Domain specific isotope labelling of membrane proteins – NMR and FTIR spectroscopy of SRII/HtrII complexes from *Natronobacterium pharaonis*

Vom Fachbereich Chemie der Universität Dortmund
zur Erlangung des akademischen Grades
„Doktor der Naturwissenschaften“
(Dr. rer. nat.)
genehmigte

Dissertation

angefertigt am Max-Planck-Institut für molekulare Physiologie in Dortmund

vorgelegt von

Dipl.-Chem. Eric Schiffer
aus Weinsheim

Betreuer der Arbeit: Prof. Dr. M. Engelhard

Dortmund, August 2005

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Arbeit selbständig und nur mit den angegebenen Hilfsmitteln angefertigt habe.

Dortmund, 15. August 2005 Eric Schiffer

Die vorliegende Arbeit wurde in der Zeit vom Februar 2002 bis August 2005 am Max-Planck-Institut für molekulare Physiologie in der Abteilung Physikalische Biochemie unter der Anleitung von Prof. Dr. M. Engelhard und Prof. Dr. R.S. Goody durchgeführt.

1. Gutachter: Prof. Dr. M. Engelhard
2. Gutachter: Prof. Dr. P. Eilbracht

Meinen Eltern

„Nature is wont to hide herself.“

Heraclitus

I Table of contents

2

2-Cl-Z
2-chlorobenzyloxycarbonyl

A

amu
atomic mass unit

APS
ammoniumpersulfate

B

bis-tris-propane
1,3-bis(tris[hydroxymethyl]methyl-amino) propane

BN-PAGE
blue-native polyacrylamide gel electrophoresis

BOC
tert-butyloxycarbonyl

BR
bacteriorhodopsin

Bzl
benzyl

C

CBD
chitin binding domain

CD
cirular dichroism

CHAPS
3-[(3-cholamidopropyl)dimethyl-ammonio]-1-propanesulfonate

Che
protein of the two-component signalling cascade

CLP
lipidic cubic phase

COSR
C-terminal thioester

D

D
dimensional

Da
Dalton

Dab
2,4-diamino buturic acid

DCM
dichloromethane

DDM
dodecyl maltoside

DEAE
diethylaminoethyl

DIEA
diisopropyl-ethylamine

DMF
dimethyl formamide

DNA
desoxyribonucleic acid

DNP
2,4-dinitrophenyl

dNTP
desoxyribonucelotide triphosphate

DTT
dithiothreitol

E

E.coli
Escherichia coli

EDTA
ethylenediamine tetraacetate

EPL
expressed protein ligation

EPR
electron paramagnetic resonance

ESI
electro spray ionisation

EtSH
ethanethiol

F

Fmoc
fluorenylmethoxycarbonyl

FPLC
fast protein purification chromatography

FTIR
Fourier-transform infrared

G

GCPR
G-protein coupled receptor
Gdm
guanidine hydrochloride

GSH
glutathion

GST
glutathion-S-transferase

H

HAMP
histidine kinases, adenylyl cyclases, methyl-accepting proteins and phosphatases

HBTU
2-(1H-benzotriazole-1yl)-1,1,3,3-tetramethyluronium hexafluoro-phosphate

HR
halorhodopsin

Hs
Halobacterium salinarum

HSQC
heteronuclear single quantum correlation

Htr
halobacterial transducer of sensory rhodopsin

I

I-AEDANS
5-(2-[iodoacetyl)amino)ethyl)amino naphthalene-1-sulfonic acid

IML
intein mediated ligation

INEPT
insensitive nuclei enhanced by polarization transfer

IPTG
isopropyl-β-D-thiogalactopyranoside

ITC
isothermal titration calorimetry

K

K_D
dissociation constant

L

LB
Luria Bertani medium

M

MALDI
matrix-assisted laser desorption ionisation

MAS
magic angle spinning

MCP
methyl accepting chemotaxis protein

MESNA
sodium 2-mercaptoethane sulfonate

MS
mass spectrometry

MWCO
molecular weight cut off

Mxe
Mycobacterium xenopi

N

NaPi
sodium phosphate

NBD
7-chloro-4-nitrobenzo-2-oxa-1,3-diazole

NMR
nuclear magnetic resonance

NOESY
nuclear Overhauser effect spectroscopy

Np
Natronobacterium pharaonis

NTA
nitrilo triacetic acid

O

OcHx
cyclohexyloxy

OD
optical density

P

PAM
phenylacetamidomethyl

PCR
polymerase chain reaction

PEG
polyethylene glycol

PM
purple membrane

pMeBzl
4-methylbenzyl

R

rcf
relative centrifugal force

RP-HPLC
reversed phase high performance liquid chromatography

rpm
rounds per minute

rt
room temperature

S

sb
Schiff base

SDS-PAGE
sodium dodecyl sulfate polyacrylamide gel electrophoresis

SEC
size exclusion chromatography

SPPS
solid phase peptide synthesis

SR
sensory rhodopsin

ss
solid state

T

Tar
aspartate chemoreceptor

TBE
tris bufferred EDTA

TCA
trichloroacetic acid

TCEP
tris(2-carboxyethyl) phosphine hydrochloride

TEMED
N, N, N, N-tetramethylethylenediamine

TEV
tobacco etch virus

TFA
trifluor acetic acid

TFE
trifluoroethanol

TM
transmembrane helix

TOCSY
total correlation spectroscopy

Tos
tosyl

TrisHCl
2-amino-2-(hydroxymethyl)-1,3-propanediol hydrochloride

TROSY
transverse relaxation-optimized spectroscopy

Tsr
serine chemoreceptor

U

u
uniformly labelled

UV
ultraviolet

Single- and three-letter amino acid code

alanine	A	Ala	leucine	L	Leu
arginine	R	Arg	lysine	K	Lys
asparagine	N	Asn	methionine	M	Met
aspartic acid	D	Asp	phenylalanine	F	Phe
cysteine	C	Cys	proline	P	Pro
glutamine	Q	Gln	serine	S	Ser
glutamic acid	E	Glu	threonine	T	Thr
glycine	G	Gly	tryptophane	W	Trp
histidine	H	His	tyrosine	Y	Tyr
isoleucine	I	Ile	valine	V	Val

1. Introduction

All organisms, including bacteria and highly complex species like mammalians and humans, own mechanisms to detect and process stimuli from their environment, which finally trigger a physiological reaction. This happens with support of complex networks, the so called signalling cascades, which were adapted to the living- and environmental conditions of the respective species during evolution and which ensured its survival and reproducibility.

The first step in these signalling cascades is the detection of the outer stimulus by a receptor that is often located in the cell membrane. The membrane is therefore often called "the interface" to the environment. A super-family of such transmembrane receptors, which is widespread in higher organisms with more than 2000 members, is the rhodopsin-like seven-transmembrane-helix-receptors or G-protein coupled receptors (GCPRs)[1]. The second name is related to the fact that these receptors interact in their active state with a guanine nucleotide-binding proteins that conveys the signal from the receptor to a variety of effectors that belong to the signalling cascade inside the cell. Other members of this family include odorant and taste sensors[1], hormone- as well as neurotransmitter receptors[2] and light-receptors[3]. Light-receptors, like rhodopsins, can be found as photosensitive pigments in all three kingdoms of life: eukaryotes, archea and eubacteria. Due to the fact that they react to experimentally easy and exact manageable light instead of chemical substances, they often are used as model systems for the whole class of rhodopsin-like seven-helix-receptors.

In contrast to rhodopsins from higher organisms, which bind the G-protein transducin and initiate an electro-physiological reaction that finally leads to nerve impulses, sensory rhodopsins from archea are simpler and therefore more accessible for investigations. These proteins exhibit the same structural features as those found in other rhodopsins: They consist of seven transmembrane helices that form a binding pocket for the covalently attached retinal-chromophore, which is bound to a conserved lysine residue via a protonated Schiff base (sb). The irradiation with visible light leads to the activation of these pigments, which subsequently undergo a cyclic reaction, the 'photo-cycle'. The primitive vision process of bacteria, phototaxis, is relatively easy to investigate in *Halobacterium salinarum* or *Natronobacterium pharaonis*. Under constant light conditions the flagellar bundle, which is responsible for locomotion, changes its rotary direction approximately every ten seconds. This leads to an undirected random motion of the

bacterium[4]. Irradiation with blue light increases the frequency of change of the rotary direction, whereas orange light decreases this frequency. As a result the bacterium directs its motion away from blue to orange light, which is used by the light-driven ion-pumps bacteriorhodopsin (BR) and halorhodopsin (HR) (figure 1.1)[5,6]. These bacterial rhodopsins, which show similar structural features, utilise sunlight to generate energy for the translocation of protons (BR) respectively chloride anions (HR) across the cell membrane. The proton-motive force so produced is used for ATP-synthesis.

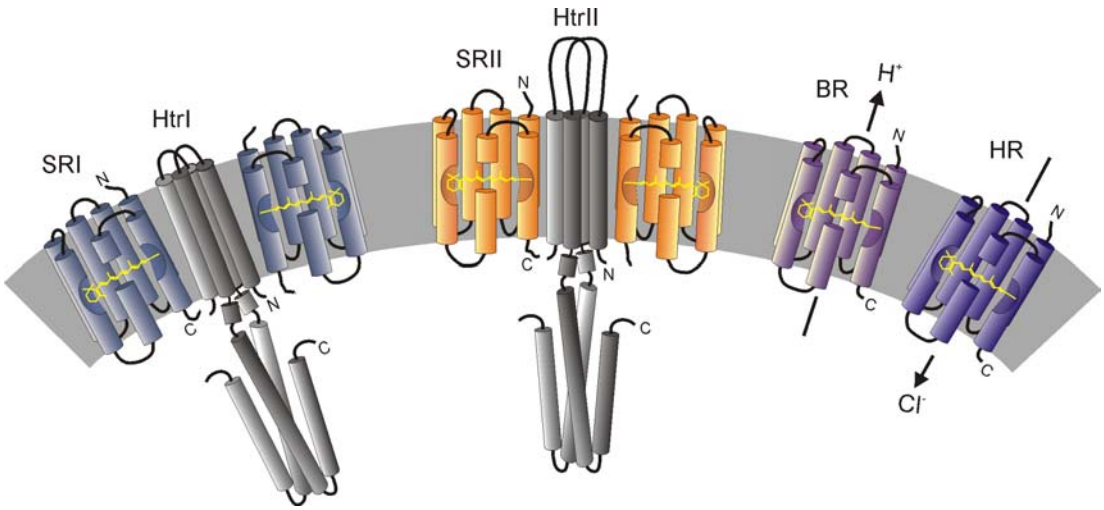


Figure 1.1. Schematic representation[7] of the four archeabacterial rhodopsins. The illustration (taken from Klare, 2002) of the sensory rhodopsins SRI and SRII with their cognate transducers HtrI and HtrII in 2:2 stoichiometry is based on publications from Chen and Spudich[8] as well as Gordeliy et al.[9]. The extra-cellular side is at the top.

In *H.salinarum* the sensory rhodopsins SRI and SRII were identified as being responsible for phototactic behaviour[10]. SRI is predominantly expressed by the organism under anaerobic conditions, mediates the photophilic response of the bacterium towards orange light as well as a photophobic answer in a two-photon process, towards blue light. Under aerobic conditions the bacterium utilises the photophobic SRII receptor to avoid blue light[11,12].

Archeabacterial as well as eucaryotic rhodopsins have the above-mentioned chromophore-retinal as prosthetic group. It is bound via a Schiffs base to helix VII (also called helix G) and on light stimulus changes its conformation. In case of eukaryotic rhodopsin this conformational change occurs from the 11-*cis* ground state to an all-*trans* excited state. However, the archeabacterial rhodopsins are different: the all-*trans* ground

state undergoes an isomerization after photon absorption to a 13-*cis* state. In addition the chromophore does not dissociate from the protein after the absorption of a photon, but thermally relaxes in a cyclic reaction to an all-*trans* ground state. This photocycle was studied in detail for the photophobic SRII receptor[13] (figure 1.2), the sensory rhodopsin I[14] as well as for the cognate ion pumps BR[15] and HR[16].

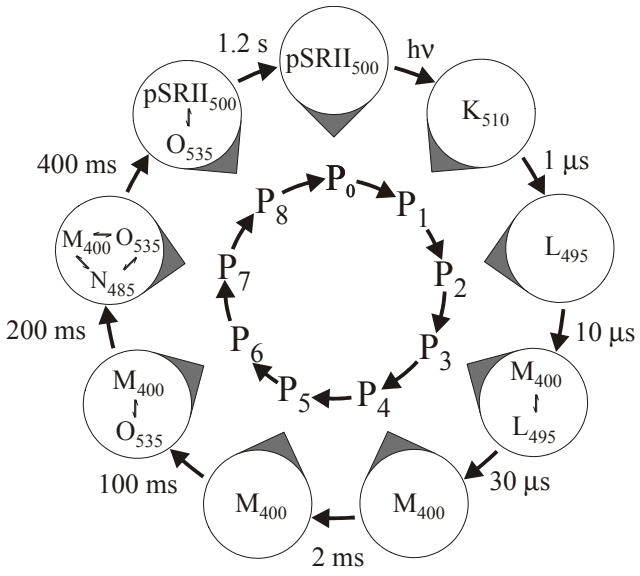


Figure 1.2. Model of the photocycle[13] (taken from Chizhov, 1998) of the photophobic receptor SRII from *N.pharaonis* based on a sequence of irreversible reactions. Absorbance maxima of the spectral intermediates K, L, M, N and O are indexed[13,17].

After photo-excitation the protein undergoes several intermediate states that can be spectroscopically distinguished and analysed by their different absorption maxima. In the case of *Np*SRII the photocycle can be described as follows: The absorption of a single photon leads to a 13-cis isomerisation of the chromophore (*Np*SRII→K). After approximately one microsecond the protein reacts with conformational changes (K→L). These changes alter the affinity of several residues towards protons leading to the donation of the Schiff base proton to Asp75 in the extra-cellular proton channel of the protein. At this point a proton is released into the medium at the extracellular site (L→M). After that, the cytoplasmic proton-channel opens and the Schiff base becomes re-protonated (M→N→O). Finally, the retinal re-isomerises and the protein relaxes to the ground state (O→*Np*SRII).

The knowledge about the response of the receptor to orange light is related to the question - 'which photo intermediate serves as signalling state of the receptor?'. Initial investigations reported that the signalling state is initiated by the deprotonation of the Schiff base and is maintained until the relaxation of the protein to the ground-sate[18-20]. Therefore the M-state and the following intermediates would be the signalling state. Later EPR experiments with higher time resolution[21] allowed to correlate the formation of the signalling state to the UV-spectroscopically silent transition between the early and the late M intermediate.

The mechanism of signal transduction from the receptor to the inner cell is an object of interest in recent investigations, which still remains unclear. The sensory rhodopsins form a stable complex with a transducer molecule in the cell membrane. Yao and Spudich (1992)[22] identified an open reading frame encoding for a putative transducer protein. This HtrI protein[23] consists of two transmembrane helices and a huge cytoplasmic domain. It was shown later that the genes encoding for *Hs*SRII and *Np*SRII as well as their respective transducer proteins are located together in one operon[23,24]. All identified transducers are highly homologous to the eubacterial chemoreceptors, like the aspartate- (Tar) or the serine-receptor (Tsr) from *E.coli*, which leads to the assumption that both classes of proteins, the methyl accepting chemotaxis proteins (MCPs) and the Htrs, are structurally related as well. Especially for the cytoplasmic domain a remarkable homology between both classes is found.

The signal transduction observed in prokaryotic chemotaxis is analogous to archeabacterial phototaxis. Rudolph and Oesterheld[25,26] were able to show that a signalling cascade similar to the eubacterial system exists in *H.salinarum*. A deletion of the archeabacterial Che-operon that was named according to the prokaryotic system proofed that the identified phosphorylation cascade is involved in the phototaxis signal transduction: This deletion led to the loss of the chemo- as well as the phototactic properties of the bacterium. A schematic representation of the signal transduction components is given in figure 1.3.

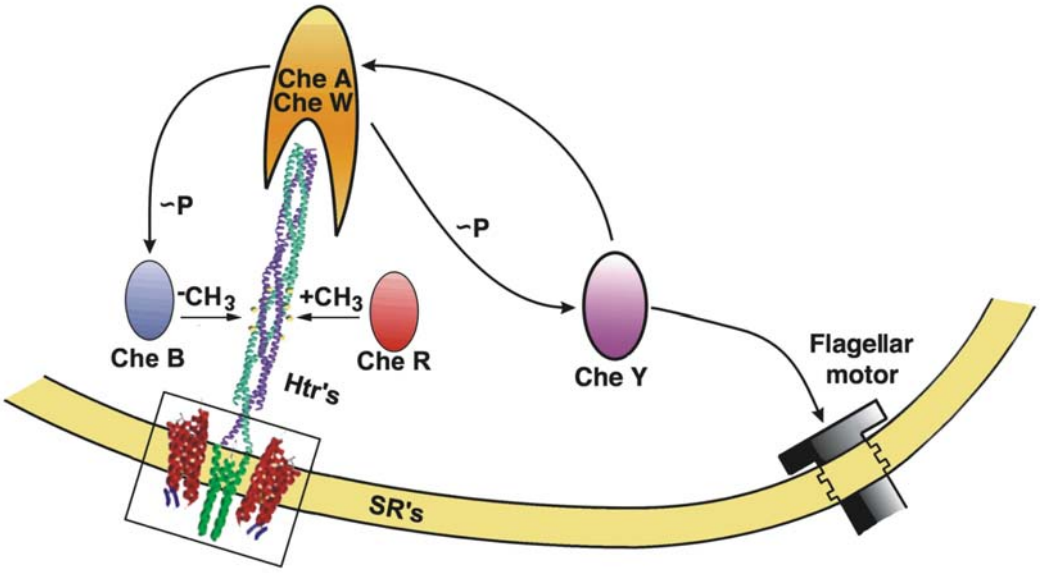


Figure 1.3. Illustration[9] (taken from Gordeliy, 2002) of the two-component signalling cascade including a model of the receptor-transducer complex, the homo-dimeric histidine kinase CheA, the scaffolding protein CheW and the receptor, the aspartate kinase Che Y and the methyltransferase/methylesterase CheR respectively CheB.

The scaffolding protein CheW mediates the interaction between the histidine kinase CheA and the cytoplasmic domain of the chemoreceptor/transducer. In case of a "negative" signal, namely binding of a repellent molecule or activation of the photophopic receptor, CheA is activated and therefore autophosphorylated. The phosphate is then transferred to the response regulator CheY, an aspartate kinase. Only phosphorylated CheY is able to bind to the FliM proteins of the flagellar motor. According to the so-called stochastic model[27] this binding increases the energy level of the clockwise turning state of the motor and simultaneously decreases the energy level of the counter-clockwise state. The more of the 30 binding sites are occupied by phosphorylated CheY the more the equilibrium is shifted to the clockwise rotation state of the flagellar motor leading to a tumbling motion of the *E.coli* bacterium. A "positive" signal inhibits the auto-phosphorylation of CheA, the phosphorylated CheY level decreases due to an auto-dephosphorylation activity of CheY and due to the activity of the phosphatase CheZ. This leads to a decreasing occupancy of the FliM proteins and a favoured counter-clockwise rotation of the flagellar motor. As a result a continuous swimming motion of the bacterium is observed.

CheA can phosphorylate as well the methylesterase CheB that subsequently demethylates the chemoreceptor protein. This demethylation decreases the auto-

phosphorylation activity of CheA and is therefore part of a feedback mechanism. This allows the bacterium to maintain the capability to direct its motion to more favourable conditions independent of the local concentration of a repellent or attractive substance or the local intensity of a repellent or attractive intensity of light.

The analogy between the two signal transduction pathways of photo- and chemotaxis, allows transferring the knowledge about the better-understood conformational changes of the chemoreceptors to the less understood conformational changes in the phototactic transducers. The question of the conformational changes that are involved in transmembrane signal transduction is intensively discussed. Chervitz and Falke preferred based on X-ray crystallography, ^{19}F NMR investigations and many disulfide cross-linking experiments with the aspartate receptor from *E.coli* the so-called "swinging-piston"-model. In this model the binding of aspartic acid induces a displacement of helix α4 (TM2) relative to the other helices and orthogonal to the plane described by the membrane. They observed a displacement of approximately 1.6 Å into the direction of the cytoplasm in combination with a tilt of the helix of about 5°[28]. This hypothesis, is supported by EPR experiments, which suggest such a translational motion as well; however without evidences for an additional tilt[29,30]. Nevertheless these investigations do not exclude an additional swinging motion.

It remains unclear, if these conformational changes are transferred to the cytoplasm. The authors of the X-ray crystallography studies[28] as well as the authors of the EPR studies[29] assume a transduction of this "signal" along the whole cytoplasmic domain due to the rigidity of α–helices. A prerequisite for influencing the histidine kinase activity is therefore a motif that links the aspartate- and the CheA binding domain. Sequence analysis[31] and further cross-linking experiments[32] produced hints for the presence of two amphiphatic helices, nevertheless they did not yield any information about the spatial arrangement of these helices.

Weis et al.[33] used electron microscopy to examine ordered assemblies of the serine receptor in membrane extracts isolated from cells engineered to overproduce the receptor. The authors report that the measured receptor length is ca. 20 % shorter than that described for the model of Kim et al.[34] and suggest a more compact arrangement of the polypeptide in the linker region (figure 1.4), which provides a plausible explanation for the discrepancy.

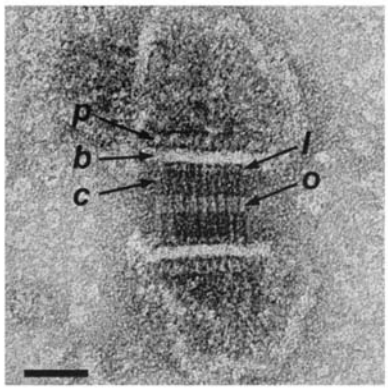


Figure 1.4 Electron micrograph[33] (taken from Weis, 2003) of an isolated serine receptor. bilayer specimen p, periplasmic ligand binding domain; b, membrane bilayer; c, cytoplasmic domain; o, zone of overlap; l, linker domain; scale bar, 25 nm. The measured receptor length is ca. 20 % shorter than that described for the model of Kim et al.[34] suggesting a more compact arrangement of the polypeptide in the linker region (l).

However, the transfer of results obtained for chemoreceptors towards Htrs is limited by the fact that sequence alignments of MCPs and Htrs revealed an up to 105 amino acid long sequence insertion into the transducers starting from the end of the second amphiphatic helix[19,31].

This insertion is reported for archeabacterial phototransducers HtrI and HtrII as well as for the Htrs IV-VI[26] from *H.salinarum* and for the MCPs from *Caulobacter crescentus*[35]. This leads to the question about how these additional amino acids influence the process of signal transduction and how signal domains with different structures undergo comparable conformational changes that regulate the activity of the bound CheA.

In summary the homology of Htr transducers to chemoreceptors and the requirement of Che proteins in the signalling cascade support a similar pathway in both haloarcheal phototaxis and bacterial chemotaxis. Nevertheless, the exact structural bases of the signal transfer from the phototaxis receptor to the transmembrane domain of its cognate transducer and hence the CheA binding domain is still unclear.

An important step in understanding the structural basis of signal transduction in archeabacterial phototaxis was the determination of the crystal structure of sensory rhodopsin II alone[36,37] and co-crystallized with a truncated transducer variant comprising the two transmembrane helices TM1 and TM2 and a short 30 amino acids part of the cytoplasmic linker domain[9].

The structure shows the receptor helices F and G tightly interacting with the transducer helices TM1 and TM2 within the membrane (figure 1.5, panel A).

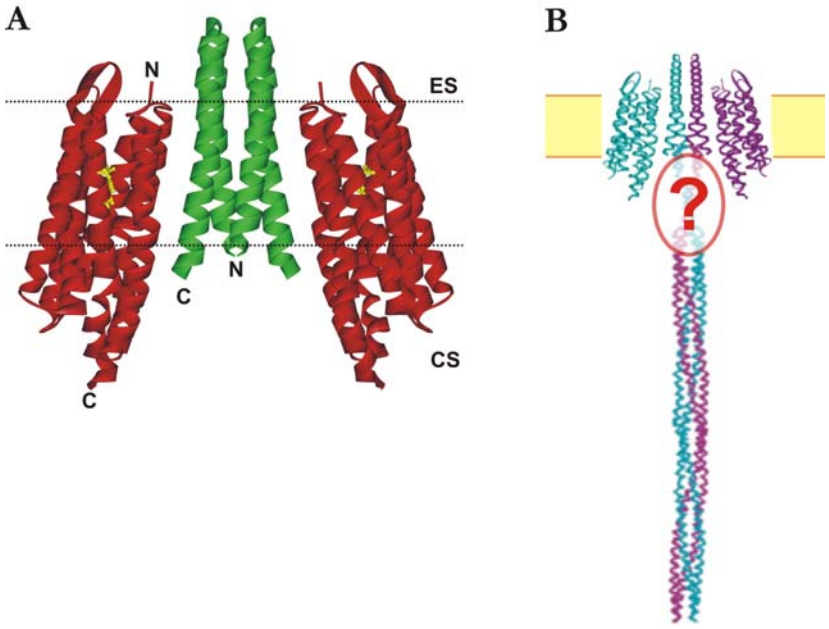


Figure 1.5. Panel A. Side view[9] (taken from Gordeliy, 2002) of the complex of SRII and HtrII from *N.pharaonis*. CS, cytoplasmic side; ES, extracellular side.
Panel B. Illustration[7] (taken from Klare, 2002) of the signalling complex from *N.pharaonis* including a modelled full-length transducer based on the X-ray crystal structure of the archeabacterial transmembrane part (Gordeliy et al.[9]) and of the cytoplasmic part of the aspartate receptor from *E.coli* (Kim et al.[34]). The unknown structure of the linker domain is hidden by the question mark.

This recent study and the observation that the SRII activation causes an outward movement of helix F[19,20] results in a displacement of transducer helix TM2[9]. This motion of TM2 is likely to be the common determinant of CheA kinase activity modulation.

This view is further supported by experiments with chimeric fusion proteins[10]. These chimeras contained the SRII fused to the transmembrane domain of the transducer HtrII, in which post-TM2 domains were replaced with cytoplasmic signalling and adaptation domains of *E.coli* chemotaxis receptors. These constructs were able to mediate a phototaxis responses in *E.coli* cells, which was shown with *in vivo* swimming experiments and with *in vitro* phosphorylation assays[38].

In summary all these studies[9,10,19,20,38] provide a detailed insight into the structural and functional organisation of the transmembrane signal-generating domain of the receptor-

transducer unit and the cytoplasmic signalling and adaptation domain. Nevertheless, until now almost no structural information is available about the linker domain between these two units. By transmitting the conformational changes of the transmembrane domain to the cytoplasmic signalling kinase and methyl-acceptor domains, the linker has an important role in signal uptake and propagation.

Interestingly, a structural linker element between a sensor and a transmitter module is widely found in histidine kinases, adenylyl cyclases, methyl-accepting proteins and phosphatases, therefore these linkers were named HAMP domains[39]. Although the primary sequence homology among predicted HAMP linkers is low, they all share similar helix-turn-helix folds based on secondary structure prediction[40]. Mutational[40] and cysteine scanning[32] analysis revealed biased signalling modes, suggesting that the linker plays an active role in signal transduction rather than being just a simple structural connecting element. The wide distribution of the HAMP linkers, not only in different sensors but also in all three kingdoms of life[39], eukaryotes, archea and eubacteria, suggests that they share a common mechanism for signal transduction.

Altogether the exact mechanism of the archeabacterial phototaxis still remains mysterious particularly in the area concerning signal propagation from the receptor unit to the signalling domain. New insights in this field are not only important for understanding phototaxis but can also contribute to the understanding of signal transduction in general.

The aim of this project was to build the foundation for further structural investigations of the photophobic receptor-transducer-complex from *N.pharaonis*. Crystallography[9] yielded insights into the structural and functional organisation of the transmembrane domains of the complex components. Nevertheless, it did not provide any information about the structure of the C-terminal cytoplasmic fragment. In order to determine its structure, magic angle spinning solid state NMR (MAS ssNMR) spectroscopy would be applied. A semi-synthetic access to a domain-specific isotope labelled HtrII protein would be established to improve the spectral resolution of applied MAS ssNMR.

In a first step special attention was paid to the generation of an isotope labelled peptide comprising the cytoplasmic part of the transducer with an N-terminal cysteine, which was subsequently fused to the recombinant transmembrane domain of the transducer protein using the Expressed Protein Ligation method. Furthermore the obtained proteins were characterised using a receptor binding assay and first NMR experiments.

In addition the influence of the formation of the receptor-transducer complex on the process of signal transduction in this functional unit would be investigated applying steady state and time-resolved rapid-scan FTIR spectroscopy.

2. Material and Methods

2.1 Chemicals

All chemicals were purchased from following companies with highest purity available:

Supplier	Compound
Amersham Pharmacia (Freiburg, Germany)	GSH-Sephadex Fast Flow, low molecular weight protein ladder, pGEX-2T1 and -4T1 vectors
Applichem (Darmstadt, Germany)	ammonium sulfate, methanol, IPTG
Baker (Griesheim, Germany)	ammonium chloride, chloroform, diethyl ether, ethanol, hydrochloric acid, magnesium sulfate heptahydrate, potassium dihydrogenphosphate, 2-propanol, sodium chloride, sodium hydroxide, urea,
Boehringer Ingelheim (Mannheim, Germany)	Expand high fidelity DNA polymerase
Calbiochem (Darmstadt, Germany)	DDM, GdmHCl
Fluka (Neu-Ulm, Germany)	DIEA, disodium molybdate dihydrate, p-cresol, L-histidine, manganese(II) chloride tetrahydrate, nickel(II) chloride hexahydrate, TCEP
euriso-top (Saarbrücken, Germany)	u-[^{13}C] D-glucose, [^{15}N] ammonium chloride, [D] chloroform, [D_3] methanol OH, [D] formic acid OH
Gibco (Eggenstein, Germany)	yeast extract, peptone 140, T4 DNA ligase, 1 kbp DNA ladder
Gerbu (Gaiberg, Germany)	ampicillin, kanamycin sulfate, EDTA
Merck (Darmstadt, Germany)	ammonium persulfate, calcium chloride dihydrate, copper(II) chloride dihydrate, disodium hydrogenphosphate, ferric(II) sulfate heptahydrate, HF, imidazole
New England Biolabs (Schwalbach, Germany)	DNA restriction enzymes, pTXB1 vector
Novabiochem (Schwalbach, Germany)	BOC protected amino acid building blocks and resins for solid phase peptide synthesis, HBTU

Promchem (Wesel, Germany)	acetonitrile
Qiagen (Hilden, Germany)	plasmid purification kit, DyeEx Spin Kit, gel extraction kit, Ni-NTA Superflow
Roth (Karlsruhe, Germany),	bis-tris, D_2O, TFA
Serva (Heiselberg, Germany)	agarose, boric acid, bromphenol blue, Coomassie blue R250, Triton X-100, xylencyanol,
Sigma (Steinheim, Germany).	ε-amino-n-caproic acid, cobalt(II) chloride hexahydrate, α-cyano-4-hydroxy cinnamic acid, ethidium bromide, D-glucose, 1-mono-oleoyl-*rac*-glycerol

Table 2.1. Chemicals.

2.2 General instrumentation

Incubation of cells was performed under aerobic conditions in an incubator from New Brunswick (Nürtingen, Germany). For electroporation a GenePulser from Bio-Rad (Munich, Germany) with electroporation cuvettes from Invitrogen (Leek, Netherlands) was used. Centrifugation usually was performed using eppendorf bench top centrifuges 5 415 C, D (Cologne, Germany) or Allegra X-15R centrifuge with SX4750A rotor (Beckman Coulter, Palo Alto, CA, USA). Preparative ultracentrifugation was carried out in a L-80 XP Ultracentrifuge with 45Ti/70Ti rotors (Beckman Coulter, Palo Alto, CA, USA). For cell harvesting an Avanti J-20 XP centrifuge with JLA 8.1 000 rotor (Beckman Coulter, Palo Alto, CA, USA) was used. Cell lysis was done in a Microfluidizer M-110S from Microfluidics Corporation (Newton, MA, USA). Polymerase chain reactions were performed using OmniGene Temperature Cycler with lid heating and in-tube temperature control from Hybaid (Heidelberg, Germany). SDS polyacrylamide gel electrophoresis was carried out on Mini-Protean 3 systems from Bio-Rad (Munich, Germany). High performance liquid chromatography was performed using System Gold facilities from Beckman Coulter (Palo Alto, CA, USA) with either analytical Prontosil 120-5 C18 column (Bischoff, Leonberg, Germany) or semi-preparative Ultrasphere 120-5 C18 from Beckman Coulter (Palo Alto, CA, USA). Measurements of pH were done with a PHM 62 Standard pH-Meter from Radiometer

(Copenhagen, Denmark). Deionized water (ddH_2O) was made with an apparatus from Milipore (Eschborn, Germany). Dialysis tubes MWCO 3.5 kDa and 12-14 kDa were from Spectrum (Gardena, CA, USA). Mass spectra were measured with a LCQ mass spectrometer from Finnigan (Bremen, Germany) or a Voyager-DE Pro MALDI-TOF mass spectrometer from Perseptive Biosystems (Weiterstadt, Germany).

2.3 Bacteria, plasmids and oligodesoxynucleotides

E.coli X1-Blue:	F'::Tn 10 *pro*A^+B^+ *lac*I^q Δ(*lacZ*)M15/*recA1 endA1* g*yrA96*(Nal^r) *thi hsdR17* (r_k^- m_k^+) *supE44 relA1 lac* [41]
E.coli BL21(DE3):	F^- *ompT*[lin] $hsdS_B(r_B^-m_B^-)$ λ(DE3) [42]
pET19btev+pARG	provided by A.J. Scheidig[43,44]
pET27bmod psopII his[19]	Wegener et al., 2000
pET27bmod phtrII114[19]	Wegener et al., 2000
pET27bmod phtrII114 G84C	provided by J.P. Klare
pET27bmod phtrII157[19]	Wegener et al., 2000
pGEX-2T-1/4T-1	Amersham Pharmacia (Freiburg, Germany)
pRPS1	provided by R.P. Seidel[23]
pTXB1	New England Biolabs (Schwalbach, Germany)[45]
pTXB1his	modified vector from New England Biolabs with a histidine-tag between Mxe intein and CBD (provided by R.P. Seidel)

The following oligodesoxynucleotides were purchased from MWG Biolabs (Ebersberg, Germany):

htr_ndeI_rev	5'-CAA TAA CAC CAT ATG TCG CTG AAC GTA TCA CG-3'
sapI_htr_for	5'-AGC GAG GCG GCT CTT CCG CAG CCC AGC GTG GCA G-3'
TEV_htr_for	5'-CGC GGA TCC GAA AAC CTT TAT TTT CAG TGT GAC ACC GCC GCC TGC CTT TC-3'
Xa_htr_rev	5'-CGG AAT TCT CAT CAC TCG CGA CGG GTC TCA AGC TCG ACA TC-3'

2.4 Solid phase peptide synthesis

The peptides HtrII(84-109)NBD G84C, HtrII(84-113) and HtrII(84-113) G84C were made by solid phase peptide synthesis[46] using the *in situ* neutralization and HBTU activation protocol for BOC protected amino acids[47]. The peptides were synthesized on a preloaded BOC-amino acid-OCH_2-phenylacetamidomethyl (PAM) solid support. Reactive side chains of tri-functional amino acids were protected as follows: Arg(Tos), Asp(OcHx), Cys(pMeBzl), Glu(OcHx), His(DNP), Lys(2-Cl-Z), Ser(Bzl) and Thr(Bzl). The side chains of all other amino acids were unprotected. The syntheses were usually performed in 0.2 mmole scale in glass vessels with filter funnel (pore-size 2) and vacuum connection. The preloaded resin was swollen for 20 minutes in DMF. The N-terminal BOC protecting group of the first amino acid was removed by rinsing the resin two times with neat TFA followed by incubation with TFA (2 x 1 min) and two subsequent washing steps with DCM (1 x 1 min) followed by DMF (1 x 1 min). The amino acids (2 mmoles, 10 equivalents) were activated with 3.8 ml of 0.5 M HBTU solution in DMF in presence of 1 ml DIEA immediately before coupling. For activation only 1.9 mmole HBTU were used to prevent blocking of the unprotected α-amino group of the N-terminal amino acid by a tetramethylguanidine hydrochloride moiety with excess HBTU[47]. All couplings were monitored by Kaiser tests[48] and repeated if necessary. For the non-natural amino acid L-BOC-Dab(Fmoc)-OH the coupling time was prolonged from 12 min to 1 hour. In order to achieve fluorescent labelling of the Dab residue, the orthogonal Fmoc-protecting group of the γ-amino moiety was removed after completion of the synthesis by rinsing the resin twice with 2 ml 20 % piperidine in DMF followed by a 3 min and 30 min treatment with 20 % piperidine in DMF. After a 1 min washing step with DMF (1 x 1 min) deprotection was monitored by the Kaiser test. The unprotected γ-amino-moiety was reacted with 1 mmole NBD chloride in the presence of 0.5 mmole DIEA for 30 min. Subsequently the resin was washed with DMF and a ninhydrine test was performed to verify successful fluorescence labelling.

2.4.1 Deprotection and HF cleavage

Before removing the peptides from the solid support, all DNP protecting groups of histidine residues were cleaved off. Therefore the resin was incubated 2 x 1 min with a solution of 21 ml DMF, 6 ml β-mercaptoethanol and 3 ml DIEA followed by two washing steps with DMF and DCM. The N-terminal BOC group was then removed according to the above-mentioned protocol and the resin was washed with DCM and dried in vacuum. In order to cleave the synthesized peptide from the solid support, the resin was treated with neat HF at 0 °C in presence of 5 % p-cresol as a scavenger[49]. After removing HF in vacuum the crude peptide was precipitated with ice-cold diethyl ether, transferred on a glass filter funnel and washed with ice-cold diethyl ether. The precipitate was dissolved in 50 % acetonitrile in water containing 0.1 % TFA, separated from the polymer support by filtration and lyophilized.

2.4.2 Peptide purification

All peptide purifications were performed using reversed phase high performance liquid chromatography (RP-HPLC). Therefore the lyophilized crude peptide was dissolved in 6 M guanidine hydrochloride, 100 mM sodium phosphate, pH 8.0 and loaded on a semi-preparative C18 column (Beckman Coulter, Palo Alto, CA, USA). Elutions were performed using linear solvent gradients starting with 25 % solvent B (acetonitrile, 0.08 % TFA) in solvent A (water, 0.1 % TFA) to 38 % B in 13 min, monitoring absorbance at 214 nm (and 480 nm for the NBD-labelled peptide). Fractions containing the desired peptide were pooled, lyophilized and stored at –20 °C in the dark.

2.5 Biological methods

2.5.1 Cell culture

E.coli XL1 blue bacteria were cultured in LB medium (10 g peptone, 5 g yeast extract, 10 g sodium chloride per litre ddH_2O, pH 7.2). Ampicillin or kanamycin was added as selection markers in 100 µg/l respectively 50 µg/l concentrations. Cell densities were determined by measuring the turbidity of the medium at 578 nm.

2.5.2 Agarose gel electrophoresis

DNA was separated by horizontal gel electrophoresis using gels with 0.7 % to 2 % agarose in TBE buffer (89 mM TrisHCl, 89 mM boric acid, 0.9 mM EDTA, 0.6 mg/l ethidium bromide, pH 8.9). Before separation, the samples were mixed with 3x sample buffer (TBE buffer, 10 % ficoll, 0.025 % bromphenol blue, 0.025 % xylencyanol).

2.5.3 Purification and isolation of DNA fragments

The desired DNA fragment was cut out of the agarose gel and extracted using the QIAquick gel extraction kit from Qiagen (Hilden, Germany).

2.5.4 Isolation of plasmid DNA

Preparation of plasmid DNA was performed using a modified plasmid midi-prep-kit (Qiagen, Hilden, Germany). After cell lysis and protein precipitation the cleared cytoplasmic supernatant was incubated with 10 µg/ml Rnase A for one hour at 37 °C and subsequently purified using the provided anion exchange column. The purified plasmid DNA was additionally precipitated with ice-cold ethanol and washed with 70 % ethanol in water.

2.5.5 Polymerase chain reaction (PCR)

The amplification of DNA fragments was carried out using the polymerase chain reaction[50] using 10 ng DNA, 10 pmole 5'- and 3'-oligodesoxynucleotides, 5 nmoles dNTPs and 1.7 U expand high fidelity polymerase in 50 µl scale. Thermocycling was performed in a thermocycler with lid heating and in tube temperature control. The polymerase was added after a first denaturing step at 95 °C for 3 min and followed by a cooling step on ice. Optimal hybridization temperature was calculated according to the

base composition of the oligodesoxynucleotides[51]. The thermocycling for the different amplifications was chosen as follows:

htr(1-83)	20x (95 °C for 15 sec, 72 °C for 70 sec)
tevhtr(84-114) G84C	5x (95 °C for 15 sec, 45 °C for 30 sec, 72 °C for 30 sec)
	15x (95 °C for 15 sec, 65 °C for 30 sec, 72 °C for 30 sec)

The first denaturing step was extended to 1 min at 95 °C.

2.5.6 DNA sequencing

DNA sequencing was performed using the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit and a DNA-Sequencer Model 373 (Applied Biosystems, Weiterstadt, Germany). Sample purification after thermocycling was purified using the DyeEx Spin Kit (Qiagen, Hilden, Germany) after the addition of Dextran blue.

2.5.7 DNA restriction

DNA (0.2 to 1 µg) was incubated in presence of the corresponding restriction enzymes in 10 µl scale for one hour at 37 °C. The amount of used enzyme was calculated according to the standard given by the manufacturer. In order to prevent unspecific DNA cleavage, glycerol concentrations above 5 % were avoided.

2.5.8 DNA ligation

DNA ligation was carried out using 100 ng digested vector and 3 molar equivalents of digested DNA fragment in the presence of 1 U T4 DNA ligase buffered with T4 DNA ligase buffer (50 mM TrisHCl, 10 mM $MgCl_2$, 1 mM ATP, 1 mM DTT, 5 % PEG 8 000, pH 7.6) in 20 µl scale. After 3 h of incubation at 37 °C the reaction mix was diluted with 3 volumes of sterile ddH_2O.

2.5.9 Transformation

For transformation of plasmid DNA into *E.coli* cells electroporation was used[52]. 75 µl suspension of electro-competent cells[52] containing 10 % glycerol were mixed with 20 µl of diluted DNA ligation mix on ice in an electroporation cuvette. After an electrostatic discharge (2 mm electrode gap, 800 Ω ohmic resistors, 1.5 kV voltages and 25 µF capacities) the mixtures were suspended in 0.5 ml antibiotic free LB medium and incubated for 1 h at 37 °C and 150 rpm. Selection was carried out on either 100 mg/l ampicillin or 50 mg/l kanamycin containing agar plates.

2.6 Biochemical methods

2.6.1 Protein expression and purification

2.6.1.1 Htr-MxeCBD fusion protein

As pre-culture 2 ml LB medium, 2 µl 100 g/l ampicillin were inoculated with a single colony BL21(DE3) pTXB1hishtr and shaken at 37 °C and 150 rpm for approximately 8 h. For expressions in 12 l-scale 250 ml overnight culture (250 LB, 0.25 ml 100 g/l ampicillin) were mixed with 0.25 ml of pre-culture (1/1000) and incubated overnight at 37 °C and 150 rpm. 6 x 2 l cultures (12 l LB, 12 ml 100 g/l ampicillin) in 5 l flasks with indications were inoculated with over night culture to a cell density of OD_{578}=0.05 and incubated at 37 °C and 150 rpm. The addition of 0.5 mM IPTG at OD_{578}=0.8-1.0 induced expression and incubation was continued at 37 °C and 150 rpm. Cells were harvested after 3 h, washed in approximately 100-150 ml cell washing buffer (150 mM NaCl, 25 mM NaP_i, 2 mM EDTA, pH 8.0) and stored at –80 °C.

2.6.1.1.1 Purification of HtrMxeCBD from membranes

Cells were suspended in cell washing buffer (150 mM NaCl, 25 mM NaPi, 2 mM EDTA, pH 8.0), 5 ml per gram cells, and lysed using a microfluidizer with 900 kPa pre-pressure. Membrane components were separated immediately by centrifugation (45Ti or 70Ti rotor, 45 000 rpm, 4 °C, 1 h). Supernatant was discarded and the obtained pellet homogenized (5 ml per gram cells) in solubilisation buffer A_{mem} (300 mM NaCl, 50 mM

NaPi, 5 mM imidazole, 2 % DDM, pH 8.0) and moderately stirred overnight at 4 °C. Insoluble components were removed by centrifugation (45Ti or 70Ti rotor, 45 000 rpm, 4 °C, 1 h).

The clear solubilisate was loaded on a 10 ml 1 x 15 cm Ni-NTA FPLC column equilibrated in B_{mem} (300 mM NaCl, 50 mM NaPi, 0.05 % DDM, pH 8.0). After a first washing step with B_{mem} the column was washed with 10 % D_{mem} (300 mM NaCl, 50 mM NaPi, 200 mM imidazole, 0.05 % DDM, pH 8.0) in B_{mem} (corresponds to 20 mM imidazole). Elution was achieved by washing with 50 ml 100 % D_{mem} with a flow-rate of 2 ml/min. HtrMxeCBD fusion protein containing fractions were pooled and concentrated up to 3 g/l using a 30 kDa MWCO and dialysed against 100 volumes B_{mem} without stirring (due to protein precipitation) using a 12-14 kDa MWCO tube at 4 °C. The yield was spectroscopically determined at a wavelength of 280 nm. Therefore a sample was diluted 1:10 with 6 M guanidine hydrochloride, 100 mM NaPi, pH 7.5 (A_{280}=0.993 corresponded to a concentration of 1 g/l, calculation: Expasy ProtParam algorithm[53])

2.6.1.1.2 Purification of Htr-MxeCBD from inclusion bodies

Cells were suspended in cell washing buffer (150 mM NaCl, 25 mM NaPi, 2 mM EDTA, pH 8.0), 5 ml per gram cells, and lysed using a microfluidizer with 900 kPa pre-pressure. Cytoplasma was removed immediately by centrifugation (45Ti or 70Ti rotor, 20 000 rpm, 4 °C, 1 h) and the obtained pellet was washed four times in approximately 50-60 ml cell washing buffer containing 0.25 % Triton X-100 followed by two washing steps with cell washing buffer to remove excessive detergent. Inclusion bodies were obtained as off-white pellets and were stored in 0.5 g portions at –80 °C.

Inclusion bodies (0.5 g) were dissolved in 10 ml denaturing buffer (6 M guanidine hydrochloride, 50 mM NaPi, 10 mM TCEP, 0.1 % DDM, pH 8.0) and stirred at RT over night. Insoluble components were removed by centrifugation (70Ti, 35000 rpm, 4 °C, 30 min). Refolding of the denatured protein was initiated by dialysis (3.5 kDa MWCO) against 500 ml folding buffer 8M (8 M urea, 300 mM NaCl, 50 mM NaPi, 20 mM L-histidine, 1 mM TCEP, 0.02 % DDM, pH 8.0) at RT for 2-3 h with moderate stirring.

