of the His-tag

Thrombin is a serine protease with highly specific cleavage site *Leu-Val-Pro-Arg-Gly-Ser* in which it cleaves between Arginine and Glycine. In nature thrombin is created by the proteolytic cleavage of its precursor prothrombin. Thrombin is a part of coagulation cascade in which it catalyzes conversion of soluble fibrinogen into insoluble fibrin giving rise to a blood clot. In biochemistry it is used very often as an enzyme for cleavage of fusion proteins or affinity tags, here it is applied to His-tagged PsbP [33].

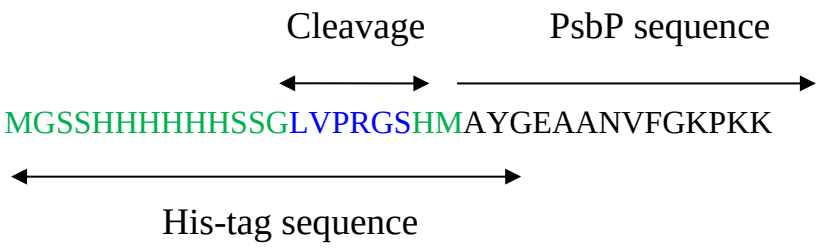


Figure 7: *His-tag sequence at N-terminus of PsbP. Thrombin cleaves the tag between arginine (R) and glycine (G) residue.*

3. 3. 4. Size-exclusion chromatography

Size exclusion chromatography known also as gel-filtration is a technique for separation of proteins according to their size. The matrix is composed of spherical particles made from dextran cross linked with agarose with average particle size of 34 µm. These particles are chemicaly and physically inert. The proteins of smaller size interact by mechanical means

with the beads which causes their longer retention in a column. Bigger proteins travel faster through the column and elute first. Gel filtration is usually used as the last purification step to clear the protein from remaining degradation products and impurities [22, 34].

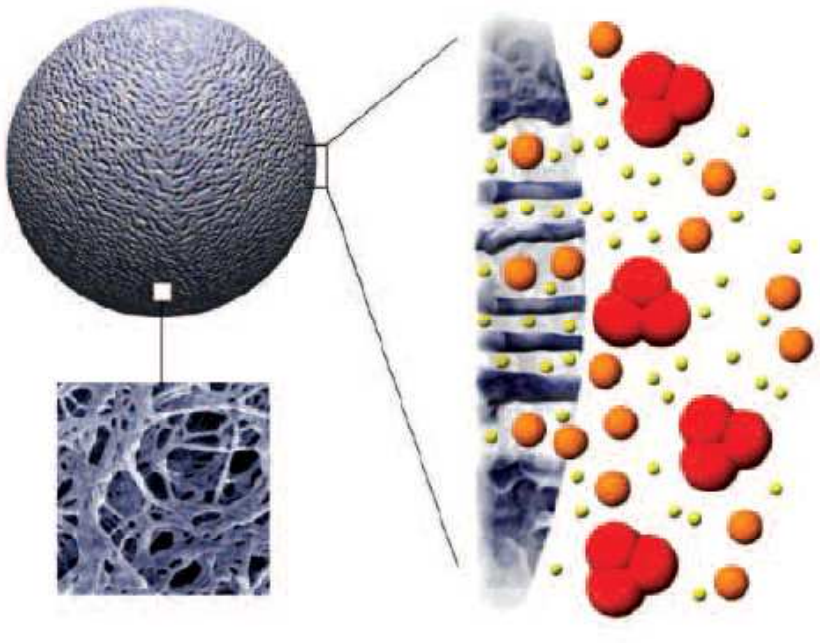


Figure 8: *Porous beads in gel filtration* [33].

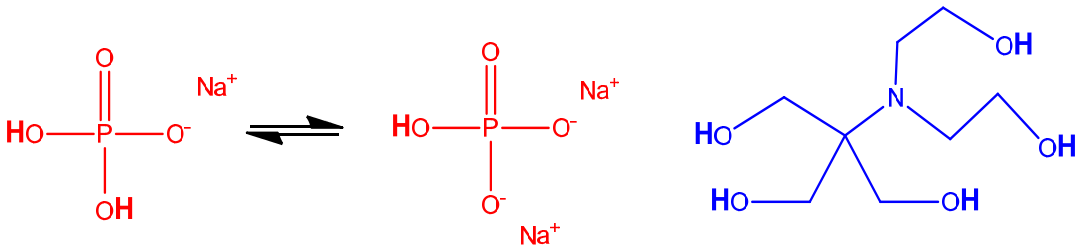
3. 4. Stability Tests

The stability of PsbP protein was one of the crucial prerequisities in order to enable characterization by NMR. There are many things which one needs to consider when it comes to protein stability in general. These factors include concentration, temperature, pH and susceptibility to proteases. Since the protein is generally rather unstable, different conditions were tested to find the optimal ones for long time measurements of NMR spectra.

3. 4. 1. Buffers

Considering the NMR requirements, phosphate and cacodylate buffers are of the best choice. Since the cacodylate buffer is cancerogenic, it was not included in the experiments. Sodium

phoshate buffer was used as the first one and stability in Bis Tris buffer was explored as well since it was proposed to be the most suitable medium for PsbP [35]. However sodium phosphate buffer is more advisable for NMR measurements compared to Bis Tris which gives a rise to solvent artefacts coming from multiple OH groups in ^{1}H dimension (Figure 9).



Na - phosphate buffer

Bis Tris

Figure 9: *Chemical structures of sodium phosphate and Bis Tris buffers. Atoms relevant in NMR measurements are bold.*

3. 4. 2. Protease Inhibitors

The PsbP protein is not bench stable and degrades rapidly if exposed to even slightly elevated temperatures. The protein solution always contained 1 mM EDTA as protease inhibitor. EDTA chelates metal ions which are necessary for the activation of metalloproteases. However EDTA alone is not able to prevent proteolysis and therefore mixture of protease inhibitors were added to the protein sample to check the stabilizing effect.

3. 4. 3. Temperature

The temperature constituted the biggest possible influence on the protein stability. In order to prevent possible degradation, the protein was handled always at 4°C as soon as it was released from the bacterial cells. Reasons for these precautions were its instability at the N-terminal part and possible attack by various proteases. The action of proteases in general can be inhibited by lower temperature to some extent. The N-terminal part however degrades very readily even at room temperature. This phenomenon, when 5-15 terminal amino acids were cleaved off, was observed by Ifuku who also suggested that the reason for it is its increased mobility which makes this site functionally important [20]. Goal of the temperature optimization

was to find the golden mean which would preserve PsbP in its native state and was suitable for NMR measurements.

3. 4. 4. His-tag

His tag is an artificial part of DNA information coding for 6 or 10 histidines followed by an amino acid sequence which can be selectively cleaved off with the tag-selective proteases. This stretch of DNA is added upstream to the sequence coding for the actual protein. Six histidines are therefore attached to the N-terminus of the protein. His tag is added to the protein to facilitate its purification with the use of Ni^{2+}, Co^{2+} or Cu^{2+} affinity chromatography. His-tag may be also introduced to label the protein for specific detection in immunoassays using anti-polyhistidine antibodies. Tags in general may increase solubility of the protein as well as its stability [36].

3. 4. 5. DTT

Sequence of PsbP contains one cysteine which may be responsible for the formation of dimers. The disulfide bridge is caused by the oxidation of thiol groups present in cysteine. Such bonds can be broken by the addition of reducing compound such as 2-mercaptoethanol or the more powerful dithiothreitol. Since the formation of such oligomers is undesired during the protein preparation and NMR experiments, DTT was added to prevent this process.

3. 5. Nuclear Magnetic Resonance

3. 5. 1. 2D NMR

3. 5. 1. 1. HSQC

A ^{1}H, ^{15}N − HSQC spectrum of a protein is the first verification of protein suitability for NMR measurements. Results tell us if the protein is properly folded and if it has been prepared with stable isotope labeling. This two-dimensional experiment gives correlation between amide protons and nitrogen atoms via one-bond J-coupling. The resulting correlation map gives resonances which should approximately correspond to the number of amino acids in the protein. This is not true for prolines which are not observable in this kind of spectrum. Nitrogen containing side chains of some amino acids like arginine, tryptophan or glutamine also give rise to additional signals observed in upper right (low frequency) part of the spectrum. There is also a ^{1}H, ^{13}C version of HSQC which shows correlations between aliphatic carbon and attached proton [37, 38].

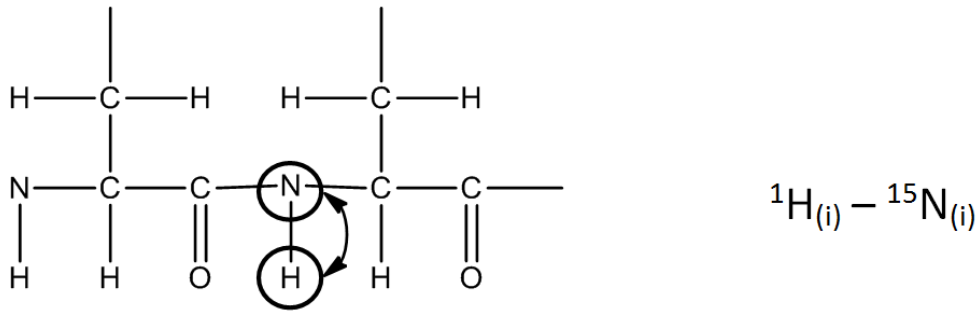


Figure 10: *Magnetization transfer and observed correlations in ^{1}H ,^{15}N-HSQC.*

3. 5. 2. 3D NMR

3. 5. 2. 1. HNCO

For every amino acid one signal is coming up in this kind of spectrum. That signals belongs to carbonyl ^{13}C from preceding amino acid. Transfer of magnetization goes from amide proton through ^{15}N (i) to ^{13}CO (i-1), with I indicating the position in the amino acid sequence. This information gives connection of the successive amino acids across the peptide bond. HNCO is used as a cross-check for the peak picking in HSQC since it has the highest sensitivity of all triple resonance experiments [37].

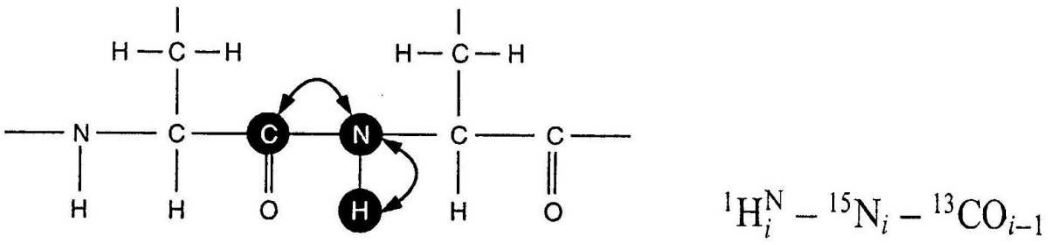


Figure 11: *Magnetization transfer and observed correlations in HNCO*[38].

3. 5. 2. 2. HNCA

The HNCA is one of the standard triple resonance experiments, which are used for backbone assignment. Magnetization is transferred from amide proton to ^{15}N and further to $^{13}C_\alpha$ of the same and the preceding amino acid. In the end the amide proton serves for the detection. In the 3D spectrum one obtains 2 correlation signals per NH pair – C_α(i) and C_α (i-1). This is due to the fact that the coupling constants from amide ^{15}N (i) to C_α(i) as well as to C_α (i-1) have similar magnitudes (~ 9 Hz and ~ 5 Hz, respectively) [37].

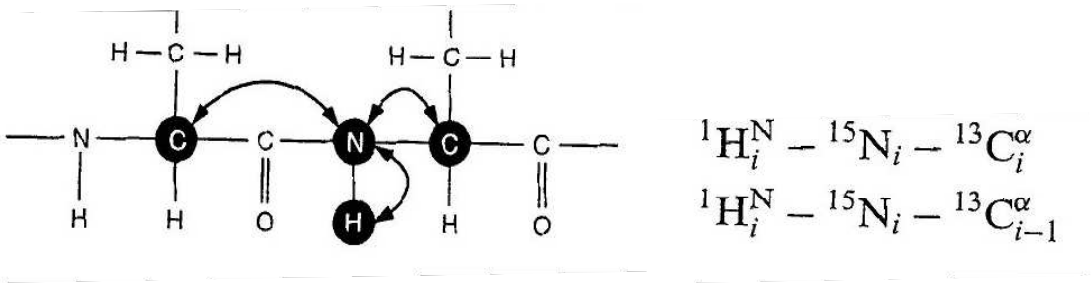


Figure 12: *Magnetization transfer and observed correlations in HNCA*[38].

3. 5. 2. 3. HNCACB

This experiment is one of the most useful designed for use in sequential protein resonance assignment. In one dimension, one obtains ^{15}N, in second ^{1}H and in the third one both $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ chemical shifts. Magnetization starts on the amide proton from where it goes via ^{15}N to $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ of the same and the preceding amino acid and back to the NH for the detection. In each resolved NH plane one can see in total four signals for each amino acid with the exception of glycine which does not possess C_{β}. The alpha carbons are usually obtained as negative peaks whereas the beta carbons are positive. Sometimes not all of the signals are observed, but generally the carbons of the ***i*** amino acid are always visible and having higher intensity than the carbons belonging to the preceding one [37, 39].

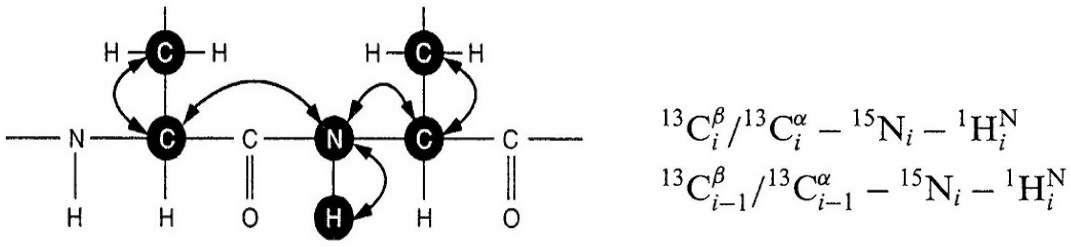


Figure 13: *Magnetization transfer and observed correlations in HNCACB* [38].

3. 5. 2. 4. CBCA(CO)NH

The CBCA(CO)NH spectrum is used to resolve ambiguities emerging from the HNCACB spectrum. This sequence gives only the correlations between ^{1}H-^{15}N (i) pair to $^{13}C_{\alpha}$ (i-1) and $^{13}C_{\beta}$ (i-1). Comparison of HNCACB and CBCA(CO)NH yields unique carbon shifts which are used for the sequential backbone assignment [37].

