Investigating the dynamic nature of the ABC transporters: ABCB1 and MsbA as examples for the potential synergies of MD theory and EPR applications

Thomas Stockner[1], Anna Mullen[2] and Fraser MacMillan[2]

[1] Institute of Pharmacology, Center of Physiology and Pharmacology, Medical University of Vienna, Vienna, Austria
[2] Henry Wellcome Unit for Biological EPR, School of Chemistry, University of East Anglia, Norwich, UK.

Corresponding Authors:
Fraser.MacMillan@uea.ac.uk and thomas.stockner@meduniwien.ac.at

Abstract:
ABC transporters are primary active transporters found in all kingdoms of life. Human multidrug resistance transporter ABCB1, or P-glycoprotein, has an extremely broad substrate spectrum and confers resistance against chemotherapy drug treatment in cancer cells. The bacterial ABC transporter MsbA is a lipid A flippase and a homolog to the human ABCB1 transporter, with which it partially shares its substrate spectrum.
Crystal structures of MsbA and ABCB1 have been solved in multiple conformations, providing a glimpse into the possible conformational changes the transporter could be going through during the transport cycle. Crystal structures are inherently static, while a dynamic picture of the transporter in motion is needed for a complete understanding of transporter function. Molecular dynamics (MD) simulations and electron paramagnetic resonance (EPR) spectroscopy can provide structural information on ABC transporters, but the strength of these two methods lies in the potential to characterise the dynamic regime of these transporters. Information from the two methods is quite complementary. MD simulations provide an all atom dynamic picture of the time evolution of the molecular system, though with a narrow time window. EPR spectroscopy can probe structural, environmental and dynamic properties of the transporter in several time regimes, but only through the attachment sites of an exogenous spin label. In this review the synergistic effects that can be achieved by combining the two methods are highlighted, and a brief methodological background is also presented.

Introduction:
Knowledge of the molecular structure of membrane proteins is essential to understand their function. Most direct structural information is derived from X-ray crystallography as well as Nuclear Magnetic Resonance (NMR) spectroscopy and currently to a lesser extent from electron microscopy. These methods suffer from challenges, which are difficult to overcome when studying proteins in their native cellular environment. All these methods typically require overexpression and purification. Additionally, X-ray crystallography measures protein structures within the confines of a regular assembly of a single crystal. The assembly of proteins into the regular raster of a crystal often comes with the burden that typically only a single conformation is selected. All other conformations may remain unobserved, which are accessible in the membrane environments. The requirements of crystal structure formation also dictates a change in the molecular environment as membrane proteins have to be extracted from the membrane and solubilized in micelle-forming detergents. There are numerous examples in the literature and in the PDB database, which demonstrate that

membrane proteins are especially susceptible to conformational changes due to environmental changes [1, 2]

NMR spectroscopy can measure proteins in solution and macromolecular complexes with interaction partners. Its biggest challenge is the size limitation for structure determination. Until recently proteins and protein complexes above 35 kDa [3] were considered difficult to study using high-resolution solution NMR techniques, although advances in solid state (SS-NMR) are starting to address this problem [4,5].

In addition, these methods quite often only observe static structures making the study of dynamic or flexible molecular machines much more difficult, although NMR can in principle obtain dynamic information. Hence biophysical methods are needed which can provide both structural and dynamic information on the molecular architectures. Both fluorescence-based (e.g. FRET) and electron paramagnetic resonance (EPR) - based techniques are very powerful in this respect, offering reliable and precise distance determination in the range 1.5 - 10 nm. The methods have the potential to be applied at the single molecule level and provide real time dynamics over several time scales. Much of the observed experimental data requires analysis and interpretation, which invariably necessitates input from computational chemists. In this short review we will briefly describe and review the EPR approach to studying the structure and dynamics of ABC transporters as well as molecular dynamics (MD) approaches to structure determination. The future synergies of both methods, especially in combination with classical structure determination methods will be highlighted.

EPR: EPR spectroscopy is a spectroscopic technique that detects unpaired electrons. This magnetic resonance technique has been around for over 70 years (since its first observation by E.K. Zavoiskii in Kazan, in the Tartar republic of the then USSR in 1944). After its development as a fundamental physical spectroscopic technique it quickly developed into a popular biophysical tool focussing especially on intrinsic paramagnetic species contained with electron-transfer proteins and metalloproteins, both soluble and membrane spanning. The method has since developed immensely in quite a similar way to NMR, albeit at a much slower pace due to the higher technical demands imposed by the intrinsic properties of an electron spin as compared to nuclear spins which today results in a broad arsenal of EPR-based methods to study and characterise paramagnetic centres [see 6 for a general review of EPR methods and 7 specifically for biological applications] Together with the recent developments in site-directed spin labelling (SDSL), an EPR technique which allows site-specific protein labelling, this biophysical tool has now expanded its scope for applicable systems to a broad range of diamagnetic proteins including ABC transporters [e.g. 8-11].

Over the past 10-15 years SDSL has become well-suited and widely used to study small soluble proteins, RNA, DNA, membrane proteins and large protein complexes providing information on protein structure, and more recently dynamics and conformation change. This has been made possible primarily by advances in both instrumentation - increases in spectrometer performance and sensitivity [13,14] - and spectral analysis and much more recently by the development of differing types of label molecule (see Figure 1)

When considering using SDSL and EPR to study ABC transporters there are a number of key questions that need addressing such as the ability to make functional variants, the choice of labelling site, the choice of paramagnetic label. The most common stable spin labels are attached to a protein through chemically modifying a cysteine residue with a sulfhydryl reactive nitroxide (Figure 1). These nitroxide moieties are relatively inexpensive and the EPR technique is very sensitive requiring sample of only a few tens of picomoles.

Continuous wave EPR at X-band frequencies (9 GHz) already offers a wide range of experiments to study macromolecular structure and dynamics. From simple spin label scanning [15], through to basic line shape analysis (figure 2a) [16] and power saturation measurements [8, 17] as well as accessibility experiments [18] and short range dipolar interactions [19, 20].

Such experiments offer detailed insight into secondary structure elements, conformational change, membrane topology as well as correlating localised site accessibility to global conformational change and measuring distances up to ca. 2nm all relevant to the mechanistic understanding of ABC transporters.

It is, however, the developments in pulsed EPR at X-band (and other) frequencies that has opened up this spectroscopy to delivering long-range (2-10 nm) structural constraint information. Being able to engineer and label at two distinct sites using SDSL allows EPR to determine distances and distance distributions (and in some cases relative orientations) between two labels at these sites using Double Electron Electron Resonance (DEER) aka Pulsed ELectron Double Resonance (PELDOR) spectroscopy [21, 22] at cryogenic temperatures. For recent reviews see [23, 24]. Double Quantum Coherence (DQC) [25, 26] and relaxation-induced dipolar modulation enhancement (RIDME) [27, 28] are other, less commonly applied, but nonetheless powerful techniques for distance determination in proteins.

