

6 Summarizing discussion

It was anticipated that the research work carried out in this thesis would lead us close to the broader goal of elucidating the bound structure of human ET-1 whilst bound to the human ET_BR. Efforts were made to establish a large-scale production system for ET_BR in insect Sf9 cells coupled with the AcNPV expression system. Stable isotope labelling of the ET-1 ligand has been achieved. The binding of the uniformly labelled ET-1 with the ET_BR was probed by means of solid state NMR through correlation spectroscopy and preliminary assignment of isotropic chemical shifts.

Based on the findings in the present study, the following was achieved: The results will lead us to the structure determination of ET-1 in the bound state.

This chapter is intended to complement the discussion sections within the previous chapters by pulling together the main findings from each chapter. It will also set out plans for future work.

6.1 Future directions in large-scale production of recombinant ET_BR using insect cultures

GPCRs, as with other membrane proteins, have very low natural levels of expression, resulting in little opportunity for structural studies of these proteins, as well as ensuring that the cost of such studies remain affordable. In the light of the emerging need for milligrams of ET_BR from insect cell culture, some future directions in receptor production can be outlined. For production of the ET_BR we started with the baculovirus infected Sf9 cells system because in our view, it is the most convenient system. The expression levels obtained in shaking cultures were not reasonable for the purpose of structural studies which would have meant labour-intensive work and high cost in production and scale up (100 € of cost incurred in production of 1 l), we therefore changed to the *P. pastoris* shaking culture and production process system established for ET_BR by Schiller *et al.* (2001, 2002). Examples demonstrated that the *Pichia* expression system can be used for the production of membrane proteins from a wide variety of organisms, to varying degrees of success. This system seemed to be a very promising expression system for the ET_B receptor, because as unusual with the majority of membrane proteins, it had been successful to obtain large enough quantities of the ET_BR in order to perform structural studies. However, the current challenge in using this expression system is the recovery of hydrophobic GPCR proteins. Many membrane proteins have

specific lipid and sterol requirements, and these lipids can be involved in the folding and stability of the protein as well as their function (Opekaroá & Tanner, 2003). The lipid content of various host cells is presented by Opekaroá & Tanner, who suggest that one of the obstacles for producing mammalian membrane protein in *P. pastoris* could be a shortage of cholesterol. *Pichia* cells might replace this with ergosterol, the main fungal sterol. However, this could affect the functionality of the membrane protein.

Because the receptor isolated from *P. pastoris* membrane was aggregation-prone which may be due to the specific lipid requirements, we return to the baculovirus expression system for receptor production. The baculovirus expression system is maturing as an industrially useful system for recombinant protein. This time, scale-up cultivation of the shear-sensitive Sf9 cells from shake flask to wave bioreactor culture (Singh, 1999) was performed (A. Srivastava, personal communication). This system consists of a disposable plastic Cellbag (100 l) placed on a rocking platform. The platform rocking motion generates waves, which in turn ensure oxygen transfer, good nutrient distribution and suspension of cells, resulting in a perfect environment for cell growth that can easily support over 20×10^6 cells/ml. Commercial experiments with this cellbag technology gave approximately 14 mg of purified functional ET_BR from a 100 l culture (A. Srivastava, personal communication), making this approach promising. This amount of receptor is sufficient for three NMR samples and to obtain ca. 4 mg of receptor, 30 l culture are required at current optimal conditions. The limitation of shake flask cultures (assuming the same expression level as in fermentation) lies in the minimal volume that can be achieved, which currently is 0.5 l. Therefore on a per sample basis, the time needed is prohibitively long.

In order to fulfill this project, significant technical challenges must be overcome, the most substantial being facile protein overexpression. The baculoviruses as vectors are exploited for heterologous protein expression using suspension insect cell cultures. The suspended cell system is convenient and relatively simple at bench scale, but can be problematic upon scale-up. For example, oxygen becomes limiting at large-scale and contamination problems can be more frequent and more costly. It is anticipated that scale-up of these processes to fermenter vessels will significantly improve the production level and thus make their use for structural studies more financially and temporally viable.