The urea concentration was lowered stepwise by buffer exchange to 6 M, 4 M, 3 M and to 2 M. These dialysis steps were performed at 4 °C for 2-3 h without stirring. The dialysis against 4 M was prolonged over night. After each dialysis step solutions were centrifuged (70Ti, 35000 rpm, 4 °C, 30 min) to remove precipitated material.
The soluble protein was concentrated up to 30-40 g/l using a 10 kDa MWCO centriprep concentrator (Millipore, Bedford, MA, USA) at 3000 rcf. The Yield was spectroscopically determined at a wavelength of 280 nm. Therefore a sample was diluted 1:10 with 6 M guanidine hydrochloride, 100 mM NaPi, pH 7.5 (A_{280}=0.993 corresponds to a concentration of 1 g/l, calculation: Expasy ProtParam algorithm[53])

2.6.1.2 Tobacco etch virus protease (TEV protease)

As pre-culture 2 ml LB medium, 2 µl 100 g/l ampicillin, 2 µl 50 g/l kanamycin were inoculated with a single colony BL21(DE3) pET19b tev + pARG and incubated over day (ca. 8 h) at 37 °C and 150 rpm. For expressions in 12 l scale, 250 ml of an over night culture (250 ml LB, 0.25 ml 100 g/l ampicillin, 0.25 ml 50 g/l kanamycin) were mixed with 0.25 ml of pre-culture (1/1000) and incubated over night at 37 °C and 150 rpm.
6 x 2 l cultures (12 l LB, 12 ml 100 g/l ampicillin, 12 ml 50 g/l kanamycin) in 5 l flasks with indentations were inoculated with over night culture to a cell density of OD_{578}=0.05 and incubated at 37 °C and 150 rpm. The Addition of 0.1 mM IPTG at OD_{578}=0.8-1.0 induced protein expression and incubation was continued at 37 °C and 150 rpm. Cells were harvested after 4 h, washed in approximately 100-150 ml cell washing buffer (150 mM NaCl, 25 mM NaPi, 2 mM EDTA, pH 8.0) and stored at -80 °C.
Cell lysis was carried out in lysis buffer (500 mM NaCl, 25 mM NaPi, 10 % glycerine, 1 mM $MgCl_2$, 1 mM benzamidine pH 7.5) with 5 ml per gram cells using a microfluidizer with 900 kPa pre-pressure. Membrane components were separated immediately by centrifugation (45Ti or 70Ti rotor, 45 000 rpm, 4 °C, 1 h). The obtained clear supernatant was loaded with a 1 ml/min flow rate on a 10 ml Ni-NTA 1 x 10 cm FPLC column equilibrated in A1 buffer (500 mM NaCl, 25 mM NaPi, 10 % glycerine, 1 mM $MgCl_2$, pH 7.5). After a washing step with 10 % A2 buffer in buffer A1 (A1, 200 mM imidazole), elution was achieved by washing the column with 50 ml A2 buffer (flow-rate of 2 ml/min).

Fractions containing the desired TEV protease were pooled and concentrated up to 3-5 g/l using a 10 kDa MWCO and afterwards dialysed against 100 volumes TEV storage buffer (500 mM NaCl, 25 mM NaPi, 10 % glycerine, 5 mM DTT, pH 7.5) with gentle stirring using a 12-14 kDa MWCO tube at 4 °C. The yield was spectroscopically determined at a wavelength of 280 nm. Therefore a sample was diluted 1:10 with 6 M guanidine hydrochloride, 100 mM NaPi, pH 7.5 (A_{280}=1.162 corresponds to a concentration of 1 g/l, calculation: Expasy ProtParam algorithm[53]).

2.6.1.3 *Np*Htr(1-114), *Np*Htr(1-114) G84C, *Np*HtrII(1-157) and *Np*SRII

As pre-culture 2 ml LB medium, 2 µl 50 g/l kanamycin were inoculated with a single colony BL21 (DE32) pET27bmod phtr114, phtr114 G84C, phtr157 or psopII respectively and shaken over day (ca. 8 h) at 37 °C and 150 rpm. For expressions in 12 l scale, a 250 ml of an over-night culture (250 ml LB, 250 µl 50 g/l kanamycin) were mixed with 0.25 ml of pre-culture (1/1000) and incubated over night at 37 °C and 150 rpm. 6 x 2 l cultures (12 l LB, 12 ml 100 g/l kanamycin) in six 5 l flasks with indentations were inoculated with over night culture to a cell density of OD_{578}=0.05 and incubated at 37 °C and 150 rpm. The addition of 1 mM IPTG at OD_{578}=0.8-1.0 induced expression and incubation was continued at 37 °C and 150 rpm. Cells were harvested after 2.5 h, washed in approximately 100-150 ml cell washing buffer (150 mM NaCl, 25 mM NaPi, 2 mM EDTA, pH 8.0) and stored at -80 °C.

Purification protocol was performed as introduced by Hohenfeld[54] and Wegener[19]. Cells were suspended in cell washing buffer (150 mM NaCl, 25 mM NaPi, 2 mM EDTA, pH 8.0) with 5 ml per gram cells, and lysed using a microfluidizer (900 kPa pre-pressure). The membrane components were separated immediately by centrifugation (45Ti or 70Ti rotor, 45 000 rpm, 4 °C, 1 h). The supernatant was discarded and the obtained pellet homogenized (5 ml per gram cells) in solubilization buffer A_{mem} (300 mM NaCl, 50 mM NaPi, 2 % DDM, pH 8.0) and moderately stirred over-night at 4 °C.

Insoluble components were removed by centrifugation (45Ti or 70Ti rotor, 45000 rpm, 4 °C, 1 h). After addition of 5 mM imidazole the obtained clear solubilisate was loaded on a 15 ml Ni-NTA FPLC column (1 x 15 cm) equilibrated in B_{mem} (300 mM NaCl, 50 mM NaPi, 0.05 % DDM, pH 8.0) with a flow-rate of 1 ml/min. After a first washing

step with B_{mem} (flow-rate of 2 ml/min), the column was washed with 10 % D_{mem} (300 mM NaCl, 50 mM NaPi, 200 mM imidazole, 0.05 % DDM, pH 8.0) in B_{mem} (corresponds to 20 mM imidazole). Elution was carried out using 50 ml 100 % D_{mem} (flow rate of 2 ml/min).

Fractions containing the desired protein were pooled and diluted with buffer E_{mem} (10 mM TrisHCl, 0.05 % DDM) 1:4 (1:10 for *Np*SRII) and loaded with 1 ml/min on a 20 ml DEAE FPLC column (1.5x10 cm) equilibrated with buffer E_{mem}. The column was washed with 16 % (6 % for *Np*SRII) buffer G_{mem} (500 mM NaCl, 10 mM TrisHCl, 0.1 %DDM, pH 8.0) in E_{mem}. Elution was achieved by a washing step with 50 ml G_{mem} (flow-rate of 2 ml/min). The fractions containing the desired protein were pooled and concentrated up to 3-5 g/l using a 10 kDa MWCO. The yield was spectroscopically determined at 228.5 nm and 234.5 nm according to Ehresmann et al.[55] (at 498 nm, ε_{498}=40 000 cm·l·mol^{-1} for *Np*SRII).

2.6.2 Isotope labelling of peptides and proteins

For isotope labelling of proteins LB medium was substituted by synthetic minimal medium M9 (7.5 g disodium hydrophosphate dihydrate, 3.0 g potassium dihydrogenphosphate, 0.5 g sodium chloride, 0.25 g magnesium sulfate heptahydrate, 14 mg calcium chloride dihydrate, pH 7.1). Medium was supplemented after autoclaving with 10 ml per litre trace elements (9 ml 10xSL4, 10 ml SL6, ad 100 ml ddH_2O, filtered sterile). One litre SL6 stock solution contained 100 mg zinc (II) sulfate, 30 mg manganese (II) chloride, 300 mg boric acid, 200 mg cobalt (II) chloride hexahydrate, 10 mg copper (II) chloride dihydrate, 20 mg nickel (II) chloride hexahydrate and 30 mg disodium molybdate dihydrate. For preparation of 100 ml 10xSL4 stock solution 500 mg EDTA were dissolved in 80 ml ddH_2O and 200 mg ferric (II) sulfate heptahydrate were added. The volume was adjusted to 90 ml. 10xSL4 was always freshly prepared due to the oxidation of Fe (II) to Fe (III).

2.6.2.1 *Np*HtrII(84-114) G84C peptide

As pre-culture 2 ml M9 medium, 20 µl 100 g/l ammonium chloride, 10 µl 40 % D-glucose, 2 µl 100 g/l ampicillin were inoculated with a single colony BL21 (DE32) pGEX-2T1-tevhrt and incubated over day at 37 °C and 150 rpm.

For expressions in 2 l scale, 100 ml overnight culture (100 ml M9, 1 ml 100 g/l ammonium chloride, 0.5 ml 40 % D-glucose, 100 µl 100 g/l ampicillin) was mixed with 0.1 ml of pre-culture (1/1 000) and incubated overnight at 37 °C and 150 rpm. A 2 litre culture (2 l M9, 4 g solid [^{13}C] D-glucose, 2 g solid [^{15}N] ammonium chloride, 100 g/l ampicillin) in a 5 l flask with indications was inoculated to a cell density of OD_{578}=0.05 and incubated at 37 °C and 150 rpm. The addition of 0.5 mM IPTG at OD_{578}=0.8-1.0 induced expression and incubation was continued at 37 °C and 150 rpm. Cells were harvested after 3 h, washed in approximately 50 ml cell washing buffer (150 mM NaCl, 25 mM NaPi, 2 mM EDTA, pH 8.0) and stored at –80 °C if necessary.

Cell lysis was carried out in cell washing buffer (150 mM NaCl, 25 mM NaPi, 2 mM EDTA, pH 8.0) with 5 ml per gram cells using a microfluidizer with 900 kPa pre-pressure. Membrane components were separated immediately by centrifugation (45Ti or 70Ti rotor, 45 000 rpm, 4 °C, 1 h). The obtained clear supernatant was loaded on a 15 ml GSH-Sephadex Fast Flow FPLC column (1.5 x 10 cm) equilibrated with cell washing buffer with a flow-rate of 1 ml/min. After a washing step with cell washing buffer, elution was achieved by washing with 50 ml cell washing buffer containing 10 mM reduced glutathion 8flow-rate of 2 ml/min).

Fractions containing the desired GST-tevHtr fusion protein were pooled and concentrated up to 10-15 g/l using a 10 kDa MWCO and afterwards dialysed against 100 volumes TEV cleavage buffer (150 mM NaCl, 10 mM TrisHCl, 1 mM TCEP, 0.5 mM EDTA) with gentle stirring using a 12-14 kDa MWCO tube at 4 °C.

The yield was spectroscopically determined at a wavelength of 280 nm. Therefore a sample was diluted 1:10 with 6 M guanidine hydrochloride, 100 mM NaPi, pH 7.5 (A_{280}=1.392 corresponds to a concentration of 1 g/l, calculation: Expasy ProtParam algorithm[53]).

TEV cleavage of the fusion protein to release the isotope labelled peptide *Np*HtrII(84-114) G84C was performed at RT using 1 mg TEV protease per 50 mg GST-tevHtr. The cleavage mixture was gently agitated. The proteolytic digest was monitored using analytical C18 reversed phase HPLC with a linear solvent gradient

starting with 5 % B (acetonitrile, 0.08 % TFA) to 65 % B in A (ddH_2O, 0.1 % TFA) in 30 min monitoring absorbance at 214 and 280 nm.

After almost quantitative cleavage, the desired peptide was separated from GST by filtration using a 10 kDa MWCO concentrator at 3000 rcf. The concentrated solution containing the GST-tag was washed four times with 15 ml TEV cleavage buffer. The flow-through was lyophilized and dissolved in approximately 5 ml buffer containing 6 M guanidine hydrochloride, 100 mM NaPi, pH 7.5.

The resulting clear solution was loaded in 2.5 ml portions on a semi-preparative C18 column (Beckman Coulter, Palo Alto, CA, USA) equilibrated with 5 % solvent B (acetonitrile, 0.08 % TFA) in A (water, 0.1 % TFA). Elution of the desired peptide was carried out using a linear solvent gradient from 5 % to 65 % B in A in 30 min monitoring the absorbance at a wavelength of 214 nm. Fractions containing the desired peptide were pooled, lyophilized and stored at –20 °C in darkness.

2.6.2.2 *Np*HtrII (1-114), *Np*HtrII (1-157) and *Np*SRII

As pre-culture 2 ml M9 medium, 20 µl 100 g/l ammonium chloride, 10 µl 40 % D-glucose, 2 µl 50 g/l kanamycin were inoculated with a single colony BL21 (DE32) pET27bmod phtr114, phtr157 or psopII respectively and shaken over day at 37 °C and 150 rpm (ca. 8 h).

For expressions in a 2 l scale 100 ml over night culture (100 ml M9, 1 ml 100 g/l ammonium chloride, 0.5 ml 40 % D-glucose, 100 µl 50 g/l kanamycin) were mixed with 0.1 ml of pre-culture (1/1000) and incubated over night at 37 °C and 150 rpm.

A 2 litre culture (2 l M9, 4 g solid [^{13}C] D-glucose, 2 g solid [^{15}N] ammonium chloride, 50 g/l kanamycin) in a 5 l flasks with indications was inoculated to a cell density of OD_{578}=0.05 and incubated at 37 °C and 150 rpm. The addition of 1 mM IPTG at OD_{578}=0.8-1.0 induced expression and incubation was continued at 37 °C and 150 rpm.

Cells were harvested after 2 h (1.5 h in case of Htr114), washed in approximately 50 ml cell washing buffer (150 mM NaCl, 25 mM NaPi, 2 mM EDTA, pH 8.0) and stored at -80 °C. Purification protocol was performed as above-mentioned (see section 2.6.1.3).

2.6.3 Isolation of polar lipids from *H.salinarum*

Samples of PM lipids were kindly supplied by A. Göppner. The applied method was introduced by Kates et al.[56] and optimised by using purified purple membranes instead of complete bacterial cells. The purple membrane (BR content 300 to 350 mg) was suspended in 200 ml 4 M NaCl. After addition of 500 ml methanol and 250 ml chloroform, the mixture was stirred under nitrogen atmosphere in darkness. The obtained suspension was centrifuged for 15 min at 3800 rcf in metal tubes. The yellow supernatant was pooled under nitrogen atmosphere and in the dark. The residual pellets were extracted once with 500 ml methanol and 250 ml chloroform.

The pooled supernatants were filtered using a paper filter (Whatman No.1) that was swollen for 2 h in methanol/chloroform 1:1. Afterwards 500 ml ddH_2O and 500 ml chloroform were added to the filtrate and extraction was performed using shaking. Phase separation was achieved over night in darkness under argon atmosphere.

The chloroform phase was separated and subsequently evaporated under reduced pressure at 30 °C in the dark using a rotary evaporator (Büchi, Essen, Germany). The obtained residue was suspended in 12 ml chloroform and centrifuged at 24 000 rcf, 0 °C for half an hour. The chloroform containing supernatant was evaporated again and the resulting residue was suspended in 45 ml acetone at 0 °C using ultrasonic pulses.

After the addition of 1 ml 10 % $MgCl_2$ in methanol, the suspension was centrifuged at 1 000 rcf (0 °C). The supernatant was discarded and the obtained pellet was extracted several times with 10 ml acetone until off-white extracts were obtained.

The residue was dried in vacuum, weighted and dissolved in methanol/chloroform 1:1 to obtain a 40 g/l solution. This solution was injected in darkness through a fine needle into 20 volumes of 50 mM NaPi, pH 8.0 under argon atmosphere. The obtained suspension was homogenized using ultrasonic pulses and subsequently lyophilized. The obtained powder was suspended in the original volume of the phosphate buffer and a stabile suspension (2 g/l) was obtained, which was stored at –80 °C.

2.6.4 Reconstitution in PM lipids

For the reconstitution in PM lipids protein solutions containing less than 0.5 % DDM were used. The solubilised protein or protein complex was mixed 1:1 (w/w) with lipids and incubated for 10 min at 4 °C with gentle agitation. Next, the NaCl concentration was adjusted to 1 M by the addition of solid sodium chloride, followed by a further incubation step of 10 min at 4 °C with gentle agitation. Subsequently absorber beads (Bio-Rad, Munich, Germany) were added (1 g/100 mg DDM) and this mixture was shaken overnight at 4 °C in darkness. A successful reconstitution was indicated by the transition of the initially clear mixture to a turbid suspension. After a short 5 sec ultrasonic pulse to achieve a desorption of reconstituted protein from the adsorber beads, the beads were separated by filtration and the turbid suspension obtained was centrifuged for 15 min (15 000 rcf, 4 °C). This pellet was washed twice and finally suspended either 20 mM NaPi pH 8.0 (MAS ssNMR samples) or in water (FTIR samples) to a final concentration of approximately 200-250 mM.

2.7 Analytical methods

2.7.1 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

For separation of proteins according to their apparent molecular weight a SDS-PAGE variant introduced by Schägger and von Jagow[57] was used. Vertical gels in 0.75 mm thickness were used in gel chambers from Bio-Rad. Samples were mixed with 1 volume 2x SDS sample buffer (6 % SDS, 35 % glycerol, 120 mM TrisHCl, 0.41 M mono-thiogylcerol, 0.05 % bromphenol blue, pH 8.0).

Samples with protein concentrations less than 0.04 g/l or containing GdmHCl were precipitated by the addition of either 1 volume of ice-cold 10 % aqueous TCA or 2 volumes of ice-cold acetone. The samples were subsequently incubated for 30 min on ice and centrifuged for 15 min at 15 000 rpm in a bench top centrifuge. The supernatant was discarded and the pellet washed with 1 volume ice-cold 70 % ethanol and dissolved in 10 µl 1x SDS sample buffer.

A mixture of phosphorylase b (97.0 kDa), albumin (66.0 kDa), ovalbumin (45 kDa), carboanhydrase (30.0 kDa), trypsin inhibitor (20.1 kDa) and α-lactalbumin (14.4 kDa) from Amersham Pharmacia Biotech (Freiburg, Germany) was applied as molecular

weight standard. Electrophoresis itself was performed using a constant electric current corresponding to starting voltage of 80 mV.

Gels were stained using Coomassie brilliant blue by boiling the gel 1 min in Coomassie staining buffer (0.1 % Coomassie Brilliant blue R250, 40 % ethanol, 10 % acetic acid) and incubated for 15 min with shaking. Destaining was performed by shaking the gel with a piece of absorbent paper tissue in Coomassie destaining buffer (10 % acetic acid, 5 % ethanol) over night. Alternatively staining was also carried out using the Silver Staining Plus Kit (Bio-Rad, Munich, Germany). For gel preparation the following solutions were used:

acrylamide bisacrylamide mixture	48 % acrylamide, 1.5 % bisacrylamide
gel buffer	3.0 M TrisHCl, 0.3 % SDS, pH 8.45
cathode buffer	0.1 M TrisHCl, 0.1 % SDS, 0.1 % tricine, pH 8.25
anode buffer	0.2 M Tris pH 8.9
stacking gel (4 %)	180 µl acrylamide bisacrylamide mixture, 0.6 ml gel buffer, 1.45 ml ddH_2O, 12.5 µl 10 % APS, 2 µl TEMED
separation gel (16.5 %)	1.25 ml acrylamide bisacrylamide mixture, 1.25 ml gel buffer, 1.25 ml ddH_2O, 19 µl 10 % APS, 10 µl TEMED

2.7.2. Blue-native polyacrylamide gel electrophoresis (BN-PAGE)

For monitoring the capability of truncated transducer variants to form a complex with SRII, blue-native-PAGE introduced by Schägger and von Jagow[57] was used. Vertical 10-20 % gradient gels in 1.5 mm thickness were used in gel chambers from Pharmacia (Upsalla, Sweden). The gradient was mixed at 4 °C using a gradient mixer with the concentrated acrylamide solution in the second reservoir. After 2 h the gel was polymerised at rt overnight.

Samples contained 25 µg SRII, the corresponding amount of transducer and less than 0.5 % DDM detergent. They were concentrated up to 25 µl using a 10 kDa MWCO, mixed with 12.5 µl 3x sample buffer and incubated overnight at 4 °C. The samples were injected below the cathode buffer that was filled into the sample pockets. Unused sample pockets were filled with 25 µl G_{mem}+12.5 µl 3x sample buffer in order to ensure a homogenous running front during electrophoresis.

Electrophoresis itself was performed using 150 V, 25 mA and 20 W until the running front reached the gradient gel and from there 500 V, 15 mA and 20 W. Gels were stained by the Coomassie Serva blue G dye contained in the cathode buffer. Destaining was carried out by replacing the cathode buffer by the corresponding buffer without Coomassie after the running front had passed half of the gradient gel. For gel preparation the following solutions were used:

acrylamide solution	48 % acrylamide, 1.5 % bisacrylamide
gel buffer	150 mM bis-trisHCl, 1.5 M ε-aminocaproric acid, pH 7.0
3x sample buffer	4.5 ml glycerol, 0.31 g bis-trisHCl, ad 10 ml ddH_2O, pH 7.0)
cathode buffer	50 mM trichine, 15 mM bis-tris, 0.02 % Coomassie Serva blue G, pH 7.0, filtered
anode buffer	50 mM bis-trisHCl, pH 7.0
stacking gel (4 %)	0.8 ml acrylamide solution, 3.33 ml gel buffer, 120 µl 10 % APS, 12 µl TEMED ad 10 ml ddH_2O
gradient solution (10 %)	3 ml acrylamide solution, 5 ml gel buffer, 3 g glycerol, 20 µl 10 % APS, 7 µl TEMED, ad 15 ml ddH_2O
gradient solution (20 %)	6 ml acrylamide solution, 5 ml gel buffer, 3 g glycerol, 70 µl 10 % APS, 7 µl TEMED, ad 15 ml ddH_2O

2.7.3 Mass spectrometry of membrane proteins

2.7.3.1 Matrix-assisted laser desorption ionisation (MALDI)

This method was adapted from Cadene et al.[58]. A saturated solution of α-cyano-4-hydroxy cinnamic acid in acetonitrile water 2:1 (sample matrix) was prepared. 20 µl of this solution were diluted with 60 µl isopropanol.

20 µl of this dilution were applied on a gold-coated sample plate and homogenously spread with the flank of a pipette tip over the complete surface and air-dried. This procedure was repeated 2-3 times and the homogenous, crystalline thin layer so obtained was wiped away with a soft paper tissue resulting a yellow shimmering ultra thin layer covering the entire plate. 1 µl detergent containing protein sample (≥1 g/l)

was diluted with 9 µl sample matrix. 1 µl of this dilution was applied onto the plate surface.

After a few seconds a yellow homogenous, crystalline layer was obtained. In case of needle formation or inhomogeneous layers the procedure was repeated. Before the sample was able to dry completely, the supernatant containing the majority of detergent and buffer salts, was removed by soaking with a soft paper tissue.

The sample was washed with a drop 1 % aqueous TFA and air-dried. The measurements were performed using the following instrument settings: linear operation mode, positive polarity, 25 000 V accelerating voltage, 93 % grid voltage, 0.3 % guide wire, 750 nsec extraction delay time, 20 Hz laser repetition rate and 1 500 Da low mass gate.

2.7.3.2 Electro spray ionisation (ESI)

A modified protocol, originally introduced by Hufnagel et al.[59] was used. 0.1 ml of micellar protein solution (≥1 g/l) was diluted with 15 ml desalting buffer (10 mM TrisHCl, 0.02 % DDM, pH 8.0) and concentrated up again to 0.5 ml using a 10 kDa MWCO membrane.

0.5 ml protein sample were mixed with 9.5 ml precipitator (10 ml acetone, 1 ml 25 % aqueous ammonium hydroxide, 100 µl 6 M aqueous TCA) and vortexed for 2 min. After incubation on ice for 30 min, the sample was centrifuged at 5 000 rpm (4 °C, 3 min) in a bench top centrifuge. The supernatant was discarded and the pellet subsequently washed twice with 90 % aqueous acetone, once with acetone and finally with n-hexane.

Finally the pellet was air-dried, dissolved in 5 µl of neat formic acid, diluted with 45 µl TFE and 200 µl ESI sample buffer (10 ml chloroform, 10 ml methanol, 3.5 ml ddH_2O).

2.8 Expressed protein ligation

2.8.1 EPL using Htr-MxeCBD intein precursor isolated from membranes

The purified Htr-MxeCBD intein fusion protein (see section 2.6.1.1.1) in 3 g/l concentration was mixed in 1.3:1 stoichiometry with the lyophilized C-terminal peptide segment containing an N-terminal cysteine using ultrasonic pulses and MESNA was added in 200 mM concentration.

The reaction mix was incubated with gentle agitation on a turning wheel. After 72 h of incubation the reaction mix was diluted 1:10 with dilution buffer (10 mM TrisHCl, 0.05 %DDM, pH 8.0) and loaded with a flow-rate of 1 ml/min on a 20 ml DEAE FPLC column (1.5 x 10 cm) equilibrated with dilution buffer. After a washing step with 16 % elution buffer (500 mM NaCl, 10 mM TrisHCl, 0.1 %DDM, pH 8.0) in dilution buffer. Elution was achieved by washing with 50 ml elution buffer with a flow-rate of 2 ml/min.

Fractions containing the desired product were pooled and concentrated up to 3-5 g/l using a 10 kDa MWCO. The yield was determined spectroscopically at 228.5 nm and 234.5 nm according to Ehresmann et al.[55].

2.8.2 EPL using Htr-MxeCBD intein precursor isolated from inclusion bodies

The refolded Htr-MxeCBD intein fusion protein (see section 2.6.1.1.2) in 30 g/l concentrations was mixed in 2:1 stoichiometry with the lyophilized C-terminal peptide segment containing an N-terminal cysteine using ultrasonic pulses and MESNA was added in 200 mM concentration. 67 mg 1-Mono-oleoyl-*rac*-glycerol portions were melted in a water bath at 40 °C and centrifuged for 5 sec at maximal speed in a bench top centrifuge. The molten lipids and the reaction mix were rapidly mixed by injecting 100 µl mix in one jet into each lipid drop. The obtained mixture became highly viscous and turbid at once and was centrifuged at 23 °C at 10 000 rpm in a bench top centrifuge for 1 min. Subsequently six 10 min centrifugation steps were carried out starting with 5 000 rpm, accelerating rotary speed after each step by 1 000 rpm ending with 10 000 rpm and turning the cups 90° clockwise after each step. The formation of the lipidic cubic phase was indicated by the transition from a turbid mixture to a homogenous, transparent gel. The lipidic cubic phase reaction mixtures were incubated

at 23 °C. After 3 days the samples were filled up to 1 ml with 10 % DDM, 10 mM TrisHCl, 5 mM DTT, pH 8.0, subsequently diluted 1:10 with dilution buffer (2 %DDM, 10 mM TrisHCl, 5 mM DTT, pH 8.0) and moderately stirred over night at 4 °C. The solubilisate was loaded with a flow-rate of 1 ml/min on a 20 ml DEAE FPLC column (1.5x10 cm) equilibrated with equilibration buffer (10 mM TrisHCl, 0.05 %DDM, pH 8.0). After a washing step with 16 % elution buffer (500 mM NaCl, 10 mM TrisHCl, 0.1 %DDM, pH 8.0), elution was achieved by washing with 50 ml elution buffer with a flow-rate of 2 ml/min. Fractions containing the desired product were pooled and concentrated up to 3-5 g/l using a 10 kDa MWCO. Yield was determined spectroscopically at 228.5 nm and 234.5 nm according to Ehresmann et al.[55].

2.9 Sample preparation for NMR spectroscopy

Lyophilized powders of isotope (un)labelled peptides were dissolved in 8 % D_2O in H_2O, 20 mM NaPi, and pH 6.5 in approximately 1 mM concentration using ultrasonic pulses and transferred into Shigemi NMR tubes (Allison Par, PA, USA) with 4.22 mm in diameter.. All measurements were performed at 25 °C using a Bruker 500 MHz (11.7 T) or a Varian 600 MHz (14.1 T) NMR spectrometer. The water signal was suppressed by selective pre-saturation during the preparation period and/or mixing time.

For MAS ssNMR experiments 5 mg membrane protein or membrane protein complex were reconstituted in PM lipids and washed with 20 mM NaPi pH 8.0 (see section 2.6.3). The membrane pellet obtained was manually transferred in several portions into a sample rotor of 4.0 respectively 2.5 mm in diameter using Transferpettors (Brandt, Oberursel, Germany). Centrifugation steps at 15 000 rpm in a bench top centrifuge for 1 min were performed between each transfer step to achieve a homogenous rotor packing. All measurements were performed using either a Bruker 400 MHz (9.3 T) or a Bruker 600 MHz (11.7 T) NMR spectrometer with MAS probe head by K Seidel of the working group of M. Baldus at the Max-Planck-Institute for biophysical chemistry (Göttingen, Germany) .

2.10 Sample preparation for FTIR spectroscopy

For FTIR experiments 1 mg membrane protein or membrane protein complex were reconstituted in PM lipids and washed with 20 mM NaPi pH 8.0 and suspended in pure ddH_2O (see section 2.6.3). The preparation of so-called sandwich-samples was performed by drying 1-1.5 nmol of reconstituted protein onto the lower sample window. The dried membranes were hydrated with 20 µl buffer (200 mM bis-tris-propane, 150 mM NaCl, pH 8.0). The second plane window squeezed the excess buffer into the circular groove, which formed a buffer reservoir ensuring a sufficient sample hydration. The thickness of the cuvette was between 2.5 and 4.5 µm. The static FTIR measurements were performed at 17 °C using a Bruker IFS 28 FTIR spectrometer with a Sens-IR Diamond Micro-ATR unit. All time-resolved rapid-scan measurements were carried out at 17 °C on a home-built FTIR spectrometer (Siebert,F., Freiburg, Germany) by I. Radu of the working group of F. Siebert at the Albert-Ludwigs-University (Freiburg, Germany).

3. Results

3.1 Expression and biochemical evaluation of *Np*HtrII(1-114) G84C

Gordeliy et al.[9] recently reported a crystallographic structure of the SRII/HtrII signalling complex. While detailed insights into the transmembrane interactions of the complex were obtained, unfortunately no electron density of the cytoplasmic HAMP linker domain of the transducer protein was observed.

In order to address the unknown structure of the cytoplasmic linker domain in *Np*HtrII, magic angle spinning solid-state NMR spectroscopy was applied. Although MAS techniques have greatly extended the sensitivity and the spectral resolution of solid-state NMR spectroscopy, it would not be possible to resolve the NMR signals of the entire SRII/HtrII-complex or even of the complete Htr114 construct with MAS ssNMR techniques due to limited spectral resolution. Therefore domain specific isotope labelling of only the cytoplasmic Htr114 C-terminus of unknown structure using the expressed protein ligation technique would reduce the problem of signal overlapping.

As ligation site in the EPL reaction Gly83-Gly84 was selected. This ligation junction was desirable in a semi-synthetic approach since all the residues of the putative cytoplasmic receptor binding site with unknown structure are located C-terminal to this site[60]. In addition Gly-83 at position –1 promised high *in-vitro* intein cleavage efficiencies[45] in combination with high ligation efficiencies[61].

In order to investigate, if a G84C mutation caused by a successful EPL reaction is functionally tolerated in a C-terminally truncated *Np*Htr114 construct, an *in vitro* binding assay using a recombinant *Np*Htr114 G84C mutant and the SRII receptor was performed. Therefore the gene encoding for the *Np*Htr114 G84C mutant with a seven-histidine purification tag at its C-terminus was cloned into the pET27bmod vector under the control of the T7 promoter. The corresponding clone was kindly supplied by J.P. Klare. This protein was expressed and purified using Ni-NTA affinity chromatography followed by DEAE ion exchange chromatography (figure 3.1, panel A).

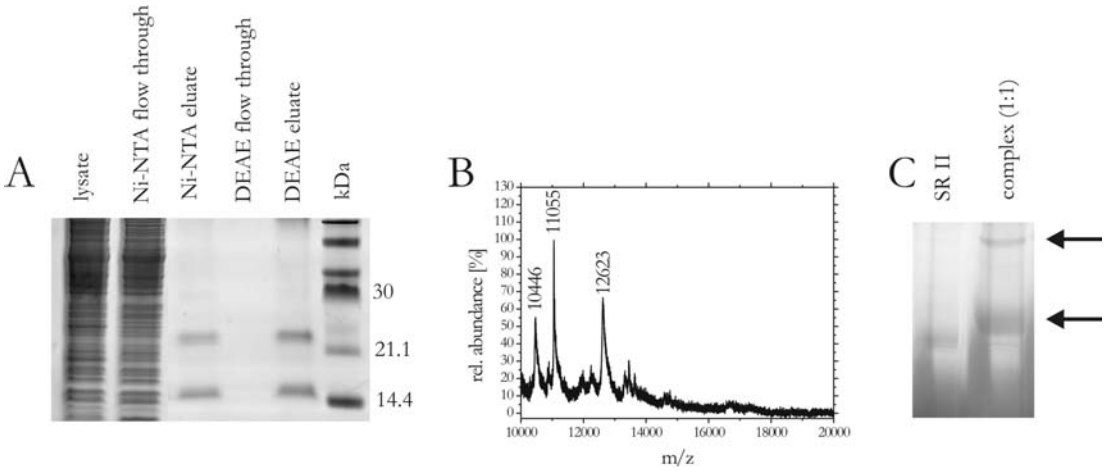


Figure 3.1. Purification and biochemical evaluation of *Np*HtrII(1-114) G84C.
Panel A: The SDS-PAGE analyzing purification procedure. The two protein bands for Ni-NTA respectively DEAE eluate indicated protein monomer (12.6 kDa) and dimer (25.2 kDa).
Panel B: MALDI-MS spectrum of the purified protein with. The signal with $(M+H)^+$=12623 Da referred to the target protein. Signals with 11055 amu and 10446 amu could not be assigned
Panel C: Blue-native PAGE monitoring the formation of a 1:1 SRII/Htr(1-114) G84C complex referring SRII alone.

The SDS-PAGE analyzing the purification of the G84C mutant showed two protein bands for the Ni-NTA and the DEAE-elution, respectively. The molecular weight indicated by the mobility of these bands is in a good agreement with the masses for a monomer (12.6 kDa) and a dimer (25.2 kDa) of the mutant. Dimer formation due to oxidative disulfide cross-linking of the Cys84 residues was excluded, since the SDS sample buffer used contained mono-thioglycerol as reducing agent. However, incubation of the purified protein with 1 mM TCEP before addition of SDS sample buffer significantly reduced the observed amount of dimerisation (results not shown). This leads to the conclusion that disulfide cross-linked Htr(1-114) G84C remains partially in the dimeric state after the addition of SDS sample buffer due to hydrophobic or ionic interactions even without intact disulfide bonds.

In order to identify Htr(1-114) G84C, MALDI-MS experiments were performed. The mass spectrum of the purified protein (figure 3.1, panel B) showed signals with 12623 amu, 11055 amu and 10446 amu. While the signals with 11055 amu and 10446 amu could not be assigned, the observed signal with $(M+H)^+$=12623 Da referred to the *Np*Htr(1-114) G84C mutant. The observed experimental mass was smaller (ΔM=-86 Da) than the calculated molecular weight of M=12709.6 Da. This

phenomenon was previously as well observed for heterologously expressed *N.pharaonis* transducers[62] and could not be explained by an N-terminal methionine deletion (ΔM=-131.2 Da). For an additional discussion concerning the observed molecular weights of different transducer variants see as well section 3.3.1.2.

Hippler-Mreyen et al. successfully characterised the SRII/HtrII complex formation of differently truncated transducer constructs using blue-native gel electrophoresis[60,62-64]. This established *in vitro* assay was used to evaluate the capability of the truncated G84C mutant to bind to its cognate receptor. The gel so obtained displayed an intensive band with a lower mobility compared to the SRII reference for a sample containing the mutant and the receptor in 1:1 stoichiometry (figure 3.1, panel C). Above this band that was assigned to a 1:1 SRII/Htr(1-114) G84C complex, a second band with lower intensity was observed. Based on the results of the above-mentioned study[60,62], this band was assigned to a 2:2 complex.

The *in vitro* binding study clearly indicated that the G84C mutant in a truncated 114 amino acids comprising form is able to interact with its native binding partner SRII. Therefore a G84C mutation caused by a successful generation of the transducer using EPL, should not prevent the access to a functional transducer variant.

3.2 Peptide generation

3.2.1 *Np*HtrII(84-113)

H-GDTAASLSTLAAKASRMGDGDLDVELETRR-OH
M=3107.4 Da

Investigations of the interaction between the SRII receptor and its cognate transducer suggested that the linker domain of the Htr comprising amino acids 84-114 is essential for the correct binding of the whole transducer to the receptor[60]. In isothermal calorimetric titrations (ITC) the interaction between the cytoplasmic linker domain itself and the SRII receptor should be characterised.

The corresponding peptide comprising amino acid residues 84-113 was chemically synthesized using SPPS[47] on a BOC-arginine pre-loaded PAM-resin in 0.1-mmole scale. Ninhydrine tests[48] displayed a low coupling efficiency of Arg99. Therefore this coupling was repeated and reaction time was extended from 12 min to 30 min. After removing

the N-terminal BOC group the synthesized peptide was cleaved from solid support and purified with C18 RP-HPLC using a linear gradient from 25 % buffer B (acetonitrile, 0.08 % TFA) in buffer A (water, 0.1 % TFA) to 38 % buffer B in buffer A in 13 min. Fractions containing the desired peptide were pooled, lyophilized and stored at –20 °C in the dark.

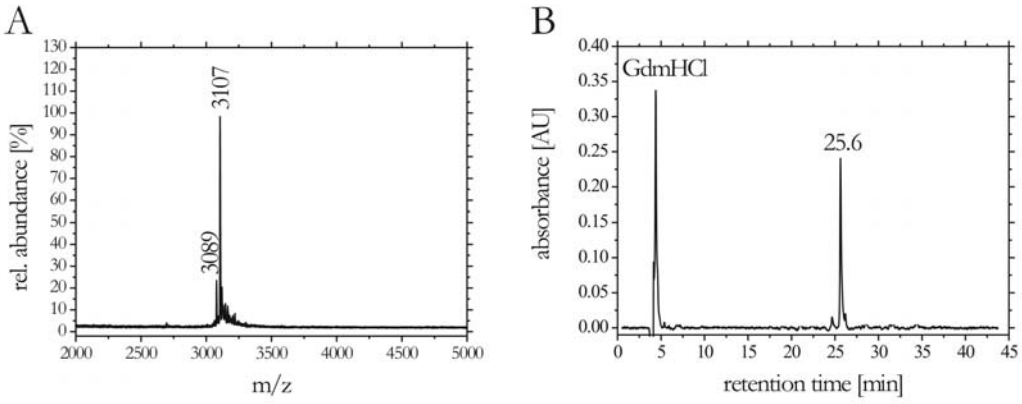


Figure 3.2. MALDI-MS (panel A) and analytical HPLC chromatogram (panel B) of purified *Np*Htr(84-113).
Signals with 3107 amu and 3089 amu corresponding to the target peptide and a water elimination product (ΔM=-18.0 Da) that could not be separated indicated by the single signal in the analytical RP-HPLC chromatogram with a retention time of 25.6 min.

The purified peptide was analyzed by mass spectrometry and analytical RP-HPLC. The mass spectrum of *Np*Htr(84-113) showed signals of two different peptide species that could not be separated with any chosen HPLC gradient. The main product with M=3107.0 Da corresponded to the target peptide with the calculated mass of 3107.4 Da. The molecular weight of the side product with M=3089.0 Da revealed a water elimination (ΔM=-18.0 Da). One possible explanation could be succinimide formation[65,66] from aspartic acid residues. The target peptide contains four of these residues and this could have led to an accumulation of the respective cyclized side product. However, without a suitable separation method the relative amount of side product in comparison with the target peptide could not be determined. The overall yield of the isolated fractions was 26.7 mg (8.7 %) based on the 0.1 mmole scale of the synthesis.

For the ITC experiments, the SRII receptor was dialyzed against ITC buffer containing 150 mM NaCl, 50 mM TrisHCl, 0.05 %DDM, pH 8.0 at 4 °C. The lyophilized peptide

was dissolved in ITC dialysis buffer. The protein concentrations of SRII and peptide were 0.055 mM and 0.47 mM, respectively. Hippler-Mreyen et al. performed the titrations of the weakly binding transducers Htr(1-101) and Htr(1-82) at room temperature. Therefore all experiments of this study were done at 22 °C on a VP-ITC Micro Calorimeter (Microcal Inc.). For control experiments, the peptide was titrated in ITC buffer.

In first titrations with 10 mM TrisHCl buffered saline strong transprotonation heats were measured due to an exhausted buffering capacity of the peptide solution: the acidic peptide (pI below 5) shifted the pH of the solution. Therefore all further titrations were performed with 50 mM TrisHCl. For the interaction of the cytoplasmic domain of the transducer with the SRII receptor the following binding curve was obtained (figure 3.4).

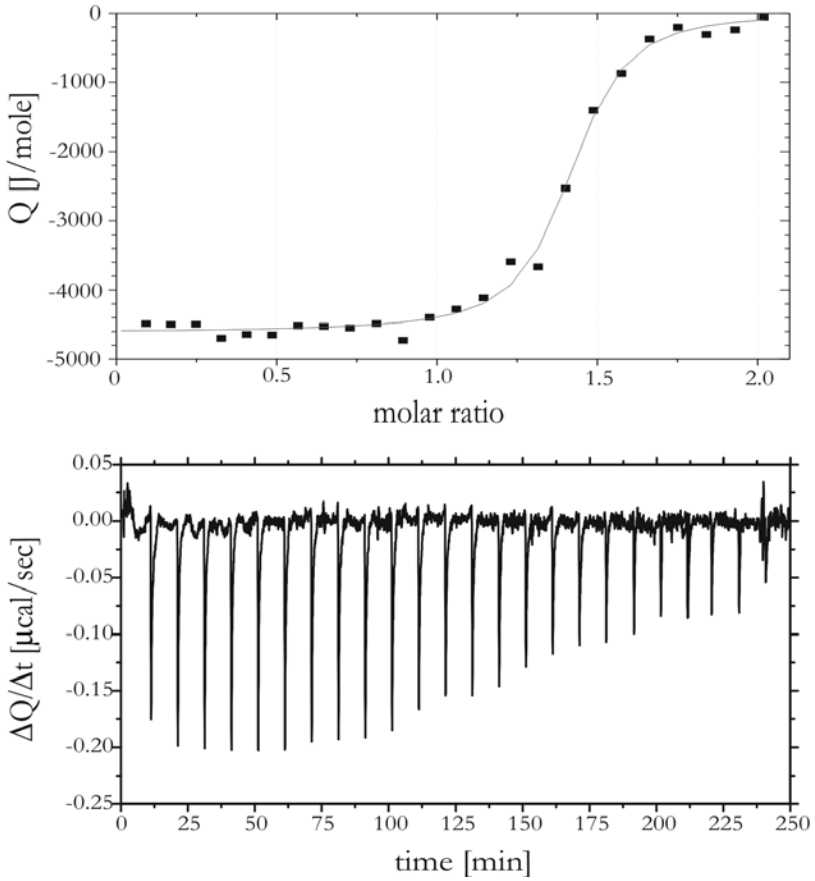


Figure 3.3. Isothermal calorimetric titration curve of SRII titrated with HtrII(84-113). The peptide was dissolved in ITC buffer containing 150 mM NaCl, 50 mM TrisHCl, 0.05 %DDM, pH 8.0 (470 μM) and added at 25 °C to the SRII (55 μM) in 25 increments of 10 μl in 12.6 sec with a time spacing of 300 sec.

The titration curve (figure 3.3) shows signals with negative amplitudes corresponding to an exothermal process. At the end of the titration a signal plateau is reached and the measured heat corresponds to the dilution heat of the peptide into the buffer (ca. -1 kJ/mole). After the time integration of raw data, this observed signal offset was mathematically corrected by subtraction. Subsequently the binding parameters, like binding enthalpy and dissociation constant, were determined using the one single binding site model and a SRII+nHtr(84-113)→SRII*nHtr(84-113) binding scheme, where n represents the observed number of peptide molecules required for the formation of a complex with SRII (fitting curve is represented as grey line, figure 3.3). Data were evaluated by employing the Origin-ITC software package. Table 3.1 summarises the obtained parameters.

	Temperature (K)	K_D (μM)	n	ΔH (kJ·mole^{-1})
Htr(84-113)	295	0.24	1.4	-4.6
Htr(83-149)	308	10	-	-
Htr(1-114)	318	0.23	1.0	-17.6
Htr(1-101)	295	10	1.0	-5.8

Table 3.1. Thermodynamic parameters of various SRII/Htr complexes. Comparison of dissociation constant (K_D), molar ratio (n) and binding heat (ΔH). Values for Htr(83-149) are taken from Sudo et al.[67] and those for Htr(1-101) and Htr(1-82) from Hippler-Mreyen et al.[60].

The Htr(84-113) peptide itself was able to interact with the receptor as was recently reported for a similar, however longer peptide[67]. Most of the above-mentioned titrations were done at different temperatures and therefore a direct comparison is difficult because of the strong temperature-dependence of binding heats; however, the dissociation constants (K_D) can be compared due to their weaker temperature-dependency.

The observed dissociation constant of 240 nM for Htr(84-113) is comparable to that obtained for the corresponding Htr(1-114) transducer construct (230 nM). The observed binding is one order of magnitude stronger than those of the longer peptide (10 μM) and the Htr(1-101) construct (10 μM). Apparently, the peptide Gly84-Arg113

comprises the essential part of the receptor binding domain, which was already postulated by Hippler-Mreyen et al.[60] based their measurements with different C-terminally truncated transducers.

The observed molar ratio (n=1.4) deviates from the value n=1.0 found for the Htr(1-114) variant suggesting a lowered effective peptide concentration compared to the determined value of 470 μM. Due to the absence of an accurate method to specify the concentration of a short peptide without aromatic residues, the concentration was determined gravimetrically. Unfortunately, this method neglects residual salt in the lyophilized peptide sample, which may explain the observed deviation.

The C-terminal extension of the peptide up to Gln149[67] resulted in a significant weakening of the SRII/Htr interaction by a factor of ca. 40 (table 3.1). However, the observed binding curve[67] showed, instead of the expected sigmoid shape, a hyperbolic one, which is unusual for a binding event. Therefore a comparison of the results obtained for both peptides seems inadvisable.

The capacity of the peptide Htr(84-113) to interact specifically with the SRII receptor raises questions about the secondary structure of the peptide comprising the receptor-binding domain. Therefore CD-spectroscopy experiments were performed under conditions used for the ITC measurements.

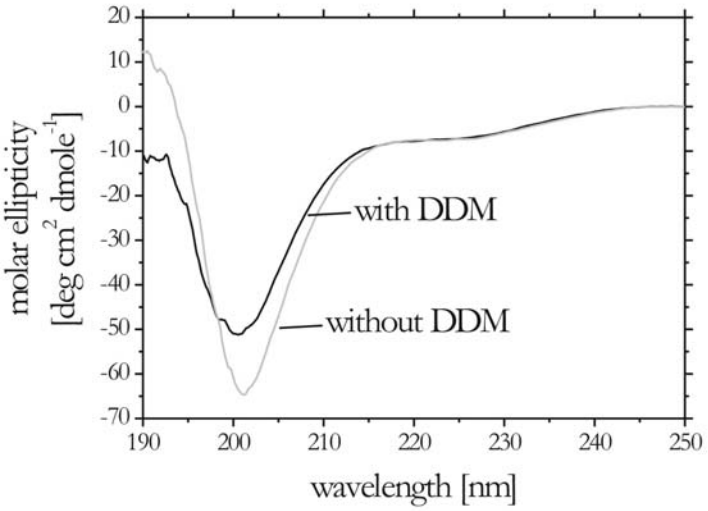


Figure 3.4. CD-spectra of wt Htr(84-113) with and without DDM at 20 °C. The negative signals at 200 nm and 220 nm indicate random coil secondary structure with α-helical tendencies. The positive signal below 195 nm was due to light scattering.

The observed CD-spectra (figure 3.4) at 20 °C were recorded in ITC buffer with and without DDM and were dominated by strong negative signals at around 200 nm that were caused by a predominant random coil structure. The addition of detergent led to a minor shift of the signal to lower wavelength. Additionally increased light scattering below a wavelength of 195 nm was observed that caused the deviation of the signals in that range. Nevertheless both spectra showed a slight negative signal at around 220 nm ($n\rightarrow\pi^*$) indicating partially α-helical folding. A signal at around 205 nm of a $\pi\rightarrow\pi^*$ transition of an α-helix could not be observed due to signal overlapping with a strong negative signal of the random coil folding. The positive signal below a wavelength of 195 nm was due to light scattering and did not indicate a β-sheet folding.
In summary the peptide Htr(84-114) comprising the main part of the receptor-binding domain is predominantly unfolded with α-helical tendencies observed with and without DDM as a detergent. Taking the specific binding capability of the peptide into account, it may be the case that the unfolded peptide folds upon receptor binding[68,69]. This hypothesis is supported by Sudo et al.[67], who observed in $^{1}H,^{15}N$ 2D HSQC NMR experiments with Htr(83-149) chemical shift changes that were interpreted as structural changes by the authors.