Compared to NMR or X-ray studies the SDSL methodology does not yet (but see 29-31] provide global structural models but rather contributes sparse structural constraints. Fluorescence-based methods such as FRET (Fluorescence resonance energy transfer) are also very powerful tools in this respect since they can provide real time dynamics over several time scales, however they lack precision when quantifying distance or determining distance changes [32, 33].

A plethora of supporting analysis packages are available to analyse these sparse constraints the most popular of which include DEERanalysis [34], a multiscale modelling of molecular systems software package (MMM) [35] and MtsslWizard [36]. Figure 2 (b) demonstrates the use of MMM to attach a spin label *in silico* to an ABC transporter. In the first instance it can report on the feasibility of a specific position to be successfully labelled. Further it calculates the number of possible stable rotamers at specific sites and can ultimately offer distance distribution profiles between pairs of labels.

Clearly there are questions to be asked with regard to the site-directed modification of proteins:- does a labelled protein still function correctly; how physiologically relevant is measuring a distance (or distances) at 20K; are purified, detergent-solubilised membrane proteins true reflections of a functional protein in a membrane. It is exactly in these areas that developments are being made, including determination of longer distances [38], distance editing [39], resolving not only distances but also relative orientations [40-42], measuring distances under physiological conditions [43] or in native environments (whole cells or oocytes) [44, 45]. To summarise, EPR and especially in combination with SDSL is able to provide a wealth of detailed information both for the mechanistic structural biologist and the computational chemist to contribute to advancing our understanding of ABC transporters.

Molecular Dynamics (MD) Simulations: MD simulations are a computational technique that allows the investigation of structure, dynamics and time evolution of a system. Typically one copy of a protein or a larger complex is studied with atomic resolution either in water or membrane embedded. The level of detail necessary for a theoretical description of the

system under study has to be answered initially, and is accompanied by several fundamental consequences. Are chemical reactions and the breaking of bonds involved, are the details of the electronic structure essential? If the answer is yes, then quantum mechanical methods are imperative. Biological systems are dominated typically by Van der Waals and Coulomb interactions, which allow for a classical description of the system, implicit electrons, and the description of atoms and molecules in terms of classical force fields. The computational effort is several orders of magnitude lower. The most widely used classical force fields are AMBER [46], CHARMM [47], GROMOS [48] and OPLS [49]. In essence molecular dynamics (MD) simulations solve Newton's equation of motion [50] to propagate the system in time. Simulations provide a trajectory of the time evolution and of the involved conformations. If sufficient sampling is achieved, probability distributions can be extracted and related to free energies [51]. In theory all properties that depend on conformations can be derived from simulations. The size of a system containing a membrane protein and the simulation length required to observe a full transport cycle do not typically allow one to obtain trajectories in which all relevant conformations are visited. Special treatments are therefore required as described below.

The most widely used EPR label MTSSL has five flexible bonds (Figure 1). It has been observed that interconversion between conformations is slow (on the MD timescale) also for unrestricted labels (in the nanosecond time regime) becoming a challenge for correct Boltzmann-weighted conformational sampling by MD [52]. The flexibility of this label has the advantage that it can be incorporated in almost any position of the protein. The disadvantage is that interpretation of the PELDOR/DEER signals in terms of structure and structural changes is challenging. Typical distance distributions determined by EPR are often broad and multimodal. The uncertainty can have several causes, e.g. (i) the protein may be very flexible or adopt several conformations; (ii) the EPR label might be rigidly positioned or might be flexible, positioned in a conformational ensemble that is not uniform, which would lead to bimodal or multimodal distances detected by EPR.

This label flexibility also challenges structural modelling, because different structural behavior (different protein conformation vs. changes in label conformational distribution) could potentially lead to the same EPR signal. It is therefore often not unambiguously possible to determine protein structure from such a limited data set as provided by EPR alone. Information from other sources could provide sufficient additional structural restraints (especially protein structures) that unambiguous structural interpretation becomes feasible.

Prediction of EPR spectra: Structural interpretation typically starts from a model of the protein. The simplest or zero order approximation normally ignores the label side chains and thus only takes into account the distances between the Cα – Cα atoms of the two EPR label carrying residues. This approach is simple, fast and allows for an order of magnitude estimation without the requirement of modelling of the EPR label.

Methods exist that start from a three-dimensional protein structure and estimate the label position and distribution using library approaches, such as PRONOX [53], the MtsslWizard [36] or the multiscale modelling of molecular systems software package (MMM) [35]. These packages typically require as input the three-dimensional structure of the protein and the site of label and then predict the ensemble of conformations the spin label can adopt. Distance distributions and EPR spectra can than be subsequently predicted. These methods have shown higher-quality results than distances estimated from Cα – Cα atoms distances and also allow to estimate if the input protein conformation is actually compatible with experimental EPR data.

Molecular dynamics or Monte Carlo simulations are also a promising approach because they have the potential to determine a Boltzmann-weighted conformational ensemble for the EPR label. The EPR label has been parameterized for the most commonly used force fields CHARMM [54], Amber [55], GROMOS [11, 56] and OPLS [57]. Literature reports indicate that the computational burden is very high for correct and complete sampling [58]. It is inherently difficult to estimate if completeness of sampling has been achieved, resulting in varying precision for the prediction of spin pair distances within the same protein. The reasons for the varying degree of sampling completeness are many. The simulation length may be too short in comparison to the energy barrier height separating multiple sub-populations of conformations: energy barriers can originate from insufficient sampling of the spin label rotamers, specific interactions with protein side chains or restrictions in the transition between sub-states due to a narrow transition point connecting them.

One of the main limitations for the application of MD simulations to generate a conformational ensemble is the required computational effort. A correct description of the protein environment (water, membrane, ionic concentration) is crucial for the correct behavior of a proteins. A large share of the resources is only used to simulate environment, however in the analysis the focus is on the protein. Clearly a simplification of the environment would be desirable and could lead to a significant speedup of the simulations. This is possible, if the protein conformation does not change and no dynamic effects are expected. The protein backbone can be restrained to the conformation observed in the crystal and the environment represented by an implicit solvent model or simulations can be carried out in vacuum using the Langevin integrator. Simulated annealing protocols or short high-temperature simulations can be applied to speed up the conformational phase space search. A comparison of experimental data of the vinculin tail with distances derived from simulations indicated that a short MD at 600 K in vacuum report experimental data better than simulations in water at 300 K [59], but both approaches showed an improved correlation over predictions calculated on the basis of Cα atom distances alone. The most likely reason for the poorer performance in explicit water is that the simulation times were still too short for proper sampling. A temperature of 600 K in simulations does mainly mean an increased speed in the conformational phase space sampling.

An interesting approach to reduce computational burden has been recently introduced [58]. It uses the concept of dummy atoms/residue, which can be chemically/physically impossible. If parametrization is done properly, they generate correct Boltzmann-weighted, conformational ensembles of the property of interest, while simplifying those interactions that hamper correct sampling. A dummy spin label has been developed which does not seem to suffer from the slow sampling speed of an all-atom representation of the EPR label. The dummy label has been parameterized against exhaustive all atom MD simulations of the same label. The advantage of the dummy EPR label consists in a smoothed energy hypersurface, thereby achieving a much faster sampling, while maintaining proper conformational sampling. Correlation between measured and predicted DEER distances improved considerably compared to the classic approaches (Cα atom based distances, application of rotamer libraries, MD simulations) [59].