Is *P. pastoris* a same good expression system as the baculovirus/insect cells system? It is difficult to draw a conclusion when there is so little understanding available of how *Pichia*-derived ET_BR differ from native protein and what these differences mean in terms of the use of the recombinant product. Yeast is not the panacea for expression problems. For the case in

which the yeast system do not suffice, it will prove necessary to move to more complex cell types.

Arguably, mammalian cells are most likely to preserve structural and functional integrity when expressing mammalian proteins. The expression levels for GPCRs using the Semliki Forest virus (SFV) system have been promising (e. g. 20 pmol/mg ET_AR/ET_BR, Lundström *et al.*, 2003; Sarramegna *et al.*, 2003). Expression of GPCRs in mammalian cells means that the recombinant protein environment is closest to the native one. Such cells have the following advantages. The lipid composition is close to native and the cells can perform all types of complex post-translational modifications required. However, production and scale up in case of the SFV system is expensive (600 € of cost incurred in production of 1 l). This is due to higher cost for the reagent used in the transcription step, which needs to be performed in large scale for bigger batches. Mammalian hosts have limited tolerance to some difficult proteins particularly membrane proteins. Its major problem therefore is poor reliability. The baculoviral expression system should be pursued further because its overall cost is lower and the number of steps between generation of construct to protein production is fewer. Moreover, the total amount of protein obtained per liter of expression culture in baculovirus is much more than in SFV. In addition, because of the high level of overexpression, some of the expressed proteins are localized in intracellular compartments (Lundström, 2003).

The challenge still is obtaining proteins in sufficient quantity for structural studies. Even commercial efforts (e. g. m-phasys), founded with mission of GPCR specific production as drug targets, still expend considerable efforts to fulfil the tremendous demand for such proteins.

6.2 ET_BR-integrated virus-like particles as an alternative to lipid-reconstituted ET_BR

Luca *et al.* (2002) investigated the interaction of the peptide [¹³C, ¹⁵N]-NT(8-13) (12.3 μg) with detergent-solubilized NTS-1 receptor (0.37 mg) [agonist/functional receptor ratio of 4.5 (M/M)] and observed occurrence of strong detergent and buffer signals and only low NT(8-13) signal intensity as a consequence of concentration limitation. To increase the amount of functional receptor and hence the amount of bound [¹³C, ¹⁵N]-NT(8-13) for solid-state NMR measurements, and to reduce the noise contributions from detergent, the receptor was reconstituted into lipid, which allows for packing of the receptor at higher density into the NMR rotor (2.5 mg of functional receptor). As a result, the amount of [¹³C, ¹⁵N]-NT(8-13) (22 μg) could be significantly increased [agonist/functional receptor ratio of 0.9 (M/M)] .

Ligand interactions with GPCRs can be performed in a number of environments and can be investigated directly on the cell surface or in membrane fragments from cells overexpressing the receptor. However, the analysis is compromised by high background levels of other membrane proteins and low concentration of receptor of interest. Alternatively, the respective receptor has to be purified. This biochemical purification however, requires the exchange of the physiological lipid/lipid-protein environment by a detergent micelle, which may modify the binding properties and stability of the receptor. In an ideal receptor preparation format, a high enrichment of the receptor of interest would be combined with the receptor's integration into its natural lipid/lipid-protein environment. To improve the receptor protein density in the lipid bilayer and to maintain the physiological lipid-protein environment, Evotec OAI established a method that selectively incorporates transmembrane proteins (e. g. GPCRs) into virus-like particles (VLPs). For this purpose, the transmembrane protein of interest is coexpressed in insect cells together with the retroviral budding protein named Gag from Moloney murine leukaemia virus, which leads to viral particle formation and budding at the cell surface. The Gag protein recruits the receptor into the budding VLP. Retroviruses possess a protein capsid enclosed by a lipid bilayer that is derived from the host cell plasma membrane in the budding process. The resulting VLPs are released into the extracellular medium. By these means, the ET_AR was integrated into VLPs (Zemanová *et al.*, 2004). These VLPs would constitute a concentrated source of receptor and thereby show great potential for solid-state NMR measurement, which requires very high concentrated amount of samples. VLP stock suspension with ~150 nM (~100 µl) ET_AR is available (Dr. R. Heilker, Boehringer Ingelheim, personal communication). The applicability of these VLPs to solid-state NMR studies is still to be validated in terms of the ability of higher enrichment of VLPs. The benefits of applicable VLPs are at least the reduction of number of experimental runs needed or the avoidance of solubilization, purification and reconstitution steps and the maintainance of the receptor in a physiological environment. Thus their use can result in great time and cost savings. In the pharmaceutical industry VLP-integrated GPCRs in miniaturized assay formats are used in the screening for novel GPCR-directed drugs (Zemanová *et al.*, 2004).