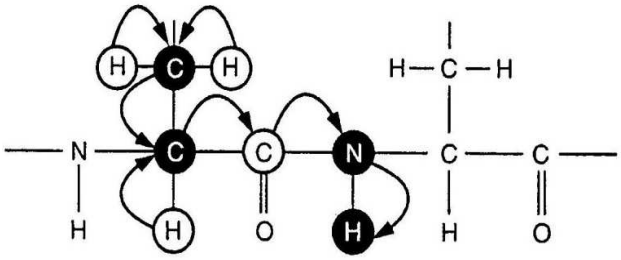


Figure 14: *Magnetization transfer and observed correlations in CBCA(CO)NH [38].*

3. 5. 2. 5. HN(CA)NNH and H(NCA)NNH

The triple resonance experiments, HN(CA)NNH and H(NCA)NNH can be used in sequential backbone assignment based only on 1H and 15N chemical shifts. The HN(CA)NNH experiment gives cross peaks between amide proton of one amino acid and the amide nitrogen of the preceding and successive amino acids. The H(NCA)NNH yield cross peaks between amide proton of amino acid i and amide protons of i+1 and i-1, respectively.

Polarization transfer starts from amide protons to directly attached ^{15}N, ^{13}C nuclei which relay the magnetization to amide nitrogen of the preceding and successive amino acid [40].

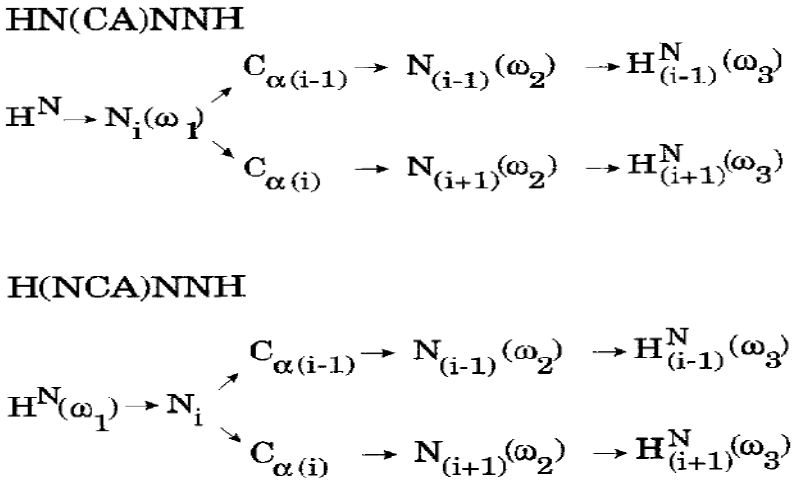


Figure 15: *Magnetization pathways in HN(CA)NNH and H(NCA)NNH experiment* [40].

Since the experiment is based only upon heteronuclear one or two-bonds J-couplings, it is independent of conformational changes thus facilitating fast backbone sequential assignment.

3. 5. 2. 6. NOESY-HSQC

The NOESY spectra are used to obtain distance constraints, which are necessary for the 3D structure calculation. Magnetization is exchanged from amide ^{1}H among all protons by NOE and is then transferred to ^{15}N from where it goes back to the hydrogens for detection. For each amide nitrogen chemical shift one obtains cross peaks coming from side chains of the selected and previous amino acid. Overall correlations through space are obtained. In this case the NOESY-HSQC was used as a validation of assignment correctness [39].

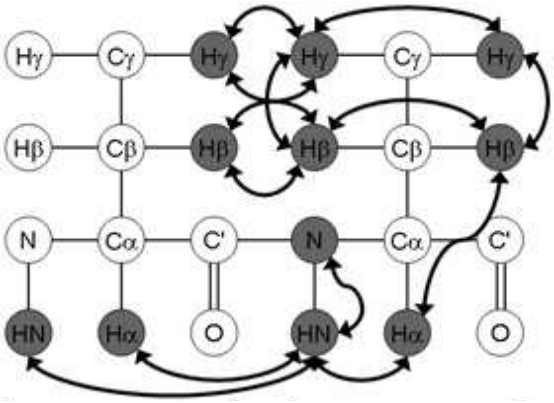


Figure 16: *Magnetization transfer pathway in NOESY-HSQC* [39].

3. 6. Resonance Assignment

Assignment of protein backbone resonances can be done using the software **CARA** ("Computer Aided Resonance Assignment" [41]) developed by Rochus Keller in the group of prof. Kurt Wütrich (ETH Zürich). All 3D spectra together with ^{15}N HSQC are loaded into the program which enables parallel mapping of two spectra.

First step is to pick peaks in 2D HSQC which would be further assigned to individual amino acids. The peak picking of each HN pair is done with the help of HNCO which has the highest sensitivity amongst the used 3D experiments. This is achieved using „Synchroscope mode", which displays discrete strips coming from the third plane for each peak in HSQC. HNCO presents the chemical shift of carbonyl carbon of the preceding amino acid. Every HN pair having a corresponding cross-peak in HNCO is assigned as a new system. In this way a group of systems waiting for further delineation is created. CBCANH or HNCACB spectra are of the next choice. They are used for the assignment of $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ resonances. Some of the amino acids have characteristic carbon chemical shifts and therefore they are easy to be recognized. These include glycine which does not possess any beta carbon or alanine with its $^{13}C_{\beta}$ shift being around 15 ppm. Other typical patterns appear with serine and threonine; both have carbon resonances shifted to 50-70 ppm region.

Sometimes it is not possible to distinguish between the carbon atoms of the given and preceding amino acid – a problem that is easily solved by looking into CBCA(CO)NH. This spectrum shows resonances coming only from the predecessor. Knowing that the alpha carbons have negative phase and the beta carbons are positive, assignment of carbon signals is

a relatively straightforward process. In the case of some ambiguities, HNCA might help to distinguish between alpha carbons of (i) and (i-1).

When the backbone chemical shifts and first side-chain carbons are assigned, one needs to link them sequentially. This is achieved in the so-called „Stripscope“ view. Strips with carbon resonances for every system are brought next to each other and by switching and comparison of the strips matching resonances are located. This step can be done manually by screening all available strips against the chosen one or with the help of CARA using the feature „Find All Predecessors/Successors“. By linking of the systems, it is possible to walk along the amino acid sequence and assign them to peaks in HSQC.

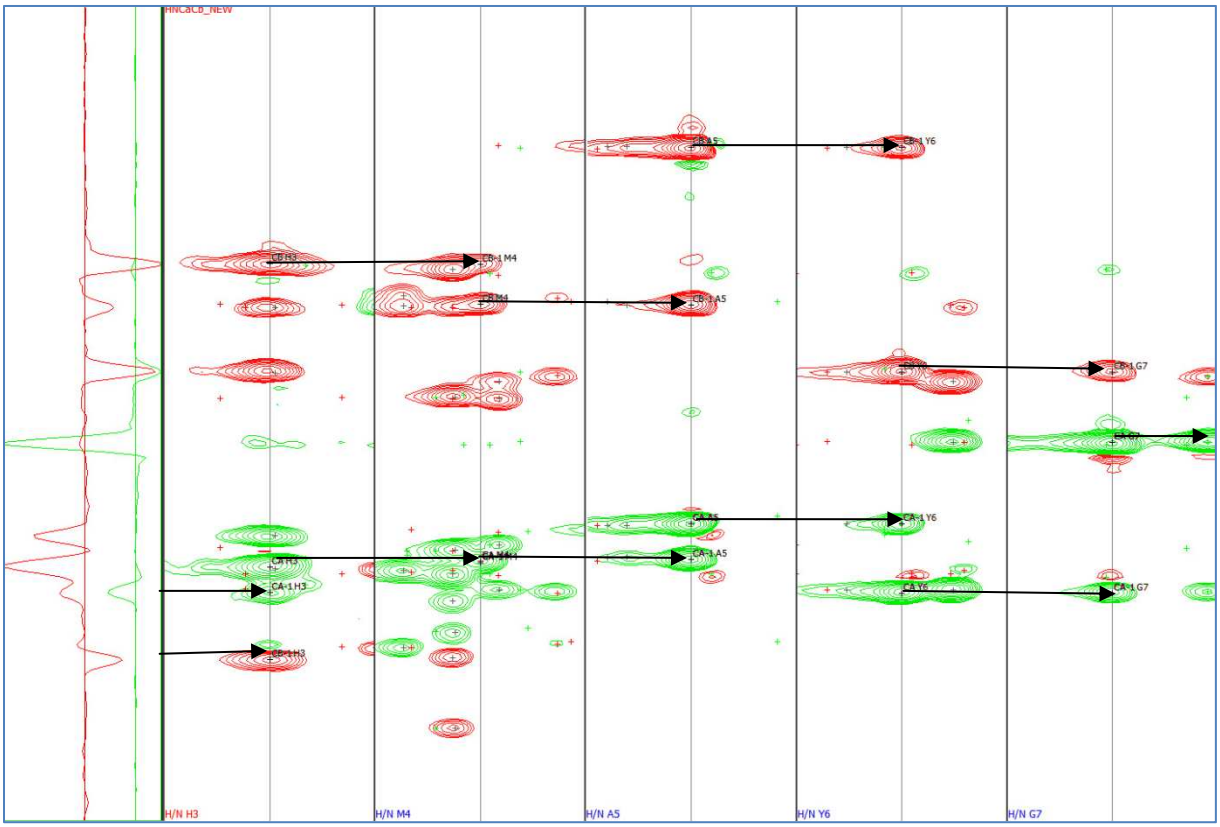


Figure 17: *Stripscope view in CARA with the HNCACB strips.*

Another approach makes use of HN(CA)NNH and H(NCA)NNH sequences. For each HN pair in HSQC, two cross-peaks appear in both spectra – one with amide chemical shifts and the other with amide proton shifts. These two peaks represent the preceding and successing amino acid relative to the one picked in HSQC. Using this method, one can link the amino acid sequentially only using the proton and nitrogen resonances.

As soon as all peaks in HSQC are assigned, side-chain assignment is started defining all atoms in the side chains. Next step is to obtain distance restraints between all assigned protons. These restraints are the basis for 3D solution structure calculation.

4. Experimental

4. 1. List of the used chemicals

Compound	Chemical formula	M [g/mol]	Purity	Supplier
Acetic acid	$C_5H_4O_2$	60.05	100%	Merck
Acrylamide/ Bis - acrylamide, 30% solution	_____	_____	___	Sigma - Aldrich
AEBSF	$C_8H_{10}FNO_2S \cdot HCl$	239.69	> 98%	neoLab
Ammonium persulfate	$(NH_4)_2S_2O_8$	228.20	> 98%	Sigma - Aldrich
Ammonium sulfate	$(NH_4)_2SO_4$	132.14	> 99%	Sigma - Aldrich
Ammonium sulfate 15N	$(^{15}NH_4)_2SO_4$	134.14	> 98%	Sigma - Aldrich
Bacteriological Agar	_____	_____	_____	Merck
Bis Tris	$C_8H_{19}NO_5$	209.24	> 98%	Sigma - Aldrich
Bromophenol blue sodium salt	$C_{19}H_9Br_4NaO_5S$	691.94	_____	ICN Biomedicals
Casein hydrolysate	_____	_____	_____	USB
Coomassie Brilliant Blue R 250	$C_{45}H_{44}N_3NaO_7S_2$	826.00	_____	Serva
Dipotassium hydrogen phosphate	K_2HPO_4	174.18	_____	Merck
Disodium hydrogen phosphate dihydrate	$Na_2HPO_4.2H_2O$	177.99	_____	Merck
Dithiothreitol	$C_4H_{10}O_2S_2$	154.25	_____	USB
Ethylenediaminetetraacetic acid disodium salt dihydrate	$C_{10}H_{14}N_2Na_2O_8.2H_2O$	372.24	> 99%	Sigma -Aldrich
Glucose D 13C	$C_6H_{12}O_6$	186.16	> 99%	Spectra Gases
Glucose D(+)	$C_6H_{12}O_6$	180.16	_____	Merck
Glycerol for molecular biology	$C_3H_8O_3$	92.09	> 99%	Sigma - Aldrich
Glycine	$C_2H_5NO_2$	75.07	>99%	Sigma – Aldrich

Compound	Chemical formula	M [g/mol]	Purity	Supplier
IsoproIsopropyl β-D-1-thiogalactopyranosidepyl	$C_9H_{18}O_5S$	238.30	>99%	Applichem
Magnesium sulfate heptahydrate	$MgSO_4.7H_2O$	246.48	_____	Merck
Potassium dihydrogen phosphate	KH_2PO_4	136.09	>99%	J. T. Baker
Protease inhibitor cocktail (EDTA-free)	_____	_____	_____	Roche Diagnostics
Sodium dihydrogen phosphate	NaH_2PO_4	119.98	_____	Fluka
Sodium dodecyl sulphate	$C_{12}H_{25}NaO_4S$	288.38	>99%	Fluka
Thrombin (human plasma)	_____	37 400	____	Sigma-Aldrich

4. 2. List of the used instruments

Instrument	Type	Manufacturer
Autoclave	CertoClav EL	CertoClav
Centrifugation filter for protein concentration	Amicon Ultracel 10K	Millipore
Centrifuge	Biofuge pico	Heraeus
Centrifuge	Biofuge stratos	Heraeus
Centrifuge	Megafuge 1.0 R	Heraeus
Electrophoresis	EasyPhor PAGE mini	Biozym
Fast purification liquid chromatography	GradiFrac system	Pharmacia Biotech
Filtration	Syringe filter 0.22 µm	GVS
Chromatographic resin	Ni^{2+} Sepharose	GE Healthcare
Chromatographic resin	SP Sepharose	GE Healthcare
Chromatographic resin	HighLoadSuperdex 75	GE Healthcare
Incubator	Shaking incubator 3033	GFL
NMR spectrometer	Avance III 700 MHz	Bruker

Instrument	Type	Manufacturer
Shaker	KS 250 Basic	IKA
Sonicator	Sonoplus HD 70	Bandelin electronic
Sterilization	Steritop Filter 0.22 µm	Millipore
Ultrasonication bath	S 10	Elmasonic
UV-VIS spectrometer	Uvikon 800	Kontron
UV-VIS spectrometer	HP 8453	Hewlett-Packard

4. 3. List of the solutions

Expression
LB medium
5 g of yeast extract, 10 g of NaCl, 10 g of casein hydrolysate were dissolved in 1 L of 18 mΩ H_2O and sterilized in autoclave at 125°C for 10 minutes. When cooled to ca 40 °C, 1 ml of Kanamycin stock solution was added.
Agar medium for Petri dishes
5 g of yeast extract, 10 g of NaCl, 10 g of casein hydrolysate and 20 g of bacteriological agar were dissolved in 1 L of 18 mΩ H_2O and sterilized in autoclave at 125°C for 10 minutes. When cooled to ca 40 °C, 1 ml of Kanamycin stock solution was added.
Minimal medium (M9)
6 g of Na_2HPO_4 . $2H_2O$, 3 g of KH_2PO_4, 1.5 g of $(NH_4)_2SO_4$, 0.5 g NaCl and 2 g of glucose were dissolved in 1 L of 18 mΩ H_2O and sterilized using Milipore filter. 1 ml of Growth elements stock solution and 1 ml of MgSO4 solution were added together with 1 ml of Kanamycin stock solution. For ^{15}N labelling, $(^{15}NH_4)_2SO_4$ was used. For double labeled sample, both $(^{15}NH_4)_2SO_4$ and ^{13}C-glucose were used.
1M $MgSO_4$
12.04 g of MgSO4 was dissolved in 100 ml of distilled water and sterilized by autoclaving at 125°C for 10 minutes. 1 µl/ml was used to have 1 mM final concentration.
10x Trace elements solution
5 g of EDTA, 0.83 g of $FeCl_3$, 0.084 $ZnCl_2$, 0.013 g of $CuCl_2.2H_2O$, 0.01 g of $CoCl_2.6H_2O$, 0.01 g H_3BO_3 and 0.0002 g of $MnCl_2.4H_2O$ was dissolved in 100 ml of distilled water and autoclaved at 125°C for 10 min.