3.2.2 *Np*HtrII(84-109)NBD G84C

H-CDTAAXLSTLAAKASRMGDGDLDVELHHHHHHHL-OH
M=3722.9 X=NBD-Dab

The first step in the semi-synthesis of the transmembrane transducer protein was the generation of the C-terminal *Np*HtrII segment. In order to improve qualitative and quantitative evaluation of the expressed protein ligation reaction, a fluorescently labelled peptide comprising the Htr residues 84-109 and an artificial C-terminal poly-histidine purification tag was synthesized. The overall length of the peptide was limited to 33 residues. In consideration of the reaction steps needed for the NBD-labelling (figure 3.5), this limitation should ensure a sufficient product yield[47].
Gly84 was substituted by a cysteine residue and a Ser89 was replaced by an NBD-labelled diaminobuturic acid residue. The peptide was chemically synthesized[46] on a BOC-leucine-PAM solid support in a 0.1 mmole scale using *in situ* neutralization and

HBTU activation protocol for BOC protected amino acids[47]. Coupling of the non-natural amino acid L-BOC-Dab(Fmoc)-OH was extended from 12 min to 1 h and monitored by Kaiser tests[48]. After sequence completion the orthogonal Fmoc protecting group of Dab was removed with 20 % piperidine in DMF. The resulting free γ-amino moiety (1) was reacted with 10 eq of NBD chloride (2) in presence of 5 eq of DIEA as base for 30 min (figure 3.5).

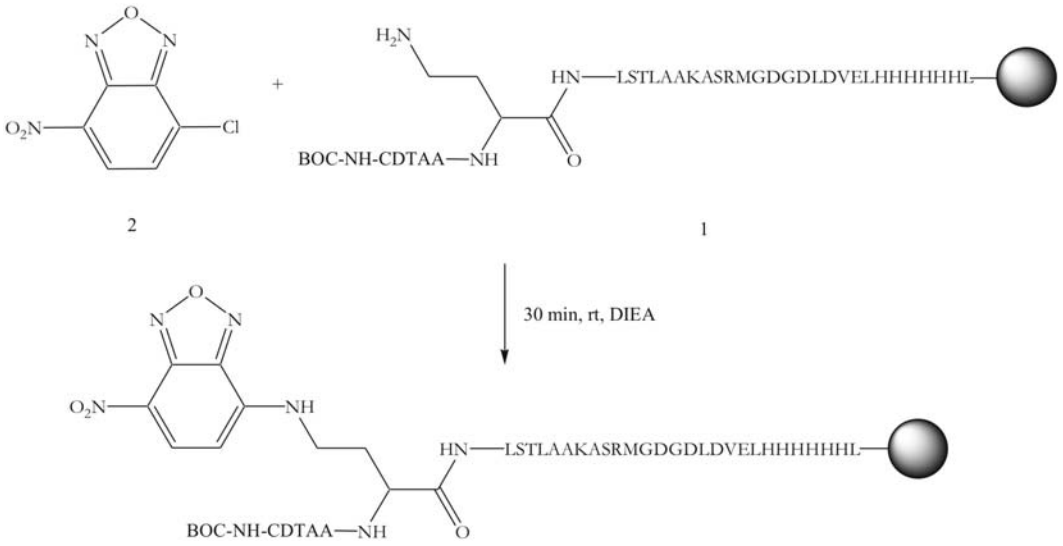


Figure 3.5. NBD labelling of *Np*HtrII(1-109)NBD G84C. All histidine residues were DNP-protected. The free γ-amino moiety of Dab (1) was reacted with 10 eq of NBD chloride (2) in presence of 5 eq of DIEA as base.

After NBD labelling of the peptide, the DNP protecting groups of the six consecutive histidines were cleaved off. Subsequently the N-terminal BOC group of the synthesized peptide was removed with TFA to avoid the risk of a possible *tert*-butylation[70] during the HF[49]cleavage from the solid support.

Purification was performed using semi-preparative C-18 RP-HPLC with a linear acetonitrile gradient in water (25 % to 38 % in 13 min). Fractions containing the desired peptide were pooled, lyophilized and stored at –20 °C in darkness.

The purified peptide was analyzed by mass spectrometry and analytical RP-HPLC. While the chromatogram of the HPLC analysis (figure 3.6, panel B) showed only one single product peak (28.0 min) besides the observed buffer peak (5.5 min), the mass spectrum of the purified peptide revealed signals of two different species, which could

not be separated with any HPLC gradient (figure 3.6). The main product with protonated states $(M+2H)^{2+}$=1861.8, $(M+3H)^{3+}$=1242.3, $(M+4H)^{4+}$=931.7 and $(M+5H)^{5+}$=745.7 amu corresponded to a deconvoluted mass of M=3723.0 Da that is in good agreement with the calculated mass of the desired peptide *Np*HtrII(84-109)NBD G84C (M=3722.9). The second peptide species with $(M+2H)^{2+}$=1793.6, $(M+3H)^{3+}$=1236.0 and $(M+4H)^{4+}$=897.3 amu displayed an experimental mass of 3585.2 Da suggesting an histidine deletion (ΔM=-137.1 Da).

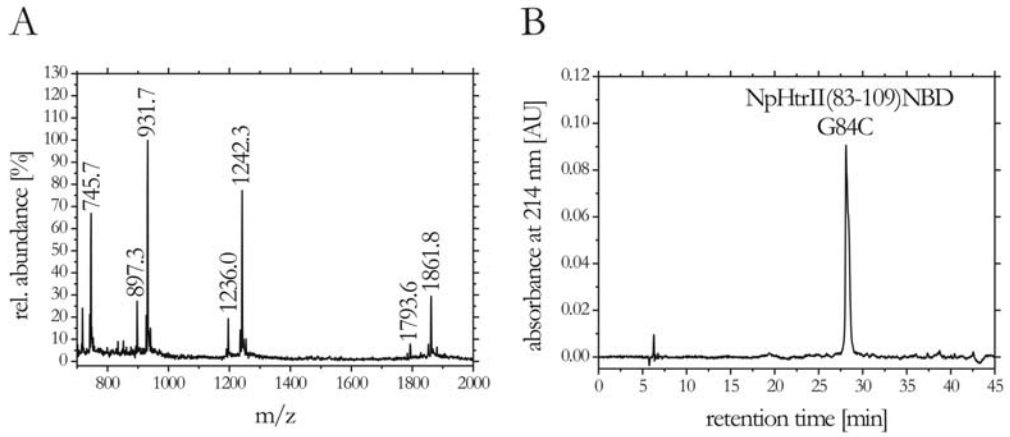


Figure 3.6. ESI-MS (panel A) and analytical HPLC chromatogram (panel B) of purified *Np*HtrII(84-109)NBD G84C.
MS signals at 1861.8 amu, 1242.3 amu, 931.7 amu and 745.7 amu corresponded to the desired peptide (M_{calc}=3722.9). Signals 1793.6 amu, 1236.0 amu and 897.3 amu suggested an histidine deletion product (ΔM=-137.1 Da), which could not be separated indicated by the single signal in the analytical RP-HPLC chromatogram (panel B).

The histidine purification tag of the histidine deletion side product would contain only five consecutive histidines. Such a deletion should only have no effect on the function of the tag; therefore no further purification steps were performed. Without a suitable separation the relative amount of side product in comparison with the target peptide could not be determined. The overall yield of the isolated fractions was 35.7 mg (9.6 %) based on 0.1 mmole synthesis scale.

3.2.3 *Np*HtrII(84-113) G84C

H-CDTAASLSTLAAKASRMGDGDLDVELETRR-OH
M=3153.4 Da

A fluorescent-labelled peptide comprising an N-terminal cysteine (see previous section) is a useful tool to ease monitoring of EPL reactions by densitometric investigations of SDS-PAGE gels. However, the used modifications, like histidine purification tag or the NBD moiety, may influence the biochemical properties of the respective product. Therefore an unmodified C-terminal segment with an N-terminal cysteine residue was chemically synthesized to facilitate the semi-synthetic access to an unmodified transducer.

The solid phase synthesis succeeded by using the *in situ* neutralization and HBTU activation protocol for BOC protected amino acids[47] on BOC-arginine pre-loaded PAM-resin in 0.1 mmole scale. Ninhydrine tests[48] displayed a low coupling efficiency of Arg99 as similar to that observed for the wild type sequence. Therefore this coupling was performed twice and prolonged to 30 min. After removal of the N-terminal BOC group to avoid *tert*-butylation[70], the synthesized peptide was cleaved from solid support with neat HF in the presence of 5 % p-cresol[49] and 1 % cysteine as scavengers. After removing HF in vacuum the crude peptide was suspended in diethyl ether resulting in a highly viscous precipitate. During extensive washing with diethyl ether the sediment became more granular. After dissolving in 50 % acetonitrile in water containing 0.1 % TFA, the peptide was separated from resin beads by filtration and lyophilized.

In order to verify successful peptide synthesis, the crude peptide was analyzed using ESI-MS. The mass spectrum of the crude peptide showed two intensive signals at 3168.3 amu and 3153.3 amu in addition to two signals of less intensity at 3015.1 amu and 2997.0 amu with low spectral resolution (figure 3.7).

The signal at 3153.3 amu was assigned to the target peptide with a calculated molecular weight of M=3153.3 Da. The signal with M=3168.3 Da (ΔM=+15 amu) suggested a possible oxidation of Met100 to the corresponding sufoxide[71]. In addition, in spite of coupling Arg99 twice, an arginine deletion product with a calculated mass of 2997.2 Da (ΔM=-156.2 Da) was detected with and without oxidized Met100 (M_{calc}=3013.2).

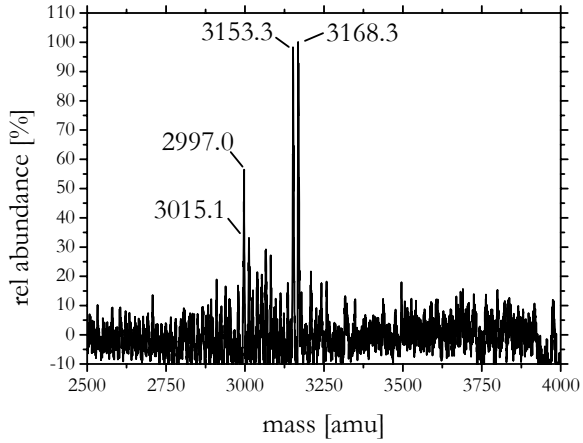


Figure 3.7. ESI mass spectrum of the crude peptide *Np*HtrII (84-113) G84C. The signal at 3153.3 amu was assigned to the target peptide (M=3153.3 Da). The signal with M=3168.3 Da (ΔM=+15 amu) suggested Met100 oxidation. In addition, the signals at 3015.1 amu and 2997.2 amu suggested an arginine deletion product with and without oxidized Met100.

In order to reduce the sulfoxide side products, the crude peptide was dissolved in 6 M guanidine hydrochloride, 50 mM Tris, 30 % TFE, pH 8.5 and incubated for 3 days in presence of 3 % DTT at rt[71]. MS experiments revealed that the reduction of the sulfoxide could not be achieved quantitatively (results not shown), however residual Met(O) containing species could be separated from the desired product with C18 RP-HPLC. Purification was performed using the above-mentioned gradient conditions (see previous section). Fractions containing the desired peptide were pooled, lyophilized and stored at –20 °C in the dark.

The purified peptide was analyzed by mass spectrometry and analytical RP-HPLC. The chromatogram of the HPLC analysis (figure 3.8, panel B) showed besides the observed buffer peak one single product peak.

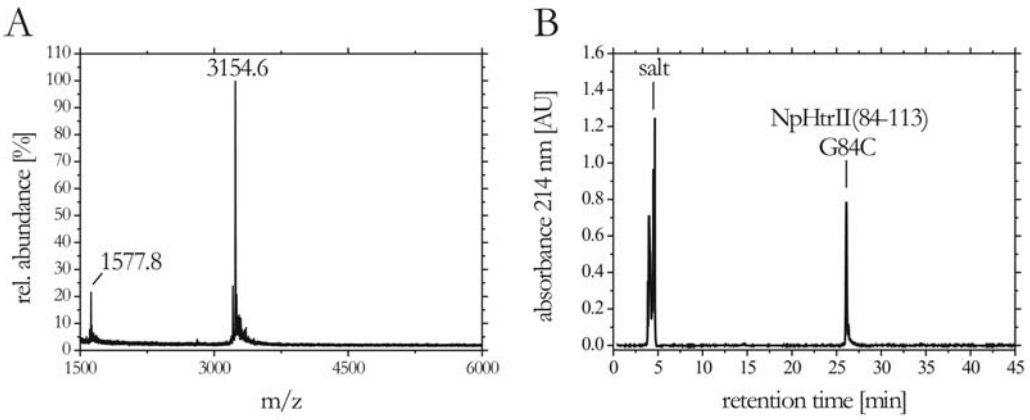


Figure 3.8. MALDI-MS (panel A) and analytical HPLC chromatogram (panel B) of purified *Np*HtrII(84-113) G84C.
MS signal at 3154.6 amu corresponded to the $(M+H)^+$=3153.4 Da signal of the desired product. Purity was confirmed by HPLC analysis showing one single product peak.

The purity of the target peptide was confirmed by the mass spectrum (figure 3.8, panel A) revealing only one signal at 3154.6 amu corresponding to the $(M+H)^+$=3153.4 Da signal of the desired product. The overall yield of the isolated fractions was 24.4 mg (7.7 %) based on 0.1 mmole synthesis scale.

3.2.4 Isotope labelled *Np*HtrII(84-114) G84C peptides

The establishment of C-terminal domain specific isotope labelling of the membrane protein *Np*HtrII in a C-terminally truncated form requires a low cost access to uniformly isotope labelled peptides containing an N-terminal cysteine residue.
Although labelled amino acid building blocks for solid phase peptide synthesis (SPPS) are commercially available, the financial outlay for the synthesis of a peptide comprising an entire protein domain would be immense.
Fortunately highly optimised, relatively inexpensive protocols for recombinant labelling of proteins using synthetic minimal media containing [^{13}C] D-glucose and/or [^{15}N] ammonium chloride as sole carbon respectively nitrogen sources are well established[72,73].
The combination of recombinant isotope labelling with the proteolytic removal of an N-terminal glutathion-S-transferase leader sequence looked promising as a suitable preparative method to generate an EPL peptide building block.

Therefore the gene encoding for Htr(84-114) G84C was genetically fused to a GST-tag with an tobacco etch virus (TEV) protease recognition site (ExxYxQ▼, where x can be every amino acid)[74] and a cysteine at the P1' position[75] (first C-terminal residue after glutamine) using PCR with the pRPS1 plasmid as template. The resulting fragment was purified using a PCR purification kit (Qiagen) and subsequently cloned into the BamHI and EcoRI sites of a dephosphorylated BamHI/EcoRI linearized pGEX-2T vector that allows high-level expression of the desired GSTtevHtr fusion protein under control of the T7-promoter.

In order to investigate the expression of the desired GSTtevHtr fusion protein under minimal medium conditions, three 50 ml cultures in minimal medium containing 2 g/l D-glucose and 1 g/l ammonium chloride were inoculated to cell densities of OD_{578}=0.05 and incubated at 37 °C. Induction was performed using 1 mM IPTG at OD_{578}=0.4, 0.6 and 0.9, respectively. Samples were taken after 2 h, 4 h and 8 h and protein expression analysed using SDS-PAGE.

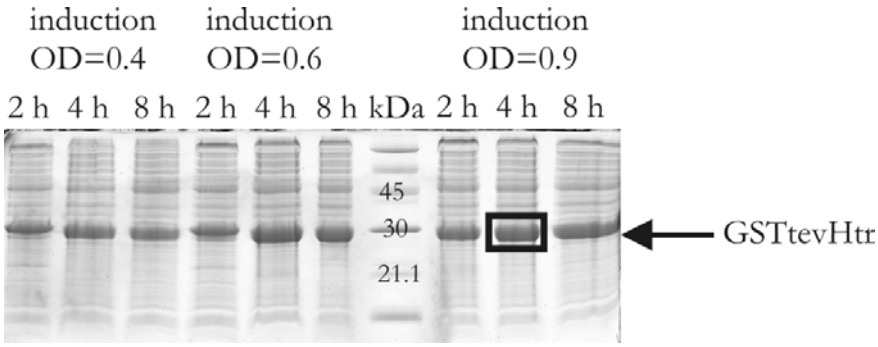


Figure 3.9. SDS-PAGE of GSTtevHtr expression optimization
Maximal GSTtevHtr (30.4 kDa) expression was observed after 3-4 hours under 1 mM IPTG at OD_{578}=0.9 conditions, indicated by black box.

The SDS-PAGE gel (figure 3.9) revealed good over-expression of the desired fusion protein upon IPTG induction as seen by the protein band appearing below the 30-kDa-reference band. Maximal GSTtevHtr (30.4 kDa) expression was observed after 3-4 hours for all different cell densities chosen as induction points. Longer incubation times up to 8 h led to a decrease of the amount of protein indicated by faint bands, suggesting 1 mM IPTG at OD_{578}=0.9 and incubation for 4 h, (indicated here by black box) as optimal expression conditions.

Isotopic labelling was hence performed on a 2 l scale using these optimized expression conditions with synthetic minimal media containing 2 g/l [^{13}C] D-glucose and/or 1 g/l [^{15}N] ammonium chloride. After protein expression, cells were lysed and the cytoplasmic fractions purified exploiting the specific binding of the glutathione-S-transferase tag to immobilized glutathion (GSH). After washing the column with cell washing buffer, the target protein was eluted with reduced glutathion in cell washing buffer and then dialyzed against 100 volumes of TEV cleavage buffer. Purification procedure was analytically monitored by SDS-PAGE (figure 3.10).

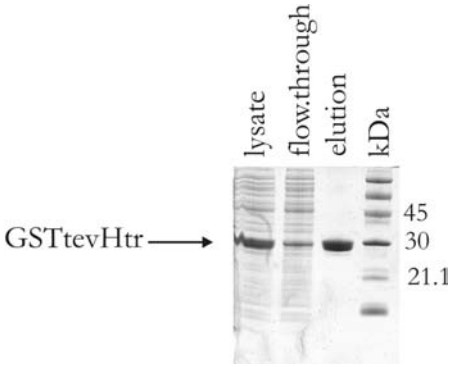


Figure 3.10. Analytical SDS-PAGE monitoring GSTtevHtr purification. Band below 30 kDa reference band indicated high over-expression of GSTtevHtr in isotope enriched minimal medium. The isotope labelled GSTtevHtr was obtained analytically pure.

The subsequent SDS-PAGE gel (figure 3.10) revealed high over-expression of the desired fusion protein upon IPTG induction in isotope enriched minimal medium as seen by the protein band below the 30-kDa-reference band. The eluate so obtained from the GSH column contained analytically pure isotope labelled GST-tevHtr fusion protein indicated by a single protein band. It was necessary to remove the used GSH from the eluted fraction by dialysis to prevent disulfide cross-linking with the N-terminal cysteine containing peptide during proteolytic digest (data not shown). The yield of the fusion protein was spectroscopically determined at a wavelength of 280 nm, where A_{280}=1.392 (calculation: Expasy ProtParam algorithm) corresponded to a concentration of 1 g/l. Generally 20-30 mg purified protein per litre culture were obtained.

In addition to the successful establishment of a suitable isotope labelling protocol for the GSTtevHtr fusion protein the proteolytic removal of the N-terminal glutathion-S-transferase leader sequence had to be optimised to liberate the isotope labelled EPL peptide building block. The conditions for the proteolytic TEV cleavage of the fusion protein were optimized by varying the substrate/protease ratios and incubation times. The progress of the cleavage was monitored by analytical SDS-PAGE (figure 3.8). Panel A summarizes the results of obtained for TEV/GSTtevHtr ratios from 1/100, 1/250, 1/500 up to 1/1000, incubated at rt for 24 h.

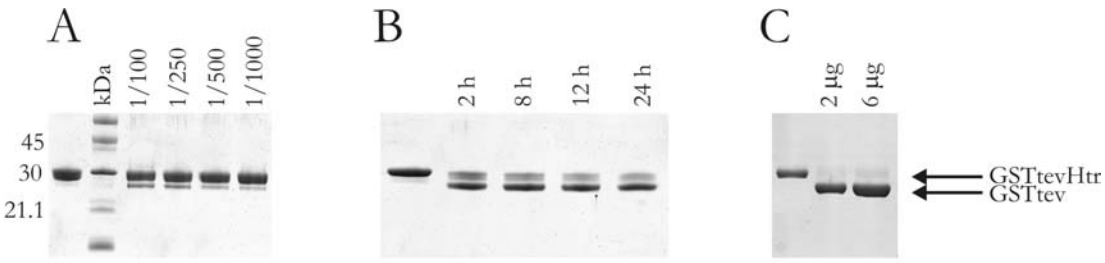


Figure 3.11. SDS-PAGE gels monitoring the optimization of proteolytic TEV digest of GSTtevHtr fusion protein.
Panel A: Variation of TEV protease/GSTtevHtr ratio in 20 µl scale. After 24 h the appearance of GSTtev protein bands below GSTtevHtr bands indicated proteolytic cleavage.
Panel B: Time course with a ratio increased to 1/50 at rt in 20 µl scale revealed 2/3 processing after 2 h followed by only minor changes.
Panel C: Up-scaling to 1 ml reaction with 1/50 ratio 2 h at rt. The absence of any starting material indicated quantitative TEV cleavage.

The SDS-PAGE gel monitoring the effect of the TEV protease/GSTtevHtr ratio on the proteolytic cleavage efficiency is given in figure 3.11, panel A. In all cases the resulting GSTtev cleavage product with a molar mass of 27.1 kDa was detected, however a quantitative cut was not achieved in any of the digests.

Therefore the TEV/GSTtevHtr ratio was increased to 1/50 and the time course of the reaction was investigated by SDS-PAGE. The observed relative intensities of the GSTtevHtr and GSTtev protein bands of around 1 to 2 (figure 3.11, panel B) revealed that within 2 h, approximately two thirds of the target protein was processed. In the following hours only minor changes occurred.

In order to consider adsorption of the TEV protease to the reaction cups in 20 µl scale reactions as a cause for incomplete digestion, an up-scaled 1 ml reaction was tested with a 1/50 ratio of TEV to GSTtevHtr for 2 h at rt. The reaction mixture was gently mixed on a spinning wheel and monitored by SDS-PAGE. Under these conditions a quantitative cleavage of the fusion protein was achieved; no more starting material was detectable even with a three times higher protein amount analyzed (figure 3.11, panel C). The cleaved Htr-peptide was not visible in the SDS-Page gels. Since the molecular weight of the cleaved peptide is only a tenth of the weight of the GSTtevHtr fusion protein, the amount of free peptide contained in a sample of 2 µg GSTtevHtr that was generally used for SDS-PAGE analysis, was below the detection limit.

After almost quantitative proteolytic cleavage, all isotope labelled peptides were separated from the GST-tag using semi-preparative C18 RP-HPLC techniques with a linear gradient of acetonitrile in water from 5-65 % in 30 min and subsequently analyzed by mass spectrometry.

Figure 3.12 summarizes the chromatograms obtained and mass spectra for unlabelled, [^{15}N] and [^{13}C, ^{15}N] *Np*HtrII (84-114) G84C. All peptides were prepared with high purity as documented by single product peaks in RP-HPLC analysis. In addition, the observed molar masses of unlabelled, [^{15}N] labelled and [^{13}C, ^{15}N] double labelled peptide were in good agreement with the respective calculated molar masses of M=3282 Da, M=3282+0.99·41=3323 Da and M=3282+0.99(132+41)=3453 Da with 0.99 representing the ^{13}C, ^{15}N enrichment of respective isotope sources. The efficiency of isotope labelling could be determined to be 99 % and corresponds very well to the isotope enrichment of the respective ^{13}C and ^{15}N sources.

The yields of the labelled peptides were usually 2.4-4.8 mg purified peptide per 2 l medium or 60-80 % based on the amounts of purified GSTtevHtr fusion proteins (yield: 20-30 mg/l).

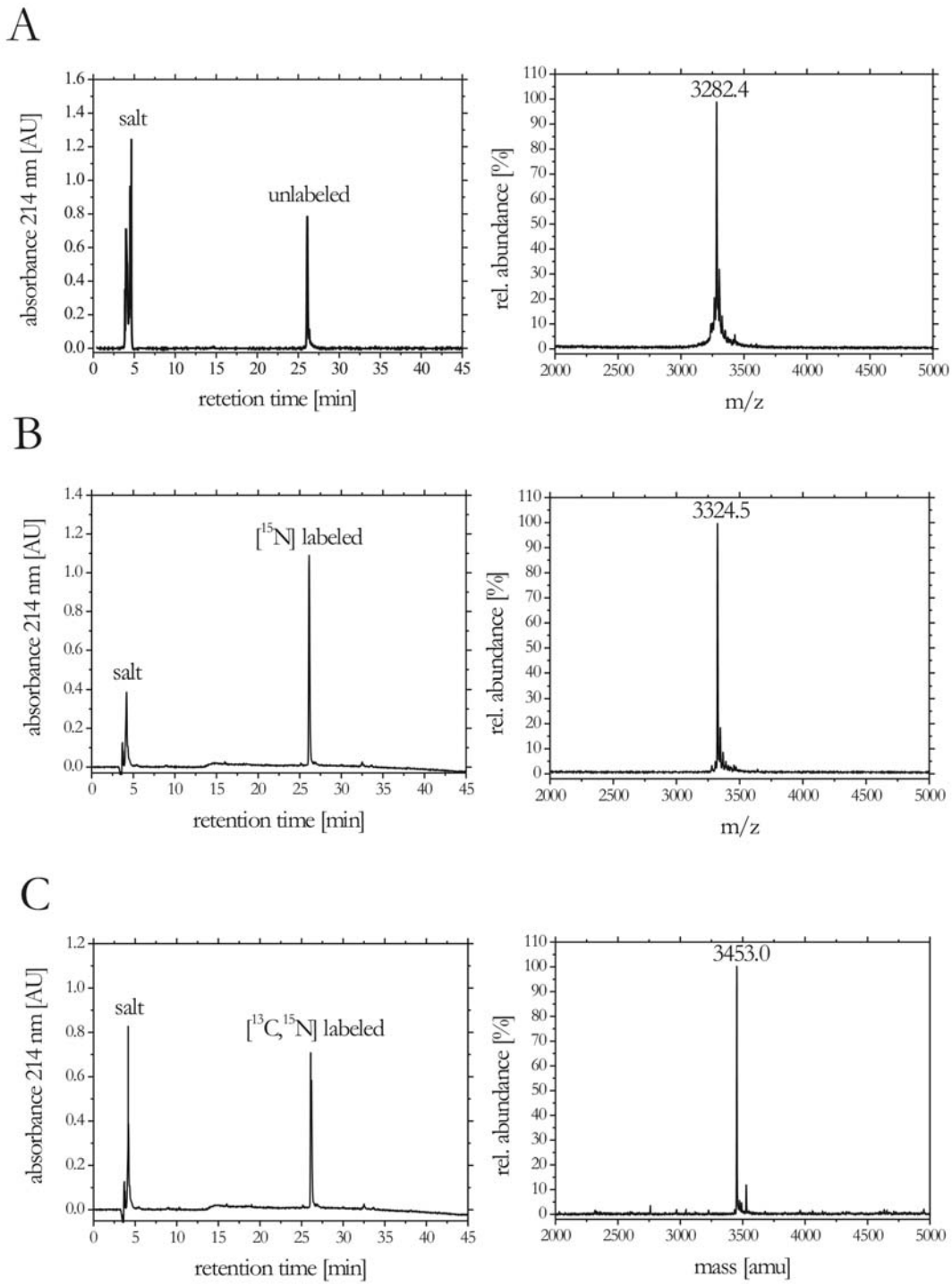


Figure 3.12. Analytical HPLC chromatograms and MALDI-MS of isotope labelled *Np*HtrII(84-114) G84C peptides. All variants were obtained analytically pure after TEV cleavage and C18 RP-HPLC with a linear acetonitrile gradient from 5-65 % in 30 min (retention time of 26.5 min in all cases).
Panel A: unlabelled peptide (M_{calc}=3282 Da)
Panel B: [^{15}N] labelled peptide (M_{calc}=3323 Da)
Panel C: [^{13}C, ^{15}N] labelled peptide (M_{calc}=3453 Da)

3.3 Expressed protein ligation

3.3.1 EPL utilising HtrMxeCBD isolated from *E.coli* membranes

3.3.1.1 HtrMxeCBD purification and evaluation

The expressed protein ligation utilises inteins to generate activated α-thioester species that are subsequently joined together with a peptide containing an N-terminal cysteine. The generation of such peptides in isotope labelled form, a prerequisite for segmental isotope labelling of the halobacterial transducer, was described in section 3.2. In the following section the utilisation of these peptides in the EPL of the transmembrane protein HtrII in a C-terminal truncated version shall be described.

For high-level expression of the transducer segment containing two transmembrane helices, a modified pTXB1 vector (New England Biolabs) was used, which encodes the GyrA intein from *Mycobacterium xenopi* N-terminally fused to a seven-histidine purification tag followed by a chitin binding domain (CBD). The modified construct was kindly supplied by R.P. Seidel. The htr(1-83) gene encoding for the two transmembrane helices of the HtrII was amplified using PCR and the pRPS1 plasmid as a template. The resulting DNA-fragment was purified with a PCR purification kit, subsequently digested and ligated into the Nde I and Sap I sites of modified pTXB1his vector. The ATG of the Nde I site is used as the start codon of the DNA-transcription. Sap I cloning allowed direct fusion of the transducer to the N-terminal cysteine residue of the intein[45]. Upon successful transformation of the plasmid into *E.coli* cells, protein expression of HtrMxeCBD was carried out at 37 °C for 3 h using 1 mM IPTG. Analytical SDS-PAGE was used to monitor the time course of the expression.

The SDS-PAGE gel (figure 3.13) displays that the fusion protein is highly over-expressed after IPTG induction as seen by the protein band above the 30 kDa reference band that appeared after 1 h of induction. After 2 h incubation no significant increase of the HtrMxeCBD fusion protein expression level could be achieved.

Subsequent analysis of protein partitioning revealed a strong HtrMxeCBD protein band derived from the inclusion body fraction (figure 3.13) suggesting the localisation of the majority of the desired protein in insoluble aggregates. A minor amount is successfully incorporated into the membrane and therefore accessible in soluble form after solubilization with DDM.

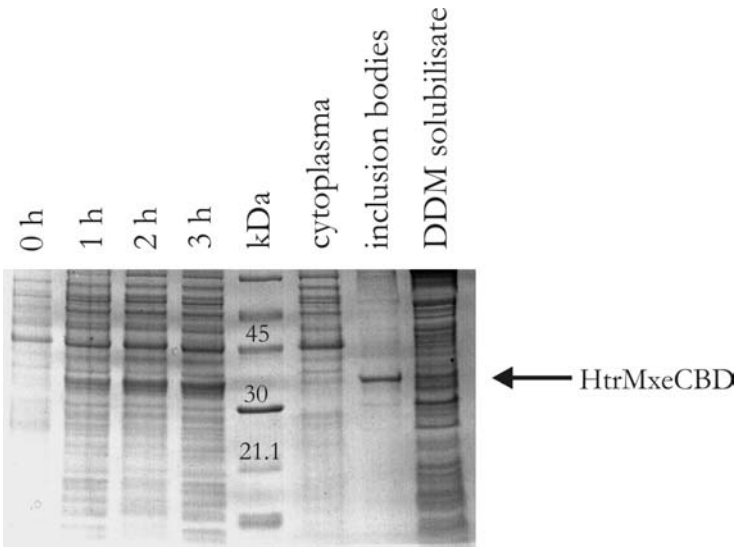


Figure 3.13. SDS-PAGE monitoring HtrMxeCBD expression and partitioning. Protein band above the 30 kDa reference documented over-expression of HtrMxeCBD ceiling after 2 h. The majority of HtrMxeCBD is located in inclusion bodies, only a minor amount is membrane incorporated.

While lowering the expression temperature from 37 °C to 25 °C or 15 °C increased the amount of membrane-localized protein, the absolute protein yields purified from the respective membrane fractions was approximately 50 % lower compared to expressions performed at 37 °C (data not shown).

Protein purification was performed using the specific affinity of the internal histidine tag to Ni-NTA matrix. After binding of the protein and a washing step with 20 mM imidazole the target protein was eluted with 200 mM imidazole and dialysed against 100 volumes of imidazole free buffer B_{mem}. The Purification was monitored by analytical SDS-PAGE (figure 3.14, Panel A). The yield was spectroscopically determined at a wavelength of 280 nm. A fusion protein concentration of 1 g/l corresponded to A_{280}=0.993 (calculation: Expasy ProtParam algorithm[53]). Typical yields were 3 mg per litre medium.

The SDS-PAGE gel documenting the Ni-NTA purification of HtrMxeCDB solubilised from *E.coli* membranes revealed two protein bands for the purified eluate.

A Comparison of gels analysing HtrMxeCDB originating from membranes (figure 3.14, panel A) or from inclusion bodies (figure 3.13) revealed an in-homogeneity of the soluble fraction: An upper band with an apparent molecular weight of 37.4 kDa, as

expected for the desired fusion protein and a second band with higher mobility appearing at approximately 32 kDa were distinguished.

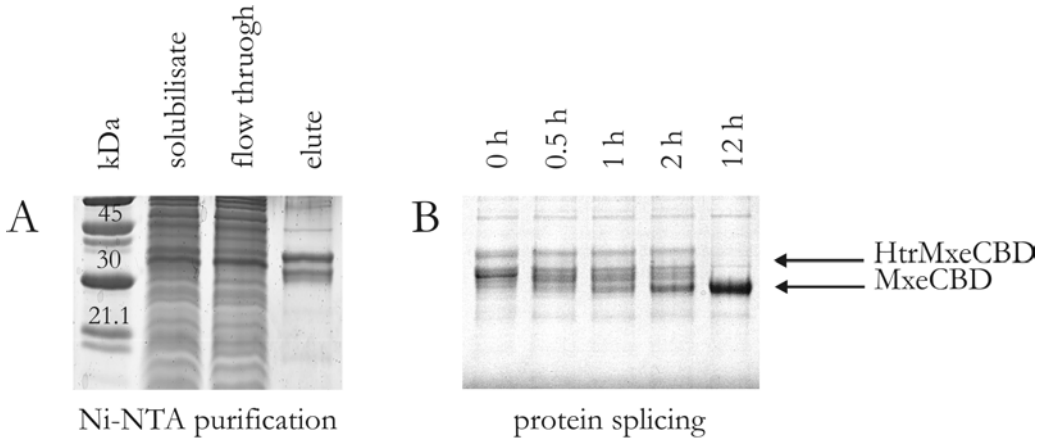


Figure 3.14. SDS-PAGE monitoring HtrMxeCBD purification and time course of its intein cleavage.
Panel A: Ni-NTA based purification of HtrMxeCDB solubilised from *E.coli* membranes. Two protein bands found for eluate revealed a full-length and a degraded intein fusion protein.
Panel B: Intein cleavage test. Disappearing of the protein bands of both the desired long and the truncated intein fusion protein species suggested quantitative intein cleavage of both intein species. After 12 h only one band with an apparent molecular weight of 29 kDa for the cleaved MxeCBD intein was observed.

The relative amount of both species varied from one protein preparation to another and was determined to vary from 33-66 % (upper band compared to lower band). The overall yield of the isolated fractions was spectroscopically determined at a wavelength of 280 nm. A fusion protein concentration of 1 g/l corresponded to A_{280}=0.993 (calculation: Expasy ProtParam algorithm[53]). Typical yields were 3 mg per litre medium.
A subsequent intein cleavage test (without peptide to ligate), which was initiated by the addition of 200 mM MESNA, indicated an almost quantitative disappearing of both corresponding intein bands (figure3.15, panel B). After 12 h only one band with an apparent molecular weight of 29 kDa, assigned to the cleaved MxeCBD intein, could be observed (figure 3.14, panel B) revealing an almost quantitative intein cleavage efficiency of both intein species.
In addition, the appearing of the single protein band corresponding to the cleaved intein, suggests a degradation of the N-terminal Htr-extein of the HtrMxeCBD fusion

protein as a reason for the observed double band. This implicated the presence of two α-thioester species: a full-length species with a calculated molecular mass of ca. 8 kDa* deriving from the desired full-length fusion protein and an N-terminally degraded species of approximately 3 kDa originating from the degraded fusion species.

In order to support this hypothesis with experimental evidence, MALDI-MS experiments were performed. The mass spectrum of the purified HtrMxeCBD showed two signals that were broadened due the presence of detergent in the sample (figure 3.15, panel A).

In agreement with the results of the SDS-PAGE analysis (figure 3.14, panel A), molecular masses of M=37329 Da for the full-length protein (calculated 37357 Da) and M=32499 Da for the N-terminally degraded species were detected.

This smaller mass signal suggests residue Ala47 as a possible degradation site in the HtrMxeCBD protein. The corresponding N-terminally degraded HtrMxeCBD(47-353) species has a calculated mass of M=32460 Da. Since Ala47 is part of a poly-alanine region forming a extra-cellular loop between TM1 and TM2[9], it is likely that proteases can cleave in this region. However, the low spectral resolution prevents an explicit elucidation of this context.

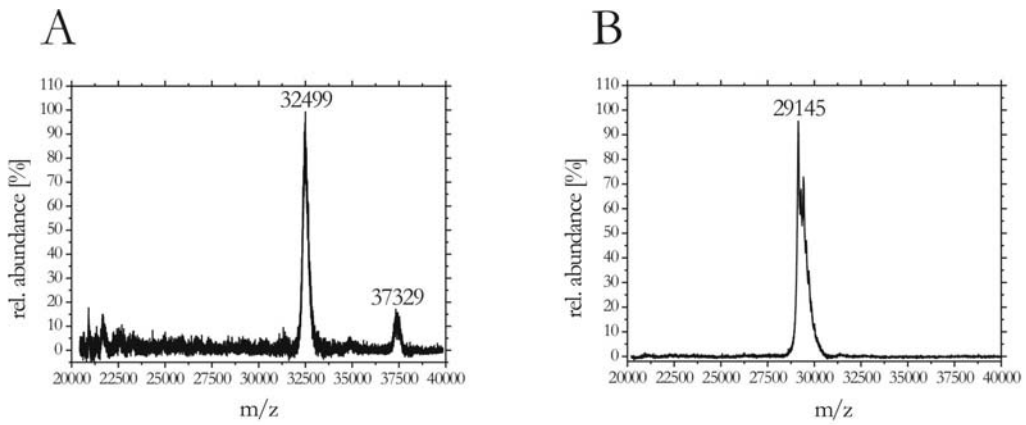


Figure 3.15. MALDI-MS experiments monitoring HtrMxeCBD cleavage reaction.
Panel A: MS of HtrMxeCBD solubilised from membranes. The two signals at 37329 amu and 32499 amu were assigned to the desired HtrMxeCBD (M=37357) and to an N-terminally degraded HtrMxeCBD(47-353) form (M=32460 Da).
Panel B: After completion of the intein cleavage reaction the signal at 29145 amu corresponded to the cleaved MxeCBD intein (M=29075).

* calculated by the subtraction of the molecular mass of the cleaved intein (29 kDa) from the molecular mass of the respective intein fusion protein species (37 kDa respectively 32 kDa).

After completion of the intein cleavage reaction only one broad signal of M=29145 Da was observed (figure 3.15, panel B). In the range of the spectral resolution this signal corresponded to the expected mass of the spliced intein (M=29075). signals of uncleaved intein species were not present confirming the quantitative intein cleavage efficiency that was already observed in the SDS-PAGE analysis (figure 3.13, panel B)
Unfortunately no experimental evidence for the presence of two α-thioester species, the full-length Htr(1-83)-COSR and the postulated degraded Htr(47-83)-COSR, could be obtained with these experiments: Besides the strong mass signal of the cleaved intein no signals for the respective thioesters could be resolved.

3.3.1.2 Expressed protein ligation

Expressed protein ligation experiments with the transducer were initially performed using the NBD-labelled fluorescent peptide *Np*HtrII(84-109)NBD G84C (see section 3.2.2) to investigate the time course of the reaction by densitometric fluorescence methods.
The Intein fusion protein HtrMxeCBD and the above-mentioned peptide, containing an N-terminal cysteine, were mixed in 1:1 ratio. Subsequently 200 mM MESNA were added to initiate the intein cleavage reaction. The ligation was carried out at rt with gentle agitation on a turntable. The progress of the reaction was monitored by SDS-PAGE with the NBD-labelled peptide that was incubated under identical conditions in the absence of an intein, as a control.
The SDS-PAGE gel was first analysed densitometrically (AIDA software, raytest, Staubenhardt, Germany) at 510 nm using 478 nm as excitation wavelength (figure 3.16, panel A) and subsequently stained with Coomassie (figure 3.16, panel B).

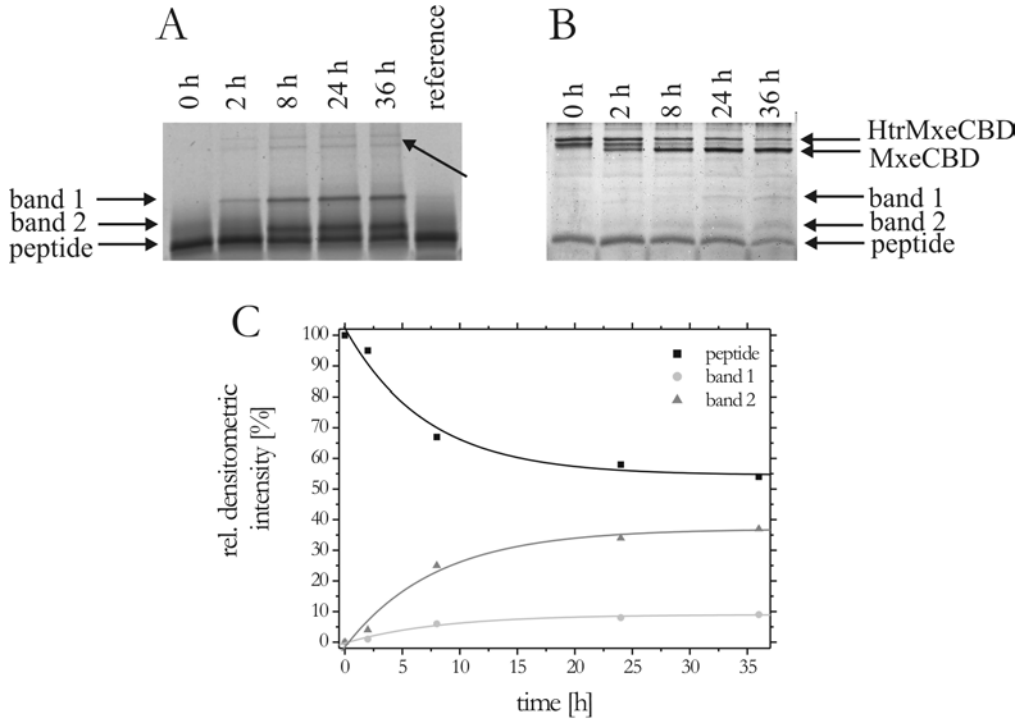


Figure 3.16. SDS-PAGE of expressed protein ligation of fluorescent Htr analogue.
Panel A: Densitogram (excitation: 473 nm, emission: 510 nm). Protein bands 1 and 2 that appeared during EPL were assigned to the desired EPL product and an N-terminally degraded product. As reference NBD-labelled peptide was used that was incubated under identical conditions.
Panel B: Coomassie staining. Disappearing of the protein bands of the full-length and the truncated intein fusion protein is slowed down referring to intein cleavage tests. Presence of HtrMxeCBD protein band after 36 h indicated incomplete intein cleavage during EPL.
Panel C: Time course of the EPL reaction based on densitometric protein band evaluation of unligated peptide (black), degraded (grey) and desired ligation product (light grey). Data traces were individually fit according to a first order reaction model revealing time constants of 7.1 h^{-1} (peptide), 7.9 h^{-1} (degraded product) and 8.2 h^{-1} (desired product).

Both analysis indicated that during the EPL reaction two new protein bands above the peptide reference band appeared (figure 3.16, panel A and B, band 1 and 2). Taking into account the results of the analysis of the previously described intein cleavage experiments using SDS-PAGE (figure 3.14, panel B) and MS (figure 3.15, panel A and B), which suggested the presence of two thioester species, these new bands were assigned to the desired target transducer and to an N-terminally degraded transducer variant.

In addition, the densitometric analysis of the obtained SDS-PAGE gel revealed two further faint bands with mobilities comparable to those observed for the cleaved and uncleaved intein species (figure 3.16, Panel A, indicated by arrow). Unspecific adsorption of the fluorescent peptide to the intein fusion proteins and to the spliced intein could be a possible explanation for this observation[76].

Coomassie staining of the SDS-PAGE gel showed a protein band of uncleaved HtrMxeCBD (starting material) after 36 h of reaction time (figure 3.16, panel B). This result suggests a decelerated intein cleavage kinetics referring to the cleavage tests without peptide (see section 3.3.1.1).

The graphic representation of the densitometric analysis of the peptide, the degraded and the desired product plotted against reaction time is given in figure 3.16, panel C. The data were normalised to the observed densitometric intensity of the peptide at the beginning of the reaction. After 36 h reaction time the amount of unligated peptide was reduced to 54 % corresponding to a ligation efficiency of 46 %. The relative amount of degraded product (37 %) is approximately four times higher than the amount of the desired product (9 %). Data traces were individually fitted according to a first order reaction model revealing time constants of 7.1 h^{-1} (peptide), 7.9 h^{-1} (degraded product) and 8.2 h^{-1} (desired product).

The ligation efficiency could be slightly improved in further experiments by prolonging the reaction time to 72 h and shifting the stoichiometry of the reaction to a 30 % HtrMxeCBD excess. However, a quantitative ligation of the peptide could not be achieved in any case (results not shown).

First purification experiments of this reaction mixture consisting of the three major compounds, the residual peptide, the long and the short ligation product were performed by analytical size exclusion chromatography (SEC) using a Superdex 200 column and 300 mM NaCl, 50 mM NaPi, 0.05 % DDM, pH 8.0 as running buffer with a flow-rate of 0.5 ml/min.

.

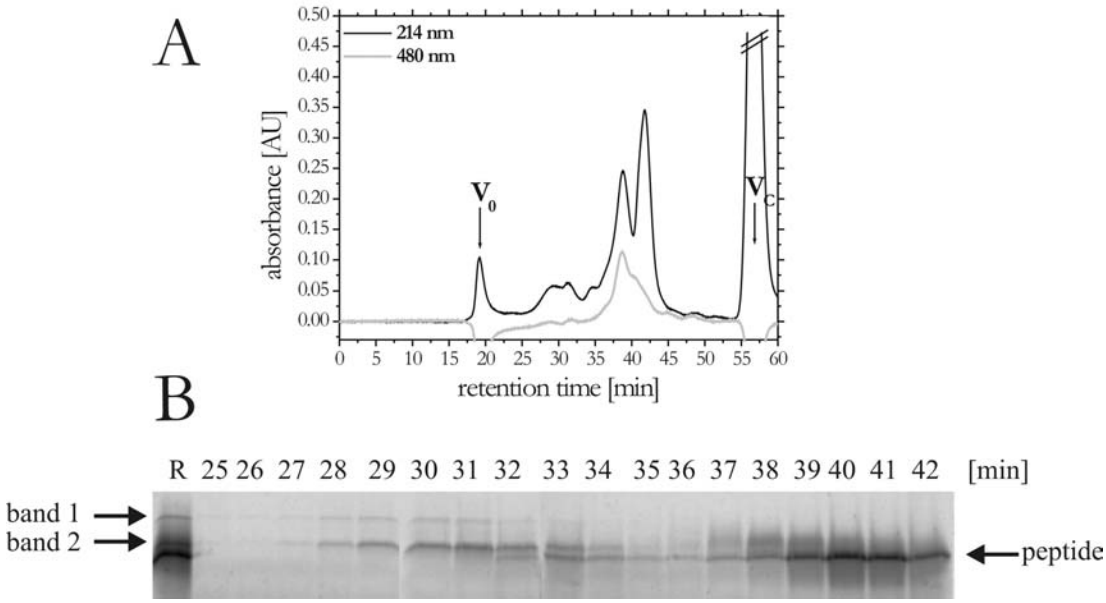


Figure 3.17. Purification of EPL reaction mix using size exclusion chromatography
Panel A: SEC chromatogram monitoring 214 nm and 480 nm (absorption maximum of NBD). Superdex 200 column with 300 mM NaCl, 50 mM NaPi, 0.05 % DDM, pH 8.0 as running buffer and a flow-rate of 0.5 ml/min. Peaks after 30 min and 39 min were assigned to the co-eluting long and short ligation products and to unligated peptide The signal after 42 min was not assigned.
Panel B: Densitogram (excitation: 473 nm, emission: 510 nm) of SDS-PAGE gel analysing SEC fractions The two protein bands (band 1 and 2) of fractions 27 min to 34 min were assigned to co-eluting full-length and degraded EPL product. Lower band (fractions 32 min to 42 min) originates from unligated peptide.