6.3 Preparation of membrane proteins for solid-state NMR

Solid-state NMR imposes some constraints on membrane protein samples, which must be structurally homogeneous and highly concentrated. Given that the available volume of a standard NMR 4 mm MAS rotor is about 50 μl , and 2-10 mg of a membrane protein has to fit into the MAS rotor, the concentration is a great hurdle, since the membrane protein has to be maintained in a natively folded state. To acquire good quality 1D ^{15}N spectra more than 20 nmol of protein is required, and 2D experiments demand even greater quantities. In the case of proteoliposomes, the difficulty of getting a large quantity of receptor into the rotor is compounded by the presence of lipids. Lorch *et al.* (2005) addressed this concentration issue by preparing lower quality crystals of a membrane protein in the presence of membrane mimetic which allow high concentration and structural homogeneity, and are easier to produce than the high-quality crystals required for X-ray diffraction structure determinations. They found that the crystalline samples produce better resolved spectra than proteoliposomes. Similar improvements upon controlled precipitation or recrystallization have been observed in other globular proteins, e. g. SH3 domain of α -spectrin (Pauli *et al.*, 2000). Another alternative would be the preparation of 2D crystals, using procedures developed for cryoelectron microscopy. Sample preparation conditions for 3D crystallisation are expected to be easier to screen than that for 2D crystals (Hiller *et al.*, 2005), since a plethora of nondiffracting crystals will be produced during the search for a diffracting crystal. If it can be generalized that crystalline preparations of membrane proteins are significant to solid-state NMR, then they represent a significant new approach to solving 3D membrane protein structures.

6.4 Detecting ET-1 at its binding site

Solid-state NMR can be used to selectively observe bound ligands in receptor proteins (Watts, 1999) by using specialized pulse sequences, which allows the investigation of ligands directly at their site of action. This is a requirement for drug design and discovery.

Solution NMR was used to assign the ET-1 signal, especially as a preliminary to a more detailed study by solid-state NMR. To obtain structural constraints from the spectra, signals must be first assigned by using intra- and inter-residues spin connectivities acquired with a suite of solid-state NMR experiments.

Since ET-1 is both ^{13}C - and ^{15}N -labelled, the assignment strategy relies on datasets from ^{13}C - ^{13}C intra-residue correlations, ^{13}C - ^{15}N intra-residue correlations (e. g. correlating N from

residue i with $C\alpha$ from the same residue) and ^{13}C - ^{15}N inter-residue (e. g. correlating N from residue $i+1$ with $C\alpha$ from residue i) correlations. This type of approach was applied to the SH3 domain (Pauli *et al.*, 2001).

It has been demonstrated that our approach (simple spin diffusion experiments combined with selective carbon labelling schemes) is sufficient to determine a moderate-resolution structure of α -spectrin. This approach is likely to be applicable to compact proteins with well-dispersed spectra. However, it is likely that when the structure of bound ET-1 is calculated in such a fashion, the root mean square deviations from loop and termini are largely apart.