Kanamycin stock solution (30 mg/ml)
30 mg of kanamycin were dissolved in 1 ml of 18 mΩ H_2O and kept at -20°C. 1 µl of the stock was used per 1 ml of the medium to have the final concentration of 30 µg/ml.
1M IPTG stock solution
238 mg of IPTG were dissolved in 1 ml of 18 mΩ H_2O and kept at -20°C. 1 µl/ml was used to be 1 mM in final solution.
100 mM AEBSF stock solution
240 mg of AEBSF were dissolved in 10 ml of 18 mΩ H_2O and kept at -20°C. 10 µl/ml was used in solutions to be 1 mM.

Purification
Phosphate buffer 1 (20 mM KH_2PO_4, 500 mM NaCl, 20 mM imidazole)
2.70 g of KH_2PO_4, 29.22 g of NaCl and 1.36 g of imidazole were dissolved in 1 l of distilled water.
Phosphate buffer 2 (20 mM K_2HPO_4, 500 mM NaCl, 20 mM imidazole)
4.56 g of K_2HPO_4, 29.22 g of NaCl and 1.36 g of imidazole were dissolved in 1 l of distilled water.
AFC Binding buffer (20 mM K-phosphate, 500 mM NaCl, 20 mM imidazole, pH 7.4)
Phosphate buffers 1 and 2 were mixed and pH was additionally adjusted using concentrated HCl to 7.4. Solution was sterilized through Millipore filter and kept at 4°C.
AFC Elution buffer (20 mM K-phosphate, 500 mM NaCl, 500 mM imidazole, pH 7.4)
32.68 g of imidazole were added to 1 l of AFC binding buffer and pH was adjusted with concentrated HCl to 7.4. Solution was sterilized through Millipore filter and kept at 4°C.
IEC Binding buffer (20 mM BisTris, 1mM EDTA, pH 6.0)
4.18 g of BisTris, 0.37 g of EDTA were dissolved in 1 l of distilled water and pH was adjusted to 6.0. Buffer was sterilized using Millipore filter and kept at 4°C.
IEC Elution buffer (20 mM BisTris, 1 mM EDTA, 1M NaCl, pH 6.0)
4.18 g of BisTris, 0.37 g of EDTA and 58.44 g of NaCl were dissolved in 1 l of 18 mΩ H_2O and pH was adjusted to 6.0. Solution was sterilized via Millipore filter and kept at 4°C.
SEC buffer (20 mM BisTris, 1 mM EDTA, 200 mM NaCl, pH 6.0)
4.18 g of BisTris, 0.37 g of EDTA and 11.68 g of NaCl were dissolved in 1 l of distilled water and pH was adjusted to 6.0. Solution was sterilized via Millipore filter and kept at 4°C.
2.5 M NaN_3 stock solution
1.6 g of NaN3 were dissolved in 10 ml of 18 mΩ H_2O and kept at 4°C. 20 µl/l of solution was used to achieve final concentration of 0.05 mM.

Polyacrylamide gel electrophoresis
10 % Ammonium persulphate solution
1 g of ammonium persulphate was dissolved in 10 ml of 18 mΩ H_2O. Solution was kept at 4°C.
5x running buffer for SDS-PAGE
15g of Tris, 75 g of glycine and 5 g of sodium dodecyl sulphate were dissolved in 1 l of distilled water. For SDS-PAGE, the buffer was diluted to 1x in distilled water.
2 M Tris solution (pH 8.9)
60.57 g of Tris were dissolved in 250 ml of distilled water and pH was adjusted to 8.9. Solution was kept at 4°C.
1 M Tris solution (pH 6.8)
30.29 g of Tris were dissolved in 250 ml of distilled water and pH was adjusted to 6.8. Solution was kept at 4°C.
Resolving buffer for SDS-PAGE
75 ml of 2 M Tris solution (pH 8.9) were mixed with 4 ml of 10% SDS and 21 ml of distilled water.
Stacking buffer for SDS-PAGE
50 ml of 1 M Tris solution (pH 6.8) were mixed with 4 ml of 10% SDS and 46 ml of distilled water.
10% SDS
1 g of SDS was dissolved in 10 ml of distilled water and kept at 4°C.
1% Bromophenol blue
50 mg of Bromophenol blue were dissolved in 5 ml of distilled water and kept at 4°C.
2.5 M DTT
0.38 g of DTT were dissolved in 1 ml of 18 mΩ H_2O and kept at -20°C.
5x Loading buffer for SDS-PAGE
0.6 ml of 1 M Tris (pH 6.8), 5 ml of 50% glycerol, 2 ml of 10% SDS, 1 ml of 1% Bromophenol blue and 0.5 ml of DTT were mixed with 0.9 ml of distilled water.
1x running buffer for native PAGE
6 g of Tris and 26.6 g of Glycine were dissolved in 1 l of distilled water, pH was adjusted to 8.9.
1.5 M Tris solution (pH 8.9)
45.43 g of Tris were dissolved in 250 ml of distilled water and pH was adjusted to 8.9. Solution was kept at 4°C.

0.5 M Tris solution (pH 6.8)
15.14 g of Tris were dissolved in 250 ml of 18 mΩ H_2O and pH was adjusted to 6.8. Solution was kept at 4°C.
2.5x Loading buffer for native PAGE
2.5 ml of 80% glycerol, 1.35 ml of 18 mΩ H_2O 1.07 ml of 0.5 M Tris solution (pH 6.8) were mixed and traces of Bromo Phenol Blue were added. Solution was kept at 4°C.
Staining solution
1 g of Coomassie Brilliant Blue was dissolved in 450 ml of methanol and mixed with 450 ml of distilled water and 100 ml of concentrated acetic acid.
Destaining solution
100 ml of acetic acid were mixed with 800 ml of distilled water and 100 ml of methanol.

Various
TE buffer (10 mM Tris, 1 mM EDTA, pH 7.0)
0.12 g of Tris and 0.037 g of EDTA were dissolved in 100 ml of distilled water and pH was adjusted to 7.0. Solution was kept at 4°C.
7x protease inhibitors
1 tablet of Protease MINI inhibitors cocktail (EDTA free) was dissolved in 1.5 ml of 18 mΩ H_2O and kept at -20°C.
11 mM DSS
2.5 mg of labeled DSS was dissolved in 1 ml of 18 mΩ H_2O and kept at -20°C.

4. 4. Analysis

4. 4. 1. SDS-PAGE

A vertical apparature for 2 parallel gels (Biozym®) was employed. Standard Laemmli procedure [44] for the preparation of stacking and resolving gel were applied (see Tab. 1). The mixture for resolving gel was poured immediately after mixing in between two glass plates and overlaid with ethanol. When the polymerization was complete (ca. 30 minutes) ethanol was poured off. Mixture for stacking gel was added on top of the resolving gel and a comb was inserted. When the gels were fully polymerized, they were transferred to the electrophoretic chamber, which was filled with 1x running buffer. Prior to the loading of the samples, all wells were washed with running buffer in order to flush out the remains of polymerized acrylamide.

20 µl of the sample were mixed with 5x loading buffer containing SDS and DTT and loaded in the well. 5 µl of the Precision Plus Protein All Blue Standards (Bio-rad) were used as marker in each gel. All experiments were run at constant voltage of 200 V for 50 - 60 min. The electrophoresis was finished when bromophenol blue dye reached the bottom of the separating gel. The gel was stained for 1 hour by gentle shaking in the staining solution at room temperature. Afterwards it was changed into the destaining solution. Destaining of the gel took at least 2 hours with iterative solution change.

Solution	**15% resolving gel**	**4.5% stacking gel**
30 % Acrylamide/Bis-acrylamide	2.5 ml	3.25 ml
H_2O	1.25 ml	1.87 ml
Resolving buffer	1.25 ml	_____
Stacking buffer	_____	1.87 ml
TEMED	5 µl	7.5 µl
10% APS	50 µl	75 µl

Tab. 1: *Protocol for the preparation of electrophoresis gels.*

4. 4. 2. Native PAGE

Standard Laemmli protocol for neutral and acidic proteins was applied. The pI value of His-PsbP 7.1, pI of PsbP is 6.19. The apparature was set up in the same way like in the case of SDS-PAGE. First, the resolving gel was prepared in between the glass plates. When polymerized, it was overlaid with stacking gel mixture. The resolving and stacking gels did not contain any SDS, which would reduce the proteins. Protein samples were mixed with loading buffer, which was free of SDS and DTT as well in order to preserve possible disulfide bonds. The buffer was a mixture of Tris (pH 6.8), glycerol and bromophenol blue traces. Purpose of this buffer was to monitor the course of electrophoresis and to ensure proper embedding of the protein mixtures in the wells. No protein marker was used in this type of electrophoresis, because it contains DTT which could diffuse through the polyacrylamide gel and skew the results. 20 µl of the protein samples were applied in individual wells and the electrophoresis was started by application of constant voltage. Running conditions comprised of 150 V and various time spans ranging from 120 – 160 minutes.

Solution	**Resolving gel**			**Stacking gel**
	15%	**12%**	**10%**	**4.5%**
30% Acrylamide/Bis-acrylamide	4.6 ml	3.8 ml	3.1 ml	0.5 ml
H_2O	2.1 ml	3.1 ml	4.1 ml	3.2 ml
Resolving buffer (1.5 M Tris, pH 8.9)	2.3 ml	2.3 ml	2.3 ml	_____
Stacking buffer (0.5 M Tris, pH 6.8)	_____	_____	_____	1.25 ml
10% APS	33.3 µl	33.3 µl	33.3 µl	50 µl
TEMED	7.7 µl	7.7 µl	7.7 µl	5 µl

Tab. 2*: Protocol for the preparation of 1 native gel.*

4. 4. 3. Concentration measurement

For a measurement, HP 8453 UV/VIS spectrometer was used. The 1 cm quartz cuvette was filled with 600 µl of degassed buffer (20 mM Bis Tris, 1 mM EDTA, 0.05 mM NaN3,pH 6.0) and the absorbance of the blank sample at 280 nm was recorded. Subsequently, 20 µl of the protein sample were pipetted into the cuvette and mixed properly. Eventual air bubbles were destroyed with the help of a plastic tip. Absorbance of the sample at 280 nm was measured. The actual concentration was calculated in the following manner:

$$A_{280} = \varepsilon \cdot c \cdot l$$

Parameter	Value
molar $\varepsilon_{His\text{-}PsbP}$	1.018
molar ε_{PsbP}	1.035
dilution factor	31
optical path length, l	1 cm

Tab. 3: *Parameters used for the calculation of protein concentration*

4. 5. Expression

4. 5. 1. Expression in vivo

E. coli cells (BL21DE3) with transformed JR3133 plasmid were used for the overexpression of His-PsbP. This construct was kindly provided by prof. Rüdiger Ettrich, Institute of Nanobiology and Structural Biology, Czech Academy of Sciences. Cells were kept as stock solution under 80% glycerol at -20°C.

MGSSHHHHHHSSGLVPRGSHMAYGEAANVFGKPKKNTEFMPYNGDGFKLL **33**

VPSKWNPSKEKEFPGQVLRYEDNFDATSNLSVLVQPTDKKSITDFGSPEDFLS **86**

QVDYLLGKQAYFGKTDSEGGFDSGVVASANVLESSTPVVDGKQYYSITVLTR **138**

TADGDEGGKHQVIAATVKDGKLYICKAQAGDKRWFKGAKKFVESATSSFSVA **190**

Figure 18: *Amino acid sequence of His-PsbP. Tag sequence is highlighted in green with blue thrombin cleavage site. The enzyme will cleave between arginine (R) and glycine (G).*

Petri dishes were filled with LB-agar medium which contained 30 mg/l of kanamycin. Stock solution of transformed *E. coli* cells was diluted in series $1{:}10^6$. Agar plates were inoculated with 200 µl of this diluted cell suspension. Cells were grown for 20 hours at 37°C in an incubator.

One colony from the plate was transferred with a loop in 25 ml of LB buffer (30 mg/l of Kanamycin) and incubated overnight at 37°C. The originated cell culture was diluted 1:100 in 1 l of LB medium (30 mg/l of Kanamycin) for the production of unlabeled protein or in 1 l of M9 medium (30 mg/l of Kanamycin) for the production of single/double labeled protein. Cells were kept at 37°C and optical density at 580 nm was measured every hour to monitor the growth of bacteria. When $OD_{580} > 0.6$, the expression of His-PsbP was initiated by addition of IPTG to 1 mM concentration. Incubation temperature was lowered to 28°C and the cells were grown for 20 more hours.

The cell suspension was afterwards centrifuged (4000 g, 4°C, 25 min.) and the supernatant was decanted. Pellets were resuspended on ice with 50 ml of phosphate buffer (20 mM KH_2PO_4, 500 mM NaCl, 20 mM imidazole, 1 mM AEBSF, pH 7.4) and thoroughly sonicated for 30 minutes. Mixture was centrifuged (4000 g, 4°C, 60 min.) and the supernatant was filtered through GVS filter (pore size: 1.2 µm). The pellet was again resuspended in a small volume of the buffer and sonicated for the second time. When centrifuged, the supernatant was filtered as well. Part of the lysate was immediately used for affinity chromatography and the rest was deep frozen in liquid nitrogen and kept at -20°C.

4. 5. 2. Expression in vitro

This experimental part was carried out within the scope of „Workshop on cell-free protein synthesis for NMR“ hosted by the Swedish NMR Centre at the University of Gothenburg.

Plasmid JR3133 was isolated from BL21DE3 *E. coli* cells using Qiagen Plasmid MIDI Prep according to the manufacturer's instructions. Its absorbance was measured at HP 8453 UV/VIS spectrometer at 260 nm. Concentration was determined using the following relationship:

$$\text{Concentration } (\mu g / ml) = \text{Abs}_{260} * 50 * \text{dilution factor}$$

Isolated plasmid DNA was mixed with other reagents (see Tab. 4) and incubated for 2 hours at 30°C and 750 rpm. After 30 minutes of incubation, 1 ml of 0.315 M creatine phosphate was added. The NiNTA resin was equilibrated with phosphate buffer (50 mM Na_2HPO_4, 300 mM NaCl, 20 mM imidazole, pH 7) for 1 hour prior to use. Expression reaction was spinned down at 4500 g for 15 min. 15 µl of the supernatant were saved for SDS-PAGE and the rest was mixed with equal amount of Tris buffer (50 mM Tris-HCl, pH 8) and incubated with NiNTA resin at room temperature for 1 hour. A disposable column was loaded with NiNTA slurry. When the resin settled, it was washed with 10 V of phosphate buffer (50 mM Na_2HPO_4, 300

mM NaCl, 20 mM imidazole, pH 7) and the His-PsbP was eluted with 3 x 800 µl of phosphate – imidazole buffer (50 mM Na_2HPO_4, 300 mM NaCl, 300 mM imidazole, pH 7). Eluates were checked using SDS-PAGE.