The chromatogram (figure 3.17, panel A) displayed a first peak after 19 min corresponding to the void volume of the column. After approximately 30 min a broad peak appeared at 480 nm. At this wavelength the absorbance maximum of the NBD-label is monitored. With analytical SDS-PAGE analysis (figure 3.17, panel B) this signal could be assigned to the co-eluting long and short ligation products, which were not separable under these conditions. Two further peaks appeared after 39 min and 42 min whereas only the first one was identified as the fluorescent peptide. The second one was not assigned. Residual MESNA that was used during EPL caused the last strong signal after 57 min. The presence of detergent in the buffers reduced the resolution of the size

exclusion chromatography column and therefore an insufficient separation of the three major components of the reaction mixture was achieved.

Analytical anion exchange chromatography using a 5x50 mm MonoQ column equilibrated under low salt conditions (10 mM Tris, 10 mM NaCl, 0.05 % DDM, pH 8.0) was also tested. The crude ligation mixture was dialysed twice against low salt buffer to facilitate binding of the analytes to the column matrix and subsequently loaded on the anion exchange column. Elution was performed using a linear gradient from 10 mM to 350 mM NaCl in 7 min and analysed using SDS-PAGE. The gel (figure 3.18.) revealed two protein bands (band 1 and 2) in fractions 4.5 min to 7.0 min that were assigned to the desired respectively the degraded ligation product, which were not separable under these conditions and co-eluted at high salt concentrations. However, unligated peptide eluted under low salt conditions indicated by a single band in fractions 0.5 min to 4.0 min and was therefore successfully separated from the two ligation products.

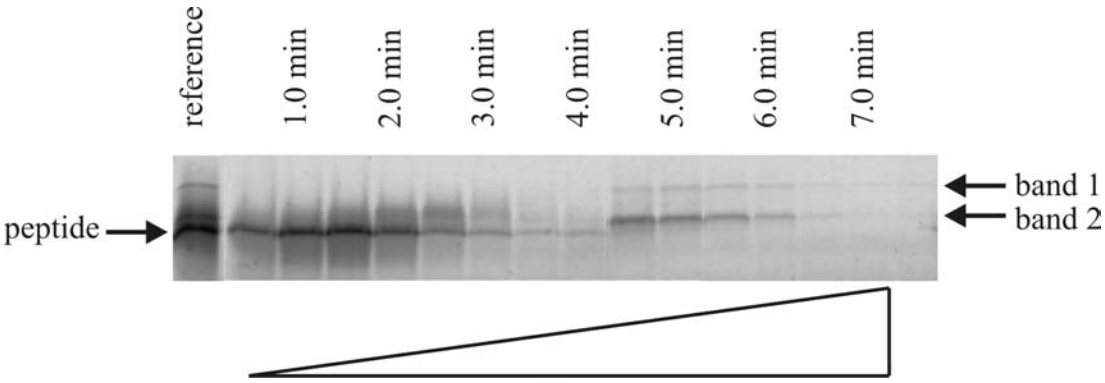


Figure 3.18. Purification of EPL reaction mix using anion exchange chromatography. 5x50 mm MonoQ column equilibrated in 10 mM Tris, 10 mM NaCl, 0.05 % DDM, pH 8.0. Elution by linear gradient from 10 mM to 350 mM NaCl in 7 min (indicated by triangle) and monitored by SDS-PAGE.
Densitogram (excitation: 473 nm, emission: 510 nm) of the obtained gel: The two protein bands (band 1 and 2) of fractions 4.5 min to 7.0 min were assigned to co-eluting full-length and degraded EPL product. Single band in fractions 0.5 min to 4.0 min indicates residual peptide successfully separated from the both ligation products.

Based on the results of these analytical purification experiments, anion exchange chromatography was chosen for preparative scale purification of EPL products. Due to the lack of a preparative MonoQ anion exchange chromatography column, this material

was replaced in accordance to the already established purification protocols for recombinant Htr variants[7,20] by DEAE anion exchange material.

In first preparative purification attempts the crude EPL reaction mix was dialyzed against low salt buffer to adjust ion strength and to remove excess MESNA. Later on, dialysis was replaced by simple 1:10 dilution with salt free buffer. This procedure was less time and material consuming and did not negatively affect the purification procedure. The diluted sample was loaded onto the DEAE column, which was then washed with a buffer containing 80 mM NaCl. Elution was performed by washing of the column with elution buffer containing 500 mM NaCl Purification was monitored with analytical SDS-PAGE.

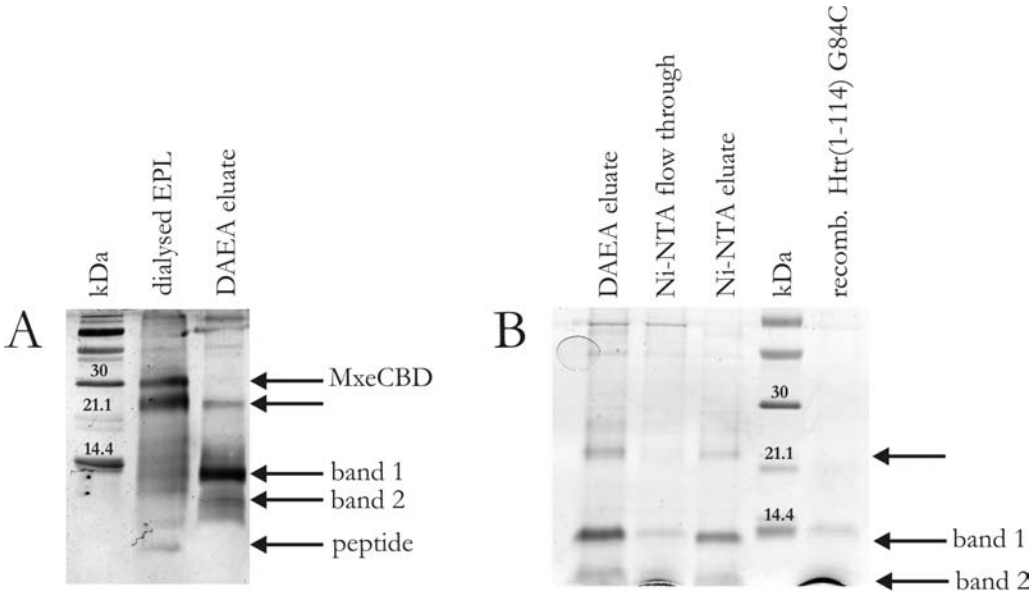


Figure 3.19. Analytical SDS-PAGE of expressed protein ligation purification
Panel A: DEAE anion exchange chromatography after 72 h reaction time. After purification no bands of cleaved intein (MxeCBD) and unligated peptide could be observed. The elution fraction revealed bands with mobilities ≥21.1 kDa (arrow) and ≤14.4 kDa (band 1) in addition to a third band with higher mobility (← band 2).
Panel B: Ni-NTA affinity chromatography. The two bands ≥21.1 kDa and ≤14.4 kDa in addition to the band with higher mobility (← band 2) correspond to dimer and monomer of the full-length ligation product and to the degraded one specifically interacting with Ni-NTA matrix.

Figure 3.19, panel A shows a silver stained SDS-PAGE gel monitoring the purification of the EPL reaction after reaction time of 72 h. After removal of MESNA by dialysis,

the EPL product appeared as an intensive band with a molecular weight ≥21.1 kDa (figure 3.19, panel A, indicated by arrow) below the protein band of the cleaved MxeCBD intein. The observed mobility of that band suggested the formation of a homo-dimer that was also observed for the recombinant G84C mutant (see section 3.1). Reasons for that may be oxidative disulfide bridging of the Cys84 residue in the absence of MESNA after dialysis or ionic interactions under the conditions with low salt concentrations.

Protein bands for uncleaved HtrMxeCBD (starting material) were not present indicating a successful improvement of the intein cleavage efficiency according to the prolonged reaction time from 36 h (see figure 3.16, panel B) to 72 h.

SDS-PAGE analysis of the DEAE elution fraction after purification revealed the appearance of a protein band below the 14 kDa reference band (figure 3.19, panel A, band 1) that was in a good agreement with the expected mass of the desired ligation product. In addition, weakening of the band with a molecular weight ≥21.1 kDa was observed (figure 3.19, panel A, indicated by arrow).

The presence of a protein band (← band 2) below the band of the monomeric target protein suggested incomplete separation of the two EPL products in the preparative purification scale. However, an enrichment of the desired full-length product was observed. The absence of protein bands of the cleaved MxeCBD intein and the unligated peptide suggest a quantitative separation of these compounds from both EPL products.

In order to investigate the possible dimer formation of the full-length EPL product under the low salt conditions of the DEAE purification protocol (figure 3.19, panel A, indicated by arrow), a Ni-NTA metal affinity chromatography was carried out.

Since the EPL ligation products are the only compounds of the DEAE elution fraction comprising an artificial histidine purification tag (the tag of the used NBD-labelled peptide), a dimer of the full-length product should specifically interact with the Ni-NTA matrix. However, a contamination with comparable molecular weight should be washed away.

Therefore a sample of the DEAE elution fraction was loaded onto an analytical Ni-NTA spin column equilibrated with 300 mM NaCl, 50 mM NaPi, 0.05 % DDM, pH 8.0, washed with the respective buffer containing 20 mM imidazole and eluted with

300 mM NaCl, 200 mM imidazole, 50 mM NaPi, 0.05 % DDM, pH 8.0. The experiment was monitored by SDS-PAGE.

A Coomassie stained gel (figure 3.19, panel B) documented two protein bands for the DEAE eluate as starting material and for the Ni-NTA eluate (indicated by arrow and band 1) in addition of a third band with higher mobility (band 2). Due to the observed specific interaction with the Ni-NTA matrix and the results for the recombinant Htr(1-114) G84C mutant (see section 3.1), these bands were assigned to the dimer and the monomer of the full-length ligation product as well as to the degraded transducer.

Compared to a recombinant (1-114)*Np*HtrII reference, the full length EPL product exhibited a slightly higher mobility in SDS-PAGE, which corresponded to the mass difference between the two proteins due to the absence of the amino acid residues 110-114 in the semi-synthetic transducer variant.

In further experiments the results of this model EPL reaction should be transferred to ligations with isotopic labelled peptides. Special attention should be paid to the question, if the desired full-length ligation product with the native C-terminus can be separated, in contrast to the ligation product with the NBD-labelled terminus, from the N-terminally degraded variant.

The ligation experiments were therefore repeated with a [^{15}N] labelled peptide under the established conditions. Yields were spectroscopically determined according to Ehresmann et al.[55] and were usually around 50 % compared to the used amount of peptide. The resulting product was purified as described using preparative DEAE anion exchange chromatography.

In first MALDI-MS experiments performed to characterise the ligation product, no valid data with sufficient spectral resolution could be obtained. Therefore the biochemical evaluation of the semi-synthetic product was performed using alternative biochemical methods.

In order to test ligation product for the free cysteine side chain as a result of a successful EPL, this residue was labelled with I-AEDANS, a thiol reactive fluorescence dye[77]. A purified EPL sample was incubated at 37 °C with 10 mM I-AEDANS. After 2 h the reaction was quenched by the addition of 40 mM DTT and further incubation for 1 h at 37 °C. Subsequently the sample was analysed using SDS-PAGE densitometric fluorescence experiments.

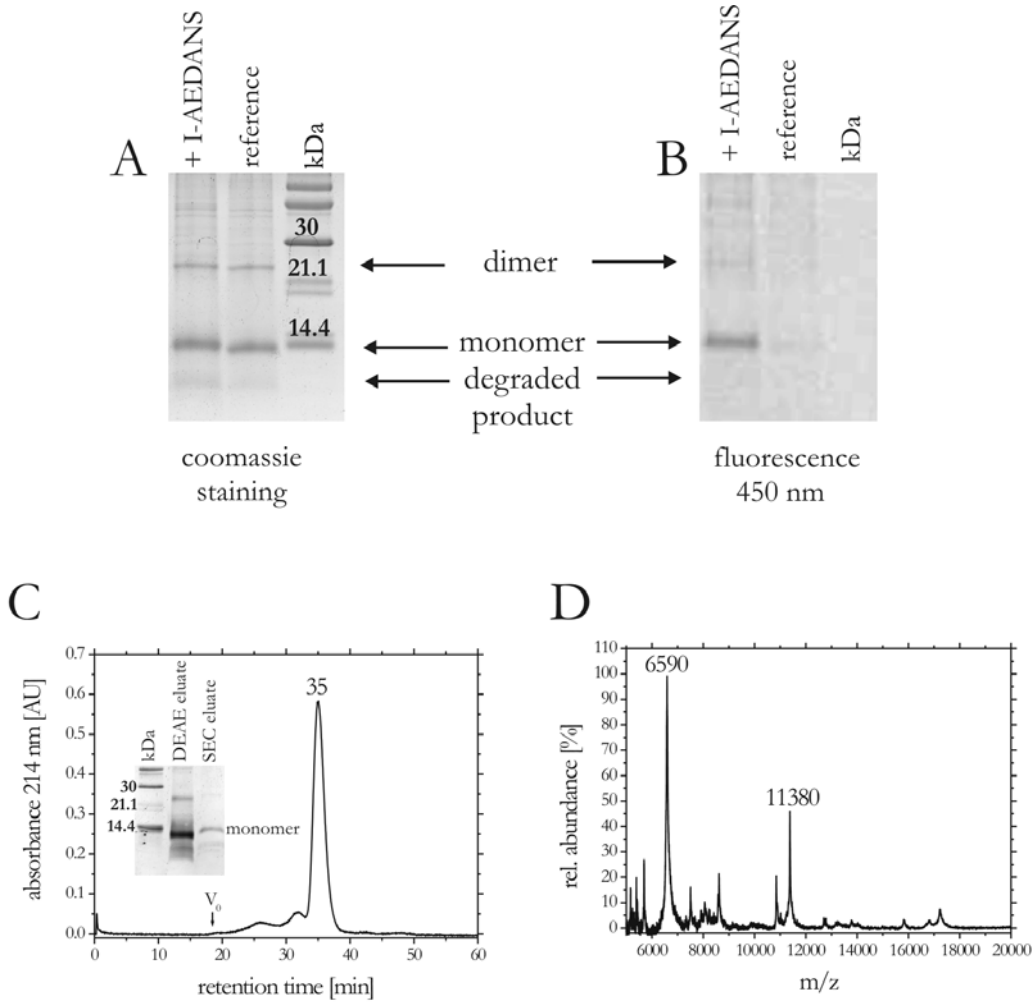


Figure 3.20. Biochemical characterisation of the purified [^{15}N] labelled EPL product.
Panel A: SDS-PAGE gel of purified and I-AEDANS labelled product. Protein bands below 14.4 kDa and above 21.1 kDa were assigned to the monomeric and dimeric full-length product. The faint band below indicated traces of N-terminally degraded by-product.
Panel B: Densitogram of the corresponding gel (excitation: 360 nm; emission at 450 nm). The intensive and the fate signal with less mobility corresponded to the full-length monomer and dimer.
Panel C: Size exclusion chromatogram. The dominant signal after 35 min was assigned to the monomeric full-length ligation product due to SDS-PAGE (inset).
Panel D: MALDI-MS. The obtained spectrum of the EPL product showed signals at 11380 amu and 6590 amu.

The Coomassie stained SDS-PAGE gel (figure 3.20, panel A) of a sample of the labelled protein displayed a protein band below the 14.4 kDa reference and a second band of less intensity above the 21.1 kDa reference. The bands were assigned according to the results of the NBD-labelled model reaction to the monomer and the dimer of the desired full-length transducer.

A mobility comparison of the labelled and the unlabelled samples (figure 3.20, panel A, lane 1 and 2) displayed a remarkable shift of the fluorescence labelled protein to higher molecular masses. A faint band below the protein band of the full-length monomer indicated traces of N-terminally degraded by-product of the EPL reaction.

The densitogram of the SDS-PAGE gel upon excitation at 360 nm revealed an intensive fluorescent signal and a faint signal with less mobility (figure 3.20, panel B, lane 1). A Comparison with the results obtained by Coomassie staining revealed the correlation of the fluorescence signals with the protein bands of the full-length monomer and dimer.

In order to exclude unspecific adsorption of the lipophilic label to the membrane proteins, the cysteines in a second sample were first capped with iodo acetamide and then incubated with the fluorescent I-AEDANS label. This reference sample showed only a minor fluorescence signals (figure 3.21, panel B, lane 2).

In summary the I-AEDANS labelling experiments confirmed the presence of the free cysteine side chain of the ligation product as a result of a successful EPL.

In further experiments samples of the purified ligation product were analysed for residual unligated peptide that would disturb subsequent NMR experiments. Size exclusion chromatography (SEC) using a Superdex 200 column was found to be a sufficient method to detect residual peptide in purification experiments of the ligation mixture with the NBD-labelled model peptide. These experiments were therefore repeated with a sample of the isotope labelled EPL product. The chromatogram of the ligation product revealed a dominant signal after 35 min (figure 20, panel C), which could be assigned to the ligation product by SDS-PAGE (figure 20, panel C, inset). No signals for residual peptide were found.

The above-mentioned low spectral resolution in MALDI-MS experiments finally could be improved by modifying the protocol introduced by Hufnagel et al. in a way, that the matrix was dissolved in 66 % acetonitrile in 1 % aqueous TFA. This led to a more homogenous crystallization of the samples on the ultra-thin matrix-layer.

The mass spectrum (figure 3.20, panel D) of the ligation product showed molecular masses of 11380 Da and 6590 Da.

The expected mass of the full-length [^{15}N] labelled product was calculated to be 11605 Da. The mass difference of ΔM=225 Da cannot be explained with a single N-terminal methionine deletion[78]. In order to address this problem, a set of reference MALDI measurements were carried out using samples of (1-157)*Np*HtrII, (1-114)*Np*HtrII and (1-114)*Np*HtrII G84C. Sample preparation and instrument settings were comparable for all measurements. The results are summarized in table 3.1.

Protein	Theoretical mass [Da]	Experimental mass [Da]	ΔM [Da]
[^{15}N] $_{84\text{-}114}$ Htr(1-114) G84C	11564+41=11605	11380[this work]	-225
Htr(1-157)	17421	17222[this work,62]	-199
Htr(1-114)	12663	12543[this work,62]	-120
Htr(1-114) G84C	12709	12622[this work]	-87

Table 3.2. Theoretical and experimental masses of different transducer variants.

In general all experimental masses were significantly smaller than the theoretically expected masses. This phenomenon was previously observed for heterologously expressed *N.pharaonis* transducers as well[62].

In the cases of the recombinant Htr(1-157) and ligated [^{15}N] Htr(1-114), the obtained masses were even smaller than an N-terminal methionine deletion (ΔM=-131.2 Da) can explain (table 3.2, column ΔM). Interestingly the experimental masses are in good agreement with theoretical masses considering a loss of the first two amino acids in both proteins: M=17219 for Htr(3-157) and M=11387 for [^{15}N] $_{84\text{-}114}$ Htr(3-114) G84C. An explanation of this phenomenon could not be found, since the low resolution of the spectra did not allow a more detailed analysis.

In summary the mass deviation of the long ligation product was similar to the one found for the reference protein from solely recombinant source and does not contradict signal assignment that was made.

According to the hypothesis of an N-terminal degradation of the HtrMxeCBD precursor at the loop between TM1 and TM2 (see section 3.2.1.1), the observed mass of 6590 amu (figure 3.20, panel D) can be explained by a [^{15}N]-labelled Htr(48-114) fragment with a theoretical mass of 6525+41=6566 Da. In the range of the obtained spectral resolution this theoretical mass matches the experimental mass quite well.

3.3.2 EPL utilising HtrMxeCBD isolated from *E.coli* inclusion bodies

3.3.2.1 HtrMxeCBD purification and folding

The expressed protein ligation utilises inteins to generate activated α-thioester species that are subsequently joined together with a peptide containing an N-terminal cysteine. The intein fusion protein HtrMxeCBD was successfully expressed in *E.coli*. However, Ni-NTA purification of the DDM solubilised membrane fraction revealed the presence of N-terminally degraded HtrMxeCBD form (see section 3.3.1). Since this intein fusion species was fully functional, in EPL reactions this contamination led to two products, the desired target protein and a respective N-terminally degraded by-product. A separation of the desired full-length semi-synthetic transducer from the shortened product or a separation of the full-length intein from the truncated one could not be achieved using size exclusion, anion exchange or Ni-NTA chromatographic purification methods.

The analysis of HtrMxeCBD protein partitioning revealed, that *E.coli* inclusion bodies consisted predominantly of the desired long precursor (figure 3.13). The establishment of a successful EPL utilising this well-defined target protein circumvents the contamination of the segmental isotope labelled transducer with an N-terminally degraded by-product. Therefore the following section describes attempts to purify, to refold the full-length HtrMxeCBD intein fusion protein from purified inclusion bodies and to use the refolded protein in EPL reactions to generate a segmental isotope labelled transducer.

In a first step the isolation and purification of inclusion bodies was performed. Therefore the inclusion bodies were separated from cytoplasma by centrifugation, washed with cell washing buffer containing 0.25 % Triton X-100 to remove residual lipids and subsequently two times with cell washing buffer to remove the detergent. Approximately 100 mg off-white pellets of inclusion bodies per gram wet cells could be isolated. The resulting product was analyzed by performing SDS-PAGE and MALDI-MS.

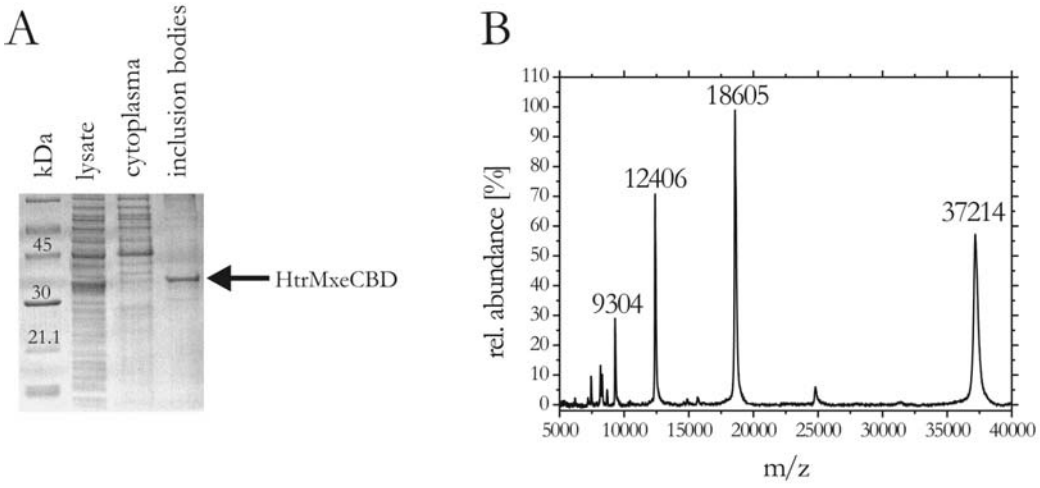


Figure 3.21. Purification and biochemical analysis of HtrMxeCBD inclusion bodies
Panel A: SDS-PAGE analysis. The single protein band above the 30 kDa reference band for the inclusion bodies fraction suggested a highly pure full-length precursor protein.
Panel B: MALDI mass spectrum. Signals at 37214 amu, 18605 amu, 12406 amu and 9304 amu corresponded to a molecular weight of M=37213 Da and was assigned to an N-terminal methionine deleted HtrMxeCBD (M_{calc}=37226).

SDS-PAGE monitoring the isolation of inclusion bodies documented a single protein band above the 30 kDa reference band for the fraction of inclusion bodies suggesting a highly pure full-length precursor protein (figure 3.21, panel A). An N-terminally degraded variant, as described for the HtrMxeCBD isolated from membranes, was detectable only in a very small amount (figure 3.21, panel A).

The mass spectrum of the purified inclusion bodies displayed signals for one protein species carrying different numbers of positive charges $(M+H)^{+}$=37214, $(M+2H)^{2+}$=18605, $(M+3H)^{3+}$=12406 and $(M+4H)^{4+}$=9304 (figure 3.21, panel B) that corresponded to an experimental mass of M=37213 Da. The mass difference of ΔM=-145 Da compared to the calculated mass of the desired intein fusion protein HtrMxeCBD (M=37357 Da) and can be explained by an N-terminal methionine deletion[78].

These results confirmed the high purity of the HtrMxeCBD protein isolated from inclusion bodies as was already indicated on the SDS-PAGE gel. Therefore no additional Ni-NTA purification step was carried out.

The process of intein cleavage requires a functionally folded intein. In order to obtain such a functional HtrMxeCBD protein isolated from inclusion bodies, the isolated aggregates had to be completely solubilised and subsequently refolded. A variety of

solubilisation and refolding conditions were tested. Figure 3.22 summarizes the tested conditions in a flowchart.

First denaturing experiments were performed using 8 M urea as denaturant (8 M urea, 25 mM NaPi, 300 mM NaCl, 0.02 %DDM). However, under these conditions complete solubilisation of the aggregates was not possible. Therefore in later experiments inclusion bodies were dissolved in 6 M guanidine hydrochloride and subsequently dialysed against the above-mentioned 8 M urea-containing buffer to lead back the new solubilisation protocol to the already established refolding conditions.

Starting from 8 M urea, the chaotrope concentration was lowered stepwise by dialysis to 6 M, 4 M, 3 M and finally 2 M urea at 4 °C. Centrifugation steps performed with high centrifugal forces after each dialysis step have been proven to be useful, because although intein solutions appeared to be free of aggregates, often a small amount of aggregated material was isolated by centrifugation. This helped to prevent further aggregation.

Dialysis in absence of a reducing agent let to almost quantitative protein aggregation at 4 M urea concentrations. An initial reduction of the protein with 10 mM TCEP over night combined with 1 mM TCEP present in all dialysis buffers facilitated decreasing the urea concentration to 2 M without significant losses due to precipitation.

Intein cleavage experiments under these conditions with 1 % MESNA gave rise to cleavage efficiencies of approximately 50 % after two days, which was documented by two protein bands of comparable intensity in SDS-PAGE analysis, which were assigned to the cleaved and the uncleaved intein precursor protein (figure 3.23, Panel A).

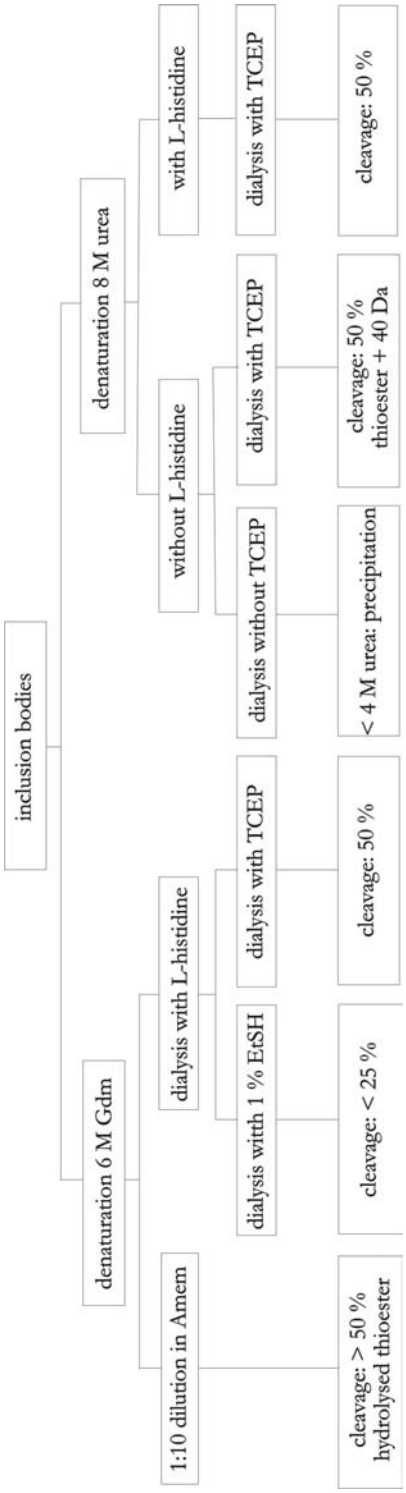


Figure 3.22. Summary of HtrMxeCBD refolding experiments. All dialyses started either directly from 8 M urea containing buffer or from 6 M guanidine hydrochloride, which was replaced in a first dialysis step by 8 M urea containing buffer. Subsequently, the chaotrope concentration was lowered stepwise by dialysis to 6 M, 4 M, 3 M and finally 2 M urea at 4 °C. Centrifugation steps with high g-forces were performed after each dialysis step.

MALDI-MS experiments performed to identify the desired Htr(1-83)-COSR thioester, revealed a signal at 8480 amu (figure 3.23, Panel B). This signal correlated to an experimental protein mass that was about 40 Da larger than the expected 8446 Da for the desired MESNA-thioester species. The observed mass difference was reproducible for different protein preparations excluding a measuring artefact due to the obtained low spectral resolution. A possible explanation for this mass difference is a carbamylated thioester protein: One of the disadvantages of using urea is that it can form cyanic acid, which can modify amine or thiol groups to yield a carbamylated protein (ΔM=+40 Da)[79].

In order to verify this hypothesis, 20 mM L-histidine was added as a scavenger for cyanic acid to all dialysis buffers[80] to prevent the postulated carbamylation reaction. The additive successfully suppressed the modification as was documented by MALDI-MS experiments. The MALDI MS spectrum (figure 3.23, Panel C) reproducibly showed a cleaved thioester species with M^+=8422 Da which corresponded nicely to the expected molar mass of the thioester without Na^+ as a counter-ion of the mercaptoethane sulfonic acid.

However, by the addition of L-histidine as a scavenger the refolding yields were reduced to ca 50 % compared to the experiments without scavenger due to protein precipitation (data not shown). It remained unclear, if the L-histidine itself caused this precipitation or if the protein had different solubility characteristics upon cabamylation compared to the unmodified protein.

In order to prevent such extensive protein precipitation, similar experiments were carried out when TCEP was replaced with ethane thiol. In contrast to TCEP, ethane thiol is able to induce the cleavage of the intein fusion protein[81]. The idea behind this substitution was to shift the equilibrium between the folded and the unfolded state of the protein by inducing the cleavage of once folded protein and thereby taking the folded protein out of the folding equilibrium. Indeed the tendency to form aggregates could be significantly reduced (data not shown).

SDS-PAGE analysis was performed to determined the cleavage efficiency of the refolded precursor in the presence of 1 % EtSH after the addition of 1 % MESNA. The gel (figure 3.23, Panel A) revealed a relative intensity of the protein bands obtained for the uncleaved compared to the cleaved HtrMxeCBD precursor of approximately 3:1 corresponding to a cleavage efficiency of ca. 25 %. In summary the replacement of TCEP by EtSH could not improve overall refolding/cleavage efficiency.

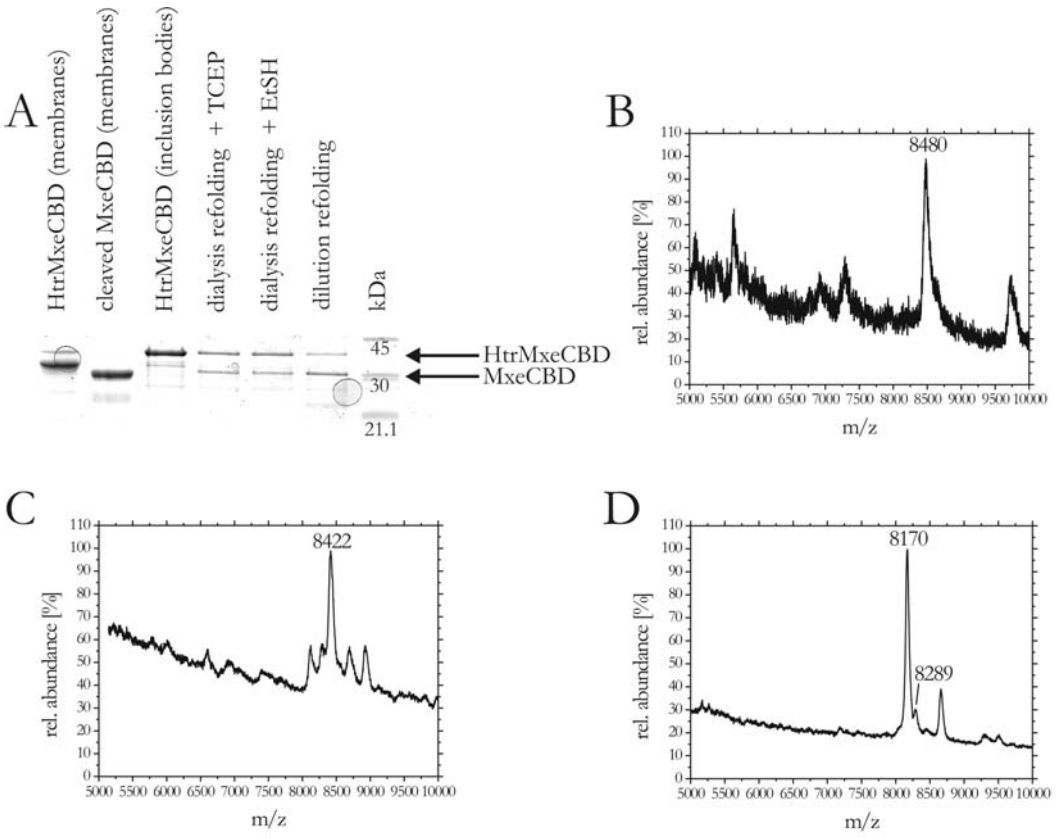


Figure 3.23. Refolding and cleavage of solubilised HtrMxeCBD from inclusion bodies
Panel A. SDS-PAGE analysis of intein cleavage of refolded HtrMxeCBD from inclusion bodies referring to samples purified from solubilised *E.coli* membranes. Intensities of the protein bands of cleaved and uncleaved HtrMxeCBD upon 1 % MESNA addition revealed cleavage efficiencies of ca. 50 %, 25 % and 75 % for refolded mature protein refolded by dialysis in presence of TCEP respectively EtSH or by fast 1:10 dilution after two days.
Panel B. MADLI-MS of cleaved HtrMxeCBD refolded by dialysis without L-histidine containing urea buffers. Signal at 8480amu was assigned to carbamylated Htr-COSR (M_{calc}=8486 Da).
Panel C. MALDI-MS of cleaved HtrMxeCBD refolded by dialysis with L-histidine containing urea buffers. Signal at 8422 amu was assigned to Htr-COSR without sodium counter-ion of the of the mercaptoethane sulfonic acid (M_{calc}=8423 Da).
Panel D. MALDI-MS of cleaved HtrMxeCBD refolded by fast 1:10 dilution into guanidine hydrochloride free buffer. Signals at 8170 amu and 8289 amu were assigned to hydrolysed Htr-COOH with and without N-terminal methionine (M_{calc}=8300 Da and 8169 Da).

As alternative to the refolding protocol using dialysis, rapid 1:10 dilution of the solubilised HtrMxeCBD protein was carried out. Starting from a solution in 6 M

guanidine hydrochloride and 2 % DDM (6 M Gdm, 300 mM NaCl, 50 mM NaPi, 2 % DDM, 1 mM TCEP, pH 8.0) dilution was performed by pumping the solution with a flow rate of 0.1 ml/min into the respective guanidine hydrochloride free buffer containing 0.05 % DDM (A_{mem}). Almost no losses due to precipitation were observed.

Protein cleavage was induced by 1 % MESNA and reached efficiencies higher than 50 % determined by the comparison of the protein bands of cleaved and uncleaved intein precursor on a SDS-PAGE gel (figure 3.24, panel A).

MALDI-MS experiments were performed to identify the desired Htr(1-83)-COSR thioester. The mass spectrum (figure 3.23, panel D) showed a major signal at M^+=8170 Da and a minor signal at M^+=8289 Da, which could be explained by an hydrolysed Htr(2-83)-COOH species with the N-terminal methionine cleaved off (calculated mass M^+=8169 Da) and an hydrolysed Htr(1-83)-COOH cleavage product (calculated mass M^+=8300 Da).

An explanation for this hydrolysis cannot be given. Sine it is not reasonable to assume altering stability properties of successfully formed Htr-MESNA thioester species under the different folding conditions performed with comparable buffers with altering chaotrope concentrations, the formation of the thioester species itself might be a critical point.

A possible explanation could be the refolded intein precursor (refolded by dilution) adopts a protein fold that facilitates water molecules to enter the active centre of the intein. As a possible consequence a nucleophilic attack of water molecules to the thioester intermediate formed upon the initial N-S acyl shift as the first step of the protein splicing mechanism[82-85] (see section 6.2.1), sets free the hydrolysed Htr-COOH N-extein. In contrast to this, the mature protein refolded by slow dialysis processes successfully prevents water entering the active centre of the intein, thus facilitating the thiol additive to trap the initial N/S acyl shift intermediate and to form an activated Htr-MESNA thioester species.

In summary, the isolated inclusion bodies were denatured and subsequently functionally refolded using different protocols. An activated α-thioester was generated with cleavage efficiencies of approximately 50 % using 1 % MESNA after refolding the HtrMxeCBD protein using dialysis in order to lower urea concentrations to 2 M in the presence of TCEP and L-histidine. Unfortunately, all attempts to isolate the resulting thioester species using HPLC purification methods failed due to irreversible adsorption of the protein to the material of the reverse phase column.

3.3.2.2 Expressed Protein Ligation

The successful refolding of the mature HtrMxeCBD protein from inclusion bodies facilitated a preparative access to a functional intein precursor without the presence of an N-terminally degraded by-product. The alternative purification protocol using denaturing and refolding of over-expressed inclusion bodies was an important prerequisite for successful segmental isotope labelling of *Np*Htr(1-114) using EPL. The following experiments should establish an EPL protocol using refolded HtrMxeCBD fusion protein and isotope labelled peptides.

In preliminary attempts, the optimized EPL protocol using the HtrMxeCBD protein isolated from membranes (see section 3.3.1.2) that describes the *in-situ* cleavage of the intein in presence of the isotope labelled peptide, was used. However, using this protocol no ligation product could be detected, either in SDS-PAGE or in MALDI-MS experiments.

Taking these conditions as a starting point, a set of EPL conditions using denaturing chaotropes or organic solvent additives was screened. In order to enhance the accessibility of the C-terminal thioester moiety in the Htr(1-83) segment, the EPL was performed under denaturing conditions. Therefore in a first step the cleavage of the refolded HtrMxeCBD protein was initiated with 3 % ethane thiol in the absence of any chaotrope[81]. In a second step the solution was adjusted to denaturing conditions by the addition of 6 M guanidine hydrochloride with or without 30 % TFE. This step rather then stopping the cleavage reaction by unfolding the Mxe intein, enhances the C-terminal accessibility of the N-terminal transmembrane segment. The ligation itself was subsequently initiated by the addition of 1 % MESNA to generate a more reactive thioester species and by the addition of the C-terminal peptide fragment (1:1 stoichiometry). After 48 h of incubation (72 h after cleavage initiation) samples were precipitated with 5 % TCA, washed and monitored with analytical SDS-PAGE and MALDI-MS. (figure 3.24).

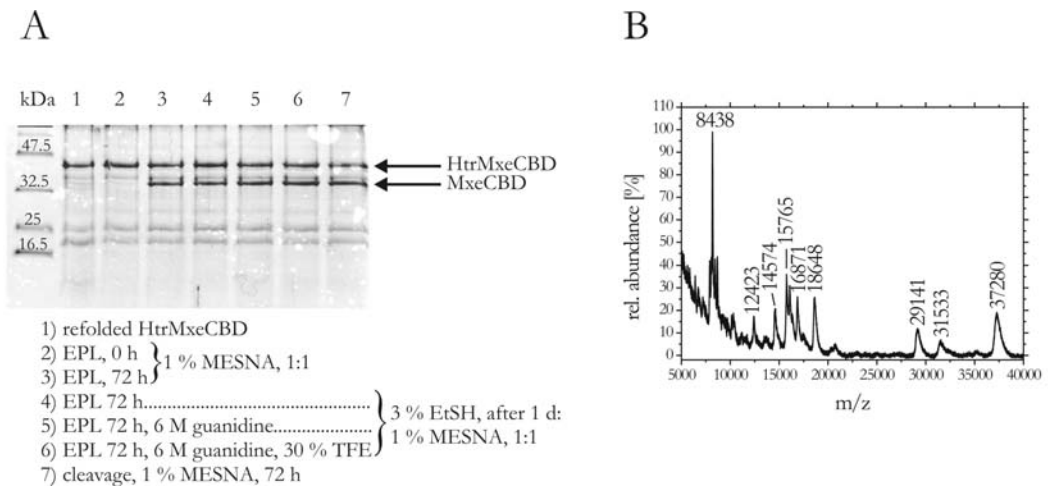


Figure 3.24. Screening of EPL conditions using chaotropic or organic solvent additives.
Panel A: SDS-PAGE gel. Cleavage of the refolded HtrMxeCBD with 3 % EtSH without chaotropes. Adjustment to 6 M guanidine hydrochloride with or without 30 % TFE after 1 d. EPL initiation with 1 % MESNA and [^{15}N] peptide (1:1). The two protein bands at 37 kDa and 30 kDa are assigned to uncleaved and cleaved precursor. No product band at 11.5 kDa is observed in any case.
Panel B: MALDI-MS of crude EPL without chaotropic or organic solvent additives after 72 h. Signals for uncleaved HtrMxeCBD at 37280 amu, 18648 amu, 12423 amu, for cleaved MxeCBD 29141 amu, 14574 amu, for degraded HtrMxeCBD 31533 amu, 15765 amu and for unligated α-thioester 8438 amu, 16871 amu. A signal at 11380 amu for the desired product is not observed.

The SDS-PAGE gel showed two protein bands of comparable intensities with mobilities corresponding to ca. 37 kDa and 30 kDa. According to the results obtained from HtrMxeCBD expression and intein cleavage experiments (see section 3.3.1) these bands were assigned to the uncleaved and cleaved intein precursor protein. This observation suggests that all investigated samples reached an intein cleavage efficiency of approximately 50 % after 72 h. This was observed, in the previously described refolding and cleavage experiments of HtrMxeCBD isolated from inclusion bodies (see section 3.3.2.1). The adjustment of the EPL reaction mixture to denaturating conditions after 24 h of intein cleavage with ethane thiol, showed no significant effect on cleavage efficiencies.

Unfortunately, no protein band with an apparent molecular weight of about 11.5 kDa as expected for a formed ligation product appeared in the SDS-PAGE gel (figure 3.22, panel A). The two visible bands that were observed in all samples including the

references without any peptide between 16.5 and 25 kDa, were protein contaminations found for this HtrMxeCBD protein preparation.

In addition to the electrophoretic investigations MALDI-MS experiments were performed to evaluate the ligation reactions. The spectra did not reveal any product formation either.. The displayed MALDI spectrum represents an EPL performed with 1 % MESNA in the absence of chaotropes and shows signals for not cleaved HtrMxeCBD $(M+H)^+$=37280 Da, $(M+2H)^{2+}$=18648 Da, $(M+3H)^{3+}$=12423 Da, for cleaved MxeCBD $(M+H)^+$=29141 Da, $(M+2H)^{2+}$=14574 Da, for N-terminally degraded HtrMxeCBD protein $(M+H)^+$=31533 Da, $(M+2H)^{2+}$=15765 Da and for the not ligated α-thioester $(M+H)^+$=8438 Da, $(M_2+H)^+$=16871 Da (figure 3.22, panel B). A signal around 11380 amu (see section 3.3.1.2) for the desired [^{15}N] transducer was not observed in any case.

The MALDI spectra representing the EPL reactions performed under chaotropic conditions, displayed the same signals, differing only in their signal to noise ratio. Guanidine hydrochloride containing EPL samples, even after extensive washing, often did not produce M^+-signals of the intein species (results not shown).

The absence of any signal for the desired target protein remains unclear. A successful ligation reaction of Htr-thioester was expected as it was observed in experiments with HtrMxeCBD protein isolated from *E.coli* membranes (see section 3.3.1.2). The presence of chaotropes is reported to enhance thiol mediated ligation reactions of α-thioesters with peptides containing N-terminal cysteines[86-90].

In order to exclude inhibiting effects of TCEP and L-histidine that were used during the refolding of the HtrMxeCBD protein, control experiments without TCEP and L-histidine were performed. Here, as well no ligation product was detected (results not shown). An explanation for this phenomenon cannot be given. A reasonable hypothesis might be that refolding of the HtrMxeCBD protein leads to a cleavable intein that however, liberates a misfolded transmembrane Htr(1-83) fragment, whose C-terminus even in the presence of chaotropes is not accessible for the attacking thiol-moiety of the C-terminal EPL segment.

Up to this time-point, all chosen EPL conditions contained DDM as a detergent. In order to alter this factor, a further set of experiments was carried out, where dodecyl maltoside was replaced by the mild zwitterionic detergent CHAPS. The replacement was achieved by ion exchange chromatography. A sample of refolded HtrMxeCBD protein (17 mg) in DDM containing buffer (2 M urea, 0.5 M NaCl, 10 mM TrisHCl, 20 mM L-

histidine, 1 mM TCEP, 0.02 % DDM, pH 8.0) was diluted 1:10 with TrisHCl buffered CHAPS (2 M urea, 10 mM TrisHCl, 20 mM L-histidine, 1 mM TCEP, 3 % CHAPS pH 8.0) in order to lower the salt concentration below 80 mM. After loading the sample onto the column and a washing step with the respective buffer containing 80 mM NaCl in order to remove residual DDM, the protein was eluted with the respective buffer containing 500 mM sodium chloride. Protein yield was 10 mg corresponding to a recovery of about 60 %.

With this protein sample EPL was performed as per the above-mentioned set of conditions including chaotrope-free or 6 M guanidine hydrochloride with or without TFE containing buffers. Again no product formation was detected with SDS-PAGE and MALDI-MS.

A second attempt to avoid DDM as detergent was the usage of a lipidic cubic phase environment for the EPL[91]. In recent experiments the intein fusion protein in approximately 30 g/l concentrations was mixed in 2:1 stoichiometry with the C-terminal peptide segment and 200 mM MESNA. 1-Mono-oleoyl-*rac*-glycerol was rapidly mixed with this reaction mixture (0.67 g/ml)[9,91]. The mixture so obtained became highly viscous and turbid at once and formed a transparent gel after centrifugation. The lipidic cubic phase reaction mixture was incubated at 23 °C. After 3 days MALDI-MS investigations were performed with samples that were precipitated with 2 volumes of acetone and washed twice with ice-cold 70 % ethanol.

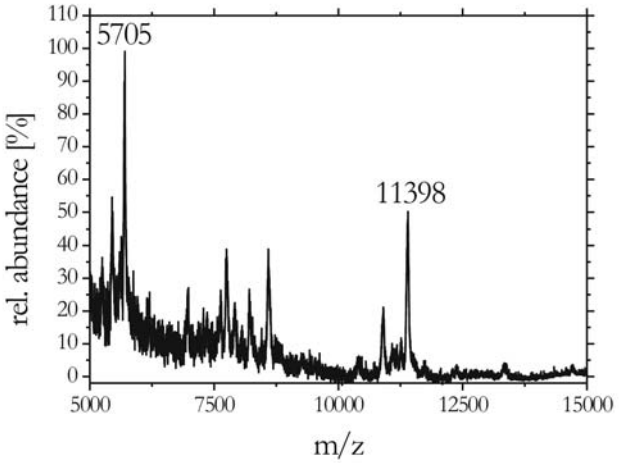


Figure 3.25. MALDI-MS of expressed protein ligation in lipidic cubic phase using a [^{15}N] labelled C-terminal peptide segment

The MALDI mass spectrum of the crude EPL reaction mixture containing a [^{15}N] labelled peptide segment displayed two major peaks in a mass range up to 20 kDa, m/z=11398 amu and m/z=5705 amu. Within the measuring accuracy the signals could be assigned to $(M+H)^+$=11398 Da and $(M+2H)^{2+}$=5705 Da signals of the $[^{15}N]_{84\text{-}114}$(3-114)Htr G84C target protein suggesting a successful ligation of the transducer in lipidic cubic phase environment. The experimental mass was in a good agreement with the calculated mass of 11387 Da and the molecular weight of 11380 Da that was detected in experiments with the HtrMxeCBD protein isolated from *E.coli* membranes (see section 3.3.1.2). A signal of unligated Htr-thioester with 8446 amu was not detected.