Structure prediction using experimental EPR data as input: Typically PELDOR or DEER data provide distances between two (or more) spin labels. These EPR-derived distances can be used directly to position domains relative to each other. As described above, unambiguous positioning of domains relative to each other is challenging, because this is an

underdetermined problem due to the EPR label's flexibility. The degrees of freedom of positioning domains are larger than the degrees of freedom derived from the EPR measurements. The prediction of protein conformation using EPR data is challenging because in this case the position of the Cα atom relative to the spin label is not known. Approaches have been developed to address the challenge by including all atom information into the modelling process.

The modelling approach RosettaEPR [29, 30], which is based on the Rosetta program, has been developed that can include EPR data as input restraints. Prediction of the conformation of the MTSSL label (side chain) is based on conformational information derived from crystal structure of T4-lysozyme and MD simulations. Comparison between crystal structures and models revealed that the rotamer prediction accuracy of the MTSSL conformation is in line with the accuracy of the 20 natural amino acids, which is above 90% in the protein interior and ~50% at the protein surface.

Haddock is a protein docking program that can include restraints from EPR data to guide the search process to predict the conformation of the protein complex [60]. A method called restrained-ensemble Molecular Dynamics Simulations [61] that include EPR distance histograms in the refinement process of protein models.

ABC transporters; We will not aim to cover all the published EPR and MD data on the family of ABC transporters since this is a huge area. Instead we aim to focus on MsbA and ABCB1 since in the former case there is great deal of structural and spectroscopic data available, while in the later there have been recent reports combining results from both EPR and MD experiments. We aim to highlight the main outcomes of a number of key publications and emphasise where and how synergies have and could be achieved in the future.

EPR investigation of the conformation of MsbA: The lipid A flippase MsbA is a multidrug exporter which has been crystallised in both the outward and the inward conformation (PDB IDs: 3B5W, 3B5X, 3B5Y, 3B5Z, and 3B60 [62]. MsbA shows the same fold as Sav1866 [63] and all other ABC exporters from the ABCB family published to date. MsbA has been well studied using EPR spectroscopy. Transmembrane helices 2, 3, 4, 5, and 6 have been extensively investigated for their accessibility to water and/or to lipid by quenching of the EPR signal by NiEDDA and molecular oxygen $ПO_2$. Transmembrane helices 2, 3, 4 and 5 revealed a clear helical pattern of membrane-exposed helices, while helix 6 was found not to be membrane- or water-exposed [8, 64]. Ref 8 indicated significant differences to the MsbA crystal structure, with which it was published back to back in Science in 2005. The crystal structure was actually later retracted and republished in its corrected form [62], implicitly confirming the original EPR results. Changes in structure were found between the nucleotide free and the ADP-vanadate trapped state both at extracellular and the intracellular sites, but not in the transmembrane helices.

Distance measurements using DEER [65, 66] revealed a large change in spin label distance, which became increasingly larger towards the NBDs. This nucleotide-free state was found to be mainly open to the inward side, while the ADP-vanadate trapped state and the AMP-PNP [10] state showed NBD-NBD contacts.

Large differences (up to 2 nm) between adjacent residues indicated that the EPR distances represented a convoluted signal between protein conformational changes and changes in label mobility. Comparison between the crystal structures and the EPR results (in detergent) indicated that the crystal structure might overestimate the conformational changes, while EPR results obtained in liposomes revealed even lower amplitudes (Figure 3).

Introduction of EPR spin labels at the NBD dimer interface into the Walker A and B, signature motif, and H loop showed limited mobility of the labels in cw-EPR spectra, which did not differ between the nucleotide-free state and in the presence of ATP [9]. Addition of vanadate or ADP revealed only a few instances of a decrease in mobility, while the mobility of labels e.g. in the Walker B motif did not differ between the various nucleotide-bound states. A large portion of EPR data has shown that the NBDs become flexible in the absence of nucleotides and dissociate to a large extent. The data from the Walker B motif are not consistent, suggesting that the two spin labels within the NBD interface might have induced a change in the distribution of conformations by potentially stabilising a closed NBD dimer already in the nucleotide free state.
Mutation of the H-loop histidine H537A and especially the catalytic E506Q strongly reduced ATP hydrolysis rates and blocked transporter function above background level, while the binding affinity of ATP and analogs remained almost unchanged [68]. EPR label mobility determined (labels were attached to Walker A and B, signature motif, and H loop) by cw-EPR line shape analysis revealed that both mutants showed similar mobilities in the nucleotide-free as well as in the nucleotide-bound state, while DEER distance analysis indicated substantial populations of NBD-NBD associated dimers already in the nucleotide-free state.

Computational investigation of the conformations and dynamics of MsbA by MD: MD simulations have been used to directly investigate the conformational transitions between the crystal structures. The transition from the outward facing to the inward facing conformations was studied using targeted MD, where an external force was added to the system to impose the transition [69]. In repeated 0.5 and 1.0 ns long simulations it was found that the NBD's are the first domains which react to the applied external force. The NBDs demonstrated increased resistance against the externally applied force in the presence of ATP, in line with crystallographic data and results from EPR. Results from targeted MD or steered DM in isolation are difficult to interpret at simulations length significantly below 100 ns, because the conversion is very fast, thereby not giving the transporter enough time to react, increasing the danger of producing artefacts. The targeted MD results could not be reproduced using steered MD simulations using a much slower rate constant [70].
Unbiased simulations [71] of 100 ns, starting from the outward facing conformation, showed that the transporter was conformationally dynamic at both the extracellular site and at the level of the NBDs in the apo state. The presence of ATP in the substrate binding site resulted in an MsbA structure which was structurally very stable at both sides of the membrane.
The application of a biasing potentials to investigate conformational changes in membrane transporters is necessary, because the processes are far too slow. The difficulty arises from the design of the bias, which must be carefully controlled and should not lead to unrealistic structural changes. Comparison with experimental data is imperative. If carried out properly, MD simulations can lead to more detailed insights, because the positions of all atoms and the associated forces and energies are known. The path of the structural conversion from the outward-facing conformation to the inward-facing conformation has been investigated using a complex and system-specific biasing protocol [70, 72] that used predefined biasing directions derived from available crystal structures. This approach identified a sequence of conformational changes consisting of periplasmic and cytoplasmic opening and closing of the TMD, NBD separation and the twisting motion of TMD vs. NBD. A low-energy path for interconversion between the conformations could also be identified. The free energy of the

transition from the inward-open to the inward-occluded conformation indicated that the conformation as observed in the crystal structure of nucleotide free MsbA is not the most stable conformation of the membrane-inserted transporter [70, 72, 73]. The energy of the NBD associated inward-closed state was only slightly higher than the free energy of the global minimum of slightly separated NBDs. A large range of conformations were accessible at room temperature, in line with the observation of a more closed and dynamic state of the nucleotide free MsbA measured using EPR.
The behavior of lipid surrounding MsbA in three conformations [74] showed that a ring of annular lipid formed, which was not homogeneous. Distortions of the membrane bilayer structure were observed and also lipid molecules reaching into the substrate binding site.