Reagent	Amount	Final concentration
Plasmid DNA	100 µg	10 µg/ml
Magnesium acetate	300 µl	10.5 mM
Low Molecular Weight Mix	2 ml	1x
Amino Acid Mix 1*: ADGKV	250 µl	0.5 mM
Amino Acid Mix 2*: ELFPT	175 µl	0.35 mM
Amino Acid Mix 3*:RNHY	80 µl	0.2 mM
Amino Acid Mix 4*: IMWC	40 µl	0.1 mM
Serine	200 µl	2 mM
Glutamine	400 µl	4 mM
T7 RNA Polymerase	100 µl	100 µg/ml
Creatine Kinase	125 µl	125 µg/ml
E. coli S12 Extract	3100 µl	31 % v/v
18mΩ water	fill up to 10 ml	_____

Tab. 4: *Reaction setup for the protein synthesis. * ^{15}N, ^{13}C – labeled compounds.*

4. 6. Purification

Fast performance liquid chromatography (FPLC) system from Pharmacia Biotech with UV detection was used for all purification steps. In order to prevent microbial growth, all buffers contained 0.05 mM sodium azide. The chromatography was run at 4°C.

4. 6. 1. Affinity chromatography

The column filled with Ni-Sepharose™ High Performance (GE Healthcare) was charged with Ni-ions prior to use according to the scheme given below. Such charged column was capable of up to 5 purification runs.

Step	Solution	Volume
1.	H_2O	5 V
2.	0.1 M $NiSO_4$	0.5 V
3.	H_2O	5 V
4.	Binding buffer	5 V
5.	H_2O / 20% ethanol	5 V

Tab. 5: *Protocol for charging of the resin („V“ stands for the volume of the column).*

Column was flushed with deionized water and afterwards equilibrated with binding buffer. The sample was loaded on a column at lowered flow rate of 1 ml/min. Afterwards the flow rate was increased again to 2 ml/min. All unbound material went through the column as a flow-through fraction. When the baseline was established, a gradient of elution buffer was applied. The protein was eluted until the gradient was finished and column was cleaned with successive washing steps.

Running conditions	
Maximal flow rate	2 ml/min
Sample volume	up to 10 ml
Binding buffer	20 mM K_2HPO_4, 500 mM NaCl, 20 mM imidazole, pH 7.4
Elution buffer	20 mM K_2HPO_4, 500 mM NaCl, 500 mM imidazole, pH 7.4
Gradient	0 – 70 % elution buffer
Gradient rate	2% per minute
Fractions	2 ml
Regeneration of the column	
1M NaCl	5 V
deionized H_2O	5 V
Binding buffer	5 V
Elution buffer	5 V
Binding buffer	5 V
20% ethanol	5 V

Tab. 6: *Setup of Ni^{2+} affinity chromatography.*

The elution fractions were analyzed via SDS-PAGE and the ones containing His-PsbP were pooled and kept at 4°C. The solution was concentrated using Amicon Ultra with the 10 kDa cut-off. Concentrated protein solution was swapped into the binding buffer for ion-exchange chromatography which followed hereafter.

4. 6. 2. Ion-exchange chromatography

The column was packed with SP Sepharose^{TM} Fast Flow (GE Healthcare) which serves as a cation exchanger. Column was properly washed with deionized water and equilibrated with binding buffer at standard flow rate. Before loading of the sample on the column, the flow rate was lowered to 1 ml/min in order to ensure proper deposition of the sample solution. The initial rate was re-established when the whole sample was applied on the column. Proteins with positive charge did not stick to the purification column and were washed out. When the

baseline was stable again, a gradient of elution buffer was initiated. The His-PsbP protein started to elute with increasing salt concentration. Fractions suspected to contain the pure protein were checked with SDS-PAGE. Positive eluates were pooled and concentrated as described in previous step. Protein was swapped into cleavage buffer (20 mM Bis-Tris, 100 mM NaCl, pH 6.0). The column was thoroughly washed and kept in 20% ethanol solution.

Running conditions	
Maximal flow rate	2 ml/min
Sample volume	up to 8 ml
Binding buffer	20 mM BisTris, 1mM EDTA, pH 6.0
Elution buffer	20 mM BisTris, 1mM EDTA, 1M NaCl, pH 6.0
Gradient	0 – 70% elution buffer
Gradient rate	2% per minute
Fractions	2 ml
Regeneration of the column	
1M NaCl	5 V
deionized H_2O	5 V
Binding buffer	5 V
Elution buffer	5 V
Binding buffer	5 V
20% ethanol	5 V

Tab. 7: *Setup of ion-exchange purification.*

4. 6. 3. Cleavage of the His-tag

Concentration of the His-PsbP was determined via UV-absorption measurement. For 1 mg of the protein, 1.5 units of thrombin from human plasma (Sigma Aldrich) were used. Desired amount of thrombin were added to protein solution and kept at room temperature under stirring for 3 hours. Reaction was stopped by addition of protease inhibitors (AEBSF) in the ratio of 10 µl per ml of reaction mixture. Time-course of the cleavage was monitored with

SDS-PAGE. Reaction was changed into the running buffer for Size-exclusion chromatography.

4. 6. 4. Size-exclusion chromatography

Column packed with Superdex 75 Prep Grade (GE Healthcare) was washed overnight with deionized water and equilibrated with running buffer.

Running conditions	
Maximal flow rate	0.3 ml/min
Sample volume	up to 5 ml
Running buffer	20 mM Bis-Tris, 200 mM NaCl, pH
Regeneration of the column	
1M NaOH	5 V
deionized H_2O	5 V
20% ethanol	5 V

Tab. 8: *Setup of size-exclusion chromatography.*

The purification column was thoroughly washed with deionized water and equilibrated with running buffer. Afterwards the protein solution was loaded onto the column. Big compounds were not retained in the purification matrix and eluted as first. PsbP fractions were collected and after verification by SDS-PAGE they were pooled and concentrated with Amicon Ultra (10 kDa cut-off) to 1-2 ml. The buffer was swapped to 20 mM Bis Tris, 1 mM EDTA, 0.05 mM NaN3, pH 6.0. The final concentration was determined by UV absorbance measurement. Such prepared protein was stored at -20°C for at least 6 months.

4. 7. Stability Tests

4. 7. 1. Buffers

The protein was diluted to 50 µM concentration in 1 ml of the buffer and kept shaking in an eppendorf tube at room temperature. Every day 15 µl of the sample were taken and kept for the final analysis by polyacrylamide gel electrophoresis.

Buffer	Time	Temperature
20 mM Na_2HPO_4, 1 mM EDTA, 0.05 mM NaN_3, pH 7.0	10 days	25°C
20 mM Bis Tris, 1 mM EDTA, 0.05 mM NaN_3, pH 6.0	7 days	25°C

Tab. 9: *Setup of stability experiment with two buffers.*

4. 7. 2. Temperature

PsbP was diluted to be 50 µM in 1 ml of buffer. The sample was kept for 7 days at different temperatures in order to find a suitable compromise between proper and stable folding of the protein and its eligibility to give good signal in NMR spectra. Every day, a sample for later SDS-PAGE analysis was taken.

Temperature	Time
22 °C	7 days
25°C	7 days
30°C	7 days
35°C	7 days
40°C	7 days

Tab. 10: *Temperature range applied in the experiments.*

4. 7. 3. Protease Inhibitors

7x stock solution of Complete Protease MINI Inhibitors cocktail (Roche Applied Sciences) were added to 1 ml of 50 µM protein in a buffer and kept for 7 days at 25°C. 15 µl samples were taken each day and analyzed by SDS-PAGE after the experiment has finished.

4. 7. 4. DTT and His-tag

2.5 M DTT was diluted to be 0.5 mM in His-PsbP and PsbP sample, respectively. The mixture was incubated both, at 25°C and 30°C for 7 days. Aliquot of 15 µl was taken form each sample and saved for SDS-PAGE analysis.

4. 8. Nuclear Magnetic Resonance

All NMR experiments were recorded at 700 MHz Bruker Ascend spectrometer equipped with TCI cryoprobe at The NMR Research Centre funded by European Union through EFRE INTERREG IV ETC-AT-CZ programme (Project M00146 „RERI-uasb“).

The spectra were processed using Topspin 3.1 (Bruker) and loaded into CARA. 2D spectra were Fourier transformed using command *xfb*. The F1, F2 and F3 datasets in 3D experiments were individually Fourier transformed using commands *tf1*, *tf2* and *tf3*, respectively. Automatic baseline correction was done in all directions separately with the commands *tabs1*, *tabs2* and *tabs3*. For a better illustration, typical processing parameters for 2D 15N HSQC and 3D HNCACB are given below (Tab. 11).

^{15}N HSQC			
Frequency axis	F2	F1	
Nucleus	^{1}H	^{15}N	
Size of real spectrum	2048	512	
Spectrometer frequency [MHz]	700.33	70.96	
Offset [ppm]	11.44	134.00	
Spectral resolution [Hz]	2.02	7.07	
Number of raw data points used by ft	1024	0	
HNCACB			
Frequency axis	F1	F2	F3
Nucleus	^{13}C	^{15}N	^{1}H
Size of real spectrum	256	128	2048
Spectrometer frequency [MHz]	176.10	70.96	700.33
Offset [ppm]	75.39	134.00	11.66
Spectral resolution [Hz]	49.53	28.28	4.78
Number of raw data points used by ft	1024	0	0

Tab. 11: *Typical processing parameters used for 2D and 3D spectra.*

4. 8. 1. 2D NMR

All samples were measured in 5 mm Shigemi tubes. 200 µl of the protein sample were mixed with 20 µl of D_2O and 1 µl of labeled DSS. Additionally, 30 µl of protease inhibitors stock solution and 1 µl of 2.5 M DTT were added to increase stability of the protein in long-time experiments. Final protein concentration ranged from 0.25 mM to 0.5 mM.

^{15}N HSQC was recorded to obtain ^{1}H – ^{15}N correlation map and was used as a cross-check of protein stability. This experiment was recorded at 20°C. The following parameters were employed for the measurement.

^{15}N HSQC		
Frequency axis	F2	F1
Nucleus	^{1}H	^{15}N
SW	13.99	80.00
TD	2048	256
DS	4	
NS	8	
Pulse program	hsqcetf3gpsi2	
^{13}C HSQC		
Frequency axis	F2	F1
Nucleus	^{1}H	^{13}C
SW	13.32	86.00
TD	1024	256
DS	2	
NS	4	
Pulse program	hsqcetgp	

Tab. 12: *List of the used parameters for 2D NMR.*

4. 8. 2. 3D NMR

Set of HNCO, HNCA, HNCACB and CBCACONH were recorded first with short 15N HSQC in between every experiment to check protein state. The spectra were recorded at 20°C with the same sample preparation like in the case of 2D experiments. Protein concentration of 0.5 mM was used for the measurement. The following parameters were employed for the measurements.

HNCO			
Frequency axis	F1	F2	F3
Nucleus	^{13}C	^{15}N	^{1}H
SW	22.08	51.00	13.99
TD	80	80	2048
DS	8		
NS	2		
Pulse program	hncogpwg3d		
HNCA			
Frequency axis	F1	F2	F3
Nucleus	^{13}C	^{15}N	^{1}H
SW	32.06	51.00	13.99
TD	128	80	2048
DS	32		
NS	8		
Pulse program	hncagpwg3d		
HNCACB			
Frequency axis	F1	F2	F3
Nucleus	^{13}C	^{15}N	^{1}H
SW	72.00	51.00	13.99
TD	132	80	2048
DS	32		
NS	32		
Pulse program	hncacbgpwg3d		
CBCA(CO)NH			
Frequency axis	F1	F2	F3
Nucleus	^{13}C	^{15}N	^{1}H
SW	72.00	51.00	13.99
TD	132	80	2048
DS	32		
NS	16		
Pulse program	cbcaconhgpwg3d		

Tab. 13: *The acquisition parameters for 3D HNCO, HNCA, HNCACB and CBCA(CO)NH experiments.*

HN(CA)NNH			
Frequency axis	F1	F2	F3
Nucleus	^{15}N	^{15}N	^{1}H
SW	51.00	51.00	13.99
TD	104	56	2048
DS	32		
NS	32		
Pulse program	hncannhgpwg3d		
H(NCA)NNH			
Frequency axis	F1	F2	F3
Nucleus	^{1}H	^{15}N	^{1}H
SW	12.00	51.00	12.02
TD	128	56	2048
DS	32		
NS	32		
Pulse program	hncannhgpwg3d.2		
^{15}N NOESY - HSQC			
Frequency axis	F1	F2	F3
Nucleus	^{1}H	^{15}N	^{1}H
SW	13.99	51.00	13.99
TD	140	64	2048
DS	32		
NS	8		
Pulse program	noesyhsqcf3gpsi3d		

Tab. 14: *The acquisition parameters for 3D HN(CA)NNH, H(NCA)NNH and NOESY-HSQC experiments.*

5. Results and Discussion

5. 1. Expression

PsbP was expressed in *E. coli* cells without any problems yielding sufficient amounts of unlabeled, singly (^{15}N) and doubly (^{15}N, ^{13}C) labeled protein. Fastest growth of bacteria was observed in LB medium in the case of unlabeled protein, however the biggest OD_{580} was measured for single labeled PsbP in minimal medium M9. This phenomenon was not correlated with the higher protein production, because the bacteria were probably already on the interface of stationary and death phase.

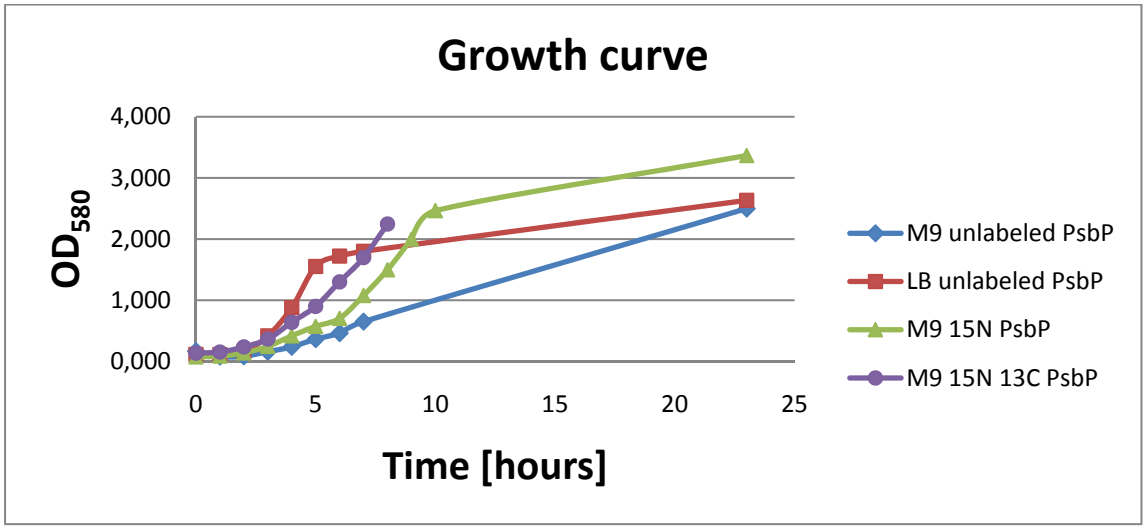


Figure 19: *Growth curve of BL21DE3 expressing His-PsbP in different growth media*

It was proven that single run of sonication is not sufficient to lyse all the cells. Only after the second sonication round, the prevailing fraction of the protein was present in the lysate (Figure 21).