Since our applied method requires the production of several isotope-labelled receptor complex variants, its application is of limited use associated with difficulties of producing adequate quantities of isotope labelled receptor complex samples. Such a procedure is time consuming, expensive and hardly practical. Therefore, it is appropriate to develop an approach that permits the determination of the 3D structure with a single uniformly isotope-labelled receptor complex sample.

Further challenges will be to exploit the power of solid state methods to define the ligand structure at its site of action, define the residues at the binding site, as well as the known conformational changes that occur upon binding, such as helix orientational changes in ET_BR , when it become available for study.

6.5 Secondary structure from chemical shifts and conformation of bound ET-1

Chemical shifts obtained by the signal assignments are easily accessible in comparison with internuclear distances. Luca *et al.* (2003) have predicted the backbone structure of neurotensin from the ^{13}C and ^{15}N chemical shifts with the program TALOS. The prediction by TALOS is based on the similarity in chemical shifts and amino acid sequences for selected proteins stored in the BioMagResoBank (BMRB) database. The resulting dihedral angles was used to construct a 3D model of the backbone conformation of receptor-bound NT(8-13). Recent developments of the theoretical study of chemical shift and an increase in the number of NMR protein structures in databases improve the reliability of proteins structure analysis by the chemical shifts. However, structural information obtained from the chemical shifts is limited to backbone dihedral angles. Thus the experimental methods for obtaining the information on side chain conformation and long-range order are necessary for the complete structure determination.

Determinants in the ET-1 C-terminal tail are critical for the interaction with ET_BR. The ET-1 α -helix may be important for ET_BR binding because the binding/activity is sensitive to replacement of residues positioned on the same side of the helix. Simply assuming that the receptor-bound conformation of ET-1 represents the pharmacologically active form, the insight gleaned from the chemical shifts so far suggests that the bound ET-1 structure seen in the crystal structure may not represent the active conformation. In this study, the solid state NMR measurement was carried out for the ET_BR/ET-1 complex in the absence of G-protein. This state represents the initially bound conformation of ET-1 to the ET_BR before association with a G-protein. An interesting question to be answered is whether G-protein association accompanies a conformational change/shift of ET-1.

6.6 Locating the binding pocket

Intramolecular structural details for a bound ligand alone do not give a direct insight into the binding site, although they might help in locating and understanding putative ligand interactions. Having established the bound structure of ET-1, the next question is ‘where does it bind?’. An approach to answering this question is that of ‘chemical shift mapping’. This requires a knowledge of the structure of the receptor (or at least a model based on homology) and also the assignment of at least the backbone resonances of the receptor. This latter requirement is now and in the near future difficult to meet, using ¹³C and ¹⁵N labelling.

Precise identification of the binding site are provided by methods yielding interatomic distances, because they can be used not simply to locate the binding site but also to dock the ligand into a known or modelled receptor structure to obtain a model of the ET_BR/ET-1 complex. With labels in both the binding site and ET-1, geometric triangulation that use distance measurements can reveal structural, as well as electronic and orientational details. Mapping experiments will need to be led by a modelling approach to determine which sites to label.

Ligand recognition presupposes a match between ligand and binding site not only of shape but also of physicochemical interactions such as charge and dipolar interactions. Hydrogen bonds, e. g., are important in this regard. They provide directionality.

NMR-generated structural information can be included in docking programs and molecular dynamics. A computational approach in form of homology modelling that is based on the resolved high-resolution structure of rhodopsin can be used to generate a structural model of ET_BR and utilized for interpretation of the findings. The limitations of molecular modelling,

however, especially regarding the low homology and the inactive state of bovine rhodopsin compared to ET_BR are recognized.

However, the most detailed information can be obtained by determining the full 3D-structure of the complex, which requires essentially complete assignment of the resonances of the receptor.

6.7 Dynamics of the complex

One of the strengths of NMR spectroscopy is its ability to provide a detailed picture of the dynamics of the receptor complex.