ABCB1: The multidrug resistance protein ABCB1 has been studied extensively and there have already been several EPR approaches that have used spin labelled moieties such as labelled ATP [75] as well as labelled transport substrates (e.g. spin labelled verapamil [76] to probe mechanism. In addition several papers have used MD, but the difficulty associated with purification and reconstitution of functional ABCB1 in high amounts has limited the applications of EPR to the study of ABCB1. The focus of this review is on the synergistic application of EPR and MD and we have limited ourselves to recent publications which compare results from these methods and/or synergistically use them.

Conformational dynamics and separation of the NBDs in the mouse ABCB1 transporter have been investigated by EPR and directly compared to MD simulations [77]. Three cysteine residue pairs (K615C-S1276C, A627C-K1260C and R613C-A1258C) were introduced in an otherwise cysteine-less ABCB1. The mutant transporters were labelled with MTSSL and reconstituted into micelles. DEER measurements of the three label pair showed consistently a broad distribution of distances in the nucleotide free state. Distances were both longer and shorter than found in the crystal structure of mouse ABCB1 (PDB ID: 3G5U [78]. The large range of distances pointed towards protein mobility and a wide variation of interdomain distances of the NBDs. MD simulations were initiated from the inward-open structure of the mouse ABCB1 (PDB ID: 3G5U) [78]. The simulation data were in line with the DEER measurements and showed that the transporter adopted a wide range of conformations. Interdomain separation ranged from 3 to 5 nm. Addition of ATP changed the range of NBD separation, preferring shorter separations. Each system was simulated three times. Convergence could not be achieved, indicating that the time scales of NBD motion are much longer than that accessible by MD.
Reaching orders of magnitude longer trajectories will require faster computer systems. Moore's law predicts that the computer power double every 17-18 months, indicating that it might be possible to carry out these unbiased simulations in 10 years from now. If the path of the conformational change can be defined correctly, then free energy path methods as the Potential of Mean Force calculations [79] (see above simulations of MsbA [70, 72, 73]) can derive the conformational ensemble from the energy profile along the reaction path. The information is equivalent with respect to a Boltzmann weighted distribution of protein conformations necessary to compare to EPR data, but requires much less computational effort. Correctly defining the path of conformational changes at the atomic level remains the biggest challenge.
A recent investigation into the alternating access mechanism in ABCB1 has been performed using SDSL with cw-EPR spectroscopic methods and by rationalising this data with MD

simulations to describe the conformational changes that the central cavity undergoes during translocation [11]. This investigation was driven by the recent structural data from the murine and *Caenorhabditis elegans* homologues of ABCB1. It should be noted that while these structures were obtained in the absence of nucleotide, a common feature was the presence of a central cavity that is inaccessible from the extracellular face of the protein. Several residues, proposed to line, or reside proximal to, the central cavity in the transmembrane helices TM6 (331, 343 and 354) and TM12 (980) were mutated to cysteine and labelled with the MTSSL paramagnetic probe enabling the use of cw-EPR spectroscopic methods. Power saturation experiments were performed in the presence of hydrophobic (O2) or hydrophilic (NiEDDA) quenching agents to report on the local environment of each residue. These experiments were repeated on ABCB1 trapped in both its nucleotide-bound and post-hydrolytic conformations. Rationalization of the data with molecular dynamic simulations indicated that the cavity is converted to a configuration closed to the intracellular side of the protein following nucleotide binding, thereby suggesting a possible intermediate step in an alternating access to the cavity on opposite sides of the membrane during translocation. To further support this proposal and indicate whether this cavity is open to the extracellular side of the membrane during translocation PELDOR experiments on doubly labelled ABCB1 using these residues were proposed [80]

Conclusions:

The knowledge of the protein structure is essential to understand its function. Crystallography is the mainstream method to obtain structural information, but crystal structures of membrane proteins are difficult to derive, and often only provide static snapshots. Only a few membrane proteins have been crystallised in more than one conformation. Knowledge of such structural and conformational changes is key to understand how membrane transporters translocate substrates across the membrane

EPR provides complex dynamic information over a large range of distances, allowing for measurement of distances, environment and protein dynamics. Structural interpretation of the results is an underdetermined problem; Analysis of EPR data relays strongly on the structural information of protein snapshots from crystals. Similar restrictions apply to FRET.

MD simulations are complementary to EPR and FRET and generate high resolution structural and dynamic information. MD simulations require structures as starting points and the method is limited in the time domain, because the time scale of the transport cycle of a typical membrane transporter is much longer.

Our goal is to develop a fuller understanding of transporter function at the atomic level. This needs an integration of information from all methods to overcome inherent limitations. We have focused on the integration of EPR and MD to highlight the synergies. Biochemical data add a new dimension to the analysis of transporter function, providing orthogonal information from transporter function and kinetics.

Figure Legends:

Figure 1: Structure of commonly used nitroxide spin labels and their reaction with a cysteine residue. (A) the most commonly used methanthiosulfonate spin label (MTSSL), which is highly selective to thiol groups and offers a good compromise between rigidity of the N-O group and flexibility to ensure minimal perturbation of the protein structure (B) the iodoacetamido-proxyl spin label (IAP). The disadvantages of attachment by a disulphide bond, which is prone to reductive cleavage, mean that under reducing conditions (e.g. in the reductive environment of living cells) maleimido or iodoacetamido labels like IAP should be

used. (C) The rigid cross-linked side chain RX spin label. The link is formed between any elements with properly spaced cysteine residues (see ref. 12)

Figure 2: (a) Examples of simulated X-band EPR spectra of a nitroxide spin label. The lowest spectrum corresponds to a spin label with slow isotropic motion of 100 ns (labelled), The separation of the two outer hyperfine extrema may be taken as a measure of label mobility for spin labels undergoing slow motion. In this case this value is close to the static limit. As motion is increased thus the EPR line shape alters quite significantly. The uppermost trace corresponds to fast isotropic motion of 0.1 ns. (b) Using ABCB1 (PDB: 4Q9H) as an example three sites have been selected and *in-silico* spin labelling has been performed using MMM (35). The figures represent the number of predicted spin label rotamers (depicted in red, 134 for the extracellular surface site, 10 for protein buried site and 71 for the intracellular aqueous cavity site) and hence flexibility and degree of freedom of the attached label which will influence its EPR characteristics.

Figure 3: Crystal structures of MsbA (panel A, C, and D) [62] have been obtained in several conformations and shaped our view of how ABC transporters translocate substrates across the membrane. Panel B shows the recently crystallized McjD structure observed in the intermediate occluded transition state [67]. ABC transporters are assumed to rest in the inward facing state (panel D). Transition to the outward-facing state is believed to be initiated by substrate and ATP binding. The structural conversion path moves through an inward-closed (panel C) and the occluded conformation (panel B) and reaches the outward-open conformation (panel A), in which substrate is released. The green stars symbolize attached EPR labels. Measured distances showed that the transporter might not reach the extreme conformation as observed in the crystal structures [70].