Doubly labeled His-PsbP was produced using cell-free expression in high amounts but it was not possible to purify the protein. His-PsbP was not present in any of the elutions which were done in a similar way to standard *in vivo* protocol. However the Western blot of the synthesis reaction revealed a vast amount of the protein being overexpressed.

5. 2. Purification

5. 2. 1. Affinity chromatography

His-PsbP was purified from most of the impurities during the affinity chromatography. The His-PsbP started to wash out at 40% of elution buffer, i.e. at 200 mM imidazole concentration. At the end of the imidazole gradient (350 mM), the protein was fully recovered from the column. Total of 7 fractions contained purified His-PsbP which were concentrated and prepared for subsequent purification on a cation-exchange column.

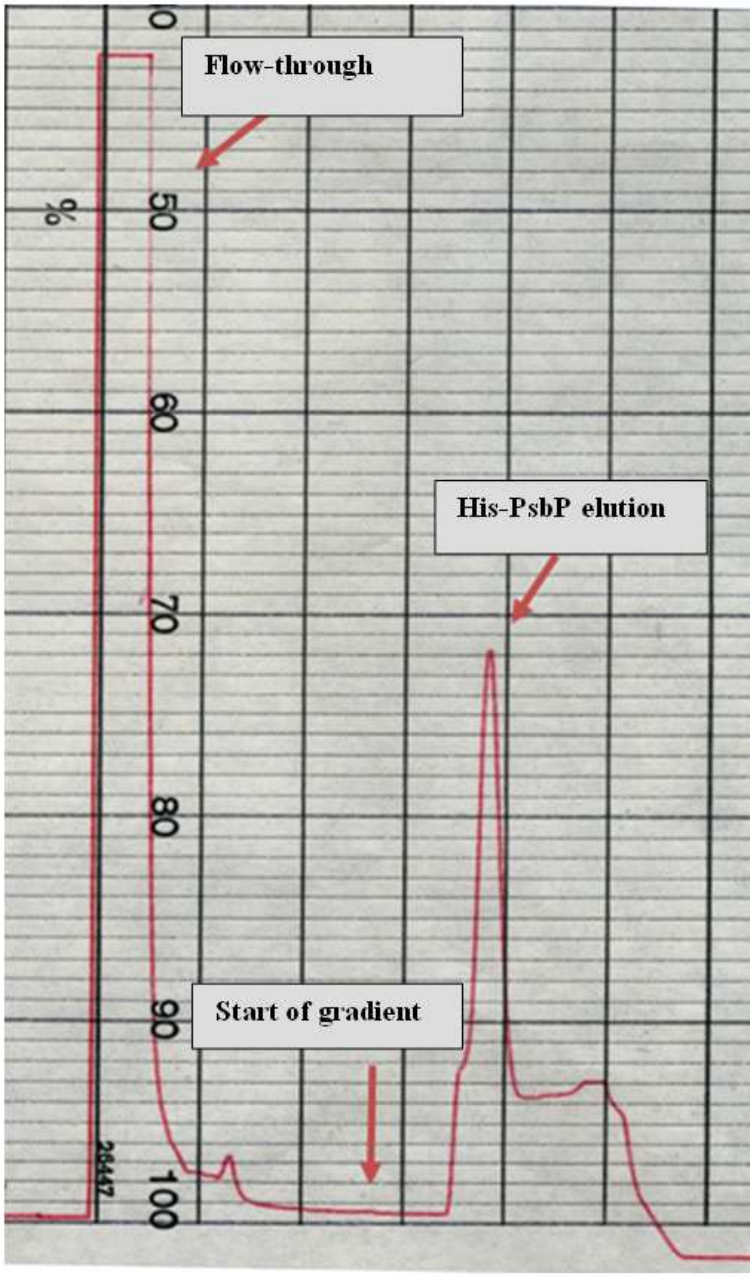


Figure 20: *Elution profile of His-PsbP during the affinity chromatography*

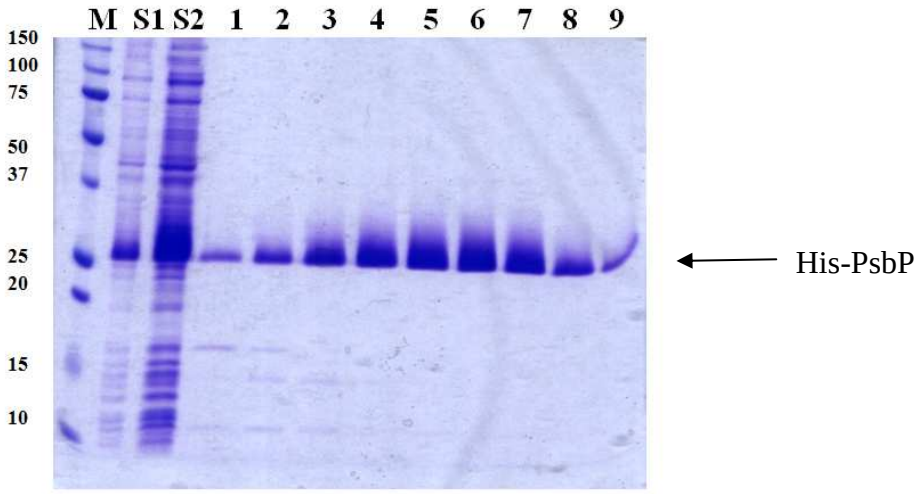


Figure 21: *SDS-PAGE of His-Ni^{2+} affinity chromatography. M = marker, S1= cell lysate after first sonication, S2 = cell lysate after the second run of sonication, 1 – 9 = elutions of His-PsbP*

5. 2. 2. Ion-exchange chromatography

A second purification step was included to further purify His-PsbP from E. coli proteins and to prepare it for the tag cleavage. His-PsbP was retained on a cationic exchanger column from which it was washed out increasing Na^+ concentration. Elution of His-PsbP started at 440 mM NaCl and was finished at 700 mM NaCl. His-PsbP was present in roughly 10 eluates. The collected fractions were concentrated using an Amicon (centriprep) filter (cut-off 10 kDa) and used for the subsequent proteolysis by thrombin.

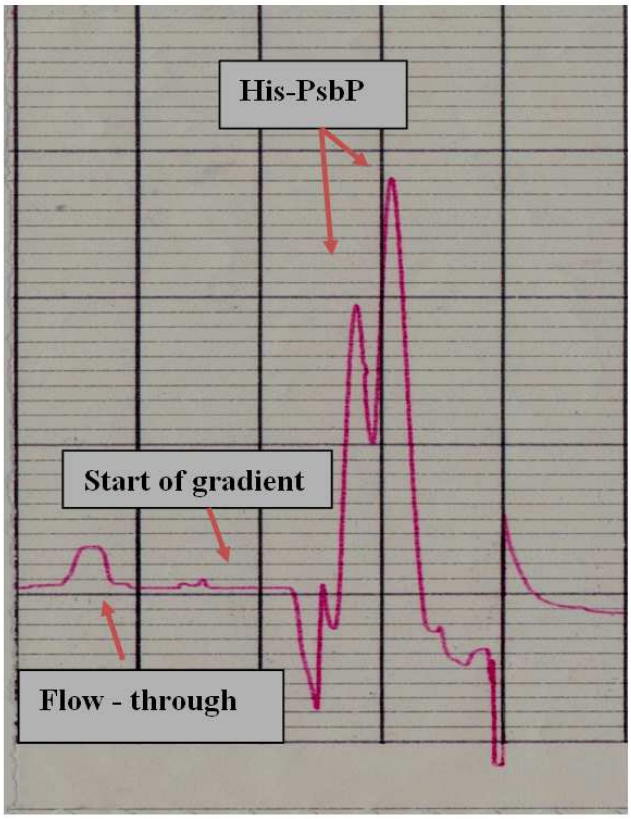


Figure 22: *Elution of the His-PsbP during Ion-exchange chromatography*

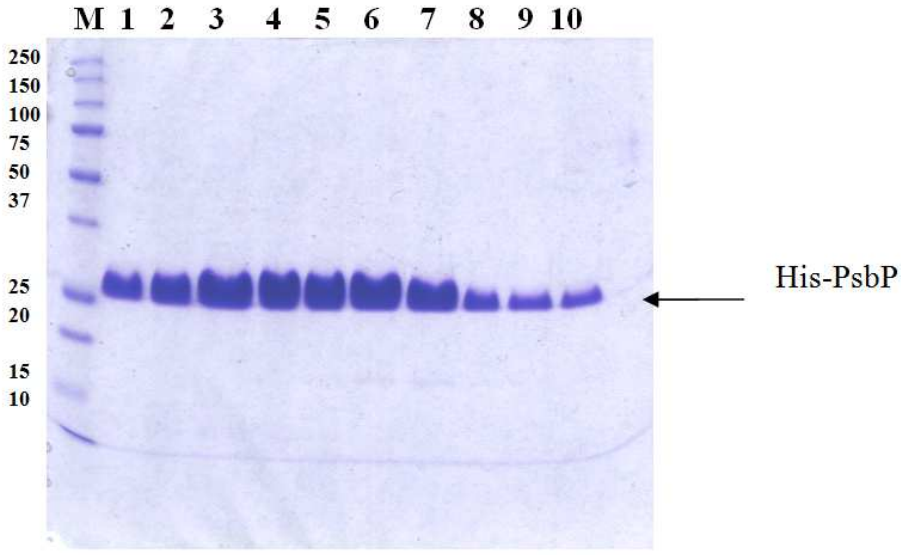


Figure 23: *SDS-PAGE of Ion-exchange chromatography. M = marker, 1 – 10 = elutions of His-PsbP*

5. 2. 3. Cleavage of the His-tag

The protease thrombin was used to cleave off the His-tag which was used for facile purification of PsbP and improved its stability. This practice was necessary prior to use of the PsbP for NMR measurement. The His-tag would interfere with signals from the mature protein and might have an effect on PsbP folding.

1.5 units of the protease/mg of protein were used to guarantee cleavage of all protein molecules. Amount of concentrated His-PsbP prepared for the cleavage was in the range of 20 mg for single and doubly labeled protein. The reaction was nearly over after 90 minutes, but it was carried out for 90 more minutes. The reaction mixture was then loaded onto a size exclusion column.

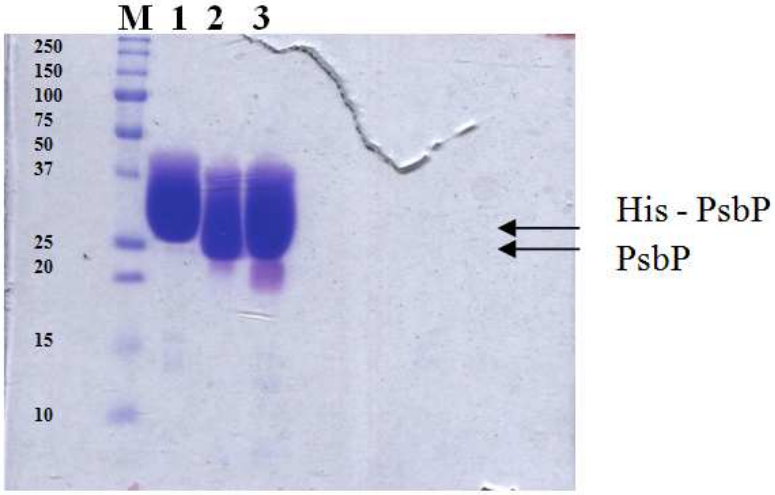


Figure 24: *Time course of thrombin cleavage. M = marker, 1 = start of reaction, 2 = 90 minutes of cleavage, 3 = 180 minutes of cleavage*

5. 2. 4. Size-exclusion chromatography

Final purification method was gel filtration which was supposed to filter out thrombin, protease inhibitors and the His tag. This was achieved and pure PsbP was recovered from the column after approximately 3 hours. PsbP eluted usually in 5 fractions. After the concentration, singly and doubly labeled PsbP were used for NMR measurements and unlabeled PsbP was applied in stability experiments.

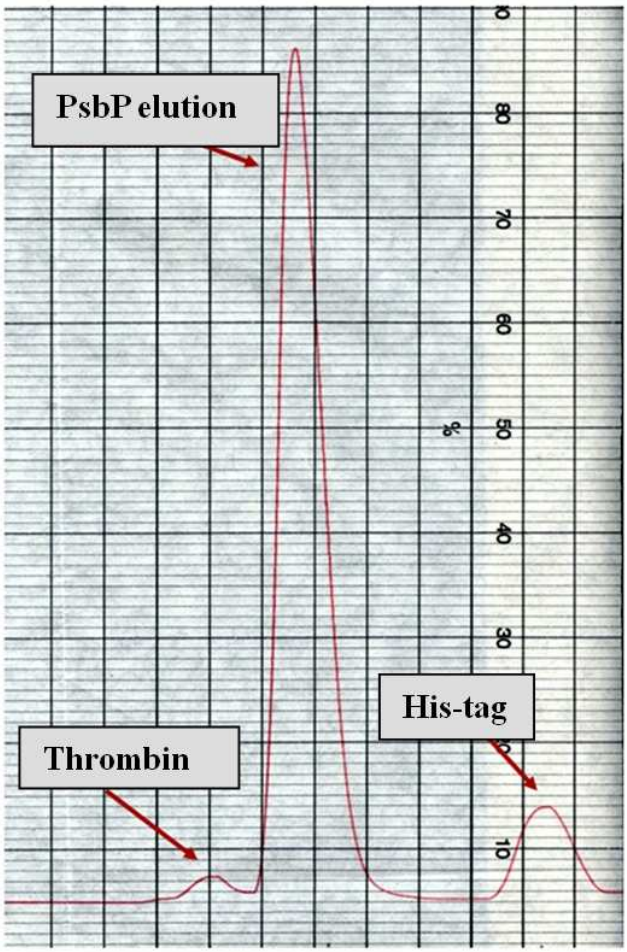


Figure 25: *Time course of Size exclusion chromatography.*

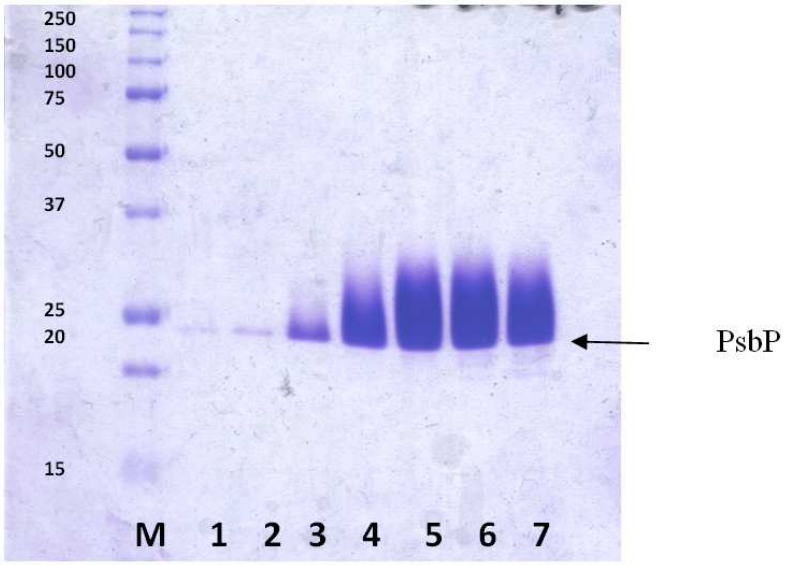


Figure 26: *SDS-PAGE of SEC fractions. M = protein marker, 1 – 7 = PsbP eluates*

Sample	Concentration in mol/l	Concentration in mg/ml	Total amount in mg
^{15}N PsbP	0.50 mM	9.9 mg/ml	14.85
^{15}N, ^{13}C PsbP	0.25 mM	5.1 mg/ml	20.82
unlabeled PsbP	0.29 mM	6.5 mg/ml	28.00

Tab. 15: *Protein yields in different expressions*

5. 3. Native PAGE

This technique was applied to reveal possible oligomers formed by PsbP or His-PsbP. Since some of the samples showed dimer formation even on the SDS-PAGE gel, it was very interesting to see how the protein would behave under non-denaturing conditions. Because no optimized protocol for native PAGE of His-PsbP exists, different conditions of the electrophoresis were investigated. In Figure 27 there is an overview of results obtained from different density of the separation gels and running times.