Ligands are not so tightly bound as to undergo no molecular motions within their binding sites, since exchange must occur, as well as specific receptor-ligand binding. It cannot be assumed that the rigid structure of a bound ligand is representative of the whole ligand. One part might be mobile, whereas a linked part might be motionally restricted. On the other hand, the ligand binding site could be the mobile part of the structure, and becomes more rigid when ET-1 binds. Molecular motion in the binding site is of fundamental importance for the kinetics of association and dissociation, and for determining the structural range of ligands that can bind. In the case where the residues that are mobile are directly involved in ligand binding, a degree of active site mobility could be required to enable the receptor to bind a range of ligands, and a knowledge of this mobility could be exploited in the design of new structurally diverse ligands.

One solid-state NMR approach that can be used to gain insight into ligand motion is labelling with the ²H nucleus, which gives rise to NMR spectra with sensitivity to motions in the functionally important millisecond to picosecond timescale (exchange rates are in the millisecond range or faster). The line-width changes which occur due to dynamic effects are greater than those caused by other factors, such as inhomogeneous effects. Narrow spectra can imply fast (ms to ps) motion and broad spectra can indicate slow (greater than ms) motion. Fast motion of the ligand while at its binding site in the receptor was shown for an acetylcholine agonist at its target site (Williamson *et al.*, 2001).

6.8 Endothelin as test system for solid state NMR

One motivation for studying small proteins in the solid state is the development of solid-state NMR methodology that would be applicable to membrane proteins. In this context, more

methods are needed for efficient determinations of structure and dynamics of proteins in the solid state. As a new test system for optimizing line width, assignment protocols and probing new pulse sequence methods for structure and dynamics, the 2.5 kDa or 21 residues small ET-1 peptide with mixed secondary structure may be an excellent biological sample. Extensive information about the structure of ET-1 is available from X-ray crystallography and solution NMR studies. Usually, NMR lines of lyophilized ^{13}C (Met7, Val12, Phe14), ^{15}N (Ile19, Ile20) ET-1 sample was inhomogeneously broadened and hydration of lyophilized ET-1 was shown to improve the spectral resolution in ^{13}C CP/MAS spectrum (Veniamin Chevelkov, FMP, personal communication). A prerequisite for any high-resolution structural studies are microcrystalline or precipitated samples. Indeed, ^{13}C (Met7, Val12, Phe14), ^{15}N (Ile19, Ile20) ET-1 was observed to form needle-like crystals under the conditions of 20 mM Tris, pH 7.4 after a few days at 4°C and room temperature (personal observation). An optimized sample crystallization protocol still to be developed would open up possibilities for further structural and dynamical characterization of ET-1 using solid-state NMR methods.

6.9 Concluding remarks

Previous development of ET-1 agonists and antagonists have provided some information on the structural requirement for ET-1 binding and selectivity. However, the nature of the dynamic changes in the receptor that take place in ET-1 binding is unknown. Undoubtedly, the availability of a high-resolution structure of the receptor complex will give information on receptor domains responsible for agonist or antagonist binding and contribute to enhanced understanding of general principles underlying receptor functions at molecular level. The availability of more GPCR structures would allow us to compare these structures bound to various ligands.

The ET system seems to provide an exciting and promising therapeutic target for the development of new treatments for a number of cardiovascular diseases. However, current clinical trials by Lee *et al.* (2003), who conducted a systematic review and meta-analysis of controlled trials of endothelin receptor blockade in animal models of heart failure, after clinical trials in humans had found no evidence of benefit, showed that the animal data provided no evidence of benefit overall and showed a tendency towards increased mortality with early administration.

Therefore, the purpose of this and future work is rather to encourage basic research on the structure of receptor protein at (or near) atomic resolution. Further objectives are to encourage additional research to further develop NMR methods for studying the structure of membrane

proteins at atomic resolution. The large amount of receptor for solid state NMR necessitates an efficient expression system.