Acknowledgements:
FM would like to thank the Wellcome Trust [094392/B/10/Z] and the Royal Society Wolfson Research Merit Award for financial support. ASM would like to that the Biochemical Society for a summer studentship bursary and COST [CM1306] for a STSM financial support. TS would like to thank financial support of the Austrian Science Fund (FWF) [P23319 and SFB-3524]. All authors would like to acknowledge the COST Action (CM1306) “Understanding Movement and Mechanism in Molecular Machines” for continued financial support.

References:

[1] Krishnamurthy, H. and Gouaux, E. (2012) X-ray structures of LeuT in substrate-free outward-open and apo inward-open states. Nature **481**, 469-474

[2] Jiang, Y., Lee, A., Chen, J., Ruta, V., Cadene, M., Chait, B.T. and MacKinnon, R. (2003) X-ray structure of a voltage-dependent K^+ channel Nature **423**, 33-41

[3] Wüthrich, K. (1998) The second decade — into the third millenium. Nat. Struct. Biol. **5**, 492–495

[4] Wang, S., Munro, R.A., Shi, L., Kawamura, I., Okitsu, T., Wada, A., Kim, S.Y., Jung, K.H., Brown, L.S. and Ladizhansky, V. Solid-state NMR spectroscopy structure determination of a lipid-embedded heptahelical membrane protein. Nat. Methods. **10**, 1007-1012

[5] Comellas, G. and Rienstra, C.M. (2013) Protein structure determination by magic-angle spinning solid-state NMR, and insights into the formation, structure, and stability of amyloid fibrils. Annu. Rev. Biophys. **42**, 515-536

[6] Schweiger, A. (1991) Pulsed electron spin resonance spectroscopy: Basic principles, techniques, and examples of applications. Angew. Chem. Int. Ed. Engl. **30**, 265–292
[7] Prisner, T. F., Rohrer, M. and MacMillan, F. (2001). Pulsed EPR spectroscopy: biological applications. Annu. Rev. Phys. Chem. **52**, 279–313
[8] Dong, J., Yang, G. and McHaourab, H.S.: Structural basis of energy transduction in the transport cycle of MsbA. Science 2005, **308**, 1023-1028
[9] Westfahl, K.M., Merten, J.A., Buchaklian, A.H. and Klug, C.S. (2008) Functionally important ATP binding and hydrolysis sites in Escherichia coli MsbA. Biochemistry **47**, 13878-13886
[10] Mittal, A., Böhm, S., Grütter, M.G., Bordignon, E. and Seeger, M.A. (2012) Asymmetry in the homodimeric ABC transporter MsbA recognized by a DARPin. J. Bio.l Chem. **287**, 20395-20406
[11] Van Wonderen, J.M., McMahon, R.M., O'Mara, M.L., McDevitt, C.A., Thomson, A.J., Kerr, I.D., MacMillan, F. and Callaghan, R. (2014) The central cavity of ABCB1 undergoes alternating access during ATP hydrolysis. FEBS J. **281,** 2190-2201
[12] Fleissner, M.R., Bridges, M.D., Brooks, E.K., Cascio, D., Kalai, T., Hideg, K. and Hubbell, W.L (2011) Structure and dynamics of a conformationally constrained nitroxide side chain and applications in EPR spectroscopy. Proc. Natl. Acad. Sci. USA **108**, 16241-16246
[13] Polyhach, Y., Bordignon, E., Tschaggelar, R., Gandra, S., Godt, A. and Jeschke, G. (2012) High Sensitivity and Versatility of the DEER Experiment on Nitroxide Radical Pairs at Q-band Frequencies. Phys. Chem. Chem. Phys. **14**, 10762-10773
[14] Spindler, P.E., Glaser, S.J., Skinner, T.,E. and Prisner, T.F. (2013) Broadband Inversion PELDOR Spectroscopy with Partially Adiabatic Shaped Pulses. Angew. Chem. Int. Ed. **52**, 3425-3429
[15] Cuello, L.G., Cortes, D.M. and Perozo E: (2004) Molecular architecture of the KvAP voltage dependent K+ channel in a lipid bilayer. Science **306**, 491-495.
[16] Crane, J.M., Mao, C., Lilly, A.A., Smith, V.F., Suo, Y., Hubbell, W.L. and Randall, L.L. (2005) Mapping of the docking of SecA onto the chaperone SecB by site-directed spin labeling: insight into the mechanism of ligand transfer during protein export. J. Mol. Biol. **353**, 295-307
[17] Altenbach, C., Greenhalgh, D.A., Khorana, H.G. and Hubbell, W.L (1994) A collision gradient method to determine the immersion depth of nitroxides in lipid bilayers: application to spin-labeled mutants of bacteriorhodopsin. Proc. Natl. Acad. Sci. USA **91**, 1667-1671
[18] Nelson, W.D., Blakely, S.E., Nesmelov, Y.E. and Thomas, D.D. (2005) Site-directed spin labeling reveals a conformational switch in the phosphorylation domain of smooth muscle myosin. Proc. Natl. Acad. Sci. U.S.A. **102**, 4000-4005
[19] Rabenstein, M.D. and Shin, Y.K. (1995) Determination of the distance between two spin labels attached to a macromolecule. Proc. Natl. Acad. Sci. USA **92**, 8239-8243.
[20] Steinhoff, H.-J., Radzwill, N., Thevis, W., Lenz, V., Brandenburg, D., Antson, A., Dodson, G.G. and Wollmer, A. (1997). Determination of interspin distances between spin labels attached to insulin: comparison of electron paramagnetic resonance data with the X- ray structure. Biophys. J. **73**, 3287-3298
[21] Milov, A. D., Salikhov, K.M., Shirov, M.D. (1981) Application of the double resonance method to electron spin echo in a study of the spatial distribution of paramagnetic centres in solids. Sov. Phys. Solid State **23**, 565-569
[22] Martin, R.E., Pannier, M., Diederich, F., Gramlich, V., Hubrich, M. and Speiss, H.W. (1998) Determination of end-to-end distances in a series of TEMPO diradicals of up to 2-8 nm length with a new four-pulse double electron-electron resonance experiment. Angew. Chem. Int. Ed. **37**, 2834-2837
[23] Jeschke, G. (2012) DEER Distance Measurements on Proteins. Annu. Rev. Phys. Chem. **63**, 419-446

[24] Borbat, P.P. and Freed, J.H. (2013) Pulse Dipolar Electron Spin Resonance: Distance Measurements. In Structure and Bonding; Timmel C. R., Harmer J., Eds.; **152**, pp 206–210, **152**, pp 1-82. Springer: Heidelberg

[25] Borbat, P.P., McHaourab, H. and Freed, J.H., (2002) Protein Structure determination using long-distance constraints from double-quantum coherence ESR: Study of T4 lysozyme. J. Am. Chem. Soc. **124**, 5304-5314

[26] Borbat, P.P., Davies, J.H., Butcher, S.E. and Freed, J.H. (2004) Measurement of large distances in biomolecules using double-quantum filtered refocused electron spin-echoes. J. Am. Chem. Soc. **126**, 7746-7747

[27] Kulik, L.V., Dzuba, S.A., Grigoryev, I.A. and Tsvetkov, Y.D. (2001) Electron dipole–dipole interaction in ESEEM of nitroxide biradicals. Chem. Phys. Lett. **343**, 315–324.