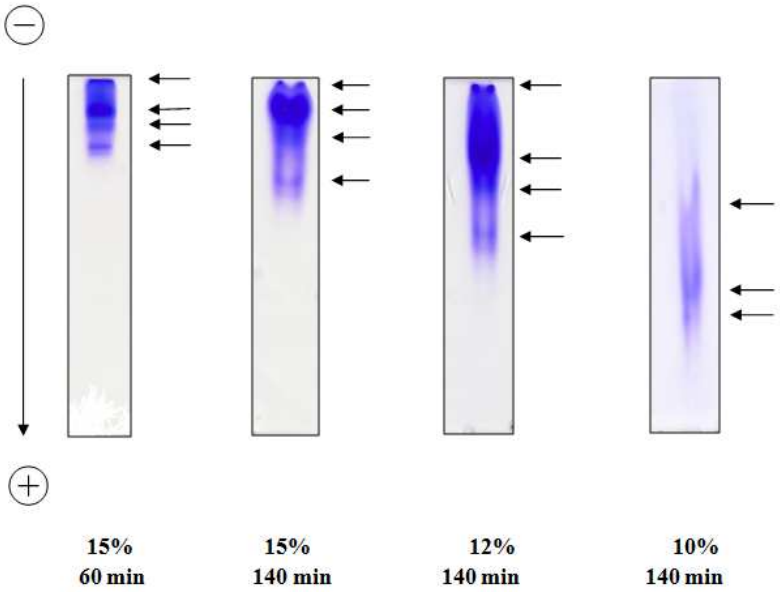


Figure 27: *Comparison of different conditions for native PAGE*

In the four upper lanes, concentrated His-PsbP sample from a stock solution was present. The arrows point to four distinctive species found in the sample. The uppermost band might represent the dimer followed by His-PsbP monomer. A band slightly smaller than His-PsbP is most probable the protein void of the N-terminal part. This would explain presence of the very small band, which would be the cleaved terminal amino acid stretch. All four species are present in His-PsbP sample, with increasing resolution upon the decreasing of the gel concentration. The sample in the very right lane is diluted His-PsbP for which only three bands were observed. This difference can be explained by very low resolution of native PAGE for diluted samples.

5. 4. Stability Tests

An effect of different factors including temperature, buffer composition, protease susceptibility and reducing agents were investigated. Bis Tris buffer was proved to be more suitable than the phosphate buffer, although the use of either of these is conditioned with the use of protease inhibitors. Both of the buffers contained 1 mM EDTA which alone was not sufficient to prevent proteolytic attack. Addition of a protease inhibitor mixture rapidly increased the PsbP stability and enabled its use even at higher temperatures.

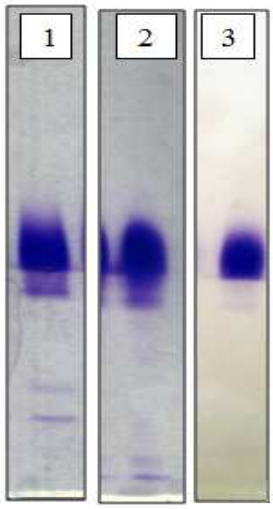


Figure 28: SDS-PAGE. *Influence of buffer type and protease inhibitors on the stability of PsbP. Lane 1 = PsbP in phosphate buffer after 7 days at RT. Lane 2 = PsbP in Bis Tris buffer after 7 days at RT. Lane 3 = PsbP in Bis Tris buffer with 1x protease inhibitors cocktail after 7 days at RT.*

Possible formation of PsbP dimers or oligomers was always a puzzle during the work with PsbP. When PsbP and His-PsbP stability at elevated temperature was compared, surprisingly it was the mature PsbP which was more stable. Fig. 29 shows the behavior of both, uncleaved and cleaved protein, at 30°C over one week of incubation. All samples were kept in Bis Tris buffer including the protease inhibitors. Whereas the His-PsbP shows dimerization at 30°C, PsbP is not forming dimers even in absence of DTT. This result suggests that the driving force of dimerization might be the His-tag sequence at the N-terminus together with elevated temperature.

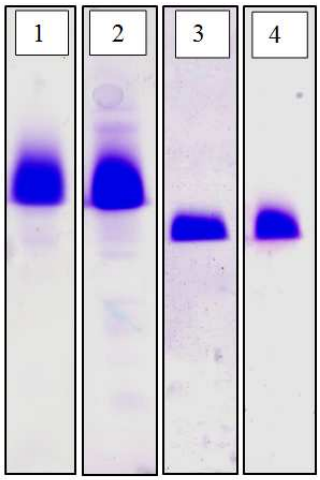


Figure 29: SDS-PAGE. *Stability of PsbP and His-PsbP samples at different temperatures. Lane 1 = His-PsbP with protease inhibitors and DTT after 7 days at 30°C. Lane 2 = His-PsbP with protease inhibitors after 7 days at 30°C. Lane 3 = PsbP with protease inhibitors and DTT after 7 days at 30°C. Lane 2 = His-PsbP with protease inhibitors after 7 days at 30°C. Lane 4 = PsbP with protease inhibitors after 7 days at 30°C. Lane 2 = His-PsbP with protease inhibitors after 7 days at 30°C.*

5. 5. NMR Spectra

2D and 3D spectra were acquired on a 700 MHz Avance III spectrometer at 20°C and were processed using Topspin 3.1.

^{15}N HSQC was obtained readily and proved the PsbP to be properly folded and suitable for the NMR measurements in 0.5 mM concentration. HSQC revealed two kinds of peaks – broad intense ones in the central region of the spectrum, and the peripheral weak signals. This heterogeneity was most probably caused by the diverse dynamic properties of the PsbP regions – a highly dynamic one would account for the intense, overlapping peaks and on the opposite a structured, more rigid part would explain formation of weak signals.

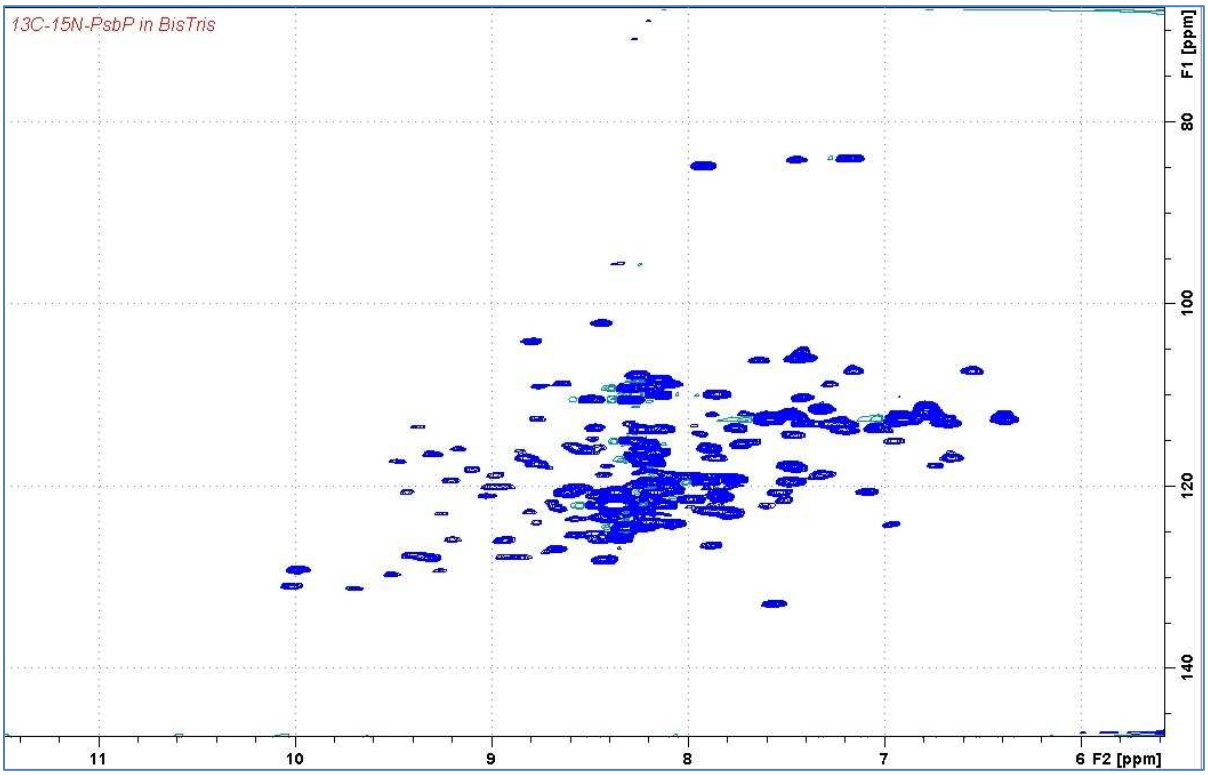


Figure 30: *^{15}N HSQC of PsbP recorded at 20°C (full spectrum).*

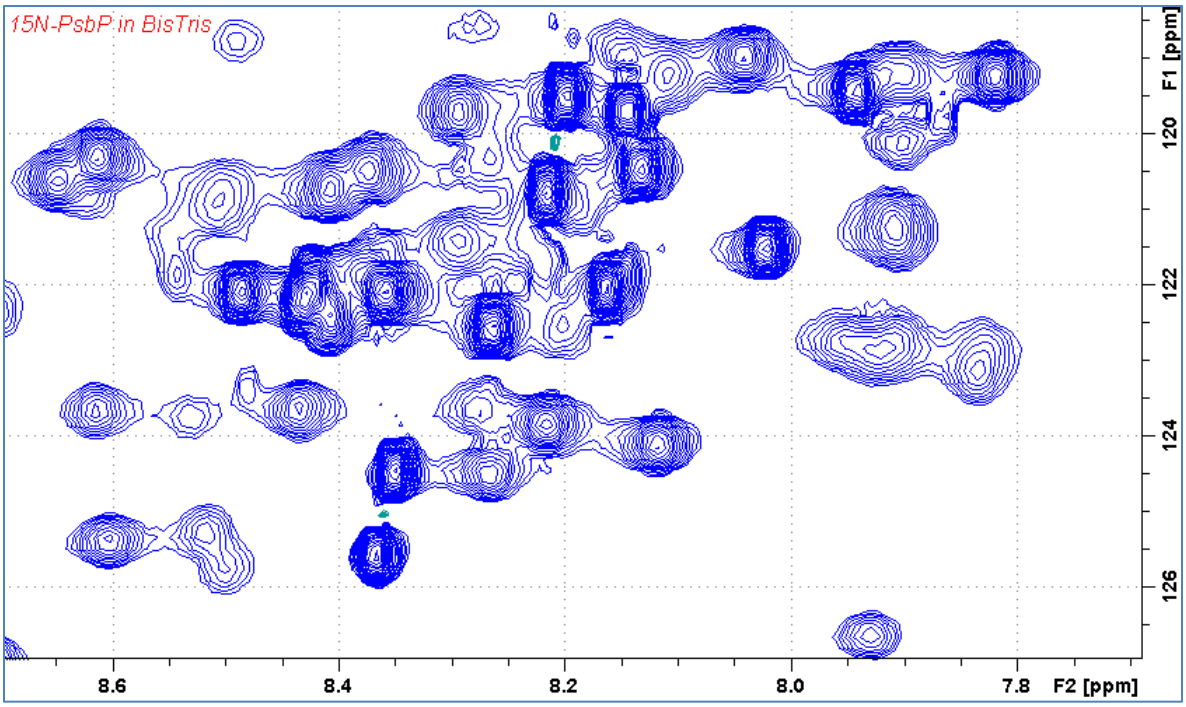


Figure 31: *Enlarged central region of ^{15}N HSQC of PsbP at 20°C.*

All of the 3D experiments were recorded at 20°C with 0.5 mM PsbP. Most of the 3D sequences resulted in fairly good spectra which were transferred to CARA for assignment.

Unfortunately, some of the experiments yielded only ca. 40% of the expected signals from which the majority was correlated to intense overlapping peaks in HSQC. Nevertheless, the signal obtained for this group of amino acids was good enough to be used for sequential backbone assignment.

The H(NCA)NNH experiment did not produce utilizable spectrum and was not included in the assignment strategy.