Attempts to purify the crude reaction mixture with RP-HPLC as described in literature[91] failed due to irreversible adsorption of the target protein to C18 and to C4 reverse phase materials. First purification experiments using DEAE anion exchange chromatography remained unsatisfactory due to incomplete solubilisation of the target protein with DDM as detergent from the lipidic cubic phase environment (results not shown). Further optimisation of the ligation reaction itself and of a suitable purification protocol has to be done in order to establish a robust and efficient semi-synthetic access to the transducer protein starting with HtrMxeCBD refolded from *E.coli* inclusion bodies.

These experiments in lipidic cubic phase environment emphasise the critical role of the correct folding of the Htr(1-83) thioester during EPL. Only the reconstitution of the membrane protein into a lipid bilayer made the C-terminal thioester moiety accessible for the peptide and thereby make a successful ligation reaction possible.

3.3.3 Biological evaluation of the semi-synthetic transducer *Np*HtrII

In the sections 3.3.1.2 and 3.3.2.2 attempts to establish a semi-synthetic access to the archeabacterial transducer from *N.pharaonis* were described. The HtrMxeCBD intein fusion protein that was used in the EPL, was either isolated from *E.coli* membranes or was refolded from insoluble inclusion bodies.

The following experiments were performed to answer the question-‘ if a semi-synthetic protein with its full biological function was obtained’. Hippler-Mreyen et al. successfully characterised the SRII/HtrII complex formation of differently truncated transducer constructs using blue-native gel electrophoresis[60,62-64]. In this work BN-PAGE was used

as well to determine the capability of the recombinant Htr(1-114) G84C mutant to bind to its cognate receptor (see section 3.1).

Due to the lack of a suitable purification method of the semi-synthetic transducer that was obtained from EPL experiments in lipidic cubic phase environment, BN-PAGE investigations were only performed with the semi-synthetic transducer obtained from membrane partitioned HtrMxeCBD protein. Similar experiments with the semi-synthetic transducer derived from the refolded intein fusion protein have to be done, when a suitable purification method is established.

The semi-synthetic transducer was mixed with its SRII binding partner in a 1:1 and 1:2 stoichiometry and was incubated at 4 °C overnight. As reference a 1:1 mixture of *Np*Htr157/*Np*SRII was used.

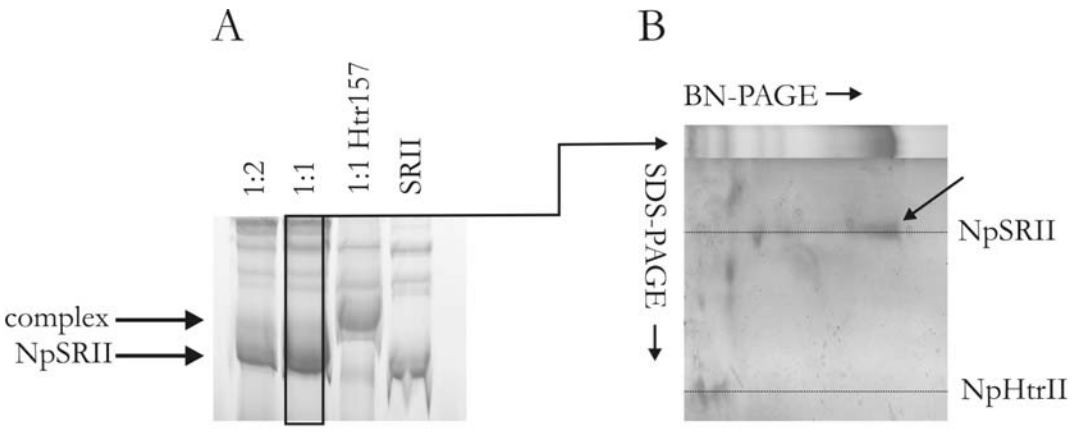


Figure 3.26. Investigation of complex formation between semi-synthetic Htr and SRII
Panel A: BN-PAGE gel with different SRII/Htr stoichiometries. Samples neither with 1:2 nor with 1:1 stoichiometry show a protein band referring to an Htr/SRII complex. The observed intensive bands showed mobilities comparable to the SRII reference indicating unbound SRII.
Panel B: Silver stained 2D-PAGE gel with BN-PAGE in 1st and SDS-PAGE in 2nd dimension (black box). The intensive BN-PAGE protein band only consists of unbound SRII indicated by a single protein band with low mobility marked by the black arrow verifying the absence of a Htr/SRII complex.

The blue-native PAGE gel (figure 3.26, panel A) did not show a protein band referring to an Htr/SRII complex. The observed protein bands with high intensity for samples with 1:2 or 1:1 stoichiometry (SRII:Htr) displayed mobilities comparable to the SRII reference. Compared to the Htr157 reference no shift of the SRII band to higher molecular masses was observed due to complex formation.

In order to exclude a SRII-shift due to complex formation that could not be observed because of a low electrophoretic signal resolution, the BN-PAGE lane of the ligation product sample (black box) was cut out and polymerised on a SDS-PAGE gel. The resulting 2D-gel (figure 3.26. panel B) showed a signal corresponding to SRII (black arrow), however revealed no crosspeak for an Htr/SRII complex and therefore confirmed the absence of an Htr/SRII complex.

Sanders and co-workers reported a method that uses reconstitution into lipids to refold the integral membrane protein diacylglycerol kinase, which is found in the cytoplasmic membrane of many bacteria[92]. They used reconstitutive refolding for the successful renaturation of 65 single-cysteine mutants of DAGK by reconstituting the mutants into POPC vesicles and by dialysing the detergent dodecylphosphocholine out of the micellar mixtures. This *in-vitro* refolding investigation revealed the important role that membrane lipids play in the folding process of membrane proteins.

According to these results, the influence of the membrane environment to correct folding of the semi-synthetic transducer and its influence to the complex formation with SRII was investigated.

The ligation product was reconstituted into polar purple membrane lipids and subsequently re-solubilised again with DDM containing buffer. After performing DEAE anion exchange chromatography to remove excessive detergent and lipids, which disturb BN-PAGE experiments, the binding assay was repeated with a 1:1 stoichiometry of SRII to semi-synthetic Htr114.

The BN-PAGE gel (figure 3.27, panel A) showed a faint band for the re-solubilised sample with a mobility higher than the Htr(1-157)/SRII complex reference band and lower than the SRII reference band. Therefore it was concluded that the re-solubilised transducer formed a complex with its binding partner SRII, although uncomplexed SRII was still detectable. A corresponding complex band of the not reconstituted reference was not observed, confirming the above-mentioned BN-PAGE results.

The ratio between functionally folded semi-synthetic transducer, capable of undergoing complex formation and inactive transducer was estimated based on band intensities to be approximately 1:5. The presence of the N-terminally degraded ligation by-product (see section 3.3.1.2) and a low refolding efficiency of the full-length semi-synthetic transducer could explain the incomplete complex formation.

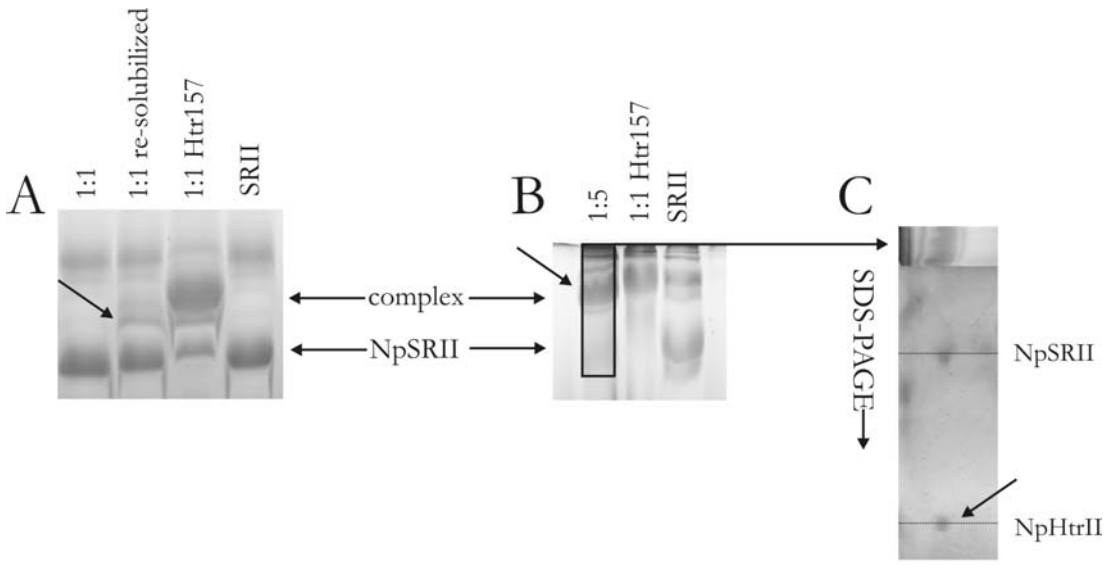


Figure 3.27. Investigation of complex formation of re-solubilised semi-synthetic *Np*Htr.
Panel A: BN-PAGE gel of reconstituted and re-solubilised Htr referring to recombinant Htr(1-157) in 1:1 stoichiometry with SRII. The faint protein band (black arrow) with a mobility higher than Htr(1-157)/SRII complex reference band and lower than SRII reference band is assigned to an Htr/SRII complex.
Panel B: BN-PAGE with 1:5 SRII/Htr stoichiometry. The intensive band (black arrow) indicates quantitative complex formation for a 1:5 stoichiometry, no band of free SRII is detected.
Panel C: 2D SDS-PAGE with BN-PAGE gel lane 1:5 as 1st dimension. The upper protein band corresponds to SRII and the signal with higher mobility (black arrow) indicates semi-synthetic Htr(1-114) revealing successful Htr/SRII complex formation.

In order to verify the estimated amount of functionally folded semi-synthetic transducer in the sample, the 5-fold amount of re-solubilised ligation product was incubated with SRII and investigated with BN-PAGE. The gel (figure 3.27, Panel B) displayed an intensive band indicating quantitative complex formation in 1:5 stoichiometry; no band of free SRII was detected. 2D-SDS-PAGE experiments were performed to verify the complex formation. The BN-PAGE lane of the ligation product (black box) was cut out and polymerised on a SDS-PAGE gel. The resulting 2D-gel (figure 3.27. panel C) showed a signal corresponding to SRII and a signal with higher mobility (black arrow) indicating semi-synthetic Htr114 revealing successful Htr/SRII complex formation.
The blue-native electrophoresis experiments revealed the ability of the semi-synthetic transducer to form a complex with its cognate receptor. In addition, evidence was found that the reconstitution of the ligation product into a lipid bilayer is necessary for the correct folding of the semi-synthetic membrane protein.

3.4 Expression of uniformly isotope labelled proteins from *N.pharaonis*

In the sections 3.3.1.2 and 3.3.2.2 attempts to establish a semi-synthetic access to the archeabacterial transducer from *N.pharaonis* were described. The HtrMxeCBD intein fusion protein that was used in the EPL reactions, was either isolated from *E.coli* membranes or was refolded from insoluble inclusion bodies.

However, purification of the DDM-solubilised membrane fraction revealed the presence of an N-terminally degraded HtrMxeCBD protein (see section 3.3.1). Since this intein fusion species showed full intein cleavage activity, this contamination led to an N-terminally degraded EPL by-product. A separation of the desired full-length semi-synthetic transducer from the degraded product or a separation of the full-length intein from the truncated one could not be achieved.

Analysis of HtrMxeCBD protein partitioning revealed that the desired long intein precursor was deposited predominantly in *E.coli* inclusion bodies. Attempts to establish a successful ligation reaction using a lipidic cubic phase environment[91] circumvented the contamination of the segmental isotope labelled transducer with an N-terminally degraded by-product (see section 3.3.2.2). However, further optimisation of the ligation reaction itself and of a suitable purification protocol has to be done in order to establish a robust and efficient semi-synthetic access to the transducer protein starting with HtrMxeCBD refolded from *E.coli* inclusion bodies.

In addition to the establishment of a semi-synthetic access to domain specific isotope labelled Htr114, uniform isotope labelled transducer variants were expressed in *E.coli*. The spectroscopic analyses of these proteins using magic angle solid state NMR and Fourier transform infrared spectroscopy would build together with data obtained in from segmental isotope labelled transducers the experimental base of a spectroscopic structure determination of the *N.pharaonis* transducer.

3.4.1 [$^{13}C,^{15}N$] *Np*Htr(1 -114)

The C-terminally truncated Htr(1-114) protein was expressed in minimal medium containing [^{13}C]-D-glucose and [^{15}N]-ammonium chloride as carbon and nitrogen source respectively, in order to obtain a uniformly isotope labelled transducer. The time course of the expression and the purification of the His-tagged protein was analysed with SDS-PAGE.

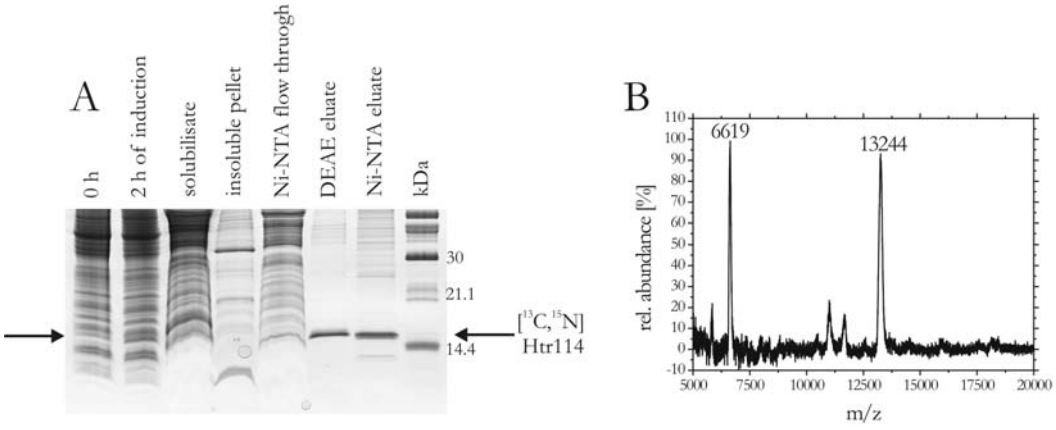


Figure 3.28. Expression, purification and evaluation of [$^{13}C,^{15}N$] labelled Htr(1-114).
Panel A. SDS-PAGE monitoring expression, Ni-NTA and DEAE anion exchange chromatography. The protein band above the 14.4 kDa reference band (black arrows) appearing after 2 h after IPTG induction is assigned to isotope labelled Htr(1-114). The absence of numerous faint protein bands in addition to this transducer band documents significant improvement of the sample purity upon DEAE anion exchange chromatography.
Panel B. MALDI-MS of [$^{13}C,^{15}N$] labelled Htr. The signals at 13244 amu and 6619 amu are assigned to the $(M+H)^+$ and the $(M+H_2)^{2+}$ signal of the isotope labelled transducer with an expected mass of 13257.8 Da.

The SDS-PAGE gel (figure 3.28, panel A) documented a protein band (black arrow) above the 14.4 kDa reference band appearing after 2 h after IPTG induction. Due to the good agreement with the calculated molecular weight of 12.7 kDa, this band was assigned to isotope labelled, recombinant Htr(1-114). Longer incubation times significantly decreased again the Htr114 expression level: After 4 h of induction no recombinant target protein could be isolated (results not shown).

Protein purification was performed exploiting the C-terminal histidine tag. After binding of the protein to Ni-NTA matrix and a washing step with 20 mM imidazole, the target

protein was eluted with 200 mM imidazole. In order to remove the competitor again, preparative DEAE anion exchange chromatography was performed yielding around 2-2.5 mg purified protein per litre minimal medium.

Purification was monitored using SDS-PAGE analysis. The gel (figure 3.28, panel A) showed a protein band above the 14.1 kDa reference band for the Ni-NTA and DEAE eluate corresponding to the desired protein. Anion exchange chromatography significantly improved the purity of the sample as was indicated by the absence of numerous faint bands of contaminating proteins.

In order to determine the isotope labelling efficiency of the purified protein, MALDI-MS experiments were performed. The mass spectrum showed a signal at 13244 amu and a second signal at 6619 amu, which were assigned to the $(M+H)^+$ and the $(M+H_2)^{2+}$ signal of the isotope labelled transducer.

The expected mass of the uniformly [^{13}C,^{15}N] labelled target protein was calculated to be 12543+x(559+163)=13243 Da with M=12543 Da representing the experimental mass of the unlabelled Htr(1-114) (see section 3.3.1.2) and with x representing an observed isotope labelling efficiency. The experimentally observed labelling efficiency of 97.0 % was in a good agreement with the content of isotope in the respective ^{13}C and ^{15}N sources of 99 %.

3.4.2 [^{13}C,^{15}N] *Np*Htr(1 -157)

[^{13}C,^{15}N]-labelling of Htr(1-157) was achieved using protein expression in synthetic minimal medium with isotope enriched D-glucose and ammonium chloride as sole carbon nitrogen sources. Purification of the protein was performed by exploiting the affinity of the C-terminal histidine-tags to Ni-NTA matrix. After a washing step with C_{mem} buffer containing 20 mM imidazole, the proteins were eluted with D_{mem} buffer containing 200 mM imidazole. Subsequently DEAE anion exchange chromatography was carried out. The protein was obtained in yields of about 2-2.5 mg per litre medium.

Protein purification was monitor with SDS-PAGE analysis. The gel (figure 3.29, panel A) showed a protein band with a mobility comparable to the 21.1 kDa reference band (black arrow). According to the achieved electrophoretic resolution this experimental molecular weight is in a good agreement with the theoretical value of 17.4 kDa.

Therefore the protein band was assigned to the isotope labelled, recombinant Htr157 transducer.

After DEAE anion exchange chromatography an analytical pure transducer was obtained. The gel showed a faint protein band above the 30 kDa reference, besides the band for the monomer (figure 3.29, panel A, black arrow). The observed mobility of this band is in agreement with the mobility of a postulated homo-transducer dimer (34.8 kDa).

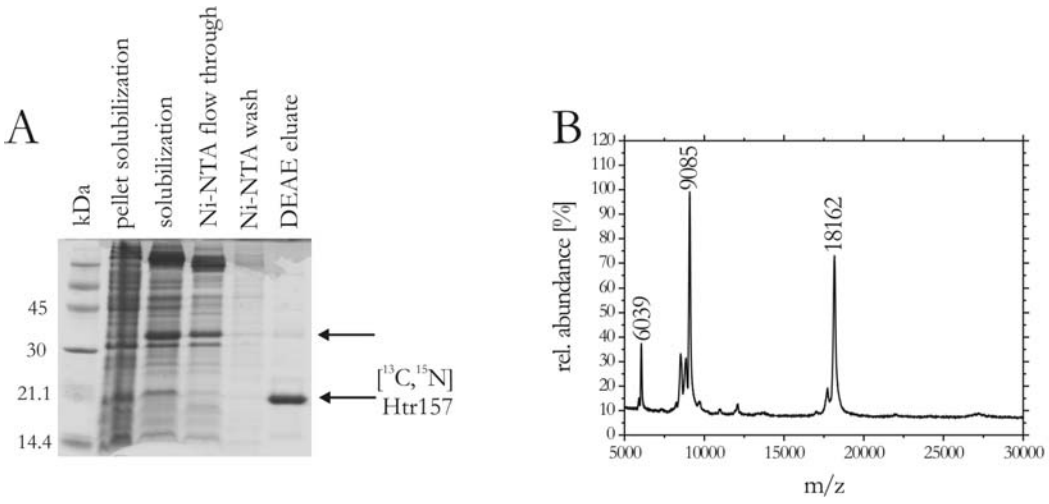


Figure 3.29. Purification of [^{13}C,^{15}N] labelled sensory rhodopsin II and Htr(1-157).
Panel A: SDS-PAGE monitoring [^{13}C,^{15}N] Htr157-purification. The protein band with mobility comparable to the 21.1 kDa reference band (black arrow) was assigned to isotope labelled, recombinant transducer. The faint above the 30 kDa reference suggests dimer formation.
Panel B: MALDI-MS of purified [^{13}C,^{15}N] labelled Htr157. The signals at 18162 amu, 9085 amu and 6039 amu were assigned to different protonated states of the isotope labelled Htr157.

In order to determine isotope labelling efficiency, MALDI-MS experiments were performed. The spectrum (figure 3.29, panel B) showed three major signals at 18162 amu, 9085 amu and 6039 amu. These signals were assigned to a series of differently protonated states of the isotope labelled Htr157. The expected mass of this protein was calculated to be M=17222+x(755+220)=18162 with 17222 Da representing the already determined experimental mass of the unlabelled protein (see section 3.3.1.2) and with x representing the observed labelling efficiency. The experimentally obtained molecular weight of 18162 Da corresponded to an observed labelling efficiency of 96.4 %, which was in a good agreement with the observed efficiency of 97.0 % for

Htr(1-114) (see 3.4.1) and corresponded to the content of isotopes in the respective ^{13}C and ^{15}N sources of 99 %.

3.4.2 [^{13}C,^{15}N] *Np*SRII

Isotope labelling of SRII was done in a manner similar to the labelling of the transducers (see section 3.4.1 und 3.4.2) in minimal medium with labelled glucose and ammonium chloride. Upon solubilisation of the protein from the *E.coli* membranes Ni-NTA-based separation of the receptor was performed. DEAE anion exchange chromatography removed the imidazole that was used for the elution of the histidine-tagged protein from the Ni-matrix.

Figure 3.30 shows a SDS-PAGE gel documenting a protein band above the 21.1 kDa reference band (black arrow) that was assigned to isotope labelled, recombinant SRII. The protein was obtained in analytically pure form in yields of about 1-1.5 mg/l.

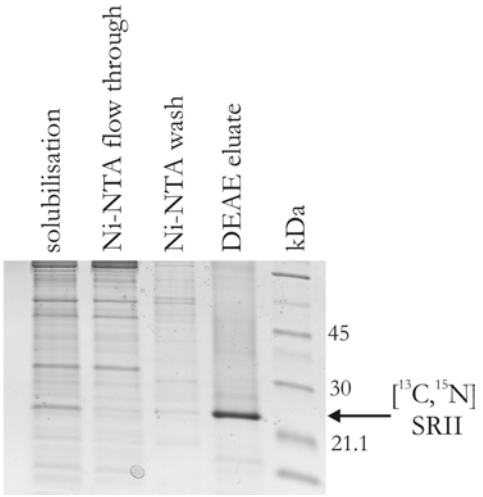


Figure 3.30. SDS-PAGE monitoring [^{13}C,^{15}N] labelled sensory rhodopsin II purification. The protein band above the 21.1 kDa reference band (black arrow) was assigned to pure, isotope labelled, recombinant SRII.

In order to determine SRII isotope labelling efficiency, MS experiments were carried out. Unfortunately no mass spectrum of the labelled receptor was obtained either with MALDI-MS according to the already described modified method from Cadene et al.[58] (see section 3.3.1.2) or with ESI-MS according to Hufnagel et al.[59].

First steady-state FTIR measurements using the labelled receptor answered the question of the obtained labelling efficiency: A qualitative downshift of the characteristic C=O stretch vibration of Asp75 from 1765 cm^{-1} to 1720 cm^{-1} indicated a quantitative isotope labelling of the receptor (for detailed discussion the section 3.6).

3.5 Nuclear magnetic resonance (NMR) spectroscopy

3.5.1 Isotope labelled *Np*HtrII(84-114) G84C peptides

The cytoplasmic domain of *Np*HtrII comprising residues 84 to 114 was soluble in aqueous buffers without any detergents or chaotrope additives. NMR experiments investigating the degree of isotope labelling of the different peptides were performed in 20 mM NaPi, 8 % D_2O, pH 6.5 at 25 °C. Under these conditions peptide concentrations of 0.85-1.65 mM were used without any problems of precipitation. A complete set of 1D and 2D spectral data was recorded using a Bruker 500 MHz (11.7 T) or Varian 600 MHz (14.1 T) NMR spectrometer. The water signal was suppressed by selective pre-saturation during the preparation period and/or mixing time. A summary of the experiments performed is given in table 3.3.

HtrII(84-114) G84C	Experiments	Nucleus	Concentration	pH
unlabelled	1D ^{1}H 2D TOCSY 2D NOESY	^{1}H	1.03	6.5
[^{15}N] labelled	1D ^{1}H 2D ^{1}H,^{15}N-HSQC 3D ^{1}H,^{15}N-NOESY	^{1}H,^{15}N	0.85	6.5
[^{13}C, ^{15}N] labelled	1D ^{1}H 2D ^{1}H,^{15}N-HSQC 2D ^{1}H,^{13}C-HSQC	^{1}H,^{15}N ^{1}H,^{13}C	1.65	6.5

Table 3.3. Summary of recorded NMR experiments of HtrII (84-114) G84C.

The assignment of proton resonances was performed iteratively with TOCSY and NOESY spectra in two steps: first the amino acid spin systems were identified in TOCSY spectra, subsequently the sequential order of the spin systems was determined by the interpretation of dipolar couplings in NOESY spectra. Software assisted data evaluation was carried out using AURELIA software (Bruker Inc.). A small experimental resonance signal dispersion of only 0.54 ppm for H^N chemical shifts and 0.77 ppm for H^α chemical shifts complicated the signal assignment of the 31-residue comprising peptide (figure 3.31).

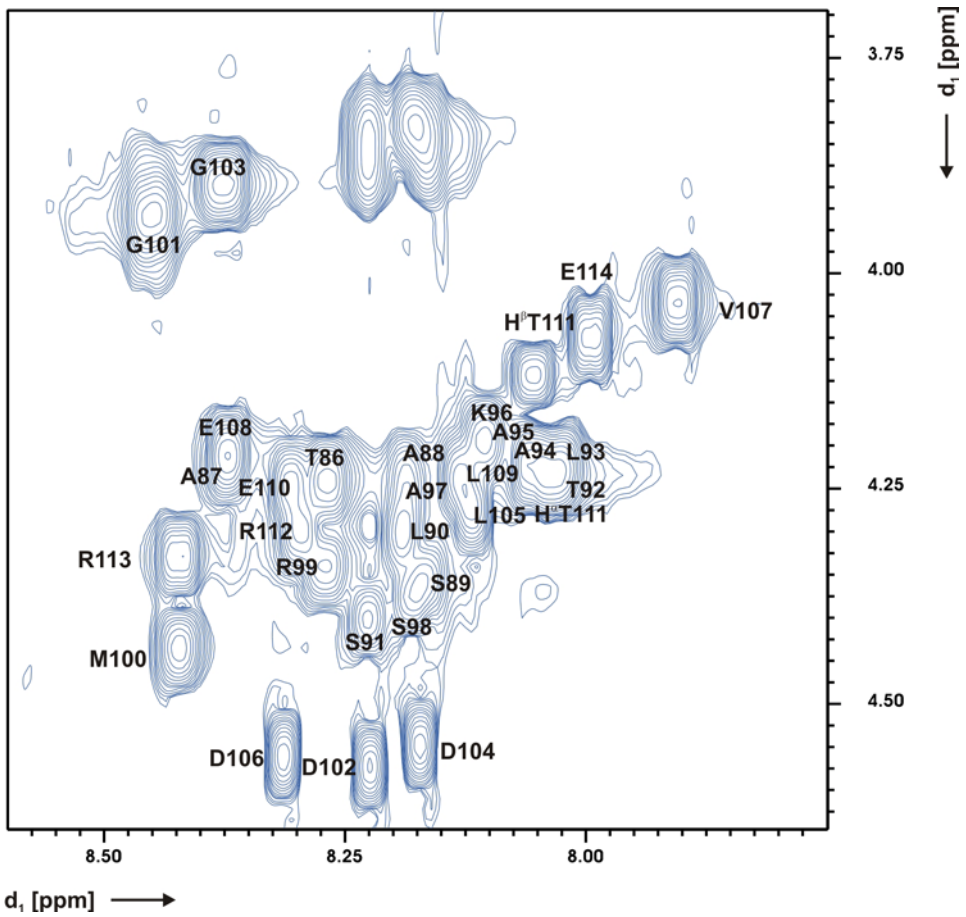


Figure 3.31. H^N/H^α region of a TOCSY spectrum of Htr (84-114) G84C.

A ^{15}N-separated ^{1}H,^{15}N-3D NOESY experiment could resolve partially overlapping resonances of five alanine and four leucine residues. All ambiguities were solved and a complete sequential assignment except for residues Cys84 and Asp85 was obtained (table 3.4).

Signal broadening due to fast proton exchange prevented NH_3^+ identification of Cys84. The Asp85 H^α resonance partially overlapped with the water signal and could not be assigned. An H^N/H^α NOE path depicts the complete sequential resonance assignment of the peptide (figure 3.32, panel A).

This small signal dispersion was especially observed for the residues up to approximately Met100. Beginning from there, a significant higher dispersion was detected due to sequential alteration of glycine and aspartic acid residues, which have highly different H^α random coil shifts. However, the signal dispersion of the residue-cluster Asp106 to Arg112 could not be explained only with differences of the random

coil shifts. The strength of the sequential NOE signals between H^N of residue i with H^N of residue i+1 in the H^N/H^N region (figure 3.29, panel B) combined with higher signal dispersion suggested the formation of an α-helical stretch within this cluster.

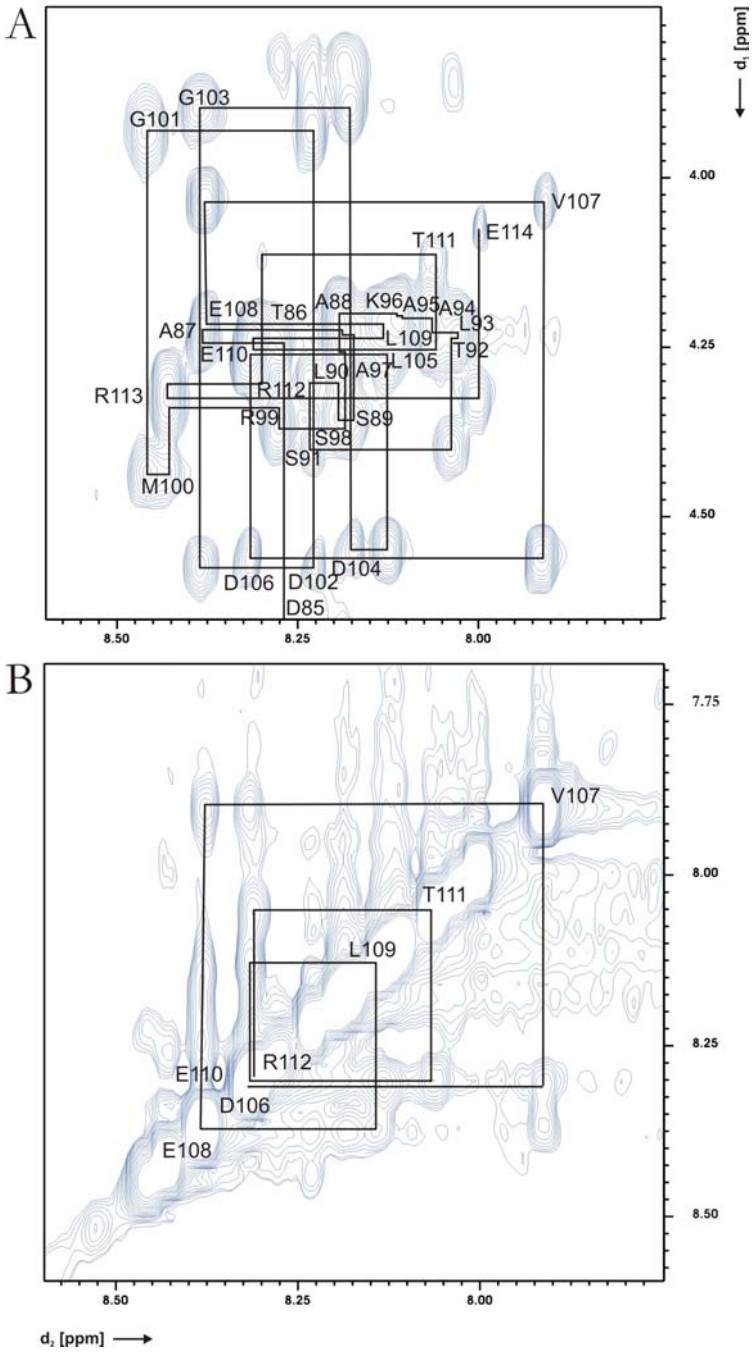


Figure 3.32. Sequential assignment in a NOESY spectrum of Htr (84-114) G84C
Panel A: H^N/H^α region
Panel B: H^N/ H^N region

Residue	^{15}N	H^N	H^α	H^β	H^γ	Others
Cys84	-	-	-	-	-	-
Asp85	117.03	8.28	4.67	2.58	-	-
Thr86	117.08	8.27	4.23	4.25	1.18	-
Ala87	128.05	8.38	4.20	1.36	-	-
Ala88	124.63	8.19	4.23	1.34	-	-
Ser89	116.66	8.16	4.36	3.82	-	-
				3.88		
Leu90	125.56	8.19	4.31	1.58	1.66	H^{δ_1} 0.82
				1.65		H^{δ_2} 0.87
Ser91	117.78	8.23	4.40	3.82	-	-
				3.89		
Thr92	117.98	8.03	4.24	4.22	1.17	-
Leu93	125.53	8.02	4.23	1.55	1.61	H^{δ_1} 0.82
				1.61		H^{δ_2} 0.87
Ala94	124.61	8.06	4.22	1.35	-	-
Ala95	125.98	8.10	4.20	1.34	-	-
Lys96	122.02	8.11	4.19	1.73	1.39	H^{δ_1} 1.63 H^{ε_1} 2.94
				1.80	1.44	H^{δ_2} 1.63 H^{ε_2} 2.94
Ala97	126.49	8.19	4.28	1.37	-	-
Ser98	116.84	8.18	4.38	3.83	-	-
				3.83		
Arg99	124.45	8.27	4.34	1.73	1.58	H^{δ_1} 3.15 H^{δ_2} 3.15
				1.85	1.58	H^{ε} 7.25
Met100	123.40	8.42	4.44	1.97	2.49	H^{ε} 2.04
				2.05	2.57	
Gly101	112.42	8.45	3.91	-	-	-
			3.95			
Asp102	122.37	8.23	4.57	2.59	-	-
				2.67		
Gly103	111.11	8.38	3.90	-	-	-
			3.90			
Asp104	122.24	8.17	4.55	2.55	-	-
				2.66		
Leu105	123.64	8.12	4.27	1.54	1.58	H^{δ_1} 0.80
				1.62		H^{δ_2} 0.86
Asp106	123.35	8.32	4.56	2.55	-	-
				2.68		
Val107	121.52	7.91	4.04	2.05	0.88	-
					0.88	
Glu108	125.51	8.37	4.21	1.90	2.18	-
				1.99	2.25	
Leu109	124.74	8.13	4.24	1.54	1.61	H^{δ_1} 0.81
				1.63		H^{δ_2} 0.88
Glu110	122.83	8.31	4.24	1.94	2.22	-
				2.00	2.22	
Thr111	117.23	8.06	4.23	4.12	1.15	-
Arg112	126.13	8.30	4.30	1.74	1.54	H^{δ_1} 3.14 H^{δ_2} 3.14
				1.80	1.60	H^{ε} 7.24
Arg113	125.82	8.42	4.33	1.71	1.59	H^{δ_1} 3.15 H^{δ_2} 3.15
				1.83	1.59	H^{ε} 7.19
Glu114	128.99	8.00	4.08	1.83	2.14	-
				1.99	2.14	

Table 3.4. ^{15}N and ^{1}H shifts of *Np*HtrII (84-114) G84C peptide at 25 °C, pH 6.5

In order to investigate the tendencies towards the formation α-helical structures that were observed in the NOESY spectra, further spectral investigations of the chemical shift perturbation were performed.

Structure predictions can be made in particular based on deviations of H^{α} and H^{N} proton resonances from random coil values. A shift of H^{α} and H^{N} resonances to higher field strengths compared to random coil values indicates tendencies towards an helical conformation[93-95].

All experimental chemical shifts did not differ much from those observed for random coils. Nevertheless two sequential clusters with characteristic chemical shift deviations could be identified. One cluster comprising residues Ser91 to Ser98 and a second cluster, in correspondence with the NOESY results, included amino acids Asp104 to Thr111 (figure 3.33).

In summary, comparison of NOESY experiments and chemical shift indices led to the conclusion that Htr(84-114) G84C is a random coiled peptide with α-helical tendencies.

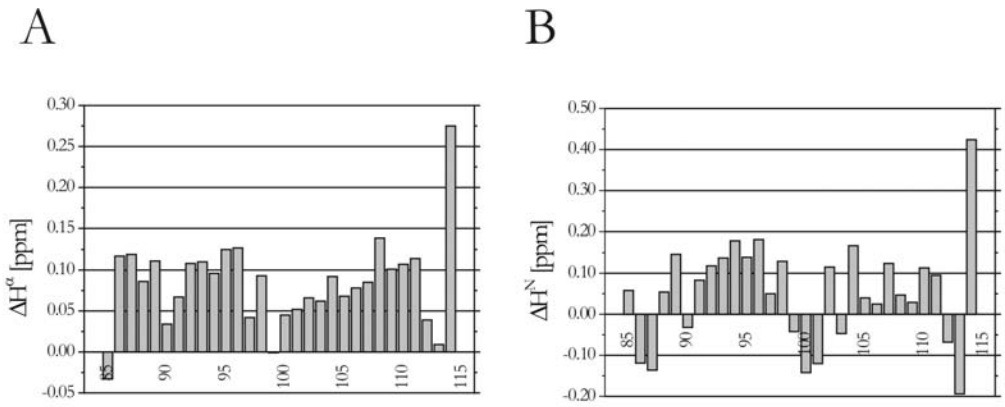


Figure 3.33. Chemical shift index of *Np*HtrII (84-114) G84C
Panel A: $\Delta H^{\alpha}=H^{\alpha}$(random coil)-$H^{\alpha}$(experimental)
Panel B: $\Delta H^{N}=H^{N}$(random coil)-H^{N}(experimental)

In an additional set of NMR experiments ^{15}N and ^{13}C labelling of the cytoplasmic transducer domain were verified. Therefore $^{1}H,^{15}N$-2D and $^{1}H,^{13}C$-2D HSQC spectra of [^{15}N]- and [$^{13}C,^{15}N$]-labelled Htr(84-114) G84C were recorded (figure 3.34).

The ^{15}N signals were completely assigned. The signals for the two glycines, six serines and threonine residues were significantly shifted to high-field as was reported in literature[93].

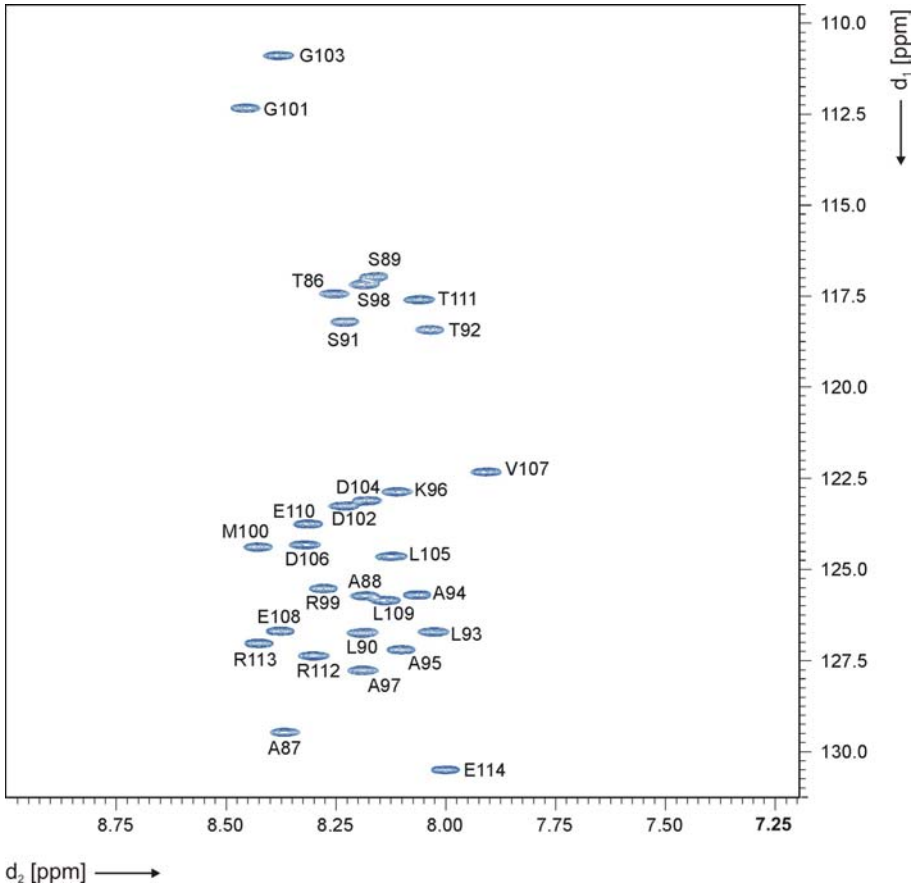


Figure 3.34. ^{15}N HSQC spectrum of [^{15}N] labelled (84-114) Htr G84C peptide.

3.5.2 Semi-synthetic, isotope labelled *Np*HtrII(1-114) G84C

Attempts to establish a semi-synthetic access to a domain specific isotope labelled archeabacterial transducer were described in section 3.3.1.2: A [^{15}N]-labelled peptide HtrII(84-114) G84C (generation and analytics: see section 3.2.4) was ligated with an *in-situ* generated unlabelled thioester protein Htr(1-83) in a 2-mercaptoethansulfonic acid (MESNA) mediated EPL reaction. The ligation product Htr(1-114) G84C was purified and characterised using SDS-PAGE and MALDI-MS. After reconstitution into polar purple membrane lipids and resolubilisation, a functional semi-synthetic membrane protein was generated (see section 3.3.3).

In initial NMR experiments the purified ligation product was investigated in aqueous buffers (100mM NaCl, 20 mM NaPi, 2 mM DTT, pH 8.0) with DDM concentrations

varying between 0.5 and 0.005 % at 25 and 42 °C. The protein concentration was kept constant at 0.6 mM. Figure 3.35 summarizes the results.

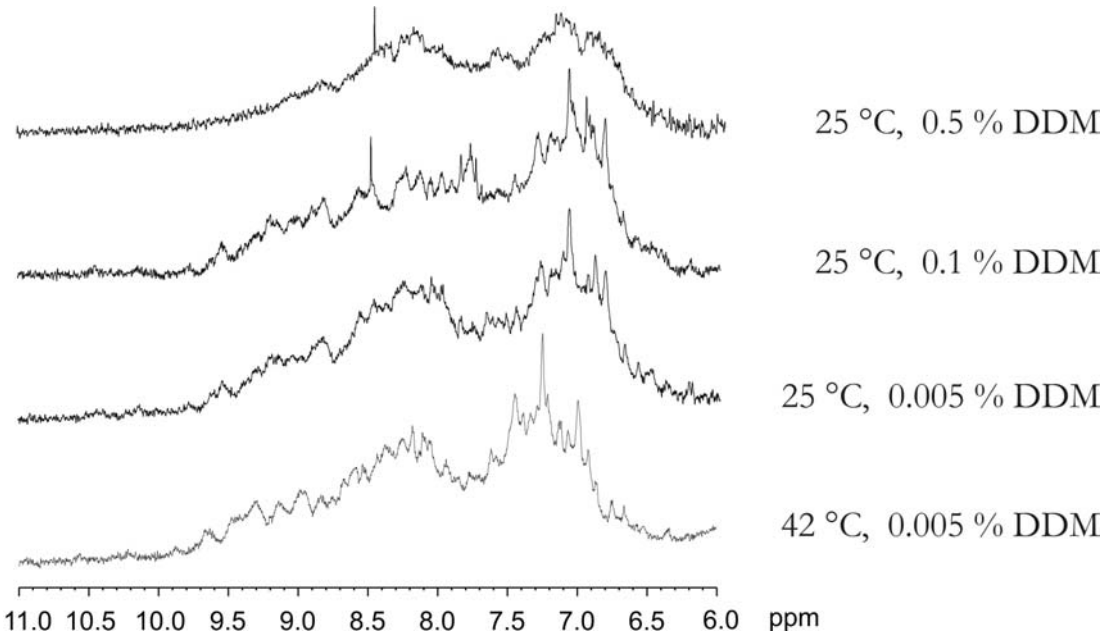


Figure 3.35. 1D ^{1}H NMR spectra of [^{15}N] $_{84\text{-}114}$Htr(1-114) G84C. The purified ligation product was investigated in aqueous buffers (100mM NaCl, 20 mM NaPi, 2 mM DTT, pH 8.0) with various DDM concentrations between 0.5 and 0.005 % at 25 and 42 °C. The protein concentration was kept constant at 0.6 mM. Lowering the DDM concentration from 0.5 % to 0.1 % improves the signal to noise ratios (S/N) from 10/1 to 15/1. A further decrease to 0.005 % has no significant influence on the S/N. Increasing the temperature from 25 °C to 42 °C corresponds to a S/R improvement to 25/1.

The signal to noise ratios (S/N) of the recorded spectra were improved from 10/1 at 25 °C and 0.5 % DDM to approximately 15/1 by lowering the DDM concentration by a factor of five to 0.1 % (10 times the critical micellar concentration, cmc). This corresponded quite well to the expected influence of the detergent on the tumbling frequency of the protein/detergent complexes: Lowering the detergent concentration leads to a lower solvent viscosity and therefore to higher tumbling frequencies and sharper resonance signals.

A further decrease of the DDM concentration below the cmc of DDM to 0.005 % had no significant influence on the S/N. The stability of the sample was not influenced by decreasing the DDM concentration below the cmc: No protein precipitation was observed. This may correspond to the fact that the dissociation constant of the

protein/detergent complex is significantly lower than the cmc of the detergent. Although the surrounding solution contains no micelles, the number of detergent molecules adsorbed to the protein and therefore the tumbling frequencies was not significantly reduced.

Increasing the temperature from 25 °C to 42 °C should enhance the Brownian motion in the system and therefore increase the tumbling frequency of the protein. This became obvious in a S/R improvement to 25/1.

The H^N region of the obtained spectra of the ligated sample displayed a significantly larger signal dispersion (9.6 to 6.6 ppm) than observed for the spectra of the unligated peptides. While the strong signals below 8 ppm might have been caused by the two transmembrane α-helices of the semi-synthetic transducer, the origin of the signals above 8.5 ppm remained unclear. In principle their H^N shifts correspond to β-sheet domains or to aromatic amino acid residues. However, both possibilities can be eliminated for the α-helical transducer that comprises no aromatic amino acid residues.

Although in the NMR spectra of the ligated transducer the signal to noise ratio was improved by a factor of 2-3 by variation of detergent concentration and temperature, no 2D $^1H,^{15}N$-HSQC spectrum in aqueous solution was obtained.

Therefore, experiments in organic solvents without detergent as solubilising additive were carried out. These experiments were not performed to gain any structural information under these strong denaturing, artificial conditions, but rather to collect further evidence for a successful labelling and ligation reaction.

In contrast to the unligated peptide, the ligation product was expected to show a signal for residue Cys84. Due to the formation of the native peptide bond during the EPL, the H^N signal of Cys84 is not suppressed by fast proton exchange as it was the case for the N-terminal Cys84 NH_3^+-moiety of the unligated peptide.

In order to replace the water by organic solvents, the protein was dialysed against ammonium bicarbonate buffer containing 0.1 % β-mercaptoethanol and subsequently dialysed several times against pure water to remove salt and detergent. The partially precipitated protein sample was lyophilized and afterwards dissolved in 100 µl DCOOH and diluted with 200 µl $D_3COH/CDCl_3$ (1:1). In order to avoid a fast H/D exchange of the H^N and H^α protons that would drastically decrease the signal intensity in the NMR spectra, only partially deuterated solvents were used

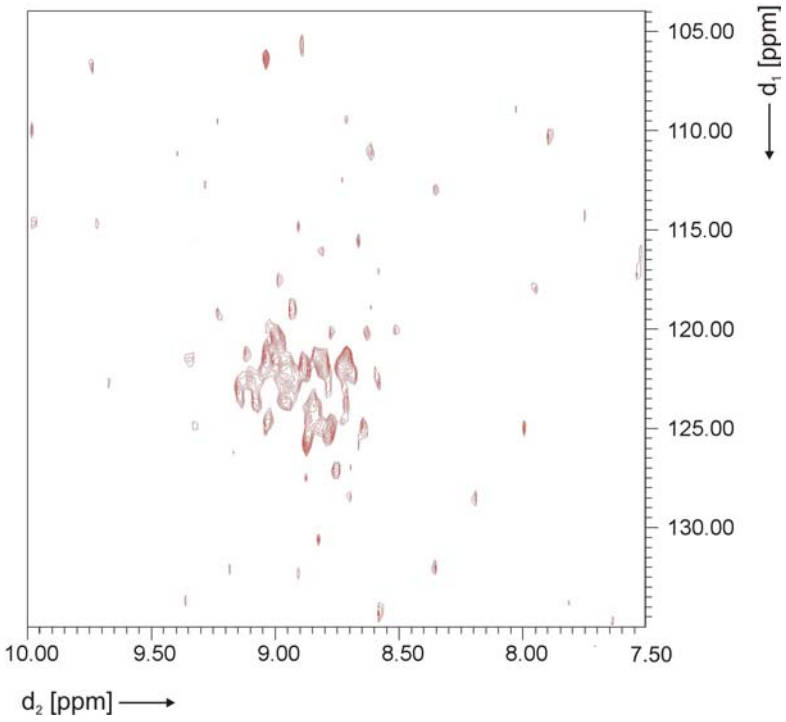


Figure 3.36. $^{1}H,^{15}N$-HSQC spectrum of the semi-synthetic transducer $[^{15}N]_{84\text{-}114}$Htr(1-114) G84C in $DCOOH/D_3COH/CDCl_3$ (1:1:1) at 25 °C (600 MHz Varian NMR spectrometer).