[28] Kulik, L.V., Grishin, Y.A., Dzuba, S.A., Grigoryev, I.A., Klyatskaya S.V., Vasilevsky S.F. and Tsvetkov, Y.D. (2002) Electron dipole–dipole ESEEM iN Field-step ELDOR of Nitroxide Biradicals, J. Magn. Reson. **157**, 61-68

[29] Kazmier, K., Alexander, N.S., Meiler, J. and McHaourab, H.S. (2011) Algorithm for selection of optimized EPR distance restraints for de novo protein structure determination. J. Struct. Biol. **173**, 549-557.

[30] Hirst, S.J., Alexander, N., McHaourab, H.S. and Meiler, J. (2011) RosettaEPR: an integrated tool for protein structure determination from sparse EPR data. J. Struct. Biol. **173**, 506-514.

[31] Hilger, D., Polyhach, Y., Padan, E., Jung, H. and Jeschke G. 2007 High-resolution structure of a Na+/H+antiporter dimer obtained by pulsed electron paramagnetic resonance distance measurements. Biophys. J. **93**, 3675–3683

[32] Agafonov, R.V., Negrashov, I.V., Tkachev, Y.V., Blakely, S.E., Titus, M.A., Thomas, D.D. and Nesmelov, Y.E. (2009) Structural dynamics of the myosin relay helix by time-resolved EPR and FRET. Proc. Natl. Acad. Sci. USA **106**, 21625–21630

[33] Sisamakis, E., Valeri, A., Kalinin, SS, Rothwell, P.J., and Seidel, C.A.M. Accurate Single-Molecule FRET Studies Using Multiparameter Fluorescence Detection, Methods in Enzymology **475**, 455-514

[34] Jeschke, G., Chechik, V., Ionita, P., Godt, A., Zimmermann, H., Banham, J., Timmel , C.R., Hilger, D. and, Jung, H. 2006 DeerAnalysis2006: A comprehensive software package for analyzing pulsed ELDOR data. Appl. Magn. Reson. **30**, 473– 498.

[35] Polyhach, Y., Bodignon, E. and Jeschke, G. (2011) Rotamer libraries of spin labelled cysteines for protein studies. Phys. Chem. Chem. Phys. **13**, 2356-2366.

[36] Hagelüken, G., Ward, R., Naimsith, J.H. and Schiemann, O. (2012) MtsslWizard: In silico Spin- Labelling and Generation of Distance Distributions in PyMOL. Appl. Magn. Reson. **42**, 377-391

[37] Szewczyk, P., Tao, H., McGrath, A.P., Villaluz, M., Rees, S.D., Lee, S.C., Doshi, R., Urbatsch, I.L., Zhang, Q. and Chang, G. (2015) Snapshots of ligand entry, malleable binding and induced helical movement in P-glycoprotein. Acta Crystallogr., Sect.D **71**, 732-741

[38] Ward, R., Bowman, A., Sozudogru, E., El-Mkami, H., Owen-Hughes, T., Norman D.G. (2010) EPR distance measurements in deuterated proteins. J. Magn. Reson. **207**, 164-167

[39] van Wonderen, J.H., Kostrz, D.N., Dennison, C. and Macmillan, F. (2013) Refined distances between paramagnetic centers of a multi-copper nitrite reductase determined by pulsed EPR (iDEER) spectroscopy. Angew. Chem. Int. Ed. Engl. **52**, 1990-1993

[40] Abe, C., Klose, D., Dietrich, F., Ziegler, W.H., Polyhach, Y., Jeschke, G. and Steinhoff, H-J. (2012) Orientation selective DEER measurements on vinculin tail at X-band frequencies reveal spin label orientations. J. Magn. Reson. **216**, 53-61

[41] Reginsson, G.W., Hunter, R.I., Cruickshank, P.A., Bolton, D.R., Sigurdsson, S.T., Smith, G.M. and Schiemann, O. (2012) W-band PELDOR with 1 kW microwave

power: molecular geometry, flexibility and exchange coupling. J. Magn. Reson. **216**, 175-182
[42] Tkach, I., Pornsuwan, S., Hobartner, C., Wachowius, F., Sigurdsson, S.T., Baranova, T.Y., Diederichsen, U., Sicoli, G. and Bennati M (2013) Orientation selection in distance measurements between nitroxide spin labels at 94 GHz EPR with variable dual frequency irradiation. Phys. Chem. Chem. Phys. **15**, 3433-3437
[43] Yang, Z., Liu, Y., Borbat, P., Zweier, J.L., Freed, J.H. and Hubbell, W.L. (2012) Pulsed ESR dipolar spectroscopy for distance measurements in immobilized spin labeled proteins in liquid solution. J. Am. Chem. Soc., **134**, 9950-9952.
[44] Krstić, I., Hänsel, R., Romainczyk, O., Engels, J.W., Dötsch, V. and Prisner, T.F. (2011) Long-Range Distance Measurements on Nucleic Acids in Cells by Pulsed EPR Spectroscopy Angew. Che. Int. Ed. **50**, 5070–5074
[45] Azarkh, M., Okle, O., Singh, V., Seemann, I.T., Hartig, J.S., Dietrich, D.R. and Drescher, M. (2011) Long-Range Distance Determination in a DNA Model System inside Xenopus laevis Oocytes by In-Cell Spin-Label EPR and Chem. Bio. Chem. **12(13)**, 1992–1995
[46] Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J.L., Dror, R.O. and Shaw D.E. (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins **78**, 1950-1958
[47] Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., Darian, E., Guvench, O., Lopes, P., Vorobyov, I. and Mackerell A.D. Jr (2010) CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. J. Comput. Chem. **31**, 671-690
[48] Oostenbrink, C., Villa, A., Mark, A.E. and van Gunsteren, W.F. (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. J. Comput. Chem. **25**, 1656-1676
[49] Kaminski, G.A., Friesner, R.A., Tirado-Rives, J. and Jorgensen, W.L. (2001) Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. J. Phys. Chem. B **105**, 6474–6487
[50] van Gunsteren, W.F. and Weiner P.K. (1989) Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications, Springer Verlag, Berlin.
[51] Allen, M.P. and Tildesley, D.J. (2006) Computer Simulation Of Liquids, Oxford Science Publications.
[52] Stoica, I. (2005) Force field impact and spin-probe modeling in molecular dynamics simulations of spin-labeled T4 lysozyme. J. Mol. Model. **11**, 210-225
[53] Hatmal, M.M., Li, Y., Hegde, B.G., Hegde, P.B., Jao, C.C., Langen, R. and Haworth, I.S. (2012) Computer modeling of nitroxide spin labels on proteins. Biopolymers **97**, 35-44
[54] Sezer, D., Freed, J.H. and Roux, B. (2008) Parametrization, molecular dynamics simulation, and calculation of electron spin resonance spectra of a nitroxide spin label on a polyalanine alpha-helix. J. Phys. Chem. B **112**, 5755-5767
[55] Hermosilla, L., Prampolini, G., Calle, P., de la Vega, J.M.G., Brancato, G. and Barone V. (2013) Extension of the AMBER Force Field for Nitroxide Radicals and Combined QM/MM/PCM Approach to the Accurate Determination of EPR Parameters of DMPO-H in Solution. J. Chem. Theory Comput. **9**, 626–636
[56] Garay, A.S. and Rodrigues, D.E. (2008) Effects of the Inclusion of the Spin Label 10-Doxyl-Stearic Acid on the Structure and Dynamics of Model Bilayers in Water: Stearic Acid and Stearic Acid/Cholesterol (50:20) J. Phys. Chem. B **112**, 1657-1670
[57] Murzyn, K, Róg, T., Blicharski, W., Dutka, M., Pyka, J., Szytula, S. and Froncisz, W. (2006) Influence of the disulfide bond configuration on the dynamics of the spin label attached to cytochrome c. Proteins: Structure, Function, and Bioinformatics **62**, 1088–1100
[58] Islam, S.M. and Roux, B. (2015), Simulating the distance distribution between spin-labels attached to proteins. J Phys Chem B **119**, 3901-3911