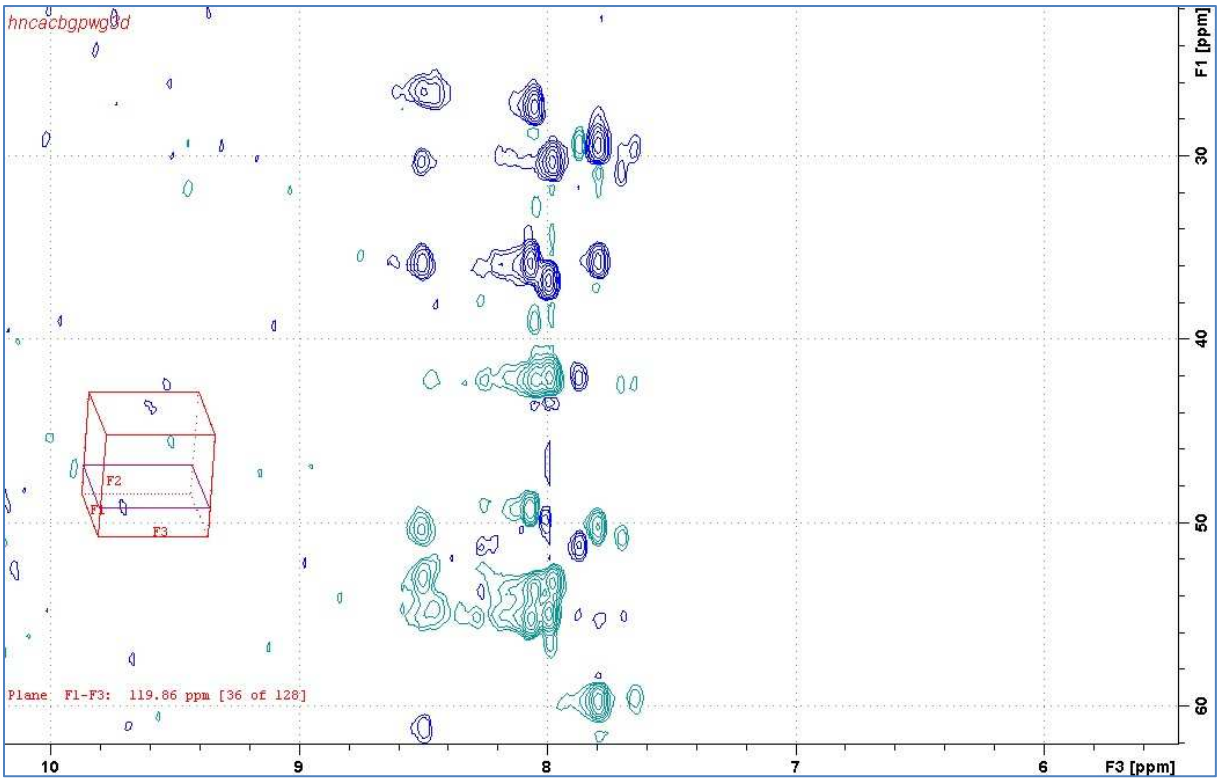


Figure 32: *A zoom-in of the ^{13}C-^{1}H plane in HNCACB with C_α and C_β cross-peaks.*

5. 6. Assignment

A sequential backbone assignment of PsbP protein (190 amino acids) was done in CARA using the 3D NMR experiments mentioned above with the exception of H(NCA)NNH. A total of 150 peaks were picked in ^{15}N HSQC with the help of HNCO. This step was followed by peak picking in HNCACB, CBCA(CO)NH and HNCA. Assignment of carbon resonances from the backbone and first side chain atom were accomplished for ca. 40% of HN pairs in HSQC. Stripscope was used for the alignment of matching chemical shifts to link amino acids with respect to PsbP sequence. The assignment was at the same time improved using HN(CA)NNH which identified preceding and successive amino acids for 40% of HSQC peaks. Validity of the assignment was controlled also in ^{15}N NOESY-HSQC spectrum.

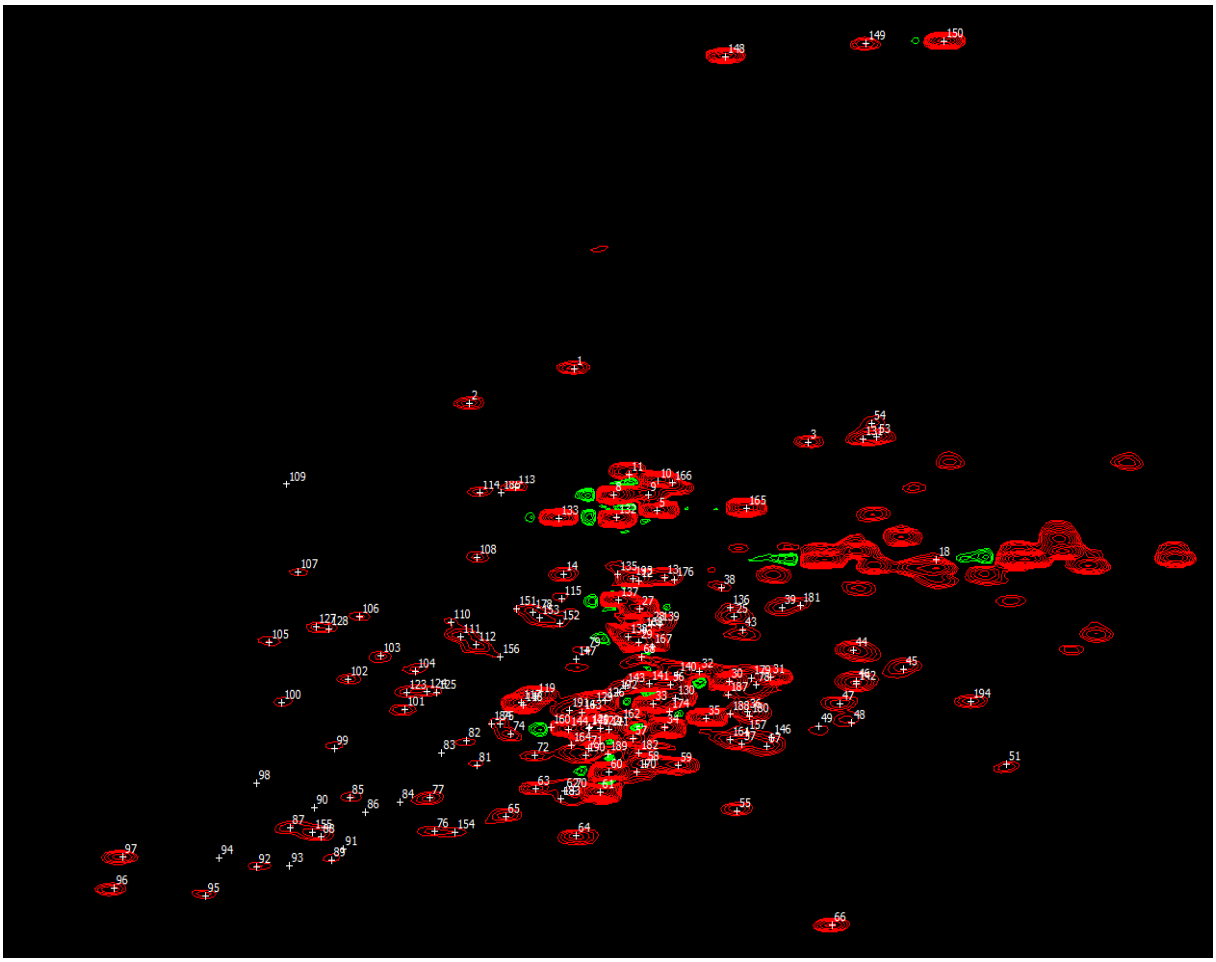


Figure 33: *Picked peaks in ^{15}N HSQC (CARA view).*

In the first run a total of 42 peaks were assigned sequence specifically out of 190 amino acids present in the PsbP and 150 peaks appearing in HSQC. That accounts for 28% (22%) backbone assignment of PsbP. PsbP sequence contains 8 Prolines – this means that total amount of HN cross peaks visible in HSQC should be around 180 while only 150 peaks showed up in the spectrum. As assumed, all assigned peaks are located in the middle region of the spectrum.

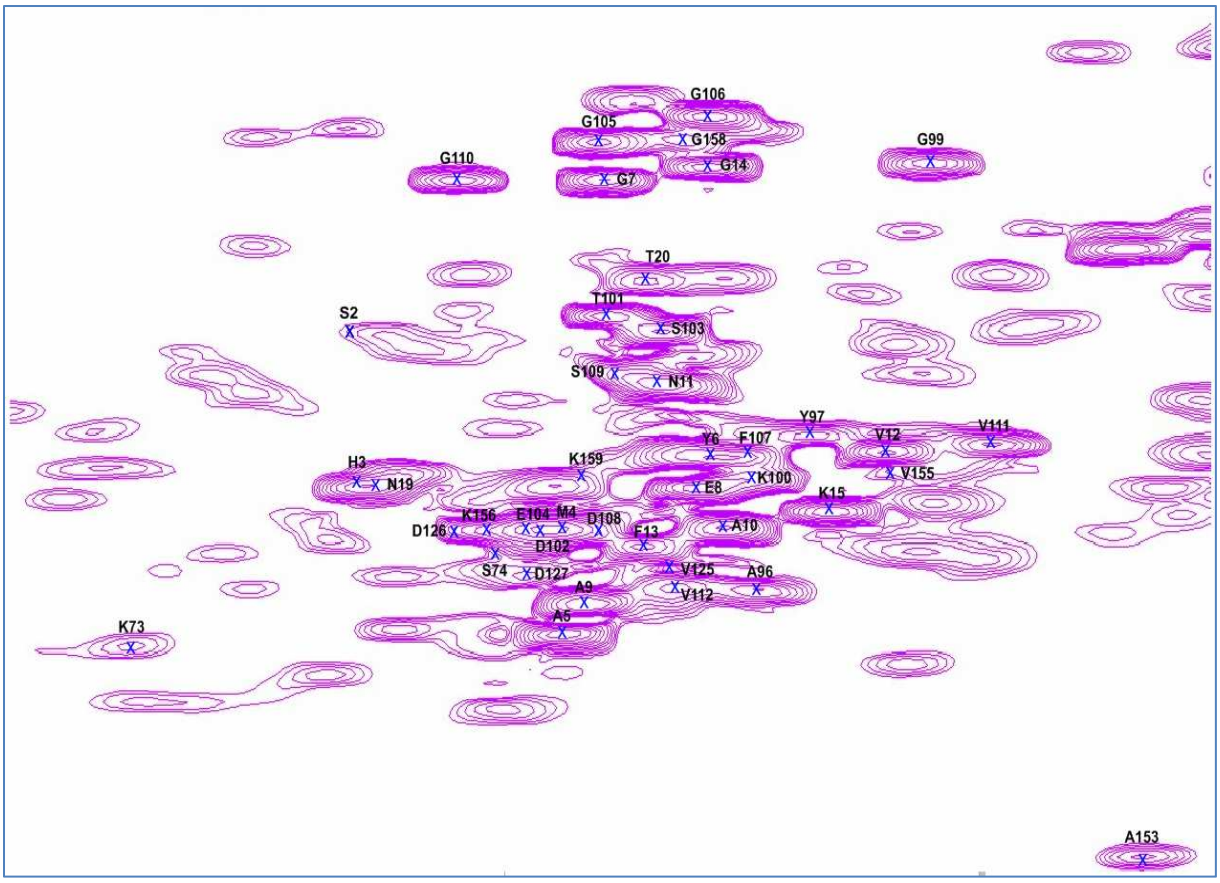


Figure 34: *^{15}N HSQC with the overview of assigned amino acids (CARA).*

The assigned amino acids all showed nice cross-peaks in HNCACB and CBCA(CO)NH. The assignment strategy for these amino acids was a straight forward process of matching selected HNCACB strips and supporting the hypothesis with the use of HN(CA)NNH and NOESY-HSQC spectra. The strips in HNCACB/HNCA(CO)CB corresponding to dispersed peaks in the outer parts of HSQC did not contain resonances of $^{13}C_{\alpha}$ (i-1), $^{13}C_{\beta}$ (i-1) and sometimes not even the information related to $^{13}C_{\alpha}$ (1) and $^{13}C_{\beta}$ (1). Such complication could be overcome with the use of HNCANNH, but this unfortunately did not work for the weak HN-pairs as well. As a result, reliable assignment of only part of the resonances was possible.

Amino acid	Spin System	Chemical shifts in ppm			
		^{1}H - N	^{15}N	$^{13}C_{\alpha}$	$^{13}C_{\beta}$
S 2	151	8.52	115.36	55.57	60.91
H 3	117	8.50	120.56	53.03	26.34
M 4	122	8.23	121.94	52.46	29.96
A 5	61	8.23	125.48	49.24	16.20
Y 6	141	8.07	119.49	55.28	35.89
G 7	132	8.18	110.31	42.12	_____
E 8	33	8.05	120.62	53.63	27.21
A 9	60	8.20	124.38	49.83	16.05
A 10	34	8.01	121.89	49.67	16.07
N 11	29	8.10	117.22	50.17	35.71
V 12	30	7.79	119.38	59.79	29.40
F 13	57	8.12	122.54	55.27	36.16
G 14	5	8.04	109.95	42.08	_____
K 15	35	7.87	121.42	51.06	29.40
N 19	118	8.49	120.65	50.26	35.91
T 20	193	8.12	113.68	58.95	66.53
K 73	77	8.81	125.79	53.15	29.41
S 74	164	8.33	122.89	53.43	60.07
A 96	59	7.97	123.99	49.10	16.49
Y 97	32	7.89	118.83	54.73	36.43
G 99	165	7.73	109.83	42.06	_____
K 100	130	7.98	120.29	53.32	30.25
T 101	137	8.17	114.90	58.65	67.10
D 102	120	8.27	121.87	51.49	38.12
S 103	27	8.09	115.36	56.13	60.91
E 104	145	8.27	121.96	53.59	26.88

Amino acid	Spin System	Chemical shifts in ppm			
		^{1}H - N	^{15}N	$^{13}C_{\alpha}$	$^{13}C_{\beta}$
G 105	8	8.18	109.07	42.35	_____
G 106	10	8.03	108.26	42.06	_____
F 107	56	7.99	119.57	54.99	36.70
D 108	121	8.20	122.05	51.06	38.14
S 109	138	8.14	116.89	56.12	60.91
G 110	133	8.37	110.37	42.34	_____
V 111	31	7.65	119.16	59.50	29.69
V 112	58	8.08	123.94	59.51	29.71
V 125	182	8.10	123.32	59.52	29.67
D 126	160	8.39	121.89	53.59	38.13
A 153	66	7.44	132.82	51.34	16.75
V 155	187	7.79	120.09	59.65	29.54
K 156	144	8.34	122.01	53.31	30.11
D 157	71	8.27	123.09	51.60	38.14
G 158	9	8.07	109.07	42.33	_____
K 159	126	8.22	120.44	53.58	27.26

Tab. 16: *Assigned amino acids with their corresponding location in PsbP sequence and obtained chemical shifts*

Given below is the sequence of PsbP with highlighted amino acids, which were successfully assigned. Prolines are marked in red, since they are not visible in HSQC spectrum and were skipped in the assignment procedure. Interestingly, the sequence stretches, which were assigned by NMR, fall into the unresolved parts of the most novel X-ray structure. This result corroborates the previous idea of flexible regions in PsbP available for possible interactions.