The 2D $^{1}H,^{15}N$-HSQC spectrum (figure 3.36) showed a poor signal to noise ratio of approximately 2:1. A comparison of this spectrum with that of the labelled peptide led to the following conclusions: The spectral signal dispersion from 0.5 ppm in the ^{1}H dimension was similar in both spectra, although the signals in organic solvents were significantly shifted to lower field compared to the peptide spectrum in aqueous solution. This could be explained by the influence of the lower polarity and the higher acid content of the organic solvents that were used.

Signals for threonines and serines could not be resolved, however resonance signals corresponding to the residues Gly101 or Gly103 were observed (figure 3.36, H^N: 9.05 ppm; ^{15}N: 106 ppm).

The spectrum verified a successful ^{15}N-labelling of the semi-synthetic transducer. However, due to low signal intensities, an assignment was not possible. Therefore the experiment did not confirm a successful ligation reaction by resolving resonance signals of the cysteine residue at position 84.

3.5.3 ssNMR of u-[^{13}C,^{15}N] labelled Htr114 in complex with SRII in PM lipids

The secondary structure elements of transmembrane proteins differ in terms of their flexibility. Membrane spanning helices or β-sheets for example are well-structured regions that are relatively immobile compared to loops at both the N- and C-terminus of a protein. These differences in the mobility of divers regions of a protein can be used in solid state NMR to separate the signals of flexible regions from those of more immobile regions.

A recently established method (M. Baldus, MPI Göttingen, personal correspondence) able to distinguish between signals of flexible and immobile residues, is based on the variety of the J-coupling in the distinct regions. Under MAS conditions the J-coupling in immobile protein domains is effectively averaged out. However, in more dynamic regions it is still present. This leads to the possibility to transfer magnetization from protons to hetero-atoms either with cross-polarization-type methods via dipolar coupling in immobile domains or via J-coupling with INEPT-type methods (insensitive nuclei enhanced by polarization transfer) in mobile regions of a protein. The resulting ^{1}H,^{15}N- or ^{1}H,^{13}C-single quantum spectra show either predominately signals of mobile or predominately signals of immobile regions of the investigated protein. These methods were used to separate the signals of the two transmembrane helices of the transducer from the signals of the N- and above all of the C-terminus, that is important for the capability of the protein to form a stable complex[60,96] with its cognate receptor.

A C-terminal truncated variant of the transducer comprising the two transmembrane helices and the cytoplasmic domain until amino acid residue Glu114 in complex with the SRII receptor was used for MAS ssNMR measurements. A similar complex was already investigated in the crystallographic studies of Gordeliy et al.[9].

For NMR measurements the truncated Htr114 protein was expressed in minimal medium containing [^{13}C]-D-glucose and [^{15}N]-ammonium chloride as carbon and nitrogen source respectively, in order to obtain a uniformly isotope labelled transducer (see section 3.4.1).

The purified protein was mixed in a ratio of 1:1.05 with unlabelled SRII. The excessive amount of receptor (5 %) would ensure a quantitative complex formation of all transducer molecules. The complex was reconstituted into purple membrane lipids isolated from *H.salinarum* with a constant protein/lipid ratio of 1:45. The reconstituted sample was washed with buffer containing 20 mM NaPi, pH 8.0 to remove excess salt

that was used during the reconstitution process. Finally the sample was transferred into a MAS spinning rotor of 4 mm-diameter and an approximate volume of 50 µl.

In first experiments that were performed in cooperation with M. Baldus and C. Seidel at the MPI in Göttingen (Germany), a cross-polarization-type $^{1}H,^{13}C$-single-quantum NMR-spectrum at 5 °C with a MAS spinning frequency of 12.5 kHz was recorded (figure 3.37, panel A). A preliminary assignment of the observed signals to intra-residual spin-systems could be achieved.

In a second step a set of INEPT-type spectra at –30 °C, –20 °C and -10 °C was recorded (figure 3.37, panel B).

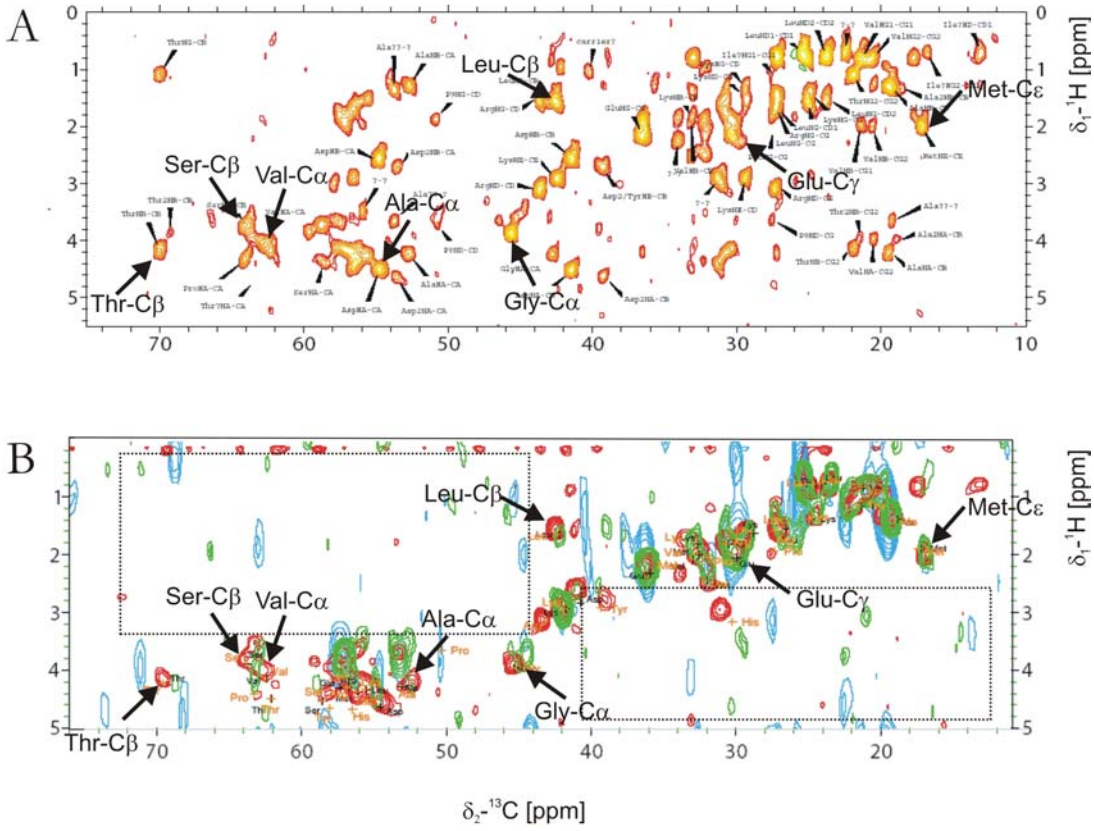


Figure 3.37. MAS ssNMR spectra of the SRII/[$^{13}C,^{15}N$]-Htr114-complex in PM lipids. Black arrows indicate signals whose chemical shifts are summarized in table 3.4.
Panel A: Cross-polarization-type single quantum spectrum at 5 °C
Panel B: INEPT-type single quantum spectrum at -10 °C (red), -20 °C (green) and -30 °C (blue). Black boxes indicate absent signals due to the absence of a carbon/carbon polarization transfer by a C/C mixing time.

At –30 °C only a few signals of amino acid residues with a low spectral resolution were observed indicating a few residues that are mobile at this low temperature. The dynamic of most of the amino acid residues was frozen.

At higher temperature (–20 °C) a number of additional spectroscopic signals were observed suggesting a rising number of mobile residues. This increased residual mobility improved the spectral resolution as well.

At –10 °C the highest number of signals of mobile amino acid residues combined with the best spectral resolution was observed.

Since the INEPT-type of experiments were performed without a carbon/carbon polarization transfer, no corresponding cross peaks (figure 3.37, panel B indicated by dotted black boxes) were observed in contrast to the cross-polarisation-type spectrum (figure 3.37, panel A).

The temperature-dependency of the spectroscopic signals in the INEPT type of experiments is in a good agreement with the expected signal behaviour of mobile protein clusters, like both N- and C-terminus and the loop between the two transmembrane helices.

In addition to the information about the flexibility of these clusters, the recorded spectra contained information about their secondary structure. A comparison of the observed C_α–shifts in the INEPT-type spectrum (table 3.5) with those of random coil chemical shift values allow to predict secondary structure elements of the mobile protein domains[93-95].

Residue	Experimental shift	Random coil Shift
Thr-C_β	69.7	69.8
Ser- C_β	63.7	63.8
Val-C_α	62.4	62.2
Ala- C_α	52.2	52.5
Gly- C_α	45.2	45.1
Leu- C_β	42.3	42.5
Glu- C_β	29.8	29.9
Met-C_ε	17.0	16.9

Table 3.5. Comparison of chemical shifts of exemplary residues (indicated with arrows in figure 3.35) with random coil shifts.

The chemical shifts did not differ significantly from those observed for random coils (table 3.5). Combined with the information about the mobility of the amino acid residues at different temperatures, these results suggest that the mobile protein domains, like the loop between TM1 and TM2 and both termini, are highly mobile, random coil amino acid clusters under the chosen low salt conditions.

At higher temperatures a rising number of amino acid residues are included into these flexible clusters. Any signs for a defined secondary structure of these domains, like significant chemical shift deviations, were not found. Due to the lack of a sequential signal assignment, at the moment no information about the exact sequence of the mobile clusters is available.

3.6 Fourier-Transform Infrared Spectroscopy (FTIR)

The recently reported crystallographic structure of SRII in complex with its cognate transducer protein[9] revealed that the observed structure is almost identical with the crystal structure of SRII itself[36,37]. In fact, the backbone structures are similar, while there are some deviations in the orientations of some side chains.

In addition, the photocycle of SRII is not much influenced by the presence of the transducer[13,97]. However, it remains unclear how HtrII influences the structural changes of the receptor upon photo-isomerization of the retinal-chromophore.

Steady state and rapid-scan FTIR spectroscopy were applied in cooperation with F. Siebert and I. Radu at the University of Freiburg (Germany) to investigate the structural changes in atomic details of both complex partners during the photocycle of the receptor.

In initial measurements the unlabelled protein complex was investigated under steady-state conditions. For sample preparation the SRII receptor was mixed with two different transducer variants, Htr114 and Htr157 respectively, in 1:1 stoichiometry and was reconstituted into purple membrane lipids. After a washing step with buffer containing 20 mM NaPi, pH 8.0 included to remove excess salt that was used during the reconstitution process, the membranes were dried onto the lower window of the measurement cuvette. Hydration of samples was achieved by overlaying them with approximately 20 µl of measuring buffer (200 mM bis-tris-propane, 150 mM NaCl, pH 8.0). The second plane window with a circular groove at its border squeezed excess buffer into this groove serving as a buffer reservoir. Sample thickness was either 2.5 or 4.5 µm. This newly established sandwich-sample preparation[98] guaranteed a sufficient hydration of the sample during several measurements and improved therefore the yields of the late M/O photo intermediate.

Figure 3.38 shows the steady-state FTIR difference spectra of the photo-intermediate M with contributions of the O-intermediate minus the respective unphotolysed groundstates (M/O minus groundstate) of samples comprising SRII in complex with different transducer variants referring to the receptor alone.

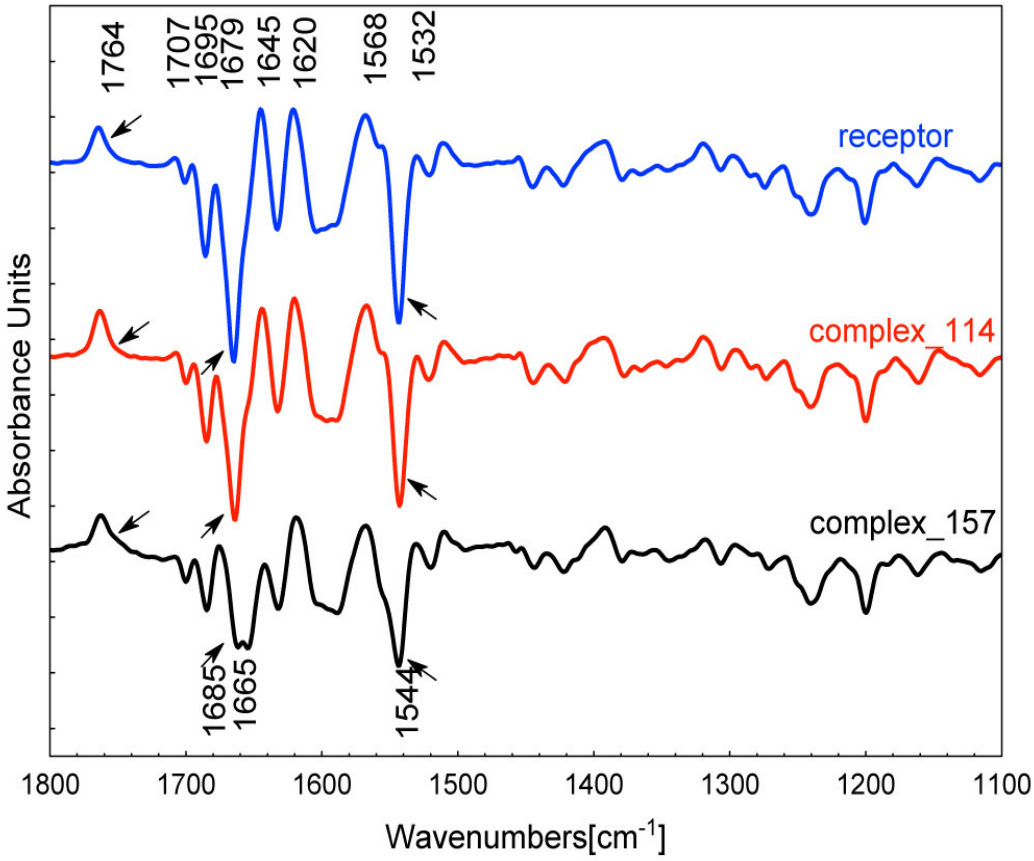


Figure 3.38. Steady-state spectra of the M/O signalling state of different SRII/Htr complexes with black arrows indicating observed spectral changes.
Blue trace: SRII reference without any transducer.
Red trace: SRII complexed with Htr114.
Black trace: SRII complexed with Htr157.

The spectra of the complexes SRII/Htr(1-114) (figure 3.38, red trace) and SRII/Htr(1-157) (figure 3.38, black trace) showed spectral changes due to the binding of the transducer referring to the receptor spectrum (figure 3.38, blue trace).

The signal at 1764 cm^{-1} is broadened in the spectrum of the complex with Htr114 (figure 3.38, red trace). A similar effect is more pronounced for the Htr157-complex (figure 3.38, black trace). According Engelhard et al.[99] this signal was assigned to the C=O moiety of the counter-ion Asp75 that becomes protonated upon M-formation.

In addition, the amplitude of the negative band at 1675 cm^{-1} was significantly larger in the spectrum of the receptor (figure 3.38, blue trace) compared to the spectra of the complexes (figure 3.38, red and black traces). For the longer transducer fragment a splitting of this signal was observed (figure 3.38, black trace). This band in the amide I region of the spectra was assigned to an α_I-helical C=O stretch vibration of the protein

backbone of the complex and indicated distortions of the complex partners upon transducer binding. These distortions were more distinct in the complex with the longer transducer fragment (figure 3.38, black trace).

A further difference between the spectrum of the receptor and the spectra of the receptor in complex formed with the transducer variants was a larger signal amplitude of the ethylenic C=C stretch vibration of the retinal-chromophore at 1544 cm^{-1}. Again a more pronounced effect was observed for the longer transducer variant.

The spectral changes upon SRII/Htr complex formation suggest alterations of the retinal binding pocket affecting the protonation of the Asp75 counter-ion and conformational distortions of the protein backbones of both complex partners. The observed changes were more pronounced for the longer transducer.

In order to facilitate an assignment of the spectral changes to contributions of either the receptor or its transducers, uniform isotope labelling of the transducer variants was performed (see section 3.4).

For initial rapid-scan FTIR measurements isotope labelled Htr(1-114) and Htr(1-157) were complexed with unlabelled receptor in a 1:1.05 stoichiometry to ensure quantitative complex formation of the transducers. Sample preparation was performed as above-mentioned.

Figure 3.39 shows the time-resolved difference spectra of the M_2 intermediate minus the groundstate (M_2 minus groundstate) of the labelled SRII/Htr157-complex referring to the unlabelled complex. The spectra were measured at -17 °C to improve the yield of the signalling state.

Interestingly, both spectra showed only minor differences. Especially the amide I/II spectral ranges were not influenced by labelling indicating that the observed spectral changes represent predominately conformational changes of the SRII receptor. Indications for conformational changes of the transducer protein backbone upon signal transfer from the receptor were not obtained. This phenomenon was recently reported by Kandori and co-workers as well [100].

Due to the low spectral resolution, in the ranges between 1500-1460 cm^{-1} and 1380-1340 cm^{-1} non-reproducible, minor changes were observed (indicated by arrows). These observations led to the conclusion that the contributions of the transducer to the observed spectral changes might be hidden by much stronger contributions of the receptor.

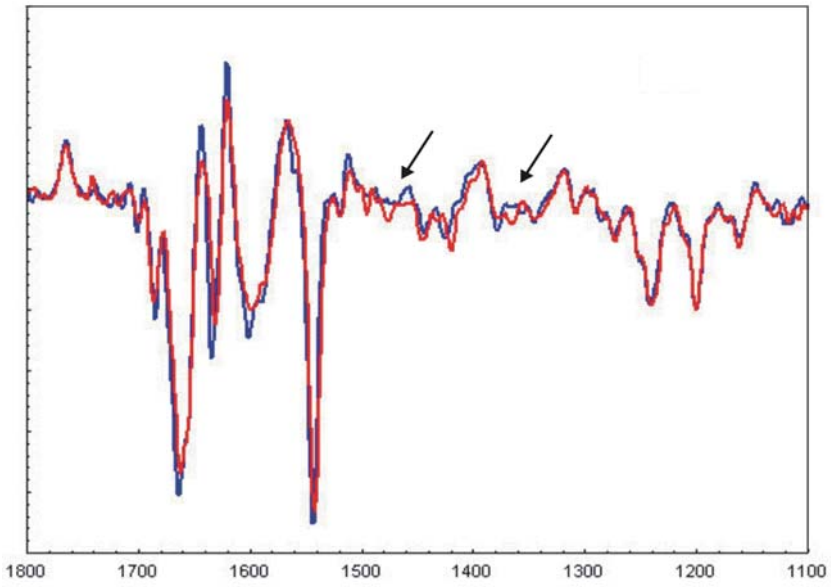


Figure 3.39. Overlay of the rapid-scan M_2-difference spectra of SRII/[^{13}C,^{15}N] Htr157 (red trace) and unlabelled SRII/Htr157 (blue trace) measured at -17 °C. The amide I/II spectral ranges were not influenced by isotope labelling. Due to the low spectral resolution, in the ranges between 1500-1460 cm^{-1} and 1380-1340 cm^{-1} irreproducible, minor changes were observed (indicated by arrows).

While the investigated protein complexes consisting of isotope labelled transducer and unlabelled receptor, revealed no detectable isotope shift, investigations with complexes of isotope labelled receptor with unlabelled transducer were performed. These complexes are expected to exhibit a downshift by approximately 30 cm^{-1} of the amide I and by approximately 10 cm^{-1} of the amide II bands of the receptor. This effect may facilitate the spectroscopic resolution of vibrational transducer bands.

For initial rapid-scan FTIR measurements isotope labelled SRII (protein expression and purification see section 3.4.2) was complexed with either unlabelled Htr(1-114) or unlabelled Htr(1-157) in a 1:1.05 stoichiometry to ensure quantitative complex formation of the labelled receptor. Sample preparation was performed as mentioned above.

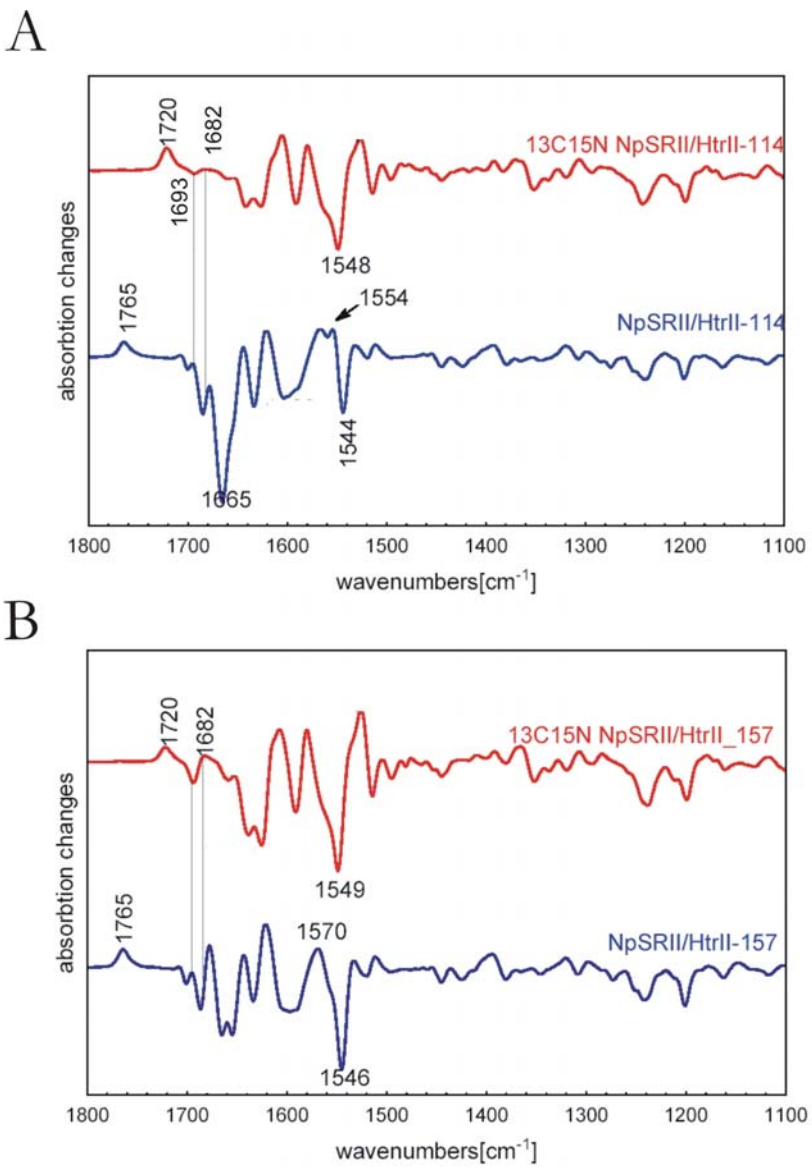


Figure 3.40. Steady-State difference spectrum of the M/O state minus the groundstate of [^{13}C,^{15}N] SRII in complex with different unlabelled transducers referring to the respective unlabelled complexes.
Panel A: Spectrum of [^{13}C,^{15}N] SRII/Htr114 (red) and SRII/Htr114 (blue).
Panel B: Spectrum of [^{13}C,^{15}N]-SRII/Htr157 (red) and SRII/Htr157 (blue).

A comparison of the M/O-state difference spectra of the labelled complexes (figure 3.40, panel A and B, red traces) with those of the unlabelled complexes (figure 3.40, panel A and B, blue traces) revealed a quantitative downshift of the characteristic C=O stretch vibration of SRII-Asp75 from 1765 cm^{-1} to 1720 cm^{-1} in the spectra of labelled

receptor in complex with both different transducers (figure 3.40, panel A and B, red traces). This indicates an almost quantitative isotope labelling of the receptor.

In addition to that, a significant isotope shift of the amide I and amide II vibrational bands of the labelled complexes (figure 3.40, panel A and B, red traces) was observed. Interestingly, the signal amplitudes of the labelled receptor in complex with the longer Htr157 transducer (figure 3.40, panel B, red trace) are improved compared to the corresponding complex with the short Htr114 transducers. This phenomenon was observed as well for the corresponding unlabelled complexes (figure 3.38, red and black trances).

As a consequence of the isotope shift of the amide I/II receptor vibrations, two bands appeared in the spectra of the labelled complexes at 1693 cm^{-1} and 1682 cm^{-1} (figure 3.40, red traces, indicated by vertical lines). These bands were insensitive to receptor labelling and their amplitudes were significantly enlarged for the longer transducer Htr157 compared to the shorter one. However, the origin this spectral shift from 1693 cm^{-1} to 1683 cm^{-1} remained unclear.

4. Discussion

4.1 Semi-synthetic access to domain specific isotope labelled *Np*HtrII

Sensory rhodopsin II in complex with its transducer protein HtrII from *N.pharaonis* initiates the photophobic reaction of the bacterium to green-blue light. Recently the X-ray structure of this complex yielded insights into the transmembrane structure of the complex components. In the membrane the complex is a 2:2 heterodimer: Two receptor/transducer units with tightly packed receptor helices F, G and the transducer helices TM1 and TM2.

Unfortunately, the crystallographic structure did not provide information about the C-terminal cytoplasmic fragment of the transducer including amino acids 83 to 114[9]. This region of the transducer is essential for the receptor–transducer interaction resulting in a low dissociation constant (K_d < 100 nM)[60] and the capability of the transducer to inhibit basal proton pump activity of the receptor upon complex formation[17,96].

Different reasons can explain the missing electron density information for this protein region: The C-terminal poly-histidine purification tag used for the Htr114 construct may disturb a correct folding of this domain. The tight packing of the complexes in the crystal may as well prevent a correct folding of the cytoplasmic domain.

In order to address this problem of the unknown structure of the cytoplasmic linker domain of *NpHtrII* other methods had to be used. Besides X-ray crystallography and electron microscopy, NMR is the only method to determine structures of proteins with atomic resolution[95]. In the case of membrane proteins the protein structure is strongly influenced by the native membrane environment. Magic angle spinning solid-state NMR spectroscopy (MAS ssNMR) offers the possibility to investigate membrane protein structures, such as the SRII/HtrII-complex, in their natural lipid environment[101].

Although MAS techniques have extended the sensitivity and the spectral resolution of solid-state NMR spectroscopy, the still limited spectral resolution could not permit a widespread application like high-resolution liquid ^{1}H NMR spectroscopy. Therefore it is yet not possible to resolve the NMR signals of the entire SRII/HtrII-complex or of the complete Htr114 construct alone with MAS ssNMR techniques and complete signal assignment for structural calculations cannot be achieved.

A method that allows domain specific isotope labelling of only the cytoplasmic Htr114 C-terminus, would considerably reduce the problem of signal overlapping: Unlabelled regions of the protein can be filtered out using suitable heteronuclear correlation

experiments leaving only the signals from the labelled part of the protein. In combination with the crystallographic data it would be possible to build a detailed model of the C-terminally truncated transducer on the atomic level.

The feasibility of domain specific isotope labelling has been demonstrated using two different protein-engineering methods, protein trans-splicing[102-104] and expressed protein ligation[105,106]. In a first study, Nakamura and co-workers[102] exploited the trans-spicing phenomenon for segmental [^{15}N] labelling of the C-terminal domain of the *E.coli* RNA polymerase α subunit. In a complementary study[106] EPL was used to achieve a domain specific [^{15}N] labelling within a Src-homology domain pair derived from the Abl protein tyrosine kinase. NMR experiments illustrated the power of segmental isotope labelling for studying large proteins.

The trans-splicing method convinces with methodical elegance. However, its requirement for denaturants and elevated temperatures at various points is a potential drawback. Although successful refolding of denatured integral membrane proteins has been reported, e.g. the halobacterial proton-pump bacteriorhodopsin[107], the *E.coli* outer membrane protein OmpA[108] or the bacterial potassium channel KcsA[109], functional refolding of integral membrane proteins remains a challenging task.

The method of expressed protein ligation offers the possibility to perform ligation reactions under mild physiological conditions. This renders EPL as a promising procedure to extend domain specific isotope labelling to the class of integral membrane proteins.

EPL was extensively used for the semi-synthesis of soluble proteins,[110,111] however, at the beginning of this project no example for the semi-synthesis of membrane proteins was reported. In the meanwhile Muir and co-workers described the use of EPL in the synthesis of a truncated form of the potassium channel KcsA[112]. This truncated form corresponded to the entire membrane-spanning region of the protein. The N-terminal peptide was fused with the *Mycobacterium xenopi* GyrA intein (Mxe) and a MESNA-thioester of this N-terminal KcsA segment was generated under mild conditions in the presence of the detergent Triton X-100. Native chemical ligation of the N-terminal α-thioester and the synthetic C-terminal peptide was performed. The ligation product was folded into the native tetrameric form by incorporation into lipid vesicles.

The extension of C-terminal domain specific isotope labelling of proteins to the truncated membrane protein *Np*HtrII requires access to uniformly isotope labelled peptides containing an N-terminal cysteine residue. Labelled amino acid building blocks

for solid phase peptide synthesis (SPPS) are commercially available and could be used to synthesise such peptides. However, the costs for the synthesis of a peptide comprising an entire protein domain would be immense.

A visible alternative are highly optimised protocols for recombinant labelling of proteins using synthetic minimal media containing [^{13}C] D-glucose and/or [^{15}N] ammonium chloride as sole carbon respectively nitrogen sources are well established[72,73]. The combination of recombinant isotope labelling with recombinant methods for the generation of peptide building blocks for an EPL-based approach offers an access to domain specific isotope labelled membrane proteins .

Two approaches for the generation of recombinant peptide building blocks with N-terminal cysteine are reported. The first approach makes use of proteolytic removal of an N-terminal leader sequence from a precursor protein by specific proteases that cleave C-terminal at their recognition site.

The tobacco etch virus nuclear inclusion A protease (TEV) fulfils the requirements to be a suitable tool for the preparative generation of peptides containing an N-terminal cysteine. It is an inexpensive, highly specific cysteine protease that cleaves C-terminal at its recognition site (ExxYxQ$^{\blacktriangledown}$, where x can be any amino acid)[74] and can be heterologously expressed in *E.coli*.[113].

An alternative approach does not require a separate proteolysis step and utilises N-terminal *M.xenopi* Gyr A intein fusion proteins that on temperature induction cleaves off the required peptide carrying an N-terminal cysteine[45,114]. A disadvantage of this approach is the possibility of *in-vivo* auto-cleavage or low *in-vitro* cleavage efficiencies of the purified precursor, diminishing the peptide yields or even preventing any yield of the peptide[45].

On the other hand a proteolytic digest of a precursor protein may lead to incomplete or unspecific cleavage of the precursor fusion protein. However, these risks can be minimized by the choice of the protease and a careful optimisation of the digest conditions.

The studies of Blaschke et al.[105] and Tolbert et al.[74] showed the generation of peptides with N-terminal cysteine residues by the proteolytic removal of a leader-sequence from a precursor protein. The function of these leader-sequences was to facilitate a recombinant access to the precursor protein and to serve as a purification tag thus facilitating a straightforward purification of the precursor protein.

The *Schistosoma japonicum* glutathione-S-transferase (GST) gene fusion system combines these features in an excellent manner[115]. It is a versatile system for the expression and purification of fusion proteins produced in *E.coli*. The system is based on inducible, high-level expression of genes or gene fragments as fusions with GST and yields fusion proteins with the GST moiety at the N-terminus of the protein of interest. The fusion protein accumulates within the cytoplasm of the cells and is purified in high yields exploiting the specific interaction of the GST with glutathion (GSH).

A fusion protein consisting of the desired cytoplasmic transducer fragment and the glutathione-S-transferase (GST)[115] linked by a TEV protease recognition site in a way that the N-terminal cysteine of the peptide is in the P1' position of the TEV recognition site promised to be a suitable construct to generate a building block for the EPL-based approach to an segmental isotope labelled transducer.

Based on this components, a semi-synthetic strategy was developed that allowed *Np*HtrII to be generated from two protein segments by a single ligation reaction in a C-terminal truncated form (residues 1–114) (figure 4.1). This transducer construct comprises two unlabelled transmembrane helices and an additional isotope labelled cytoplasmic fragment representing a transducer that possesses the ability to form a native-like receptor–transducer complex showing a low dissociation constant (K_d < 100 nM)[60] and that possesses the capability to inhibit the proton pump activity of the receptor[17,96].

Gly83-Gly84 was selected as the ligation site, because all the residues of the putative cytoplasmic receptor binding site with unknown structure are located C-terminal to this site[60] and will allow structure determination by MAS ssNMR with isotope labelled samples. The transmembrane transducer fragment comprising amino acids 1-83 was fused to the N-terminal splice-junction of the *Mycobacterium xenopi* gyrase A intein that is C-terminally fused to the *Bacillus circulans* chitin-binding domain (MxeCBD).

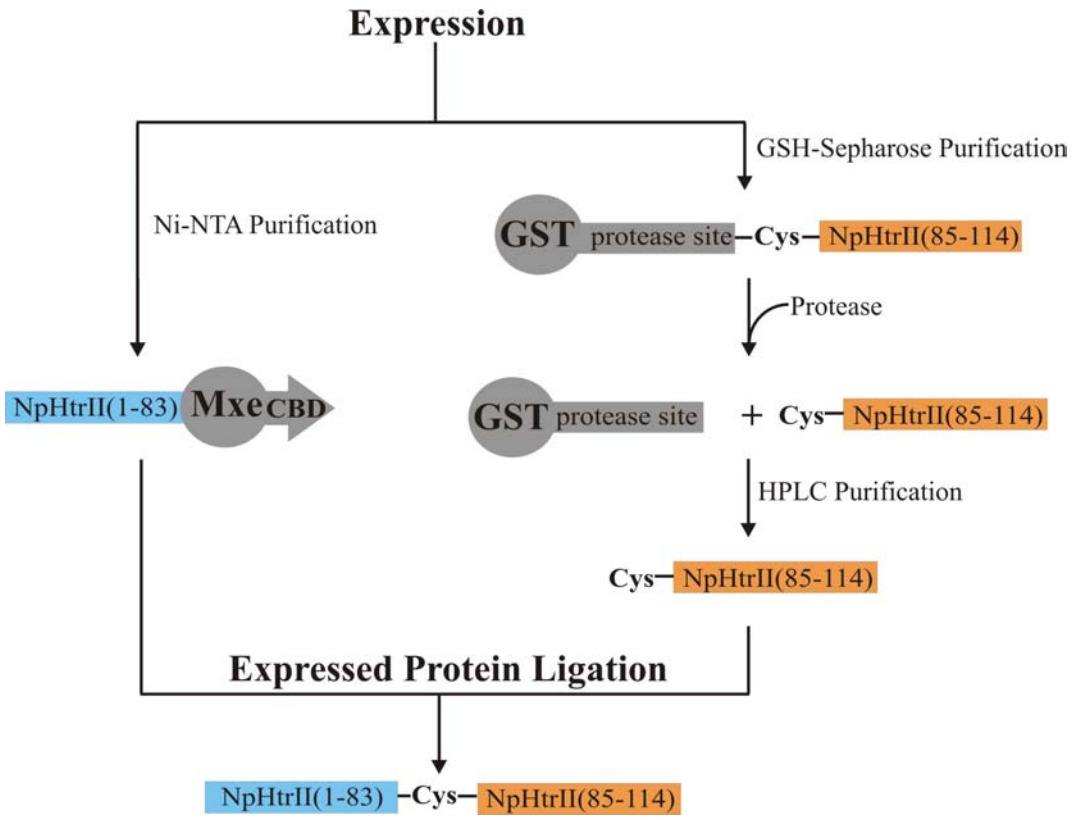


Figure 4.1. strategy for the generation of domain specific isotope labelled NpHtr114 using expressed protein ligation. The transmembrane domain comprising residues 1-83 was fused to the N-terminus of the MxeCBD mini-intein. The cytoplasmic domain comprising residues 84-114 was fused to the C-terminus of a GST-tag with a TEV protease recognition site in between. After expression in isotope enriched minimal medium, a TEV protease digest liberated the desired isotope labelled peptide containing an N-terminal cysteine residue for the subsequent ligation reaction.

The Mxe GyrA intein is a natural mini-intein (198 amino acids), which lacks a central intein endonuclease domain common to other naturally occurring inteins[45]. As a bacterial protein, it can be over expressed in *E.coli*[116]. In the absence of its native C-extein comprising a N-terminal cysteine residue, the Mxe intein remains as a precursor fusion protein at 37 °C and can then be thiol-induced to cleave at 16-23 °C[116].

Gly-83 at position –1 of the MxeCBD intein construct promises high *in-vitro* intein cleavage efficiencies[45], because Perler and co-workers evaluated the effect of the -1 residue on N-terminal Mxe intein cleavage efficiencies and determined a negligible *in-vitro* cleavage and an almost quantitative DTT-induced *in-vitro* cleavage for glycine at this position.

These observations in combination with high ligation efficiencies observed for a C-terminal glycine-thioester in a model peptide with the sequence LYRAx in a native chemical ligation study[61] led to the conclusion, that position G83-G84 of the transducer is a suitable ligation site.

A successful EPL reaction at this ligation site results in a G84C point mutation in the semi-synthetic transducer. In order to investigate, if this mutation is functionally tolerated in a C-terminally truncated *Np*Htr114 construct, an *in-vitro* binding assay based on blue native PAGE gels and a recombinant *Np*Htr114 G84C mutant was used. These gels clearly indicated that a complex is formed between the G84C mutant and its cognate transducer. This suggests that a cysteine substitution at this position does not abolish complex formation between the transducer and the receptor.

Additional evidence for this finding was provided by Spudich and co-workers with full-length *H.salinarum* transducer cysteine mutants[117]. In their study a G84C mutant exhibited a similar repellent phototaxis response to stimulating light at a wavelength of 500 nm as the wild type transducer. Both investigations suggest that position Gly84 is a suitable ligation site because the glycine to cysteine mutation at this position is functionally tolerated.

4.2 Generation of isotope labelled *Np*Htr(84-114) G84C peptides

The generation of domain specific isotope labelled semi-synthetic *Np*HtrII variants by EPL is based on the accessibility of isotope labelled peptides containing an N-terminal cysteine residue. Here the generation of [^{15}N] and [^{13}C,^{15}N] double-labelled peptides comprising the cytoplasmic region of the halobacterial transducer from residue 84-114 is presented.

An N-terminal cysteine residue was generated by TEV protease digest of a glutathion-S-transferase leader-sequence. The GST-carrier facilitated a high yield over-expression of the soluble precursor protein. 30 mg per litre minimal medium were obtained and purified by exploiting the specific interaction of the GST with glutathion (GSH) that is immobilised on a solid matrix.

Quantitative proteolytic liberation of the desired peptide was achieved with a molar ratio of TEV to GSH fusion protein of 1/45. A comparable high TEV concentration for an

effective cut was observed as well in a model study investigating the P1' specificity of the TEV protease[75].

The authors reported a 1:80 molar ratio for the processing of a fusion protein between *E.coli* maltose binding protein and *Aquifex aeolicus* NusG (MBP-NusG) with Ser, Ala and Gly at position P1', however for Cys, Phe, Gln, Tyr, Asn and Trp a 1:30 ratio was needed to achieve an efficient digest.

The target peptides in this work were obtained in analytically pure form after protease cleavage and RP-HPLC purification. The observed labelling efficiencies of 98 % corresponded well to the isotope content of the respective isotope sources. Yields were usually 1-2.5 mg per litre M9 minimal medium.

A number of publications have described the expression of isotope labelled peptides for NMR studies[118-124] that were produced together with small and highly expressed carrier proteins, like GST or ketosteroide isomerase. In some cases carrier proteins with low solubility have been exploited to direct the peptide to inclusion bodies, thereby minimising proteolytic degardation[118,121,124].

Sprules et al.[119] reported an enrichment of GST-fused peptides with ^{15}N and ^{13}C isotopes that was liberated using the PreScission Protease (Amersham Biosciences). Although this set-up was similar to the one presented in this work, the use of the commercially available PreScission protease leads to a predetermined N-terminus of the target peptide with Gly and Pro as the first two amino acid residues.

As alternative to a proteolytic digest, the application of cyanogens bromide (CNBr) for peptide release[118,121,123,124] allows the generation of peptides with any desired amino acid residue at its N-terminus. However, since CNBr cleavage occurs after each methionine residue, the method is intrinsically limited to peptide sequences without internal Met residues. Therefore it was not suitable for the generation of the cytoplasmic segment of *Np*HtrII that includes a methionine at position 100.

An interesting extension of the carrier fusion concept is the introduction of tandem repeats of peptide sequences separated by protease cleavage sites[124,125]. Takasuga et al.[125] presented the efficient production of the pentapeptide DSDGK that potently inhibits primary IgE antibody formation. The short peptide was expressed as a 28 tandem repeat construct in *E.coli*. The fused peptides were digested with trypsin that cleaved after the positively charged Lys residue and the monomeric peptide was purified by HPLC.

Kuliopulus and Walsh[124] extended this approach by using single methionine residues interspersing peptide tandem repeats. This produced isotope labelled 13mer yeast α-mating factor produced with a yield of 56 mg/L.

The method presented in this work fulfils highest demands in the generation of isotope labelled peptides without any N- or C-terminal modification. Except proline any amino acid residue can be generated at the N-terminus of the target peptide using TEV protease as a reliable tool for the generation of small isotope labelled peptides carrying an N-terminal cysteine.

4.3 Domain specific isotope labelling of *Np*HtrII(1-114) using EPL

The potential of expressed protein ligation to achieve domain specific isotope labelling of proteins has been demonstrated[105,106] in a model study with a Src-homology domain pair derived from the Abl protein tyrosine kinase. Here in this study an extension of this concept to domain specific isotope labelling of membrane proteins is reported.

For the generation of an activated α-thioester derivative of recombinant *Np*HtrII comprising its transmembrane domain up to amino acid residue Gly83, a fusion protein with the Gyr A intein was generated and over-expressed in *E.coli*.

The target HtrMxeCBD fusion protein was either found to be incorporated in into the bacterial membrane or to be deposited as insoluble inclusion bodies (90 %). From the solubilised membrane fraction two cleavable intein fusion protein species were isolated: the desired full-length protein (ca. 30-50 %) and, as the main product, an N-terminally degraded form.

MALDI mass spectrometry suggested a truncation of the HtrMxeCBD protein in the loop region between the two transmembrane helices. Since both components of this fusion protein, namely *Np*HtrII(1-83)[60] and the Gyr A mini intein[116], are reported to be functionally expressed in *E.coli*, the chosen manner of fusing both proteins together obviously caused the observed proteolytic degradation in an alanine-rich loop that connects both transmembrane helices of the transmembrane transducer domain.

Figure 4.2, panel A shows a representation of the HtrMxeCBD fusion protein lacking the chitin-binding domain. This model was built up from the crystal structures of both proteins[9,126]. Thereby Ala residue -1 of the N-extein was replaced *in-silico* by glycine and Ser1 of the crystallized Mxe intein by cysteine that was linked to Gly82 of the transducer extending the N-exteine. The dashed lines confine the major hydrophobic core of the protein according to the crystallographic data from Gordeliy et al.[9].

This model indicates that the ligation site G83-G84 may force the soluble intein domain into the hydrophobic membrane environment. This may lead to an incorrect incorporation of the fusion protein into the lipid bilayer resulting in a misfolded protein that either aggregates or is degraded by proteases. This hypothesis is in a good agreement with the observation of the degraded protein species isolated from the membrane fraction, and with the observation that the protein is deposited in insoluble inclusion bodies.

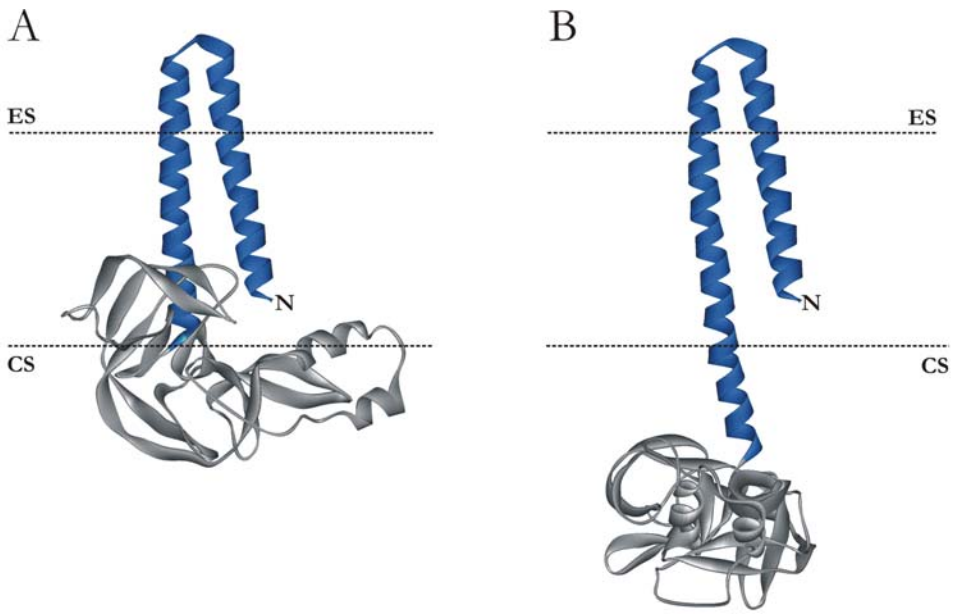


Figure 4.2. Side view of a Htr-Mxe fusion protein model lacking the CBD based on crystallographic data from Gordeliy et al.[9] (Htr in blue) and Klabunde et al.[126] (GyrA intein in grey). Cs, cytoplasmic side; ES, extra-cellular side. The dotted black lines confine the major hydrophobic of the transducer.
Panel A. The chosen G83-C84 fusion site forces the soluble GyrA intein into the membrane. This might induce misfolding of the entire fusion protein leading to insoluble aggregates (inclusion bodies) or proteolytic degradation.
Panel B. Assuming a helical continuation of TM2, a A94-C95 fusion site might facilitate a membrane independent orientation of the soluble GyrA intein. This might avoid misfolding leading to a functionally membrane incorporated conjugate.

Muir and co-workers as well reported similar problems with the expression of a transmembrane Mxe intein fusion protein[112], where the Mxe-intein was fused to the C-terminus of the KcsA potassium channel at position 73, that is located C-terminally to the end of the pore forming helix.

A variety of different expression conditions was tested. However, in no instance was the desired product detected. The authors argued that the hydrophobic N-extein targeted the intein fusion protein to the secretory pathways. To circumvent this problem, a sandwich fusion approach was chosen with a GST domain at the N-terminus of the N-extein-intein fusion protein. This construct was thus targeted to inclusion bodies.

A comparison of these results with the results presented in this work suggests that it may be important for functional incorporation of membrane protein-intein fusions into

the lipid bilayer, to limit interaction of the soluble intein domain with the bacterial membrane.

Positioning of the fusion site into flexible cytoplasmic loop regions of membrane proteins may prevent the intein to be forced into the lipid bilayer and may avoid misfolding, protein aggregation or degradation.

The work of Erni and co-workers supports these criteria for the choice of potential ligation sites. They generated a cyclized *E.coli* glucose transporter subunit II CB^{Glc} by *in-vivo trans*-splicing[127]. The N-intein of the split *recA* intein of *M.tuberculosis* was fused to the flexible N-terminus of the eight transmembrane helices containing C domain of the transporter and the C-intein was fused to the C-terminus of the cytoplasmic B domain. After expression cyclisation took place cyclized splicing product was successfully incorporated into the membrane.