[59] Gajula, P.M.N., Vogel, K.P., Rai, A., Dietrich, F. and Steinhoff H.J. (2013) How far in-silico computing meets real experiments. A study on the structure and dynamics of spin labeled vinculin tail protein by molecular dynamics simulations and EPR spectroscopy. BMC Genomics. 14 Suppl 2:S4
[60] Karaca, E. and Bonvin, A.M.J.J. (2013) Advances in integrative modeling of biomolecular complexes. Methods **59**, 372–381
[61] Roux, B. and Islam S.M. (2013) Restrained-Ensemble Molecular Dynamics Simulations Based on Distance Histograms from Double Electron–Electron Resonance Spectroscopy J. Phys. Chem. B **117**, 4733–4739
[62] Ward, A., Reyes, C.L., Yu, J., Roth, C.B. and Chang G. (2007) Flexibility in the ABC transporter MsbA: Alternating access with a twist. Proc. Natl. Acad. Sci. USA **104**, 19005-19010
[63] Dawson, R.J. and Locher, K.P. (2006) Structure of a bacterial multidrug ABC transporter. Nature **443**, 180-185
[64] Zou, P., Bortolus, M. and McHaourab, H.S. (2009) Conformational cycle of the ABC transporter MsbA in liposomes: detailed analysis using double electron-electron resonance spectroscopy. J. Mol. Biol., **393**, 586-597
[65] Borbat, P.P., Surendhran, K., Bortolus, M., Zou, P., Freed, J.H. and Mchaourab H.S. (2007) Conformational motion of the ABC transporter MsbA induced by ATP hydrolysis. PLoS Biol., **5**, e271
[66] Smriti, Zou, S.P. and McHaourab, H.S. (2009) Mapping daunorubicin-binding Sites in the ATP-binding cassette transporter MsbA using site-specific quenching by spin labels. J. Biol. Chem. **284**, 13904-13913
[67] Choudhury, H.G., Tong, Z., Mathavan, I., Li, Y., Iwata, S., Zirah, S., Rebuffat, S., van Veen, H.W. and Beis K. (2014) Structure of an antibacterial peptide ATP-binding cassette transporter in a novel outward occluded state. Proc. Natl. Acad. Sci. USA **111**, 9145-91450
[68] Schultz, K.M., Merten, J.A. and Klug C.S. (2011) Characterization of the E506Q and H537A dysfunctional mutants in the *E. coli* ABC transporter MsbA. Biochemistry **50**, 3599-3608
[69] Weng, J.W., Fan, K.N. and Wang W.N. (2010) The conformational transition pathway of ATP binding cassette transporter MsbA revealed by atomistic simulations. J. Biol. Chem. **285**, 3053-3063
[70] Moradi, M. and Tajkhorshid E. (2013) Mechanistic picture for conformational transition of a membrane transporter at atomic resolution. Proc. Natl. Acad. Sci. USA. **110**, 18916-18921
[71] Furuta, T., Yamaguchi, T., Kato, H. and Sakurai M. (2014) Analysis of the structural and functional roles of coupling helices in the ATP-binding cassette transporter MsbA through enzyme assays and molecular dynamics simulations. Biochemistry, **53**, 4261-4272
[72] Moradi, M. and Tajkhorshid E. (2014) Computational Recipe for Efficient Description of Large-Scale Conformational Changes in Biomolecular Systems. J. Chem. Theory Comput. **10**, 2866-2880
[73] Wang, Z. and Liao, J.L. (2015) Probing structural determinants of ATP-binding cassette exporter conformational transition using coarse-grained molecular dynamics. J. Phys. Chem. B **119**, 1295-1301
[74] Ward, A.B., Guvench, O. and Hills, R.D. Jr. (2012) Coarse grain lipid-protein molecular interactions and diffusion with MsbA flippase. Proteins **80**, 2178-2190
[75] Delannoy, S, Urbatsch, I.L., Tombline, G., Senior, A.E. and Vogel, P.D. (2005) Nucleotide binding to the multidrug resistance P-glycoprotein as studied by ESR spectroscopy. Biochemistry **44**, 14010-14019
[76] Omote, H., and Al-Shaw, M.K., (2002) A novel electron paramagnetic resonance approach to determine the mechanism of drug transport by P-glycoprotein J. Biol. Chem. **277**, 45688-45694

[77] Wen, P.C., Verhalen, B., Wilkens, S., Mchaourab, H.S. and Tajkhorshid E. (2013) On the origin of large flexibility of P-glycoprotein in the inward-facing state. J. Biol. Chem. **288**, 19211-19220
[78] Aller, S.G., Yu, J., Ward, A., Weng, Y., Chittaboina, S., Zhuo, R., Harrel, P.M., Trinh, Y.T., Zhang, Q., Urbatsch, I.L. and Chang G. (2009) Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. Science **323**, 1718-1722
[79] Kumar, S., Rosenberg, J.M., Bouzida, D., Swendsen, R.H. and Kollman P.A. (1992) The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. J. Comput. Chem. **13**, 1011–1021
[80] Van Wonderen, J.H., McMahon, R.M., Kerr, I.D., Callaghan, R. and MacMillan, F. (2015) unpublished data