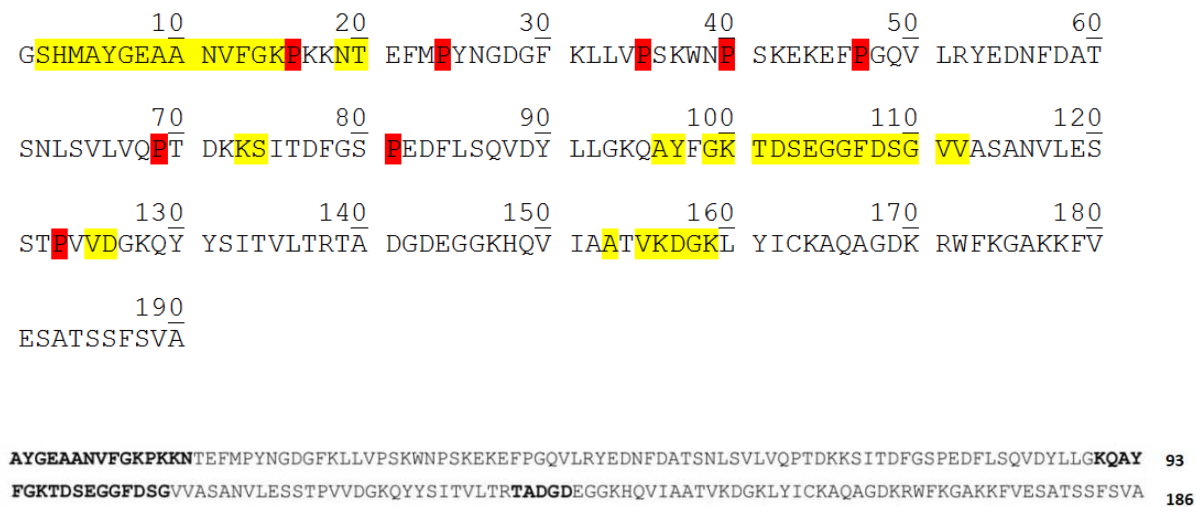


Figure 35: *PsbP sequence. Upper part shows the resulting assigned amino acids (yellow). The lower part is the sequence of PsbP used for X-ray structure determination with bold unresolved regions.*

6. Summary and Outlook

The accessory photosynthetic protein PsbP was prepared as a stable single ^{15}N and double ^{15}N, ^{13}C labeled recombinant protein by expression both in *E. coli* cells and in a cell free environment (ribosomal extract). The yield of the isotopically labeled protein prepared from 1 l of minimal (M9) medium ranged between 15 – 20 mg (0.25 – 0.5 mmol). The procedures and conditions of isolation and purification of PsbP were optimized to prevent degradation and aggregation. In particular the degradation of the functionally essential N-terminal part, where 5-15 amino acids are cleaved off very readily even at room temperature could be suppressed. This was a prerequisite for the ensuing analysis by NMR, where any possible change in the native structure of PsbP would complicate the measurements and jeopardize feasibility of resonance assignment.

Addition of a cocktail of various protease inhibitors together with EDTA subdues any proteolytic attack, even enabling the incubation of PsbP over a wide temperature range (20°C – 40°C) for at least a week without degradation or change in the 3D structure, as judged by multi-dimensional heteronuclear NMR. Dimerization of PsbP was another problem observed earlier preventing structure determination of the monomeric protein. With a series of stability experiments it could be shown that, the His-tagged PsbP, which was used in the earlier investigations, is much more prone to oligomerization than the cleaved mature protein, especially at elevated temperature. This phenomenon has been reported previously for some other proteins [42]. Addition of a reducing agent, ca. 2 mM DTT, to the working buffers as a general precaution successfully suppressed the dimerization. The pH 6 buffer used for storage and experiments with PsbP was in the end composed of 20 mM Bis Tris, 1 mM EDTA, 0.05 mM NaN_3, mixture of protease inhibitors, 1 mM DTT. As it was suggested, Bis Tris has the biggest stabilizing effect compared to the other buffers investigated [35].

A cell-free method was utilized to prepare doubly labeled His-PsbP. Although the protein could not be recovered after the purification process, it was clearly demonstrated that the plasmid construct used can be routinely employed for in vitro PsbP expression from isotopically labeled amino acids.

3D triple resonance NMR spectra were recorded to assign HN-pairs in HSQC spectra to amino acids with sequence specificity at 20°C first. Out of 190 amino acids forming the PsbP, 150 of them yielded cross peaks in ^{15}N-HSQC. For approximately 70 HN-pairs,

corresponding ^{13}C resonances were found in HNCACB and CBCA(CO)NH spectra. These were used for the sequential assignment of 42 amino acids, located mostly at the N-terminal part and central part of the protein. HN-cross peaks representing these amino acids were located in majority in the crowded middle region of the HSQC spectrum, which is probably coming from a flexible protein part. The assigned amino acids correspond closely to the unresolved regions in the newest X-ray structure of PsbP from *Spinacia oleracea* (PDB accession code 2VU4) [19]. The segments were present in the crystal however they could not be resolved using the electron density map. The missing regions located in the inner part of PsbP sequence were modeled as loops in the published structure, whereas 15 N-terminal amino acids were predicted to possess a β-type conformation. The internal loop between residues 94 and 111 is believed to play a role in the binding of PsbP to PS II[19].

The accomplished preliminary resonance assignment accounts for only one quarter of the amino acids from PsbP protein. To overcome problem with signal heterogeneity, the temperature at which the NMR experiments were recorded was increased from 20°C up to 40°C. Despite of opposite expectations, PsbP is still folded natively at this temperature as evidenced by the very minor chemical shift changes. The amplitudes and line widths of the cross-peaks in HSQC were more uniformly distributed and signal overlap was reduced significantly (Figure 36). Furthermore, nearly all PsbP amino acid residues yielded HN cross peaks in HSQC. With those promising results, a new set of triple resonance spectra will be recorded in order to complete the backbone assignment. As soon as this would be finished, near complete assignment of the side chain resonances will be possible. This is the prerequisite for deriving distance restraints, which can be used for a 3D solution structure calculation. Such structure might clarify the unresolved areas such as the N-terminus and the loop region. The solution structure of PsbP will be used as a building block in the study of binding interactions with other extrinsic proteins from PS II – PsbO, PsbQ and possibly PsbR[45]. Protein – ligand interactions might further modulate the protein-protein and protein-membrane interactions and the mechanism of binding of ions found in the thylakoids.

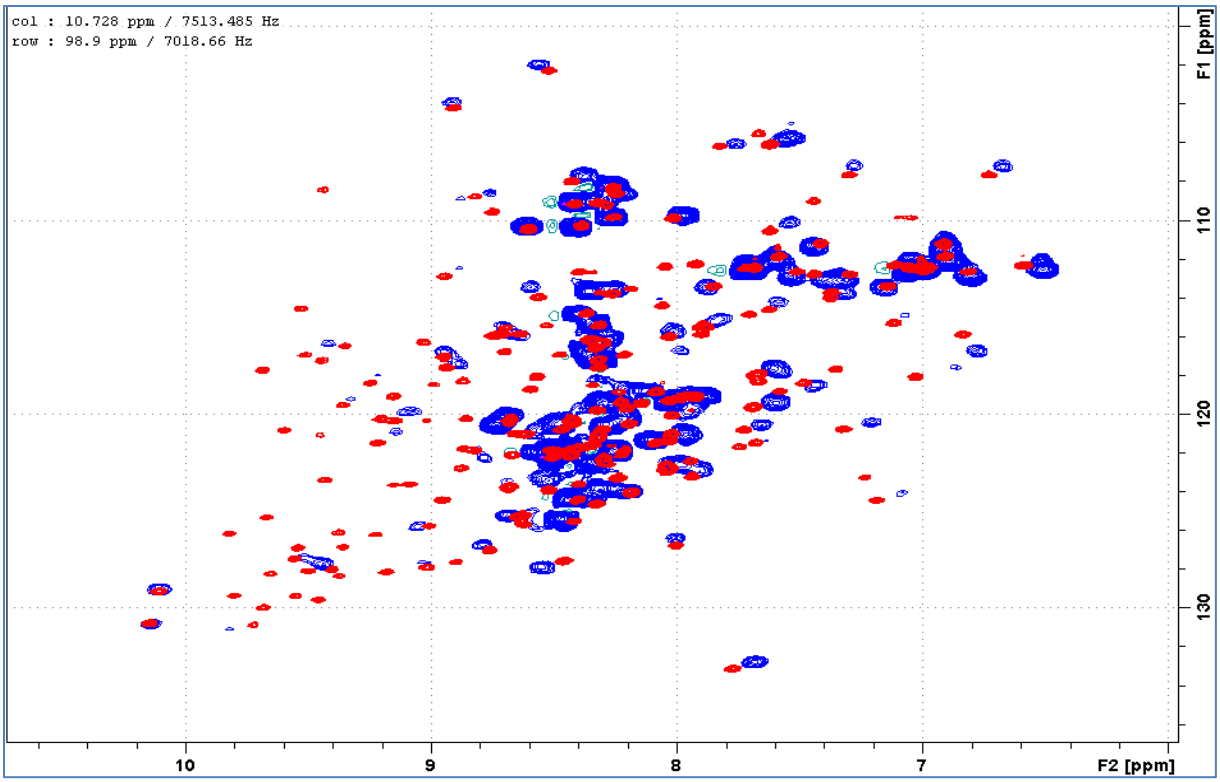


Figure 36: *Overlaid display of ^{15}N HSQC spectra. The blue spectrum was recorded at 20°C and used for the assignment. Red spectrum was very recently recorded at 40°C from a stabilized sample (see text).*

7. References and Links

1. Christopher R. Scotese, Back to Earth History : Summary Chart for the Precambrian, Paleomar Project
2. Bekker, A. et al. *Nature* 427, 117–120 (2004)
3. Tomitani, A., Knoll, A. H., Cavanaugh, C. M. & Ohno, T. *Proc. Natl Acad. Sci. USA* 103, 5442–5447 (2006)
4. Allen, J. F., Martin, W. Out of thin air. *Nature* 445/8 (2007)
5. Bryant, D. A. & Frigaard, Prokaryotic photosynthesis and phototrophy illuminated**.** *N. - U. Trends Microbiol*. 14:488–496 (2006)
6. Mulkidjanian, A. Y. et al. The cyanobacterial genome core and the origin of photosynthesis. *Proc. Natl Acad. Sci. USA* 103:13126–13131 (2006)
7. Allen, J. F. A redox switch hypothesis for the origin of two light reactions in photosynthesis. *FEBS Lett*. 579: 963–968 (2005)
8. Yano, J. et al. Where Water Is Oxidized to Dioxygen: Structure of the Photosynthetic Mn4Ca Cluster. *Science* 314: 821–825 (2006)
9. Shen J.R., Inoue Y. Binding and functional properties of two new extrinsic components, cytochrome c-550 and a 12 kDa protein, in cyanobacterial photosystem II. *Biochemistry* 32:1825–1832 (1993)
10. Enami I., Murayama H., Ohta H., Kamo M., Nakazato K., Shen J.R. Isolation and characterization of a Photosystem II complex from the red alga Cyanidium caldarium: association of cytochrome c-550 and a 12 kDa protein with the complex. *Biochim Biophys Acta* 1232:208–216 (1995)
11. Bricker, T. M. Oxygen evolution in the absence of the 33 kDa manganese stabilizing protein. *Biochemistry* 31:4623–4628 (1992)
12. Ono, T. A., Inoue, Y. Mn-preserving extraction of 33-, 24- and 16-kDa proteins from O2-evolving PS II particles by divalent salt-washing. *FEBS Lett*. 164:255-260 (1983)
13. Razeghifard, M. R., Wydrzynski, T., Pace, R. J., Burnap, R. L. YZ• reduction kinetics in the absence of the manganese-stabilizing protein of Photosystem II. *Biochemistry* 36:14474–14478 (1997)
14. Ljungberg, U. Åkerlund, H. E., Andersson, B. Isolation and characterization of the 10-kDa and 22-kDa polypeptides of higher-plant Photosystem 2. *Eur. J. Biochem.* 158:477 - 478 (1986)

15. Suorsa, M., Sirpio, S., Allahverdiyeva, Y., Paakkarinen, V., Mamedov, F., Styring, S., Aro, E. PsbR, a missing link in the assembly of the oxygen-evolving complex of plant Photosystem II. *J. Biol. Chem.* 281:145–150 (2006)
16. Jansen, T., Rother, C., Steppuhn, J., Reinke, H., Beyreuther, K., Jansson, C., Andersson, B., Herrmann, R. G. Nucleotide sequence of cDNA clones encoding the complete 23 kDa and 16 kDa precursor proteins associated with the oxygen evolving complex from spinach. *FEBS Lett*. 216:234–240 (1987)
17. Gregor, W., and Britt, D. Nitrogen ligation to the manganese cluster of Photosystem II in the absence of the extrinsic proteins and as a function of pH. *Photosynth. Res.* 65:175-185 (2000)
18. Seidler, A. The extrinsic polypeptides of Photosystem II. *Biochim Biophys Acta* 1277: 35–60 (1996)
19. Kopecky V. Jr, Kohoutova J., Lapkouski M., Hofbauerova K., Sovova Z., et al. Raman Spectroscopy Adds Complementary Detail to the High-Resolution X-Ray Crystal Structure of Photosynthetic PsbP from Spinacia oleracea. *PLoS ONE* 7 (10): e46694 (2012)
20. Ifuku et al. Structure and function of the PsbP protein of Photosystem II from higher plants. *Photosynthesis Research* 84:251–255 (2005)
21. Suzuki, T., Minagawa, J., Tomo, T., Sonoike, K., Ohta, H., Enami, I.Binding and functional properties of the extrinsic proteins in oxygen-evolving Photosystem II particles from a green alga, Chlamydomonas reinhardtii having his-tagged CP47, *Plant Cell Physiol.* 44:76–84 (2003)
22. J. C. Janson. Protein purification. 2011
23. www.bio-rad.com
24. http://www.nd.edu/~aseriann/CHAP9B.html/sld013.htm
25. D. Nelson, M. Cox. Lehninger Principles of Biochemistry. 2001
26. The Protein Protocols Handbook. 2002, pp 3-6
27. http://www.embl.de/pepcore/pepcore_services/protein_expression/ecoli/cellfree_expression_systems/,
28. Pedersen, A., Hellberg, K., Enberg, J. and Karlsson, G. Rational improvement of cell-free protein synthesis. *N Biotechnol* 28:218-24 (2011)
29. Klammt, C., Maslennikov I. et al. Facile backbone structure determination of human membrane proteins by NMR spectroscopy. *Nat Methods.* 9:834-9 (2012)
30. Affinity Chromatography – Principles and Methods. GE Healthcare

31. Ion-exchange Chromatography – Principles and Methods. GE Healthcare
32. http://mach7.bluehill.com/proteinc/tutorial/iec.html
33. http://en.wikipedia.org/wiki/Thrombin
34. Gel filtration – Principles and Methods. GE Healthcare
35. Kohoutová et al. Crystallization and preliminary crystallographic characterization of the extrinsic PsbP protein of photosystem II from Spinacia oleracea. *Acta Cryst.* F65:111–115 (2009)
36. Protein protocol handbook. GE Healthcare
37. Berger, S. 200 and more NMR experiments. 2004
38. Cavanagh, J. Protein NMR Spectroscopy. 2007
39. http://www.protein-nmr.org.uk/
40. Weisemann, R., Rüterjans, H. and Bermel, W. 3D Triple-resonance NMR techniques for the sequential assignment of NH and 15N resonances in 15N- and 13C-labelled proteins. *Journal of Biomolecular NMR*, 3:113-120 (1993)
41. http://www.cara.nmr-software.org/downloads/3-85600-112-3.pdf
42. Hom, L. G., Volkman, L. E. Nickel-Induced Oligomerization of Proteins Containing 10-Histidine Tags. *BioTechniques* 25:20-22 (1998)
43. Isaksson, L., Enberg, J., Neutze, R., Karlsson, G. and Pedersen, A. Expression screening of membrane proteins with cell-free protein synthesis.*Prot. Exp. Purif.* 82:218-25 (2012).
44. Laemmli, U. K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* 227:680 - 685 (1970)
45. Bricker et al. The extrinsic proteins of Photosystem II. *Biochim Biophys Acta* 1817: 121 – 142 (2012)