Taking these results into account, a relocation of the *Np*HtrII fusion site from Gly83-Gly84 to Ala94-Ala95, as indicated in figure 4.2, panel B, may help to avoid the N-terminal degradation of the HtrMxeCBD fusion protein. This additional amino acid stretch might provide sufficient space for the intein domain from the lipid bilayer and therefore successful incorporation of the functionally folded fusion protein into the bacterial membrane might become possible. Experiments to test this hypothesis are in progress.

In contrast to the solubilised membrane fraction, the isolated inclusion bodies of the HtrMxeCBD expression contained predominantly non-degraded intein fusion protein. Therefore the fusion protein was refolded from these aggregates.

The isolated inclusion bodies were solubilised in buffer containing 6 M guanidine hydrochloride and TCEP as a reducing agent in the presence of DDM as a detergent and were subsequently functionally refolded using different protocols. An α-thioester was generated with cleavage efficiencies of approximately 50 % after refolding of the HtrMxeCBD protein using dialysis to replace guanidine hydrochloride against 8 M urea and to lower stepwise the chaotrope concentration to 6 M, 4 M, 3 M and finally 2 M urea.

In the literature, only a few examples for the functional refolding of inteins were reported besides a large number of examples for the refolding of split-inteins[102-104]. Anraku and co-workers[128] described in 1996 that a precursor protein containing the yeast VMA intein, can undergo protein splicing upon refolding from inclusion bodies by

dialysis. The inclusion bodies were solubilised in buffer containing 6 M guanidine hydrochloride and DTT as a reducing agent.

In 1999, Muir and co-workers reported the results of studies to explore the scope and limitations of EPL. With the above-mentioned protocol only 5 % of a similar VMA precursor had been cleaved after refolding[81].

A publication by Sydor et al.[129] described the first successful renaturation of the Gyr A intein in fusion with green fluorescent protein N-extein from inclusion bodies. Their protocol used a 1:10 dilution of the solubilised inclusion bodies in buffer containing 6 M guanidine hydrochloride and DTT with a buffer containing 4 M urea followed by removing residual guanidine hydrochloride by dialysis against the chaotrope-free dilution buffer. An EPL reaction was subsequently performed under these denaturing conditions.

In the same year Muir and co-workers described the refolding of a fusion protein of the potassium channel KcsA and the MxeCBD intein by a simple dialysis step against chaotrope-free buffer. They started with a solution of the KcsAMxeCBD protein in buffer containing 6 M guanidine hydrochloride without any reducing agent.

All these experiments reported in literature and in this work show that the yeast VMA and the bacterial Gyr A intein can be functionally refolded. Noteworthy is the fact that cleavage of the Mxe intein in the presence of a certain amount of urea (up to 4 M) is possible. However, a common protocol for the renaturation of fusion proteins containing the Gyr A intein does not exist; for each construct a new refolding protocol needs to be established.

After successful intein cleavage and generation of a thioester derivative purification of the desired Htr(1-83) thioester using RP-HPLC methods failed. The target protein irreversibly adsorbed to the used RP-materials. Attempts with different elution buffers like isopropanol/acetonitrile mixtures that were successfully used for the purification of a semi-synthetic helix 1 from BR[91], or formic acid/isopropanol mixtures that allowed the purification of a synthetic F_1F_0 synthetase subunit C, remained unsuccessful.

Therefore an *in-situ* ligation strategy without intermediate thioester isolation was developed. The cleavage of the HtrMxeCBD fusion protein was initiated in the presence of the isotope labelled C-terminal transducer segment by the addition of MESNA. Cleavage and ligation were done in a one-pot reaction.

In the literature many examples have been described using this strategy[81,110,130-132], however, in most cases a huge excess of the synthetic peptide is used. Since the amount

of isotope labelled peptide including an N-terminal cysteine residue was a limiting factor in this work, a reaction stoichiometry of 1:1 was chosen. While with HtrMxeCBD fusion protein isolated from *E.coli* membranes product formation was observed under these conditions, the strategy failed for the refolded HtrMxeCBD.

Although this reaction was done in the presence of 2 M urea as denaturant, which is reported to enhance ligation rates by destabilising secondary structures and therefore improving the accessibility of the thioester moiety for the attacking cysteine residue, no ligation product was found. Even an increase of the concentration of the chaotrope to 6 M guanidine hydrochloride, after a successful generation of the Htr-thioester did not lead to a successful ligation reaction.

Hunter and Kochendoerfer[91] report a possible shielding effect of detergent molecules that has been observed which prevents the ligation reactions of hydrophobic peptides. However this effect alone cannot be the reason for the unsuccessful ligation reaction, since the reaction with the membrane isolated intein precursor was not prevented due to the presence of DDM as detergent.

This suggests an additional factor preventing the ligation reaction using refolded Htr thioester. Dawson and co-workers report that the self-assembly of two peptide fragments of the chymotrypsin inhibitor 2 enhanced their ligation rate due to the close proximity of the respective termini[133]. This conformational assistance had a bigger influence on the ligation rate than the presence of denaturants in the ligation mixture.

If a favourable thioester conformation enhances ligation reactions, it is reasonable to assume that an unfavourable thioester conformation would inhibit the ligation reaction. In combination with the shielding effect of DDM detergent, an unfavourable conformation of the Htr-thioester with an inaccessible C-terminus, in which the refolded thioester species was trapped during the refolding process, may have prevented the ligation reaction of the transducer fragments.

In order to avoid DDM as a detergent, expressed protein ligation of the refolded Htr-thioester was performed in a lipidic cubic phase environment. The successful product formation supports the postulated conformational hindrance of the ligation reaction in the presence of detergent and chaotropes. In contrast to the chaotropes used, the incorporation of the Htr-thioester into the bilayer of the lipidic cubic phase managed to facilitate the accessibility of the C-terminal thioester moiety for the isotope labelled transducer fragment. The key feature of this approach that was recently reported by Kochendoerfer and co-workers[91], is the reconstitution of the membrane polypeptide

into the lipidic cubic phase, which consists of a lipid bilayer and two aqueous channels that are both topologically unbounded, giving the aqueous phase access to both sides of an infinite lipid bilayer[134].

In this work the lipidic cubic phase concept was a useful tool to facilitate the chemical ligation reaction of the archeabacterial transducer. Nevertheless it remains unclear, if the lipidic cubic phase serves as a "solvent" that successfully prevents the precipitation of hydrophobic peptides as proposed by Kochendoerfer and co-workers, or if it acts as a "lipochaperone" that facilitates correct folding of α-helical peptides spanning the entire lipid bilayer.

Semi-synthetic access to the domain specific isotope labelled archeabacterial transducer in a C-terminally truncated variant is a successful extension of the segmental isotope labelling concept to the field of membrane proteins utilising the expressed protein ligation. The formation of the desired product was observed using either a membrane isolated HtrMxeCBD intein precursor or the respective protein refolded from insoluble inclusion bodies.

For the purification protocols of the products derived under two different reaction conditions, with and without lipidic cubic phase, different crucial points had to be considered: During the purification of the product derived from the membrane isolated HtrMxeCBD intein precursor, the observed full-length Htr114 had to separated from the N-terminally degraded by-product, the purification protocol of the lipidic cubic phase product required the successful solubilisation of the reaction product from the bilayer environment.

In both cases a RP-HPLC assisted purification was not successful due to irreversible adsorption of the membrane proteins to the reversed phase used.

Although the separation of unligated peptide, the Htr-thioester and the uncleaved intein precursor from the full-length and the degraded product, which were observed in reactions without lipidic cubic phase, was successfully carried out using DEAE anion exchange chromatography, a quantitative separation of both ligation products failed. However, the main part of the N-terminally degraded transducer by-product was removed.

The extension of the anion exchange chromatography to the transducer derived from reactions in lipidic cubic phase was not achieved due to insufficient solubilisation of the

protein from the lipids. Here a careful optimisation of the protocol has to be done in near future.

Besides ligation and purification of semi-synthetic or synthetic proteins, their successful refolding is an important step on the way to a functional protein. Numerous examples of functionally refolded synthetic or semi-synthetic soluble protein are reported[110,135]. However, membrane protein folding remains a challenging task[108].

In experiments of this work using blue native gel electrophoresis to investigate the ability of the semi-synthetic transducer to bind to the SRII receptor, no complex formation was observed. After reconstitution of the misfolded semi-synthetic Htr into purple membrane lipids from *H.salinarum*, the transducer was able to bind to the receptor.

Further BN-PAGE experiments revealed a refolding efficiency of around 20 %. This efficiency is similar to the one observed for the semi-synthetic potassium channel KcsA[112]. Muir and co-workers were able to refold the ligated monomer with an estimated efficiency of 30 % into the functional tetrameric from. Their refolding protocol included a reconstitution step of the membrane protein as well; in this case from SDS containing buffer into soybean lipid vesicles.

Sanders and co-workers published a study concerning the reconstitutive refolding of the integral membrane protein diacylglycerol kinase, which is found in the cytoplasmic membrane of many bacteria[92]. Their method for successful refolding of 65 single-cysteine mutants of DAGK is based on reconstituting the proteins into POPC vesicles.

All these *in-vitro* refolding investigations revealed the important role that membrane lipids play in the folding process of membrane proteins. Biophysical studies on model systems, like bacteriorhodopsin, are beginning to provide a sound physical basis for membrane protein folding.

For the folding of α-helical membrane proteins, like the semi-synthetic transducer or the potassium channel, a two-stage model has been proposed that breaks the processes down to two key stages[136]. In the first stage, individually stable transmembrane α-helices form and in a second, these helices pack to form a functional protein.

For bacteriorhodopsin as a model system a rate-limiting folding step has been identified that results in the key folding intermediate to which retinal binds[137]. Surprisingly, a slow formation of α–helices has been observed during this rate-limiting step[138]. It was proposed that this helix formation reflects the apparently slow folding of the ends of the

transmembrane α-helices following the formation of a helical bundle consisting of only the central regions of the α-helices within the hydrophobic bilayer interior. This model indicates that only these central helical segments are independently stable enough to fold according to the two-stage model. In the second stage the lipid bilayer seems to be necessary for the final folding events of α–helix formation[139].

The consideration of these results suggests a plausible model of the receptor binding behaviour of the ligated transducer: Folding intermediates of the semi-synthetic Htr consisting of the stable central α-helical regions of TM1 and TM2, may be trapped by the interaction with detergent micelles. At this stage the transducer would not be able to bind the receptor. Upon the interaction with the purple membrane bilayer the protein may be able to undergo conformational changes leading to the properly folded integral membrane protein. The lipidic cubic phase environment could play a similar role during the ligation reaction with the refolded intein precursor.

In this context it will be an interesting result, if the semi-synthetic transducer isolated from the lipidic cubic phase will be able to bind the SRII without an additional reconstitution step of the protein into purple membrane lipids. Such a result could help to understand the chaperone like role of lipids as an important structure-forming environment[140].

The presented work demonstrates the feasibility of expressed protein ligation for domain specific isotope labelling of membrane proteins. The results are an extension of the work of Nakamura and Muir, who established methods for segmental isotope labelling of soluble proteins based on either *trans*-splicing[102-104] or expressed protein ligation[105,106]. The results presented here may help to circumvent the limits of spectral resolution of solid state NMR or may assist signal assignment in FTIR spectroscopy supporting structural investigations of integral membrane proteins.

4.3 Investigations of the SRII/HtrII complex using NMR and FTIR

An important step in understanding the structural basis of signal transduction in archeabacterial phototaxis was the determination of the crystal structure of sensory rhodopsin II alone[36,37] and co-crystallized with a truncated transducer variant comprising the two transmembrane helices TM1 and TM2 and a short, 30 amino acids long, part of the cytoplasmic linker domain[9].

The obtained structure shows a tight packing of the receptor helices F, G and the transducer helices TM1 and TM2 within the membrane. The observation that the SRII activation causes an outward movement of helix F [19,20] resulting in a rotation of transducer helix TM2.

Time-resolved EPR experiments revealed a synchronisation of the outward movement of helix F and the rotation of TM2[21]. These synchronised movements occur during the UV spectroscopically silent $M_1 \rightarrow M_2$ transition[13]. Hein et al.[141] were able to distinguish between these two M states by static and time-resolved step-scan FTIR spectroscopy. Both states exhibited the characteristic signal at 1764 cm^{-1} of the protonated counter-ion Asp75 of the Schiff base, however, the second M state differed in the observed amide I/II signal amplitudes.

In this work the influence of complex formation on the receptor in combination with two C-terminally truncated transducer variants, Htr(1-114) and Htr(1-157), was investigated using static and time-resolved rapid scan FTIR methods. The spectra of a mixture of M and O state of the complex, so obtained were surprisingly similar to the receptor reference spectrum. Only alterations of the C=O signal of Asp75 and of the amide I region indicated complex formation and can be interpreted as distortions of the protein backbone of the complex partners upon transducer binding. No explanations for the rotary movement of the transducer helix TM2 were found. As well Kandori and co-workers[100] did not find any hints for such a movement in their previous FTIR studies.

In addition, no changes of the difference FTIR spectra of the early K[142] or the signalling M state with isotope labelled, truncated transducers in complex with unlabelled receptors were observed. The absence of transducer signals suggests that the Htr constructs may have either unstructured C-termini that show no backbone motion as a response to the transmembrane activation, or that the rotation of the transducer helix

TM2 induces "rigid body" movements of complete secondary structure elements without significant distortions of the protein backbone.

In this context the observation of a spectral downshift of the vibrational bands of Htr(1-114) respectively Htr(1-157) at 1693 cm^{-1} in complex with uniformly isotope labelled SRII due to the formation of the signalling M-state, is interesting. The frequency of these bands is characteristic for either a C=O stretching vibration of an asparagine or glutamine residue, respectively or an amide I vibration of the transducer backbone.

Kandori and co-workers suggested an assignment of these two bands to Asn74 of the transducer[100]. The X-ray structure of the complex[9] showed that this residue forms a hydrogen bond with Tyr199 of the SRII in the middle of the membrane. An N74T point mutation in Htr(1-159) caused the disappearance of these bands. From this fact it was concluded that not an amide I vibration of the transducer backbone but Asn74 was the origin of the signal.

However, taking the measurements with the different transducer constructs into account, it cannot be explained, why Asn74 showed such different signal amplitudes in the spectra obtained, as similar amplitudes for both variants were expected.

The data published by Furutani et al.[100] does not exclude the disappearance of the observed bands due to a distortion of the protein backbone upon introduction of a point mutation. The exchange of the hydrogen bond acceptor asparagine by the donor threonine, accompanied by the loss of the hydrogen bond with Tyr199 of the receptor, may cause conformational changes in the transducer. These conformational changes possibly will influence the FTIR spectrum, however, they do not cause the dissociation of the complex[100].

At the moment it remains unclear, why the observed spectral shift of the 1693 cm^{-1} band exhibits significantly increased amplitudes for the longer transducer Htr(1-157). In this context secondary structure predictions for the transducer may help to address this question. The cytoplasmic linker domain of the transducer contains two so-called heptad repeats a-b-c-d-e-f-g with hydrophobic side chains at positions a and d indicating the presence of two amphiphatic helices comprising amino acids 83-101 and 115-131 connected by an unstructured loop region. These amphiphatic helices 1 and 2 are predicted to form a coiled-coil structure.

The truncation of the shorter transducer variant Htr(1-114) after amino acid residue Glu114 that was performed by Hippler et al.[62] in order to identify a putative receptor-

binding site, removed the complete amphiphatic helix 2. Its absence may possibly lead to the loss of the predicted coiled-coil structure that could serve as a kind of "anchor" that stabilises the HAMP domain structure in the cytoplasm. These structural differences may be reflected by the differences observed in the amplitude of the negative band at 1693 cm^{-1} and the positive band at 1682 cm^{-1} in the amide I region of the isotope labelled SRII in complex with either Htr(1-157) or Htr(1-114).

While the predicted loop between both amphiphatic helices may be C-terminally fixed in the longer transducer variant by the predicted coiled coil, the unfixed and unstructured loop in the shorter Htr(1-114) variant will possibly exhibit a very high mobility.

This is in good agreement with the MAS ssNMR data obtained, of the complex comprising an unlabelled SRII and an uniformly isotope labelled Htr(1-114) presented in this study. They suggest unstructured random coil regions at both N- and C-terminus, which turn into more rigid, structured regions as more residues come close to the membrane spanning helices TM1 and TM2.

While the C-terminus of Htr(1-114) seems to be unstructured, MAS ssNMR experiments reported by Saito and co-workers[143] suggest a more structured cytoplasmic domain of the longer Htr variant. In dipolar decoupled MAS experiments with a [3-^{13}C] Ala and [1-^{13}C] Val labelled truncated transducer Htr(1-159) complexed with SRII they observed two kinds of ^{13}C NMR signals. One of these resonances, a low-field α-helical peak (α_{II}-helix), was identified as a cytoplasmic α-helix protruding from the bilayer. The other resonance was the high-field α-helical peak (α_{I}-helix) of the transmembrane helices.

The first signal was almost completely suppressed in cross-polarization MAS experiments indicating a higher structural mobility of the cytoplasmic helices compared to the more rigid transmembrane helices.

In summary the discussed FTIR spectroscopic results revealed a different receptor binding behaviour of the Htr(1-114), a minimal model for the transmembrane SRII binding domain of the transducer protein, compared to Htr(1-157) including a predicted cytoplasmic coiled-coil structure. It will be of interest in further NMR and FTIR experiments, which information about the nature of the HAMP domain can be obtained from experiments with these constructs.

The differences observed in the FTIR data of the short transducer compared to the long transducer lead to the question about the importance of the amino acid residues 83-114 for the binding of the transducer to the receptor. Calorimetric[60] and electrophysiological[96] methods showed that the Htr(1-114) variant is in contrast to Htr(1-101) able to bind to the SRII receptor suggesting an direct interaction of the amino acids Gly101 to Glu114.

The calorimetric investigations reported in this work suggest a direct interaction of SRII with the peptide comprising the amino acid residues Gly84 to Arg113. The dissociation constant of this interaction was determined to be 240 nM at 22 °C with a binding heat of –4.6 kJ/mole. Interestingly, the observed binding constant was similar to the one observed for the corresponding truncated Htr(1-114) transducer variant comprising both transmembrane helices (K_D=230 nM)[60]. This result comes as a surprise, because the van-der-Waals interactions that were observed in the transmembrane receptor-transducer binding interface[9] are expected to contribute to the receptor binding process. However, these interactions were not reflected by an alteration of the observed binding constant.

Taking into account the NMR and CD spectroscopic evaluations of the peptide, which show a predominantly random coil conformation with weak α-helical tendencies, the interaction with the receptor has to initiate a folding process of the peptide that subsequently facilitates the interaction of the folded peptide with the receptor. However, the observed binding curve or the binding constant did not indicate such a folding process.

Kamo and co-workers reported hints for a folding event in investigations with a isotope labelled peptide comprising amino acid residues Gly83 to Gln149[67]. They compared the ^{1}H,^{15}N 2D HSQC spectra of the peptide in absence/ presence of SRII and concluded a secondary structure formation in the peptide from line broadening and spectral shifts upon addition of the receptor. However, the observed signal dispersion of approximately 0.5 ppm in the ^{1}H dimension, which is comparable to the one found in this work for the short peptide, indicates random coil conformations even in the presence of the receptor contradicting a folding event of the peptide.

In summary, although the peptide comprising Gly83 to Arg113 was shown to interact with the SRII receptor exhibiting a similar dissociation constant as was observed for truncated transducer constructs Htr(1-114) and Htr(1-157), it remains unclear, why this amino acid stretch is an important binding determinant.

EPR experiments with the truncated transducer Htr(1-157) can provide important hints to answering this question. These measurements revealed a low spin label mobility for site-directed spin labelled cysteine mutants for the amino acid residues up to Ala94 and starting with Ala95, a very high spin label mobility (Klare,J.P., unpublished data).

These data suggest a structured amino acid stretch following the transmembrane helix 2 up to residue Ala95. The presence of the flexible unstructured loop region that normally links two amphiphatic helices, at the C-terminus in the short Htr(1-114) construct may prevent the structured amino acid stretch Gly83-Ala95 from unfolding. Removing this flexible loop, as in the Htr(1-101) variant, disturbs the folding of the structured stretch following TM2 and leads to the loss of the ability to bind to the receptor.

According to these arguments the amino acid stretch Gly101 to Glu114 would be no classical receptor-binding domain that supports the receptor binding with specific interactions between both binding partners. Rather its presence may prevent unfolding of the amino acid cluster Gly83-Ala95, which seems to be a very important receptor-binding determinant.

5. Summary

Crystallographic studies of the photophobic receptor-transducer-complex from *N.pharaonis*[9] yielded insights into the structural and functional organisation of the transmembrane domains of the complex components. However, they did not provide information about the structure of the C-terminal cytoplasmic domain of the transducer that is thought to play an important role in receptor binding and signal propagation.

In order to elucidate the structure of the cytoplasmic domain of a C-terminally truncated Htr(1-114) construct, magic angle spinning solid state NMR (MAS ssNMR) spectroscopy was applied. To circumvent the limits of the spectral resolution of MAS ssNMR, a semi-synthetic access to a domain-specific isotope labelled HtrII protein based on expressed protein ligation was introduced.

In a first step the generation of an isotope labelled peptide building block with an N-terminal cysteine comprising the cytoplasmic part 84-114 of the transducer was established. This peptide was C-terminally fused to a glutathion-S-transferase leading sequence linked by a TEV protease recognition site. After expression in isotope enriched minimal medium, a TEV protease digest liberated the desired isotope labelled C-terminal segment.

The N-terminal transmembrane segment comprising residues 1-83 of the transducer was fused to the N-terminus of the MxeCBD mini-intein. The resulting HtrMxeCBD fusion protein was expressed in *E.coli*. Ni-NTA purification of target protein solubilised from the bacterial membrane revealed the presence of an N-terminally degraded HtrMxeCBD form in addition to the target protein.

Since this degraded intein species showed full intein cleavage activity, in ligation reactions an N-terminally degraded by-product besides the desired target product was observed. A complete separation of the desired semi-synthetic transducer from the degraded by-product or a separation of the target intein precursor protein from the degraded one could not be achieved even by different chromatographic methods.

The ability of product so obtained to form a complex with its cognate receptor was investigated in blue-native electrophoresis experiments. After a reconstitution into purple membrane lipids, complex formation of the semi-synthetic transducer was observed indicating that the membrane bilayer contributes to the correct folding of the semi-synthetic protein.

Analysis of HtrMxeCBD protein partitioning revealed that the desired not degraded intein precursor protein was deposited predominantly in inclusion bodies. After

successful refolding of the protein, the formation of the desired semi-synthetic product without an N-terminally degraded by-product was observed in ligation reactions in lipidic cubic phase environment. However, a purification protocol including an effective solubilisation of the target protein from the lipidic cubic phase has yet to optimised.

In addition to the establishment of a semi-synthetic access to domain specific isotope labelled Htr(1-114), uniformly isotope labelled transducer constructs were expressed in *E.coli*. The spectroscopic analysis of uniformly labelled Htr(1-114) in complex with unlabelled SRII receptor using magic angle solid state NMR revealed unstructured random coil regions at both N- and C-terminus of the Htr-construct, which turn into more rigid, structured regions the more the residues come close to the membrane spanning helices TM1 and TM2.

The influence of SRII/HtrII complex formation on the structure of the receptor in combination with two C-terminally truncated transducer variants, Htr(1-114) and Htr(1-157), was investigated using static and time-resolved rapid scan FTIR methods. The observed spectra showed alterations of the C=O signal of Asp75 of the receptor and of the amide I region with increased signal amplitudes for the Htr(1-157) construct indicating complex formation. No changes in the difference spectrum of the signalling M state of both labelled transducer constructs in complex with unlabelled receptors were observed. In contrast to this, isotope labelling of the SRII receptor led to the observation of a spectral downshift of vibrational bands of unlabelled Htr(1-114) and Htr(1-157) at 1693 cm^{-1} due to the formation of the M-state. Again an increased signal amplitude for the Htr(1-157) construct was observed.

These observations led to the question about the importance of the amino acid cluster 83-114 for SRII-binding. ITC experiments with a peptide comprising amino acids Gly84 to Arg113, which was found to be predominantly random coiled with α-helical tendencies in NMR- and CD-experiments, revealed an interaction with SRII characterised by a K_D of 240 nM at 22 °C and a binding heat of -4.6 kJ/mole. The observed binding constant was similar to the one observed for a Htr(1-114) transducer construct (K_D=230 nM)[60].

The method of domain specific isotope labelling of proteins using expressed protein ligation was extended in this work to the field of membrane proteins. With the C-terminally truncated transducer Htr(1-114) as an example, a semi-synthetic strategy was introduced that may contribute to provide detailed insights into the phototactic signal

transduction process of *N.pharaonis* on atomic level in further MAS ssNMR and FTIR investigations.

5. Summary (german)

Kristallographische Untersuchungen des phototaktischen Rezeptor-Transducer-Komplexes aus *N.pharaonis*[9] ergaben Einblicke in die strukturelle und funktionelle Organisation der transmembranen Domänen der Komplexkomponenten. Jedoch konnten keine Information über die Struktur der zytoplasmatischen Domäne des Transducers erhalten werden, die eine wichtige Rolle im Prozess der Rezeptorbindung und der Signalweiterleitung spielt[60].

Magic angle spinning Festkörper-NMR (MAS ssNMR) Spektroskopie wurde zur strukturellen Untersuchung der zytoplasmatischen Domäne eines C-terminal verkürzten Htr(1-114) Konstrukts angewendet. Um die Grenzen der spektralen Auflösung der MAS ssNMR Spektroskopie zu umgehen, wurde ein halb-synthetischer Zugang auf der Basis der *expressed protein ligation* zu einem domänenspezifisch isotopenmarkierten HtrII-Protein vorgestellt.

In einem ersten Syntheseschritt wurde die Erzeugung eines isotopenmarkierten Peptids etabliert, das die Aminosäurereste 84-114 des zytoplasmatischen Teils des Transducers und ein N-terminales Cystein umfasst. Dieses Peptid wurde, verbrückt durch eine TEV Protease Erkennungssequenz, an den C-Terminus einer Glutathion-S-transferase-Leitsequenz fusioniert. Nach der Expression des Fusionsproteins in isotopenangereichertem Minimalmedium setzte ein Verdau mit TEV-Protease das gewünschte isotopenmarkierte C-terminale Segment frei.

Das N-terminale, transmembrane Segment, das die Reste 1-83 des Transducers umfasste, wurde an den N-Terminus des MxeCBD Mini-Inteins fusioniert. Das erhaltene HtrMxeCBD Fusionsprotein wurde in *E.coli* exprimiert. Eine Ni-NTA Aufreinigung des Zielproteins, das aus der bakteriellen Membran solubilisiert wurde, offenbarte die Gegenwart einer N-terminal degradierten Form des HtrMxeCBD Proteins.

Diese Inteinspezies zeigte jedoch volle Intein-Spaltaktivität, weshalb in Ligationsreaktionen ein N-terminal degradiertes Nebenprodukt zum gewünschten Zielprodukt beobachtet wurde. Eine vollständige Abtrennung des halb-synthetischen Transducers oder eine Trennung des gewünschten Inteinvorläufers vom degradierten

Fusionsprotein konnte unter Verwendung unterschiedlicher chromatographischer Methoden nicht erreicht werden.

Die Fähigkeit des so erhaltenen Produkts zur Bindung des SRII-Rezeptor, wurde in Experimenten mit Blau-nativer Gelelektrophorese untersucht. Eine Komplexbildung des halb-synthetischen Transducers konnte erst nach dessen Rekonstitution in Lipide aus Purpurmembran beobachtet werden. Dies deutet an, dass die Membrandoppelschicht zur korrekten Faltung des halb-synthetischen Proteins beiträgt.

Die Analyse der Verteilung des HtrMxeCBD Proteins ergab, dass nicht degradiertes Vorläuferprotein in Form von *inclusion bodies* vorlag. Nach erfolgreicher Faltung des auf diese Weise gewonnenen Proteins konnte in Ligationsreaktionen in kubischen Lipidphasen die Bildung des gewünschten halb-synthetischen Produkts ohne ein N-terminal degradiertes Nebenprodukt nachgewiesen werden. Die Optimierung eines Protokoll zur Produkt-Aufreinigung, das einen Schritt zur effektiven Solubilisierung des Zielproteins von der kubischen Lipidphase beinhalten muss, steht aus.

Zusätzlich zur Etablierung eines halb-synthetischen Zugangs zu domänenspezifisch isotopenmarkiertem Htr(1-114), wurden vollständig isotopenmarkierte Transducer-Konstrukte in *E.coli* exprimiert. Die MAS ssNMR spektroskopische Untersuchung eines Komplexes aus vollständig markiertem Htr(1-114) und unmarkiertem SRII-Rezeptor ergab unstrukturierte Zufallsknäuelstrukturen an beiden Enden des Htr-Konstruktes, die zu strukturierten Regionen übergehen, je mehr sie sich den Helices TM1 und TM2 nähern.

Der Einfluss der SRII/HtrII-Komplexbildung auf die Struktur des Rezeptors in Kombination mit zwei C-terminal verkürzten Transducer-Varianten, Htr(1-114) und Htr(1-157), wurde in statischen und zeitaufgelösten *rapid scan* FTIR Experimenten untersucht. Die erhaltenen Spektren zeigten durch die Komplexbildung hervorgerufene Änderungen des C=O Signals von Asp75 des Rezeptors und Änderungen in der Amid-I Region mit größeren Signalamplituden für das Htr(1-157)-Konstrukt im Vergleich zur kürzeren Variante.

Zwischen den Differentspektren des signalgebenden M-Zustandes beider markierten Transducer-Konstrukte im Komplex mit unmarkiertem Rezeptor wurden keine Änderungen beobachtet. Im Gegensatz dazu ermöglichte die Isotopenmarkierung des SRII-Rezeptors die Beobachtung einer spektralen Verschiebung von Schwingungsbanden des unmarkierten Htr(1-114) und des unmarkierten Htr(1-157) bei

1693 cm^{-1} aufgrund der Komplexbildung. Hier wurde ebenfalls eine größere Signalamplitude für das Htr(1-157)-Konstrukt beobachtet.

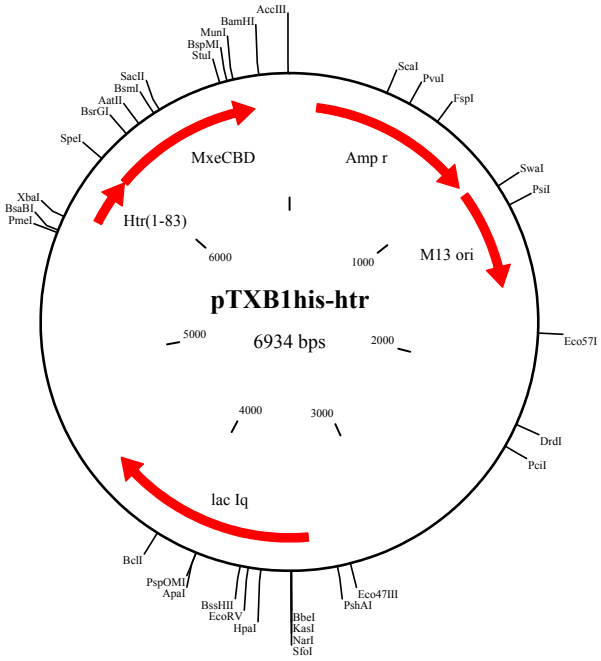
Diese Beobachtung führte zur Frage nach der Wichtigkeit der Aminosäurereste 83-114 für die Bindung des SRII-Rezeptors. ITC Experimente mit einem entsprechenden Peptid Gly84- Arg113, das sich in Experimenten mit NMR- und CD-Spektroskopie als Zufallsknäuel mit α-helicalen Tendenzen zeigte, enthüllte eine Wechselwirkung dieses Peptids mit dem SRII-Rezeptor, die durch ein K_D von 240 nM bei 22 °C und einer Bindungswärme von -4.6 kJ/mol gekennzeichnet ist. Die beobachtete Bindungskonstante war ähnlich der, die für einen Htr(1-114) Transducer beobachtet wurde (K_D=230 nM)[60].

Die Methode der domänenspezifischen Isotopenmarkierung von Membranproteinen unter Ausnutzung der *expressed protein ligation* wurde in dieser Arbeit auf das Arbeitsfeld der Membranproteine übertragen. Am Beispiel des C-terminal verkürztem Transducers Htr(1-114) wurde eine halb-synthetische Strategie vorgestellt, die in weiterführenden MAS ssNMR- und FTIR-Untersuchungen zu detaillierten Einblicke in den Prozess der phototaktischen Signaltransduktion in *N.pharaonis* auf atomarer Ebene beitragen kann.

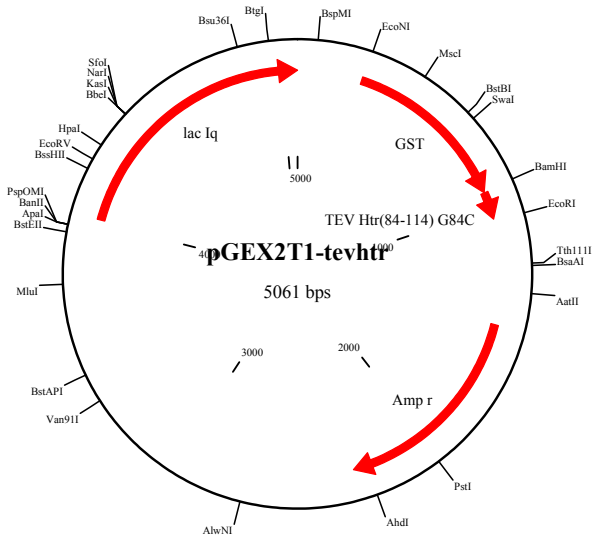
6. Appendices

6.1 Plasmid maps

6.1.1 pTXB1his-htr



6.1.2 pGEX2T-tevhtr



6.2 Protein sequences

GST-tevHtr

MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL
EFPNLPYYID GDVKLTQSMA IIRYIADKHN MLGGCPKERA EISMLEGAVL
DIRYGVSRIA YSKDFETLKV DFLSKLPEML KMFEDRLCHK TYLNGDHVTH
PDFMLYDALD VVLYMDPMCL DAFPKLVCFK KRIEAIPQID KYLKSSKYIA
WPLQGWQATF GGGDHPPKSD LVPRGSENLY FQCDTAASLS TLAAKASRMG
DGDLDVELET RRE

Number of amino acids: 263

Molecular weight: 30370.0

Formula: $C_{1373}H_{2129}N_{351}O_{396}S_{15}$

Htr-MxeCBD

MSLNVSRLLL PSRVRHSYTG KMGAVFIFVG ALTVLFGAIA YGEVTAAAAT
GDAAAVQEAA VSAILGLIIL LGINLGLVAA TLGCITGDAL VALPEGESVR
IADIVPGARP NSDNAIDLKV LDRHGNPVLA DRLFHSGEHP VYTVRTVEGL
RVTGTANHPL LCLVDVAGVP TLLWKLIDEI KPGDYAVIQR SAFSVDCAGF
ARGKPEFAPT TYTVGVPGLV RFLEAHHRDP DAQAIADELT DGRFYYAKVA
SVTDAGVQPV YSLRVDTADH AFITNGFVSH ATGLTGIHHH HHHHSGLNSG
LTTNPGVSAW QVNTAYTAGQ LVTYNGKTYK CLQPHTSLAG WEPSNVPALW
QLQ

Number of amino acids: 353

Molecular weight: 37357.5

Formula: $C_{1680}H_{2629}N_{465}O_{489}S_{6}$

*Np*HtrII(1-114)

MSLNVSRLLL PSRVRHSYTG KMGAVFIFVG ALTVLFGAIA YGEVTAAAAT GDAAAVQEAA VSAILGLIIL LGINLGLVAA TLGGDTAASL STLAAKASRM GDGDLDVELE TRRENSHHHH HHH

Number of amino acids: 123

Molecular weight: 12679.5

Formula: $C_{559}H_{907}N_{163}O_{167}S_3$

*Np*HtrII(1-114) G84C

MSLNVSRLLL PSRVRHSYTG KMGAVFIFVG ALTVLFGAIA YGEVTAAAAT GDAAAVQEAA VSAILGLIIL LGINLGLVAA TLGCDTAASL STLAAKASRM GDGDLDVELE TRRENSHHHH HHH

Number of amino acids: 123

Molecular weight: 12709.6

Formula: $C_{560}H_{909}N_{163}O_{167}S_4$

*Np*HtrII(1-157)

MAVSRLLLPS RVRHSYTGKM GAVFIFVGAL TVLFGAIAYG EVTAAAATGD AAAVQEAAVS AILGLIILLG INLGLVAATL GGDTAASLST LAAKASRMGD GDLDVELETR REDEIGDLYA AFDEMRQSVR TSLEDDAKNA EEDAEQAQKR AEEINTNSHH HHHHH

Number of amino acids: 165

Molecular weight: 17421.5

Formula: $C_{755}H_{1214}N_{220}O_{245}S_4$

*Np*SRII

MVGLTTLFWL GAIGMLVGTL AFAWAGRDAG SGERRYYVTL VGISGIAAVA
YVVMALGVGW VPVAERTVFA PRYIDWILTT PLIVYFLGLL AGLDSREFGI
VITLNTVVML AGFAGAMVPG IERYALFGMG AVAFLGLVYY LVGPMTESAS
QRSSGIKSLY VRLRNLTVIL WAIYPFIWLL GPPGVALLTP TVDVALIVYL
DLVTKVGFGF IALDAAATLR AEHGESLAGV DTDAPAVADE NSHHHHHHH

Number of amino acids: 249

Molecular weight without retinal: 26645.2

Molecular weight with retinal: 26911.6

Formula: $C_{1244}H_{1922}N_{310}O_{325}S_7$

Semi-synthetic *Np*HtrII(1-114) G84C:

MSLNVSRLLL PSRVRHSYTG KMGAVFIFVG ALTVLFGAIA YGEVTAAAAT
GDAAAVQEAA VSAILGLIIL LGINLGLVAA TLGCDTAASL STLAAKASRM
GDGDLDVELE TRRE

Number of amino acids: 114

Molecular weight: 11564.4

Formula: $C_{511}H_{849}N_{139}O_{156}S_4$

6.3 Additional material

6.3.1 Introduction to Expressed Protein Ligation

Proteins show an enormous variety of structure-function-relationships. However, they only consist out of 21 different building blocks, the canonical amino acids. The amino acid sequence of a protein determines its structure and therefore the function of the protein. In order to elucidate this complicate relationship, often methods of molecular biology are used to alter the amino acid sequence of a protein to study the observed effects of this alteration on the structure and the function of the mutant.

Most methods are using the cellular synthesis apparatus of simple organisms like bacteria or yeasts to produce the desired mutant and therefore they are restricted in the choice of the amino acids to the 21 canonical amino acids. However, for many applications a larger variety of the amino acid composition of polypeptides would be desirable.

In order to overcome this restriction Schulz and co-workers[144] introduced the cell free synthesis of proteins to facilitate the incorporation of non-coded amino acids into proteins. The method applies the suppression of the amber stop codon by a chemically amino acylated tRNA. The codon is introduced on the DNA or the RNA level by site-directed mutagenesis. The corresponding tRNA linked to a non-coded amino acid, recognizes the amber codon and this amino acid is incorporated into the desired protein. The Achilles' heel of this cell free synthesis method are the observed low yields in combination with low loading rates of the tRNA due to the labile amino acyl bond between non-coded amino acid and tRNA.

While cell free synthesis is a valuable extension to biosynthetic methods that utilise a biological synthesis apparatus, chemical peptide synthesis provides the complete freedom over the covalent structure of proteins. E. Fischer was the first to propose an in-solution coupling of individual amino acid building blocks for the synthesis of short peptides. Further development of this method facilitated the synthesis of insulin in 1966[145]. Insulin consists of two short polypeptide chains connected by disulfide bonds.

Peptide synthesis in solution is related to an enormous chemical expenditure, since the desired product has to be isolated after each coupling step. The introduction of solid phase peptide synthesis (SPPS) by Bruce Merrifield[146] represents an important improvement in the field of chemical peptide synthesis. SPPS is based on the linkage of the growing peptide chain to a solid polymer support by a removable linker. This set up allows a purification of the desired product after each coupling step by simple filtration.

This method represented a breakthrough on the way to the total chemical synthesis of proteins. However, the accumulation of a growing amount of very similar by-products with increasing chain lengths complicating the isolation of the target peptide and limits the accessible size of polypeptides to around 60 residues[147].

An important step to overcome this limitation was the introduction of the native chemical ligation of peptides[87]. Two synthetic peptide segments are joined together by the thiol-mediated formation of a native peptide bond (Figure 2.1).

A prerequisite for the ligation reaction to occur is an N-terminal peptide segment containing a carboxy-terminal α-thioester and a C-terminal peptide possessing an amino terminal cysteine residue. In a transesterfication reaction the sulfhydryl moiety of the N-terminal cysteine attacks the carboxyl carbon atom of the thioester. Subsequently an S-N acyl shift that is irreversible under the chosen reaction conditions, leads to the formation of a native peptide bond at the ligation site.

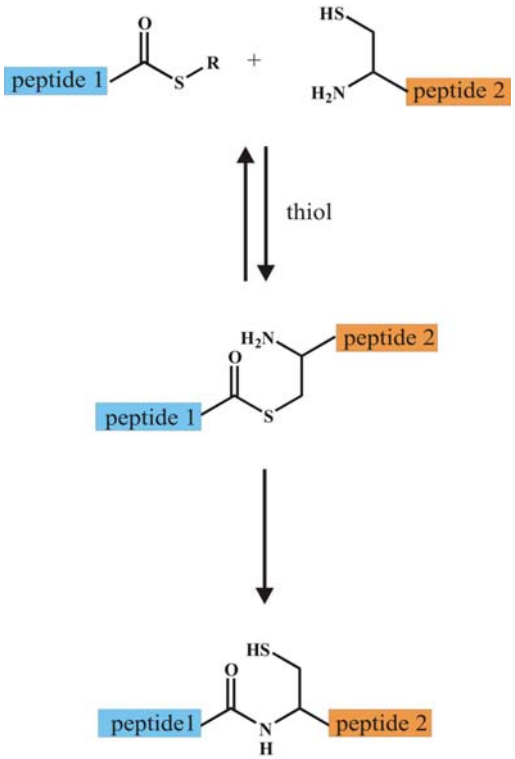


Figure 2.1. Schematic representation of the native chemical ligation.

The concept of native chemical ligation has proven to be very useful for the synthesis of small proteins or protein domains, however it could not be extended to the synthesis of proteins beyond approximately 20 kDa[88]. One possibility to further expand the scope of

this method is the chemical ligation of small synthetic peptides and large recombinant segments.

Although methods to generate proteins with N-terminal cysteines have been reported[74], until the beginning 1990s no appropriate method to generate recombinant proteins containing C-terminal α-thioesters was established[86]. Protein splicing, a phenomenon of an intra-molecular protein rearrangement leading to the extrusion of an internal fragment (intein) and the ligation of the two adjacent peptide sequences, has been shown to form an α-thioester intermediate (figure 2.2)[148].

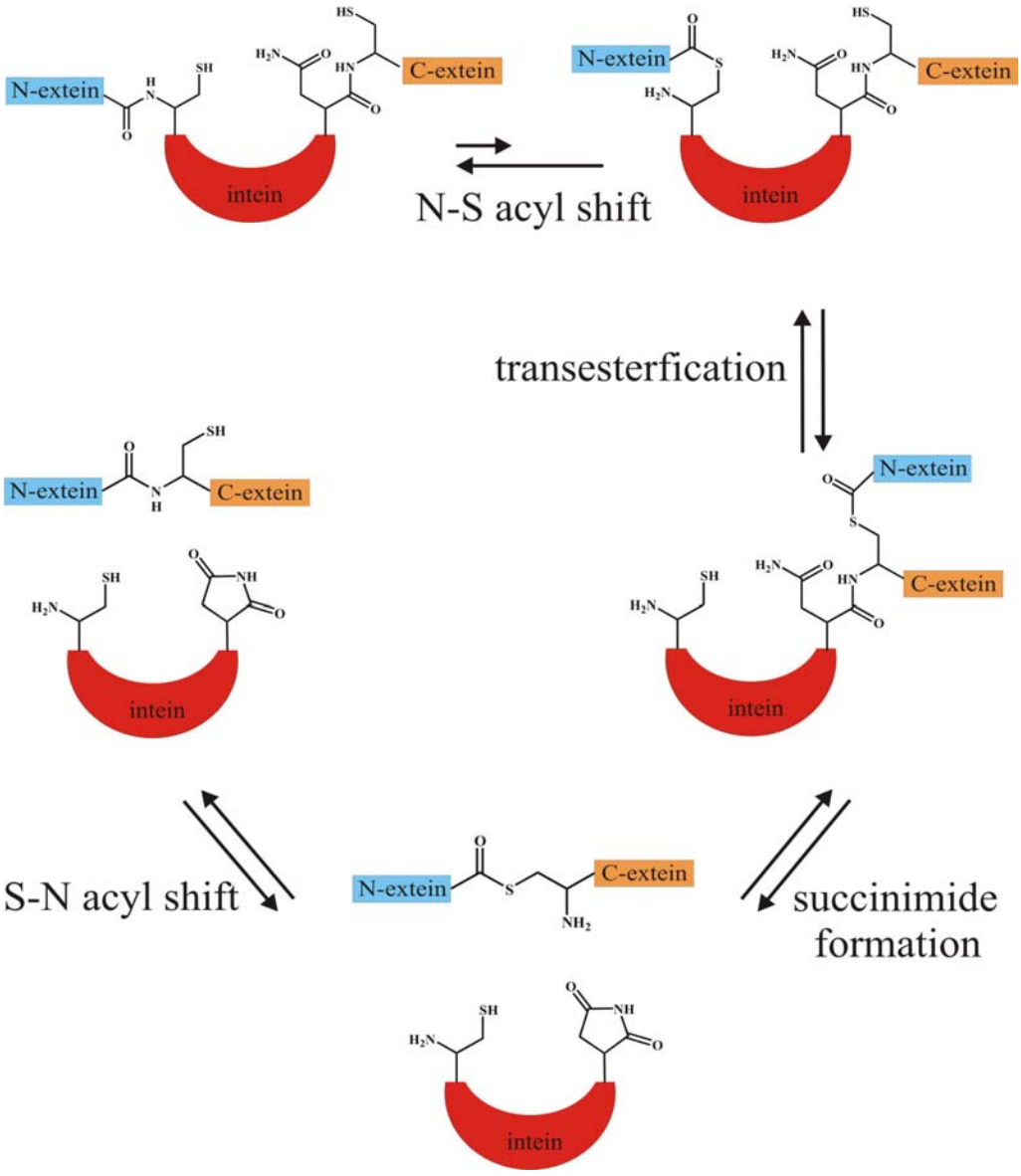


Figure 2.2. Schematic representation of the protein splicing mechanism.

Mechanistically the cleavage event is initiated by an N-S(O) acyl shift at the N-terminus of the intein which has to contain a cysteine, serine or threonine residue for this purposes[84]. The fold of the intein brings the exteins close together and enables a transesterfication reaction leading to a branched intermediate[149] with two free N-terminal amino moieties. The subsequent release of the intein by asparagine cyclisation leaves the N-extein attached to the side chain of the attacking cysteine of the C-extein (or Ser, Thr) via an (thio) ester bond, which is resolved by a spontaneous S(O)-N acyl rearrangement resulting a native peptide bond.

Intein mutants with C-terminal asparagines replaced by other amino acids like alanine have been shown to be defective in accomplishing the complete splice procedure but to be still capable of the initial α-thioester formation. This intermediate can be trapped by thiols and subsequently used reactions called expressed protein ligation (EPL) or intein mediated ligation (IML) (figure 2.3), which are analogue to the native chemical ligation reaction.