Figure 1

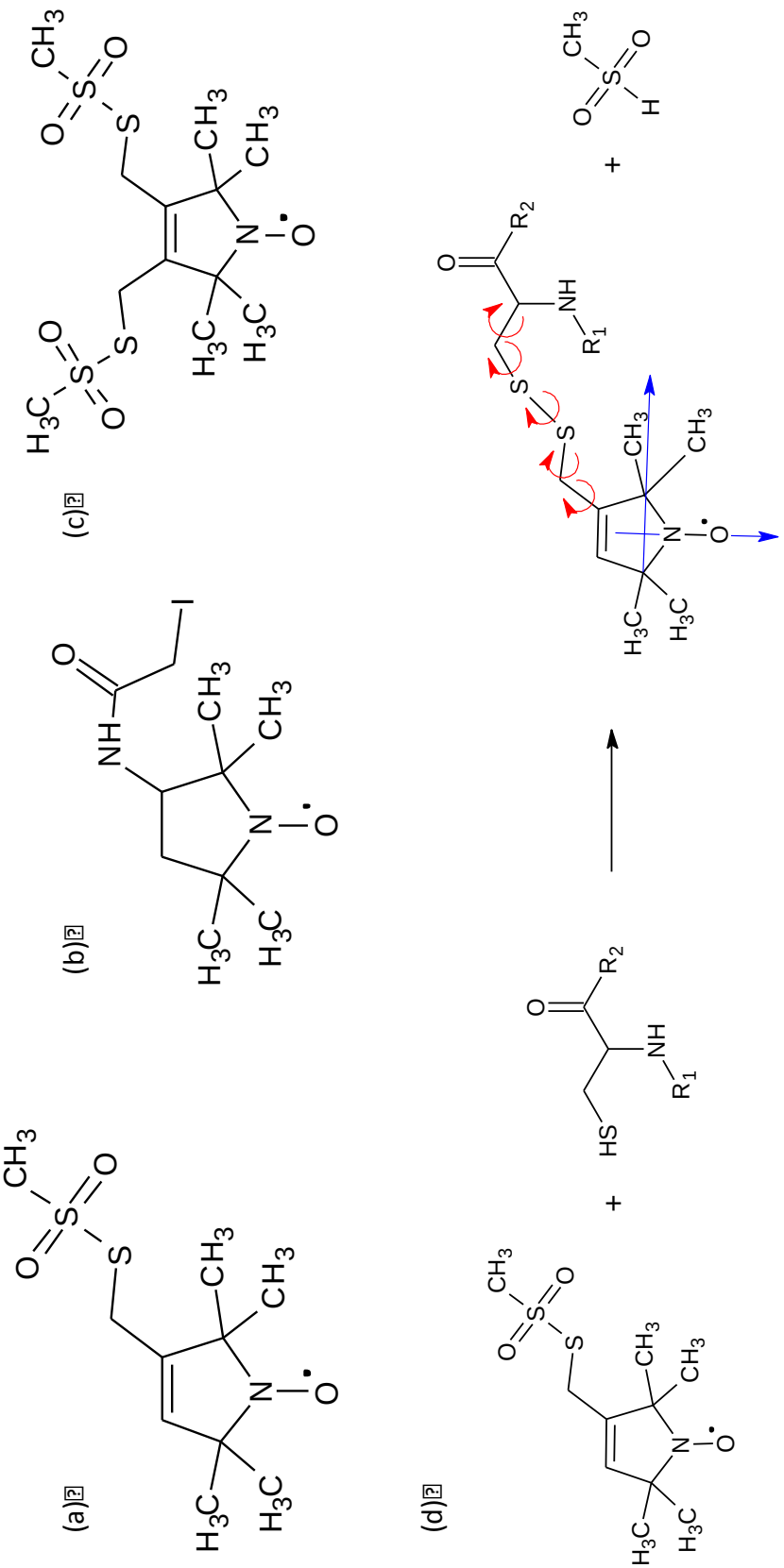

Figure 2

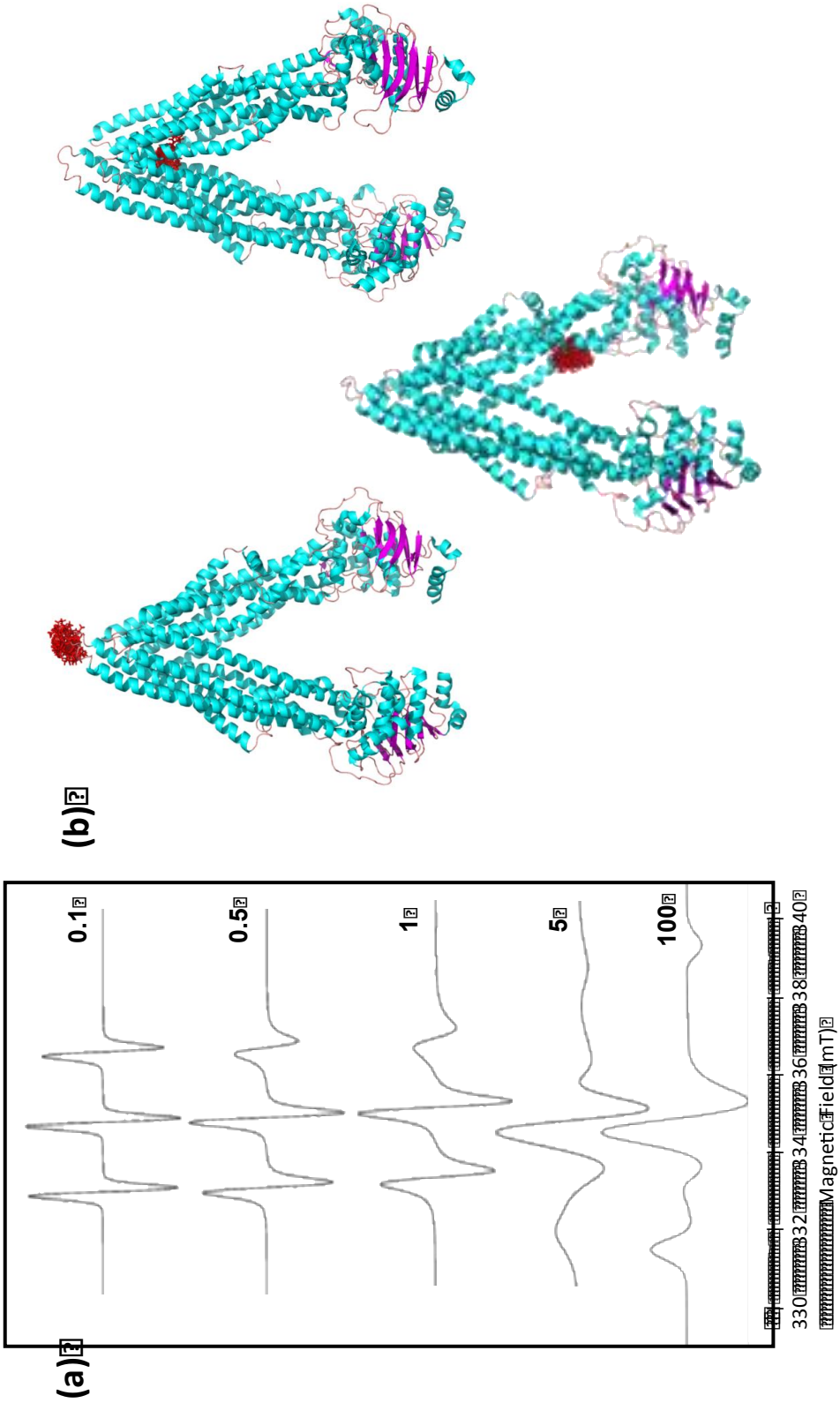

Figure 3