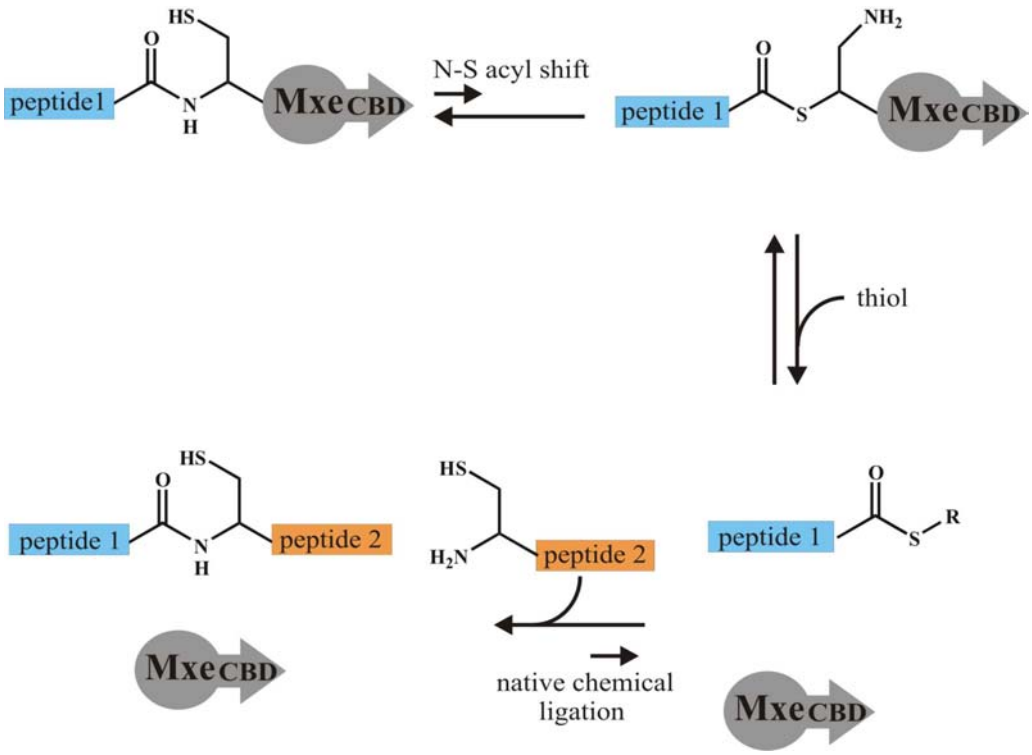


Figure 2.3. Expressed protein ligation. The commercially available pTXB[45] vectors for high-level protein expression of the short GyrA intein from *Mycobacterium xenopi* (Mxe-intein) in *E.coli* fused to a chitin binding domain facilitate the generation of the desired target protein containing a C-terminal α-thioester.

6.3.1 Introduction to Nuclear Magnetic Resonance Spectroscopy (NMR)

The phenomenon of nuclear magnetic resonance is based on the interaction of the magnetic momentum μ of a nucleus with an external magnetic field of strength B. The magnetic momentum and is proportional to the angular momentum J:

$$\vec{\mu} = \gamma \vec{J} \Rightarrow \left|\vec{\mu}\right| = \gamma\hbar\sqrt{I(I+1)}$$

The constant of proportionality γ is called gyromagnetic ratio and is a characteristic for each nucleus. The angular momentum orientates itself in the external magnetic field in that way as the vector component is parallel to the field. J_z is a whole- or half-numbered multiple of Planck's constant:

$$J_z = m\hbar \Rightarrow \mu_z = m\gamma\hbar$$

whereas the magnetic quantum number m can be –I, -I+1, …, +I. The external magnetic field splits up the energy levels of the nucleus. The energy E of these levels can be classically described by the formula of a magnetic dipole in a homogenous magnet field of strength B:

$$E = -\mu_z B = -m\gamma\hbar B$$

When a group of spins is placed in a magnetic field, each spin aligns in one of the two possible orientations. At room temperature, the number of spins in the lower energy level, N^+, slightly outnumbers the number in the upper level, N^-. Boltzmann statistics describes the ratio of both as:

$$\frac{N^-}{N^+} = e^{-\frac{E}{kT}}$$

The signal in NMR spectroscopy results from the difference between the energy absorbed by the spins which make a transition from the lower energy state to the higher energy state, and the energy emitted by the spins which simultaneously make a transition from the higher energy state to the lower energy state. The signal is thus proportional to the population difference between the states. At any instant in time, a magnetization vector can represent the magnetic field due to the spins. The size of each vector is proportional to (N+ - N-). The vector sum of the magnetization vectors from all spins is the net magnetization. At equilibrium, the net magnetization vector lies along the direction of the applied magnetic field B and is called the equilibrium magnetization Mo. In this configuration, the Z component of magnetization MZ equals Mo. MZ is referred to as the longitudinal magnetization. There is no transverse (MX or MY) magnetization

here. It is possible to change the net magnetization by exposing the nuclear spin system to energy of a frequency equal to the energy difference between the spin states. If enough energy is put into the system, it is possible to saturate the spin system and make MZ=0. The time constant, which describes how MZ returns to its equilibrium value, is called the spin lattice relaxation time (T1). The equation governing this behaviour as a function of the time t after its displacement is:

$$M_z = M_0\left(1 - e^{-\frac{t}{T_1}}\right)$$

If the net magnetization is placed in the XY plane it will rotate about the Z-axis at a frequency equal to the frequency of the photon, which would cause a transition between the two energy levels of the spin. This frequency is called the Larmor frequency ω_0, which corresponds to the resonance frequency of the spins and therefore to the transition frequency between the energy levels:

$$\gamma\hbar B = \Delta E = h\nu = \hbar\omega_0 \Rightarrow \omega_0 = \gamma B$$

In addition to the rotation, the net magnetization starts to dephase because each of the spin making it up is experiencing a slightly different magnetic field and rotates at its own Larmor frequency. The time constant, which describes the return to equilibrium of the transverse magnetization, M_{XY}, is called the spin-spin relaxation time, T_2. T_2 is always less than or equal to T_1. The net magnetization in the XY plane goes to zero and then the longitudinal magnetization grows in until M_o is along Z. It is convenient to define a rotating frame of reference (X',Y',Z) that rotates about the Z-axis at the Larmor frequency. A magnetization vector rotating at the Larmor frequency in the laboratory frame appears stationary in a frame of reference rotating about the Z-axis. A coil of wire placed around the X-axis will provide a magnetic field along the X-axis when a alternating current is passed through the coil. In a frame of reference rotating about the Z axis at a frequency equal to that of the alternating current, the magnetic field along the X' axis will be constant. In magnetic resonance, the magnetic field created by the coil passing an alternating current at the Larmor frequency is called the B_1 magnetic field. When the alternating current is turned on and off, it creates a pulsed B_1 magnetic field along the X' axis.

The spins respond to this pulse in such a way as to cause the net magnetization vector to rotate about the direction of the applied B_1 field inducting a voltage in a detector coil, which is registered. The rotation angle depends on the length of time the field is on, τ,

and its magnitude B_1

$$\theta = 2\pi\gamma\tau B_1$$

A 90 ° pulse is one which rotates the magnetization vector clockwise by 90 degrees about the X' axis and rotates the equilibrium magnetization down to the Y' axis. After a 90° pulse the free induction decays (FID) in time due to T_2 relaxation processes. The time-resolved signal is a superposition of all exited frequencies and is converted from time to frequency domain by Fourier transformation (FT).

In a molecule the electrons surrounding the nuclei produce themselves a weak magnetic field and insulate therefore the nucleus marginally from the external field. As a result the Lamor frequencies of the nuclei differ due to different chemical environments. This effect is called chemical shift. Nuclei experiencing the same chemical environment or chemical shift are called equivalent. Nuclei, which are close to one another, exert an influence on each other's effective magnetic field. If the distance between two non-equivalent nuclei A and B is less than or equal to three bond lengths, this effect is observable and called spin-spin coupling or J coupling. For two nuclei, A and B, three bonds away from one another in a molecule the spin of each nucleus can be either aligned with the external field or opposed to the external. The magnetic field at nucleus A will be either greater than B_o or less than B_o by a constant amount due to the influence of nucleus B. In NMR, only transitions between these states are allowed where the spin of one nucleus changes from spin up to spin down, or spin down to spin up. Absorptions of energy where two or more nuclei change spin at the same time are forbidden. Therefore two absorption frequencies for the nucleus A and two for the B nucleus can be observed centred on δ_A and δ_B, respectively The distance between two split absorption lines is called the J coupling constant or the spin-spin splitting constant and is a measure of the magnetic interaction between two nuclei. The coupling constant is depending on the strength of the external magnetic field and of the torsion angle between the two nuclei. Therefore coupling constants contain structural information and especially the torsion angle ϕ of the protein backbone can be determined by means of $^3J(H^N\text{-}H^\alpha)$ coupling constants.

The interpretation of a 1D NMR spectrum of more complex molecules becomes more or less impossible due to signal overlapping. This restriction can be overcome by the introduction of a second spectral dimension. In a 2D NMR experiment after a first 90 ° pulse the spins start to depahse. After an evolution time (t_1) the magnetization is transferred with a second pulse to an other nucleus. Finally the FID of the transferred

magnetization is recorded as a function of time (t_2). In the consecutive 1D experiments of a 2D experiment the evolution time is stepwise varied. Therefore the indirect time domain is screened. This data is Fourier transformed first in the t_2 direction to give a set of one-dimensional spectra modulated in t_1, the δ_2 dimension. A second Fourier transformation in the t_1 direction gives a δ_1 dimension.

In NMR structure determination three kinds of homo-nuclear 2D experiments become more and more important: 2D-COSY, 2D-TOCSY and 2D-NOESY. The 2D-COSY (Correlation Spectroscopy) experiment transfers magnetization via J-coupling from one nucleus to the other and yields therefore structural information in form of $^3J(H^N\text{-}H^\alpha)$ coupling constants. In a 2D-TOCSY experiment (Total Correlation Spectroscopy) the magnetization is successively transferred via J-coupling to all spin systems of a amino acid residue. On this basis the amino acids of a protein can be identified. Decisive for protein structure analysis is the NOESY experiment (Nuclear Overhauser and Exchange Spectroscopy). It is based on dipolar coupling of nuclear spins through space, the Nuclear Overhauser Effect, and is in first order reciprocal proportional to the sixth order of magnitude of the distance between both spins. Therefore in the spectrum spins are correlated, which can be are far away from each other in primary structure, however are close together in ternary structure. This is the most important structural information in NMR structure determination.

The number of signals in a 2D-^{1}H spectrum disproportionately rises with rising protein sizes. The resulting limited the spectral resolution can be enhanced by hetero-nuclear NMR experiments. Besides hydrogen a protein contains other magnetically active nuclei, namely ^{13}C and ^{15}N. However, their low natural abundance and low gyromagnetic ratios compared to ^{1}H limit the spectral sensitivity. Therefore mainly two strategies are used to enhance the sensibility: isotopic labelling during protein expression on [^{15}N] ammonium chloride and/or [^{13}C] glucose containing minimal media and effective transfer of huge proton magnetization to hetero-nuclei and back. The HSQC experiment (Hetero-nuclear Single Quantum Coherence) is the most important Experiment, which utilises the magnetization transfer to a heteroatom. In a 2D experiment the hetero nucleus frequency is correlated to the bounded proton frequency. Therefore overlapping homo-nuclear signals can be displayed in the heteroatom dimension with improved signal resolution. Homologue experiments can be performed either with ^{15}N or with ^{13}C nuclei.

The relevance of NMR for structural biology enormously rose since the end of 1960s with the introduction of pulsed Fourier-Transform-NMR[150] and multi-dimensional NMR spectroscopy[151] accompanied with the development of super-induction magnets and computational methods. Besides X-ray crystallography and electron microscopy NMR spectroscopy is the only method to determine structures of proteins and nucleic acids on atomic level[95]. In addition, time-resolved investigations of intra-molecular dynamics, molecular recognition processes, reaction kinetics or of protein folding are the unique strength of NMR[152-154].

Although NMR spectroscopy based on observation of only a small number of spins yielded biologically relevant information on human haemoglobin[155] (M=65 000 Da) as early as 1969, the use of NMR for de novo structure determination was limited until the middle 1990s to relatively small, soluble proteins (<30 000 Da). At high magnetic fields, typically used for studies of proteins in solution, chemical shift anisotropy interaction forms a significant source of relaxation in addition to dipole-dipole relaxation. This leads to increasing overall transverse relaxation rates, therefore to poor spectral resolution, and a low signal-to-noise ratio. For studies of membrane protein, the size of the mixed protein-detergent micelles needed to solubilize membrane proteins in water is beyond the limit of conventional NMR experiments. Recently the introduction of transverse relaxation-optimized spectroscopy (TROSY)[156] at high magnetic fields, combined with refined isotope labelling strategies[157] extended NMR studies to molecular sizes up to 100 kDa. As an example the structure of the integral membrane protein OmpX from *E.coli* reconstituted in 60 kDa DHPC micelles could be calculated[158]. However, a significant barrier to membrane protein research is retaining the structure and function of the protein during solubilization in detergents[159]. Some proteins are soluble only in one detergent species; others are soluble in many detergents, however are only functionally active in one of them. Therefore the structural investigation of membrane proteins in their native membrane environment is in the centre of scientific interest. In order to study solid-phase systems special NMR techniques were designed. In an asymmetric molecule, the chemical shift is dependent on the orientation of the molecule relative to the external magnetic field, thus each possible orientation has its own resonance frequency, leading to a broad lin width, the so-called "powder pattern" under static conditions. Another reason for broad spectral lines is dipolar broadening. The magnitude of the interaction of two spins varies with angle θ of their relative orientation and is a function of $(3\cos^2\theta-1)$.

When the angle in the above term is 54.7 °, 125.3 °, 234.7 °, or 305.3 °, the dipole interaction vanishes and the spectral line width is narrowed to the fast tumbling limit. The angle 54.7 ° is called the magic angle, θ_m. In magic angle solid-state NMR (MAS ssNMR) experiments samples are oriented at θ_m with respect to the B_0 magnetic field and spun at a rate comparable to the solid-state line width. Such conditions have, thus far, not permitted a widespread application of high-resolution ^{1}H NMR, however they have greatly extended the possibilities to study multiply or uniformly [^{13}C,^{15}N] labelled polypeptides in high sensitivity and adequate spectral resolution[101]. The first example for which almost complete sequential resonance assignments were obtained providing the basis of 3D structure calculations[160], was the SH3 domain of α-spectrin (62 residues)[161]. A further example is the elucidation of inter-atomic distances in an uniformly [^{13}C,^{15}N] labelled light-driven proton pump bacteriorrhodopsin (BR)[162]. Frequency-selective methods allowed distance measurements between N_ε of Lys216, which binds the retinal chromophore via a protonated Schiff base, and the C_γ of surrounding aspartic acid residues. Advances in sample preparation including for example modular labelling and improvements in NMR hardware instrumentation could open up the opportunity to complete 3D characterization of larger membrane proteins.

7. References

1. Buck,L. & Axel,R. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell* **65**, 175-187 (1991).

2. Birnbaumer,L. & Brown,A.M. G proteins and the mechanism of action of hormones, neurotransmitters, and autocrine and paracrine regulatory factors. *Am. Rev. Respir. Dis.* **141**, S106-S114 (1990).

3. Hargrave,P.A. & McDowell,J.H. Rhodopsin and phototransduction: a model system for G protein-linked receptors. *FASEB J.* **6**, 2323-2331 (1992).

4. Spudich,J.L. & Bogomolni,R.A. Sensory rhodopsins of halobacteria. *Annu. Rev. Biophys. Biophys. Chem.* **17**, 193-215 (1988).

5. Hildebrand,E. & Dencher,N. Two photosystems controlling behavioural responses of Halobacterium halobium. *Nature* **257**, 46-48 (1975).

6. Marwan,W. & Oesterhelt,D. Quantitation of photochromism of sensory rhodopsin-I by computerized tracking of Halobacterium halobium cells. *J. Mol. Biol* **215**, 277-285 (1990).

7. Klare,J.P. Strukturelle und funktionelle Untersuchungen des photophoben Rezeptor/Transducer-Komplexes aus *Natronobacterium pharaonis*. 1-137. 1-12-2002. Universität Dortmund.
   Ref Type: Thesis/Dissertation

8. Chen,X. & Spudich,J.L. Demonstration of 2:2 stoichiometry in the functional SRI-HtrI signaling complex in Halobacterium membranes by gene fusion analysis. *Biochemistry* **41**, 3891-3896 (2002).

9. Gordeliy,V.I. *et al.* Molecular basis of transmembrane signalling by sensory rhodopsin II-transducer complex. *Nature* **419**, 484-487 (2002).

10. Jung,K.H., Spudich,E.N., Trivedi,V.D. & Spudich,J.L. An archaeal photosignal-transducing module mediates phototaxis in Escherichia coli. *J. Bacteriol.* **183**, 6365-6371 (2001).

11. Bogomolni,R.A. & Spudich,J.L. Identification of a third rhodopsin-like pigment in phototactic Halobacterium halobium. *Proc. Natl. Acad. Sci. U. S. A* **79**, 6250-6254 (1982).

12. Spudich,J.L. & Bogomolni,R.A. Mechanism of colour discrimination by a bacterial sensory rhodopsin. *Nature* **312**, 509-513 (1984).

13. Chizhov,I. *et al.* The photophobic receptor from Natronobacterium pharaonis: temperature and pH dependencies of the photocycle of sensory rhodopsin II. *Biophys. J.* **75**, 999-1009 (1998).

14. Schmies,G., Chizhov,I. & Engelhard,M. Functional expression of His-tagged sensory rhodopsin I in Escherichia coli. *FEBS Lett.* **466**, 67-69 (2000).

15. Chizhov,I. *et al.* Spectrally silent transitions in the bacteriorhodopsin photocycle. *Biophys. J.* **71**, 2329-2345 (1996).

16. Chizhov,I. & Engelhard,M. Temperature and halide dependence of the photocycle of halorhodopsin from Natronobacterium pharaonis. *Biophys. J.* **81**, 1600-1612 (2001).

17. Schmies,G. *et al.* Sensory rhodopsin II from the haloalkaliphilic natronobacterium pharaonis: light-activated proton transfer reactions. *Biophys. J.* **78**, 967-976 (2000).

18. Yan,B., Takahashi,T., Johnson,R. & Spudich,J.L. Identification of signaling states of a sensory receptor by modulation of lifetimes of stimulus-induced conformations: the case of sensory rhodopsin II. *Biochemistry* **30**, 10686-10692 (1991).

19. Wegener,A.A., Chizhov,I., Engelhard,M. & Steinhoff,H.J. Time-resolved detection of transient movement of helix F in spin- labelled pharaonis sensory rhodopsin II. *J. Mol. Biol.* **301**, 881-891 (2000).

20. Wegener,A.A., Klare,J.P., Engelhard,M. & Steinhoff,H.J. Structural insights into the early steps of receptor-transducer signal transfer in archaeal phototaxis. *EMBO J.* **20**, 5312-5319 (2001).

21. Klare,J.P. *et al.* The archaeal sensory rhodopsin II/transducer complex: a model for transmembrane signal transfer. *FEBS Lett.* **564**, 219-224 (2004).

22. Yao,V.J. & Spudich,J.L. Primary structure of an archaebacterial transducer, a methyl-accepting protein associated with sensory rhodopsin I. *Proc. Natl. Acad. Sci. U. S. A* **89**, 11915-11919 (1992).

23. Seidel,R. *et al.* The primary structure of sensory rhodopsin II: a member of an additional retinal protein subgroup is coexpressed with its transducer, the halobacterial transducer of rhodopsin II. *Proc. Natl. Acad. Sci. U. S. A* **92**, 3036-3040 (1995).

24. Zhang,W., Brooun,A., McCandless,J., Banda,P. & Alam,M. Signal transduction in the archaeon Halobacterium salinarium is processed through three subfamilies of 13 soluble and membrane-bound transducer proteins. *Proc. Natl. Acad. Sci. U. S. A* **93**, 4649-4654 (1996).

25. Rudolph,J. *et al.* A family of halobacterial transducer proteins. *FEMS Microbiol. Lett.* **139**, 161-168 (1996).

26. Rudolph,J. & Oesterhelt,D. Deletion analysis of the che operon in the archaeon Halobacterium salinarium. *J. Mol. Biol* **258**, 548-554 (1996).

27. Scharf,B.E., Fahrner,K.A., Turner,L. & Berg,H.C. Control of direction of flagellar rotation in bacterial chemotaxis. *Proc. Natl. Acad. Sci. U. S. A* **95**, 201-206 (1998).

28. Chervitz,S.A. & Falke,J.J. Molecular mechanism of transmembrane signaling by the aspartate receptor: a model. *Proc. Natl. Acad. Sci. U. S. A* **93**, 2545-2550 (1996).

29. Ottemann,K.M., Xiao,W., Shin,Y.K. & Koshland,D.E., Jr. A piston model for transmembrane signaling of the aspartate receptor. *Science* **285**, 1751-1754 (1999).

30. Ottemann,K.M., Thorgeirsson,T.E., Kolodziej,A.F., Shin,Y.K. & Koshland,D.E., Jr. Direct measurement of small ligand-induced conformational changes in the aspartate chemoreceptor using EPR. *Biochemistry* **37**, 7062-7069 (1998).

31. Le Moual,H. & Koshland,D.E., Jr. Molecular evolution of the C-terminal cytoplasmic domain of a superfamily of bacterial receptors involved in taxis. *J. Mol. Biol* **261**, 568-585 (1996).

32. Butler,S.L. & Falke,J.J. Cysteine and disulfide scanning reveals two amphiphilic helices in the linker region of the aspartate chemoreceptor. *Biochemistry* **37**, 10746-10756 (1998).

33. Weis,R.M. *et al.* Electron microscopic analysis of membrane assemblies formed by the bacterial chemotaxis receptor Tsr. *J. Bacteriol.* **185**, 3636-3643 (2003).

34. Kim,K.K., Yokota,H. & Kim,S.H. Four-helical-bundle structure of the cytoplasmic domain of a serine chemotaxis receptor. *Nature* **400**, 787-792 (1999).

35. Alley,M.R., Gomes,S.L., Alexander,W. & Shapiro,L. Genetic analysis of a temporally transcribed chemotaxis gene cluster in Caulobacter crescentus. *Genetics* **129**, 333-341 (1991).

36. Luecke,H., Schobert,B., Lanyi,J.K., Spudich,E.N. & Spudich,J.L. Crystal structure of sensory rhodopsin II at 2.4 angstroms: insights into color tuning and transducer interaction. *Science* **293**, 1499-1503 (2001).

37. Royant,A. *et al.* X-ray structure of sensory rhodopsin II at 2.1-A resolution. *Proc. Natl. Acad. Sci. U. S. A* **98**, 10131-10136 (2001).

38. Trivedi,V.D. & Spudich,J.L. Photostimulation of a sensory rhodopsin II/HtrII/Tsr fusion chimera activates CheA-autophosphorylation and CheY-phosphotransfer in vitro. *Biochemistry* **42**, 13887-13892 (2003).

39. Aravind,L. & Ponting,C.P. The cytoplasmic helical linker domain of receptor histidine kinase and methyl-accepting proteins is common to many prokaryotic signalling proteins. *FEMS Microbiol. Lett.* **176**, 111-116 (1999).

40. Zhu,Y. & Inouye,M. The HAMP linker in histidine kinase dimeric receptors is critical for symmetric transmembrane signal transduction. *J. Biol Chem.* **279**, 48152-48158 (2004).

41. Bullock,W., Fernandez,J. & Short,J. XL1-Blue: a high efficiency plasmid transforming *recA Escherichia coli* strain with beta-galactosidase selection. *Bio Tech* **5**, 376-378 (1987).

42. Studier,F.W. & Moffatt,B.A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol* **189**, 113-130 (1986).

43. Alexandrov,K. *et al.* Characterization of the ternary complex between Rab7, REP-1 and Rab geranylgeranyl transferase. *Eur. J. Biochem.* **265**, 160-170 (1999).

44. Brinkmann,U., Mattes,R.E. & Buckel,P. High-level expression of recombinant genes in Escherichia coli is dependent on the availability of the dnaY gene product. *Gene* **85**, 109-114 (1989).

45. Southworth,M.W., Amaya,K., Evans,T.C., Xu,M.Q. & Perler,F.B. Purification of proteins fused to either the amino or carboxy terminus of the Mycobacterium xenopi gyrase A intein. *Biotechniques* **27**, 110-120 (1999).

46. Merrifield,B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **85**, 2149-2154 (1963).

47. Schnolzer,M., Alewood,P., Jones,A., Alewood,D. & Kent,S.B. In situ neutralization in Boc-chemistry solid phase peptide synthesis. Rapid, high yield assembly of difficult sequences. *Int. J. Pept. Protein Res.* **40**, 180-193 (1992).

48. Sarin,V.K., Kent,S.B., Tam,J.P. & Merrifield,R.B. Quantitative monitoring of solid-phase peptide synthesis by the ninhydrin reaction. *Anal. Biochem.* **117**, 147-157 (1981).

49. Miranda,L.P., Alun Jones,A., Meutermans,W.D.F. & Alewood,P.F. p-Cresol As a Reversible Acylium Ion Scavenger in Solid-Phase Peptide Synthesis. *J. Am. Chem. Soc.* **120**, 1410-1420 (1998).

50. Mullis,K.B. & Faloona,F.A. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol.* **155**, 335-350 (1987).

51. Suggs,S.V., Wallace,R.B., Hirose,T., Kawashima,E.H. & Itakura,K. Use of synthetic oligonucleotides as hybridization probes: isolation of cloned cDNA sequences for human beta 2-microglobulin. *Proc. Natl. Acad. Sci. U. S. A* **78**, 6613-6617 (1981).

52. Dower,W.J., Miller,J.F. & Ragsdale,C.W. High efficiency transformation of E. coli by high voltage electroporation. *Nucleic Acids Res.* **16**, 6127-6145 (1988).

53. Gasteiger,E. *et al.* The Proteomics Protocols Handbook. Walker,J.M. (ed.), pp. 571-607 (Humana Press,2005).

54. Hohenfeld,I.P., Wegener,A.A. & Engelhard,M. Purification of histidine tagged bacteriorhodopsin, pharaonis halorhodopsin and pharaonis sensory rhodopsin II functionally expressed in Escherichia coli. *FEBS Lett.* **442**, 198-202 (1999).

55. Ehresmann,B., Imbault,P. & Weil,J.H. Spectrophotometric determination of protein concentration in cell extracts containing tRNA's and rRNA's. *Anal. Biochem.* **54**, 454-463 (1973).

56. Kates,M., Kushawa,S.C. & Sprott,G.D. Lipids of Purple Membrane from Extreme Halophiles and of Methanogenic Bacteria. *Methods in Enzymology* **88**, 98-111 (1982).

57. Schagger,H. & von Jagow,G. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**, 368-379 (1987).

58. Cadene,M. & Chait,B.T. A robust, detergent-friendly method for mass spectrometric analysis of integral membrane proteins. *Anal. Chem.* **72**, 5655-5658 (2000).

59. Hufnagel,P., Schweiger,U., Eckerskorn,C. & Oesterhelt,D. Electrospray ionization mass spectrometry of genetically and chemically modified bacteriorhodopsins. *Anal. Biochem.* **243**, 46-54 (1996).

60. Hippler-Mreyen,S. *et al.* Probing the sensory rhodopsin II binding domain of its cognate transducer by calorimetry and electrophysiology. *J. Mol. Biol* **330**, 1203-1213 (2003).

61. Hackeng,T.M., Griffin,J.H. & Dawson,P.E. Protein synthesis by native chemical ligation: expanded scope by using straightforward methodology. *Proc. Natl. Acad. Sci. U. S. A* **96**, 10068-10073 (1999).

62. Hippler-Mreyen,S. Funktionelle Untersuchungen der Wechselwirkung des archaebakteriellen Photorezeptors NpSRII mit seinem Transducer NpHtrII durch heterologe Expression in E.coli. 1-133. 2003. Universität Düsseldorf. Ref Type: Thesis/Dissertation

63. Schagger,H. & von Jagow,G. Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal. Biochem.* **199**, 223-231 (1991).

64. Schamel,W.W.A. & Schamel,K.R. Die Blaue Native Gelelektrophorese zur Untersuchung von Proteinkomplexen unter nativen Bedingungen. *Immunologie Aktuell* **2**, 33-39 (2001).

65. Geiger,T. & Clarke,S. Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. *J. Biol Chem.* **262**, 785-794 (1987).

66. Stephenson,R.C. & Clarke,S. Succinimide formation from aspartyl and asparaginyl peptides as a model for the spontaneous degradation of proteins. *J. Biol Chem.* **264**, 6164-6170 (1989).

67. Sudo,Y. *et al.* Linker Region of a Halobacterial Transducer Protein Interacts Directly with Its Sensor Retinal Protein. *Biochemistry* **44**, 6144-6152 (2005).

68. Fink,A.L. Natively unfolded proteins. *Curr. Opin. Struct. Biol* **15**, 35-41 (2005).

69. Dyson,H.J. & Wright,P.E. Unfolded proteins and protein folding studied by NMR. *Chem. Rev.* **104**, 3607-3622 (2004).

70. Schnolzer,M., Jones,A., Alewood,P.F. & Kent,S.B. Ion-spray tandem mass spectrometry in peptide synthesis: structural characterization of minor by-products in the synthesis of ACP(65-74). *Anal. Biochem.* **204**, 335-343 (1992).

71. Houghten,R.A. & Li,C.H. Reduction of sulfoxides in peptides and proteins. *Methods Enzymol.* **91**, 549-559 (1983).

72. Goto,N.K. & Kay,L.E. New developments in isotope labeling strategies for protein solution NMR spectroscopy. *Curr. Opin. Struct. Biol* **10**, 585-592 (2000).

73. McIntosh,L.P. & Dahlquist,F.W. Biosynthetic incorporation of 15N and 13C for assignment and interpretation of nuclear magnetic resonance spectra of proteins. *Q. Rev. Biophys.* **23**, 1-38 (1990).

74. Tolbert,T.J. & Wong,C.-H. New Methods for Proteomic Research: Preparation of Proteins with N-Terminal Cysteines for Labeling and Conjugation. *Angew. Chem. Int. Ed.* **41**, 2171-2174 (2002).

75. Kapust,R.B., Tozser,J., Copeland,T.D. & Waugh,D.S. The P1' specificity of tobacco etch virus protease. *Biochem. Biophys. Res. Commun.* **294**, 949-955 (2002).

76. Becker,C.F. Chemische Synthese von Proteinen: Generierung biologischer Aktivität ohne Nutzung des zellulären Syntheseapparats. 1-147. 2001. Universität Dortmund.
Ref Type: Thesis/Dissertation

77. Weimbs,T. & Stoffel,W. Proteolipid protein (PLP) of CNS myelin: positions of free, disulfide-bonded, and fatty acid thioester-linked cysteine residues and implications for the membrane topology of PLP. *Biochemistry* **31**, 12289-12296 (1992).

78. Ben Bassat,A. *et al.* Processing of the initiation methionine from proteins: properties of the Escherichia coli methionine aminopeptidase and its gene structure. *J. Bacteriol.* **169**, 751-757 (1987).

79. Stark,G.R. Modification of Proteins with Cyanate. *Methods Enzymol.* **11**, 591-594 (1967).

80. Wan,M. & Ropp,P. Reagents for protection of peptide/proteins carbamylation in urea solutions utilizing non-ethylene-diamine like compounds. *United States Patent Application* **20040166572**, 1-12 (2004).

81. Ayers,B. *et al.* Introduction of unnatural amino acids into proteins using expressed protein ligation. *Biopolymers* **51**, 343-354 (1999).

82. Perler,F.B. & Adam,E. Protein splicing and its applications. *Curr. Opin. Biotechnol.* **11**, 377-383 (2000).

83. Southworth,M.W., Benner,J. & Perler,F.B. An alternative protein splicing mechanism for inteins lacking an N-terminal nucleophile. *EMBO J.* **19**, 5019-5026 (2000).

84. Xu,M.Q. & Perler,F.B. The mechanism of protein splicing and its modulation by mutation. *EMBO J.* **15**, 5146-5153 (1996).

85. Xu,M.Q. *et al.* Protein splicing: an analysis of the branched intermediate and its resolution by succinimide formation. *EMBO J.* **13**, 5517-5522 (1994).

86. Dawson,P.E. & Kent,S.B. Synthesis of native proteins by chemical ligation. *Annu. Rev. Biochem.* **69**, 923-960 (2000).

87. Dawson,P.E., Muir,T.W., Clark-Lewis,I. & Kent,S.B. Synthesis of proteins by native chemical ligation. *Science* **266**, 776-779 (1994).

88. Kochendoerfer,G.G. *et al.* Total chemical synthesis of the integral membrane protein influenza A virus M2: role of its C-terminal domain in tetramer assembly. *Biochemistry* **38**, 11905-11913 (1999).

89. Kochendoerfer,G.G. & Kent,S.B. Chemical protein synthesis. *Curr. Opin. Chem. Biol* **3**, 665-671 (1999).

90. Low,D.W., Hill,M.G., Carrasco,M.R., Kent,S.B. & Botti,P. Total synthesis of cytochrome b562 by native chemical ligation using a removable auxiliary. *Proc. Natl. Acad. Sci. U. S. A* **98**, 6554-6559 (2001).

91. Hunter,C.L. & Kochendoerfer,G.G. Native chemical ligation of hydropobic peptides in lipid bilayer systems. *Bioconjug. Chem.* **15**, 437-440 (2004).

92. Gorzelle,B.M. *et al.* Reconstitutive refolding of diacylglycerol kinase, an integral membrane protein. *Biochemistry* **38**, 16373-16382 (1999).

93. Wishart,D.S. & Sykes,B.D. The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data. *J. Biomol. NMR* **4**, 171-180 (1994).

94. Wishart,D.S., Sykes,B.D. & Richards,F.M. The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry* **31**, 1647-1651 (1992).

95. Wuthrich,K. NMR of Proteins and Nucleic Acids. *John Wiley & Sons, Inc. , New York* (1986).

96. Schmies,G., Engelhard,M., Wood,P.G., Nagel,G. & Bamberg,E. Electrophysiological characterization of specific interactions between bacterial sensory rhodopsins and their transducers. *Proc. Natl. Acad. Sci. U. S. A* **98**, 1555-1559 (2001).

97. Sudo,Y., Iwamoto,M., Shimono,K. & Kamo,N. Pharaonis phoborhodopsin binds to its cognate truncated transducer even in the presence of a detergent with a 1:1 stoichiometry. *Photochem. Photobiol.* **74**, 489-494 (2001).

98. Vogel,R. & Siebert,F. Conformations of the active and inactive states of opsin. *J. Biol Chem.* **276**, 38487-38493 (2001).

99. Engelhard,M., Scharf,B. & Siebert,F. Protonation changes during the photocycle of sensory rhodopsin II from Natronobacterium pharaonis. *FEBS Lett.* **395**, 195-198 (1996).

100. Furutani,Y. *et al.* Structural Changes of the Complex between pharaonis Phoborhodopsin and Its Cognate Transducer upon Formation of the M Photointermediate. *Biochemistry* **44**, 2909-2915 (2005).

101. Luca,S., Heise,H. & Baldus,M. High-resolution solid-state NMR applied to polypeptides and membrane proteins. *Acc. Chem. Res* **36**, 858-865 (2003).

102. Yamazaki,T. *et al.* Segmental Isotope Labeling for Protein NMR Using Peptide Splicing. *J. Am. Chem. Soc.* **120**, 5591-5592 (1998).

103. Otomo,T., Ito,N., Kyogoku,Y. & Yamazaki,T. NMR observation of selected segments in a larger protein: central-segment isotope labeling through intein-mediated ligation. *Biochemistry* **38**, 16040-16044 (1999).

104. Otomo,T., Teruya,K., Uegaki,K., Yamazaki,T. & Kyogoku,Y. Improved segmental isotope labeling of proteins and application to a larger protein. *J. Biomol. NMR* **14**, 105-114 (1999).

105. Blaschke,U.K., Cotton,G.J. & Muir,T.W. Synthesis of Multi-Domain Proteins Using Expressed Protein Ligation: Strategies for Segmental Isotopic Labeling of Internal Regions. *Tetrahedron* **56**, 9461-9470 (2000).

106. Xu,R., Ayers,B., Cowburn,D. & Muir,T.W. Chemical ligation of folded recombinant proteins: segmental isotopic labeling of domains for NMR studies. *Proc. Natl. Acad. Sci. U. S. A* **96**, 388-393 (1999).

107. Allen,S.J., Curran,A.R., Templer,R.H., Meijberg,W. & Booth,P.J. Folding kinetics of an alpha helical membrane protein in phospholipid bilayer vesicles. *J. Mol. Biol* **342**, 1279-1291 (2004).

108. Booth,P.J. & Curran,A.R. Membrane protein folding. *Curr. Opin. Struct. Biol.* **9**, 115-121 (1999).

109. Cortes,D.M. & Perozo,E. Structural dynamics of the Streptomyces lividans K+ channel (SKC1): oligomeric stoichiometry and stability. *Biochemistry* **36**, 10343-10352 (1997).

110. Muir,T.W. Semisynthesis of proteins by expressed protein ligation. *Annu. Rev. Biochem.* **72**, 249-289 (2003).

111. Tolbert,T.J. & Wong,C.-H. Intein-Mediated Synthesis of Proteins Containing Carbohydrates and Other Molecular Probes. *J. Am. Chem. Soc.* **122**, 5421-5428 (2000).

112. Valiyaveetil,F.I., MacKinnon,R. & Muir,T.W. Semisynthesis and folding of the potassium channel KcsA. *J. Am. Chem. Soc.* **124**, 9113-9120 (2002).

113. Parks,T.D., Howard,E.D., Wolpert,T.J., Arp,D.J. & Dougherty,W.G. Expression and purification of a recombinant tobacco etch virus NIa proteinase: biochemical analyses of the full-length and a naturally occurring truncated proteinase form. *Virology* **210**, 194-201 (1995).

114. Evans,T.C., Jr., Benner,J. & Xu,M.Q. The cyclization and polymerization of bacterially expressed proteins using modified self-splicing inteins. *J. Biol Chem.* **274**, 18359-18363 (1999).

115. Smith,D.B. & Johnson,K.S. Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. *Gene* **67**, 31-40 (1988).

116. Telenti,A. *et al.* The Mycobacterium xenopi GyrA protein splicing element: characterization of a minimal intein. *J. Bacteriol.* **179**, 6378-6382 (1997).

117. Yang,C.S. & Spudich,J.L. Light-induced structural changes occur in the transmembrane helices of the Natronobacterium pharaonis HtrII transducer. *Biochemistry* **40**, 14207-14214 (2001).

118. Osborne,M.J., Su,Z., Sridaran,V. & Ni,F. Efficient expression of isotopically labeled peptides for high resolution NMR studies: application to the Cdc42/Rac binding domains of virulent kinases in Candida albicans. *J. Biomol. NMR* **26**, 317-326 (2003).

119. Sprules,T., Green,N., Featherstone,M. & Gehring,K. Lock and key binding of the HOX YPWM peptide to the PBX homeodomain. *J. Biol Chem.* **278**, 1053-1058 (2003).

120. Koenig,B.W., Rogowski,M. & Louis,J.M. A rapid method to attain isotope labeled small soluble peptides for NMR studies. *J. Biomol. NMR* **26**, 193-202 (2003).

121. Majerle,A., Kidric,J. & Jerala,R. Production of stable isotope enriched antimicrobial peptides in Escherichia coli: an application to the production of a 15N-enriched fragment of lactoferrin. *J. Biomol. NMR* **18**, 145-151 (2000).

122. Tugarinov,V. *et al.* NMR structure of an anti-gp120 antibody complex with a V3 peptide reveals a surface important for co-receptor binding. *Structure. Fold. Des* **8**, 385-395 (2000).

123. Sharon,M., Gorlach,M., Levy,R., Hayek,Y. & Anglister,J. Expression, purification, and isotope labeling of a gp120 V3 peptide and production of a Fab from a HIV-1 neutralizing antibody for NMR studies. *Protein Expr. Purif.* **24**, 374-383 (2002).

124. Kuliopulos,A. & Walsh,C.T. Production, Purification and Cleavage of Tandem Repeats of Recombinant Peptides. *J. Am. Chem. Soc.* **116**, 4599-4607 (1994).

125. Takasuga,A. *et al.* Efficient production of a small peptide by expression as a multimeric form fused with the dihydrofolate reductase affinity handle. *J. Biochem. (Tokyo)* **112**, 652-657 (1992).

126. Klabunde,T., Sharma,S., Telenti,A., Jacobs,W.R., Jr. & Sacchettini,J.C. Crystal structure of GyrA intein from Mycobacterium xenopi reveals structural basis of protein splicing. *Nat. Struct. Biol* **5**, 31-36 (1998).

127. Siebold,C. & Erni,B. Intein-mediated cyclization of a soluble and a membrane protein in vivo: function and stability. *Biophys. Chem.* **96**, 163-171 (2002).

128. Kawasaki,M. *et al.* Folding-dependent in vitro protein splicing of the Saccharomyces cerevisiae VMA1 protozyme. *Biochem. Biophys. Res. Commun.* **222**, 827-832 (1996).

129. Sydor,J.R., Mariano,M., Sideris,S. & Nock,S. Establishment of Intein-Mediated Protein Ligation under Denaturing Conditions: C-Terminal Labeling of a Single-Chain Antibody for Biochip Screening. *Bioconjug. Chem.* **13**, 707-712 (2002).

130. Blaschke,U.K., Silberstein,J. & Muir,T.W. Protein engineering by expressed protein ligation. *Methods Enzymol.* **328**, 478-496 (2000).

131. Hofmann,R.M. & Muir,T.W. Recent advances in the application of expressed protein ligation to protein engineering. *Curr. Opin. Biotechnol.* **13**, 297-303 (2002).

132. Muir,T.W., Sondhi,D. & Cole,P.A. Expressed protein ligation: a general method for protein engineering. *Proc. Natl. Acad. Sci. U. S. A* **95**, 6705-6710 (1998).

133. Beligere,G.S. & Dawson,P.E. Conformationally Assisted Protein Ligation Using C-Terminal Thioester Peptides. *J. Am. Chem. Soc.* **121**, 6332-6333 (1999).

134. Lindblom,G. & Rilfors,L. Cubic phases and isotropic structures formed by membrane lipids — possible biological relevance. *Biochim. Biophys. Acta* **988**, 221-256 (1989).

135. David,R., Richter,M.P. & Beck-Sickinger,A.G. Expressed protein ligation. Method and applications. *Eur. J. Biochem.* **271**, 663-677 (2004).

136. Popot,J.L. & Engelman,D.M. Membrane protein folding and oligomerization: the two-stage model. *Biochemistry* **29**, 4031-4037 (1990).

137. Booth,P.J. Folding alpha-helical membrane proteins: kinetic studies on bacteriorhodopsin. *Fold. Des* **2**, R85-R92 (1997).

138. Riley,M.L., Wallace,B.A., Flitsch,S.L. & Booth,P.J. Slow alpha helix formation during folding of a membrane protein. *Biochemistry* **36**, 192-196 (1997).

139. Booth,P.J. *et al.* Evidence that bilayer bending rigidity affects membrane protein folding. *Biochemistry* **36**, 197-203 (1997).

140. Bogdanov,M. & Dowhan,W. Lipid-assisted protein folding. *J. Biol Chem.* **274**, 36827-36830 (1999).

141. Hein,M., Wegener,A.A., Engelhard,M. & Siebert,F. Time-resolved FTIR studies of sensory rhodopsin II (NpSRII) from Natronobacterium pharaonis:

implications for proton transport and receptor activation. *Biophys. J.* **84**, 1208-1217 (2003).

142. Furutani,Y., Sudo,Y., Kamo,N. & Kandori,H. FTIR spectroscopy of the complex between pharaonis phoborhodopsin and its transducer protein. *Biochemistry* **42**, 4837-4842 (2003).

143. Yamaguchi,S. *et al.* Conformation and dynamics of the [3-(13)C]Ala, [1-(13)C]Val-labeled truncated pharaonis transducer, pHtrII(1-159), as revealed by site-directed (13)C solid-state NMR: changes due to association with phoborhodopsin (sensory rhodopsin II). *Biophys. J.* **86**, 3131-3140 (2004).

144. Cornish,V.W., Kaplan,M.I., Veenstra,D.L., Kollman,P.A. & Schultz,P.G. Stabilizing and destabilizing effects of placing beta-branched amino acids in protein alpha-helices. *Biochemistry* **33**, 12022-12031 (1994).

145. Zahn,H. [Structure and synthesis of insulin]. *Med. Klin.* **61**, 876 (1966).

146. Merrifield,B. Solid phase synthesis. Nobel lecture, 8 December 1984. *Biosci. Rep.* **5**, 353-376 (1985).

147. Sakakibara,S. Chemical synthesis of proteins in solution. *Biopolymers* **51**, 279-296 (1999).

148. Cooper,A.A. & Stevens,T.H. Protein splicing: self-splicing of genetically mobile elements at the protein level. *Trends Biochem. Sci.* **20**, 351-356 (1995).

149. Xu,M.Q., Southworth,M.W., Mersha,F.B., Hornstra,L.J. & Perler,F.B. In vitro protein splicing of purified precursor and the identification of a branched intermediate. *Cell* **75**, 1371-1377 (1993).

150. Ernst,R.R. Sensitivity enhancement in magnetic resonance. *Adv. Magn. Reson.* **2**, 1-135 (1966).

151. Aue,W.P., Bartholdi,E. & Ernst,R.R. Two-dimensional spectroscopy. Application to nuclear magnetic resonance. *J. Chem. Phys.* **64**, 2229-2246 (1976).

152. Raschke,T.M. & Marqusee,S. Hydrogen exchange studies of protein structure. *Curr. Opin. Biotechnol.* **9**, 80-86 (1998).

153. Eaton,W.A. *et al.* Fast kinetics and mechanisms in protein folding. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 327-359 (2000).

154. Kumar,T.K. & Yu,C. Monitoring protein folding at atomic resolution. *Acc. Chem. Res* **37**, 929-936 (2004).

155. Shulman,R.G. *et al.* The absence of "heme-heme" interactions in hemoglobin. *Science* **165**, 251-257 (1969).

156. Pervushin,K., Riek,R., Wider,G. & Wuthrich,K. Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. U. S. A* **94**, 12366-12371 (1997).

157. Gardner,K.H. & Kay,L.E. The use of 2H, 13C, 15N multidimensional NMR to study the structure and dynamics of proteins. *Annu. Rev. Biophys. Biomol. Struct.* **27**, 357-406 (1998).

158. Fernandez,C., Hilty,C., Wider,G., Guntert,P. & Wuthrich,K. NMR structure of the integral membrane protein OmpX. *J. Mol. Biol* **336**, 1211-1221 (2004).

159. Seddon,A.M., Curnow,P. & Booth,P.J. Membrane proteins, lipids and detergents: not just a soap opera. *Biochim. Biophys. Acta* **1666**, 105-117 (2004).

160. Castellani,F. *et al.* Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature* **420**, 98-102 (2002).

161. Pauli,J., Baldus,M., van Rossum,B., de Groot,H. & Oschkinat,H. Backbone and side-chain 13C and 15N signal assignments of the alpha-spectrin SH3 domain by magic angle spinning solid-state NMR at 17.6 Tesla. *Chembiochem.* **2**, 272-281 (2001).

162. Jaroniec,C.P. *et al.* Measurement of dipolar couplings in a uniformly (13)C,(15)N-labeled membrane protein: distances between the Schiff base and aspartic acids in the active site of bacteriorhodopsin. *J. Am. Chem. Soc.* **123**, 12929-12930 (2001).

8. Acknowledgements

I am taking this opportunity to thank all those who have assisted me in one way or the other with my Ph.D. project. First, I would like to express my gratitude to my supervisors Prof. Dr. Roger Goody and Prof. Dr. Martin Engelhard. They have always been extremely generous with their time, knowledge and ideas and allowed me great freedom in my research.

I am also grateful to all who have been part of the Engelhard research group. I have enjoyed working with skilful colleagues and friends, in a very inspiring and helpful atmosphere. I am particularly grateful to Annika Göppner for preparing the many proteins I consumed throughout my studies and to Dr. Ralf Seidel, Dr. Matthias Geyer and Dr. Christian Becker for their great contribution to the structural part of this thesis.

I would like to thank Prof. Dr. Fritz Siebert, Ionela Radu (University Freiburg, Germany) and Dr. Marc Baldus and Karsten Seidel (MPI Göttingen, Germany) who have collaborated with me and contributed to various parts of this research.

Thanks are also due to Dr. Petra Janning for fruitful discussions on the mass spectroscopic analysis.

I would like to thank the many people from the Goody department who have contributed to an extraordinarily friendly and helpful working environment and made working on this thesis a fun and pleasurable experience.

I would also like to thank Dr. Christian Becker, Dr. Johann Klare and Ekta (Nici) Ahuja for their constructive comments on the manuscript.

Last but not least, I am greatly indebted to my fiancé Sarah, to my family and friends for their understanding, patience and support during the entire period of my study. _