



OPEN

## In silico method for selecting residue pairs for single-molecule microscopy and spectroscopy

Hendrik R. Sikkema & Bert Poolman

Obtaining (dynamic) structure related information on proteins is key for understanding their function. Methods as single-molecule Förster Resonance Energy Transfer (smFRET) and Electron Paramagnetic Resonance (EPR) that measure distances between labeled residues to obtain dynamic information rely on selection of suitable residue pairs for chemical modification. Selection of pairs of amino acids, that show sufficient distance changes upon activity of the protein, can be a tedious process. Here we present an in silico approach that makes use of two or more structures (or structure models) to filter suitable residue pairs for FRET or EPR from all possible pairs within the protein. We apply the method for the study of the conformational dynamics of the substrate-binding domain of the osmoregulatory ATP-Binding Cassette transporter OpuA. This method speeds up the process of designing mutants, and because of its systematic nature, the chances of missing promising candidates are reduced.

Years of X-ray crystallography, NMR and more recently CryoEM<sup>1,2</sup> have made available a wealth of structural data for (membrane) proteins. Even though all these techniques can give information on protein dynamics, for instance using caged compounds and free-electron lasers in serial crystallography<sup>3</sup>, or imaging under turnover conditions in CryoEM<sup>4</sup>, transient states and continuous dynamics are not readily obtained. Double electron–electron resonance (DEER) or Pulsed Electron–electron double resonance (PELDOR) can be used to probe distance changes upon changing conditions<sup>5</sup> and single-molecule Förster Resonance Energy Transfer (smFRET) can be used to obtain single molecule dynamics of proteins and other macromolecular assemblies<sup>6–10</sup>. The latter two techniques make use of two labels that are introduced for instance by attaching them to cysteine residues via maleimide<sup>11</sup> or methanethiosulfonate chemistry<sup>12</sup> or introducing them as non-natural amino acids<sup>13</sup>. SmFRET even enables study of protein dynamics in vivo<sup>12</sup>. In vitro studies can provide a wealth of dynamics information, for instance when smFRET is applied to determine the dynamics of single-surface attached proteins upon addition of a ligand or e.g. photoactivation<sup>10</sup>.

There are a few difficulties in the application of the optical microscopy or electron spin resonance techniques. A key challenge is the selection of the labeling sites and the selectivity of the modifications. Both smFRET and EPR detect a change in distance between the labels, therefore naturally, the label should report a change in conformation. But even when they do, there are more restrictions. For instance, FRET typically occurs at spacings smaller than 8–10 nm, and the signal is the strongest when the distance is close to the Förster radius ( $R_0$ ), which is a property of the used FRET-pair<sup>14</sup>. On the same note, the larger the distance change, the larger the change in FRET signal, thus when selecting FRET pairs, one typically looks for pairs making large movements. Similarly, PELDOR is sensitive for distance changes in the range from 1.6 to 8 nm<sup>5,15</sup>. Furthermore, the site of labeling should be solvent-accessible, the label should be able to rotate freely to prevent anisotropic artifacts and the modification should not affect the functionality of the protein. In FRET, where a fluorescence donor and acceptor label are required, the labeling efficiency makes the procedure even more complicated. A typical labeling with donor (D) and acceptor (A) yields 25% of DD, 25% of AA and therefore only half of the particles are useful (25% of DA + 25% of AD). Alternating Laser EXcitation (ALEX), TIRF-based smFRET<sup>16</sup> and Pulsed Interleaved Excitation (PIE) spectroscopy<sup>17,18</sup> are techniques that allow FRET measurements corrected for the unwanted DD and AA populations.

All these restrictions make the selection of labeling sites challenging. Hand picking is tedious and one easily misses potentially useful sites because the approach is easily biased. We have developed a systematic in silico method that makes use of two or more protein structures that differ in conformation and applies distance and accessibility restraints to all possible residue pairs. If only one high-resolution structure is available, the

Department of Biochemistry, Groningen Biomolecular Sciences and Biotechnology Institute & Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands. email: b.poolman@rug.nl

combination with a carefully designed homology model of the other conformation should also work. Our method reduces the number of possibilities drastically and allows to focus on the biological restraints rather than the technical ones to obtain the best possible residue pairs. We tested the method by designing pairs for labelling in the substrate-binding domain (OpuAC) of the ABC transporter OpuA and performed single-molecule FRET measurements and functionality assays of the full-length protein complex.

## Results

**The ABC-transporter OpuA.** The protein that we use to showcase the *in silico* approach is the osmoregulatory ABC transporter OpuA. Its substrate-binding protein (OpuAC) undergoes a conformational change upon binding of glycine betaine. Manually selected labelling positions are already available for this protein, which have been used in previous smFRET studies<sup>10</sup>. Even though the used labelling sites (V360C/N423C) report large differences in distance upon glycine betaine binding and do not affect the binding process, they affect the transfer of substrate from the SBD to the membrane domain of OpuA. In fact, the V360 and N423 are present in the lobes of OpuAC that interact with the transmembrane domain (TMD) of OpuA<sup>19</sup>. Therefore, we aimed to find new residue pairs for smFRET that do not interfere with the docking of OpuAC. We present a general procedure for selecting labelling sites based on a minimum of two protein structures. This method provides a screening of all possible residue pairs and allows smart filtering, prior to performing the actual experiments. The results can be inspected manually, for instance by using knowledge of the activity and structure of the entire protein complex. In our case we manually filtered out regions that would affect the interaction of the substrate-binding domain (here OpuAC) with the TMD of the OpuA complex<sup>19</sup>.

**In silico distance mapping.** Crystal structures of OpuAC in the open (PDB: 3L6G) and ligand-bound closed (PDB: 3L6H) conformation<sup>20</sup> were used as a starting point for the *in silico* distance mapping. In short, a distance map plots the distance between each possible pair of residues, in this case between the two centers of mass of the side chains of the amino acids (Ca in case of glycine). The center of mass of the sidechain was chosen, because it is closer to the site of labelling than Ca and it takes the direction of the sidechain into account; in the script (<https://github.com/MembraneEnzymology/ResiduePairs>) the center of mass is easily changed into Ca, if preferred. This ( $d_{\text{conformationA}}$   $d_{\text{conformationB}}$ ) leads to a symmetrical ( $d_{1,2} = d_{2,1}$ ) area plot (Fig. 1a,b). Next, a distance change map is generated by subtracting the distance map of the second conformation (here the closed state of OpuAC) from the first one (open state) (Eq. 1). This difference map shows the distance shift for each residue pair upon the conformational change that is elicited by the binding of glycine betaine (Fig. 1c).

$$d_{\text{conformationA}} - d_{\text{conformationB}} = d_{\text{conformationAB}} \quad (1)$$

**Filtering of the results.** The three obtained maps (Fig. 1a–c) contain all possible pairs of residues, which can now be used to apply restraints in a mathematical way to the residue pairs. First, we select for distances within the predefined range (*e.g.* as set by the  $R_0$  value of the FRET pair), by discarding pairs with distances that are larger or smaller than two threshold values (Eqs. 2 and 3). The resulting pairs are shown in Fig. 1d.

$$d_{\text{min}} < d_{\text{conformationA}} < d_{\text{max}} \quad (2)$$

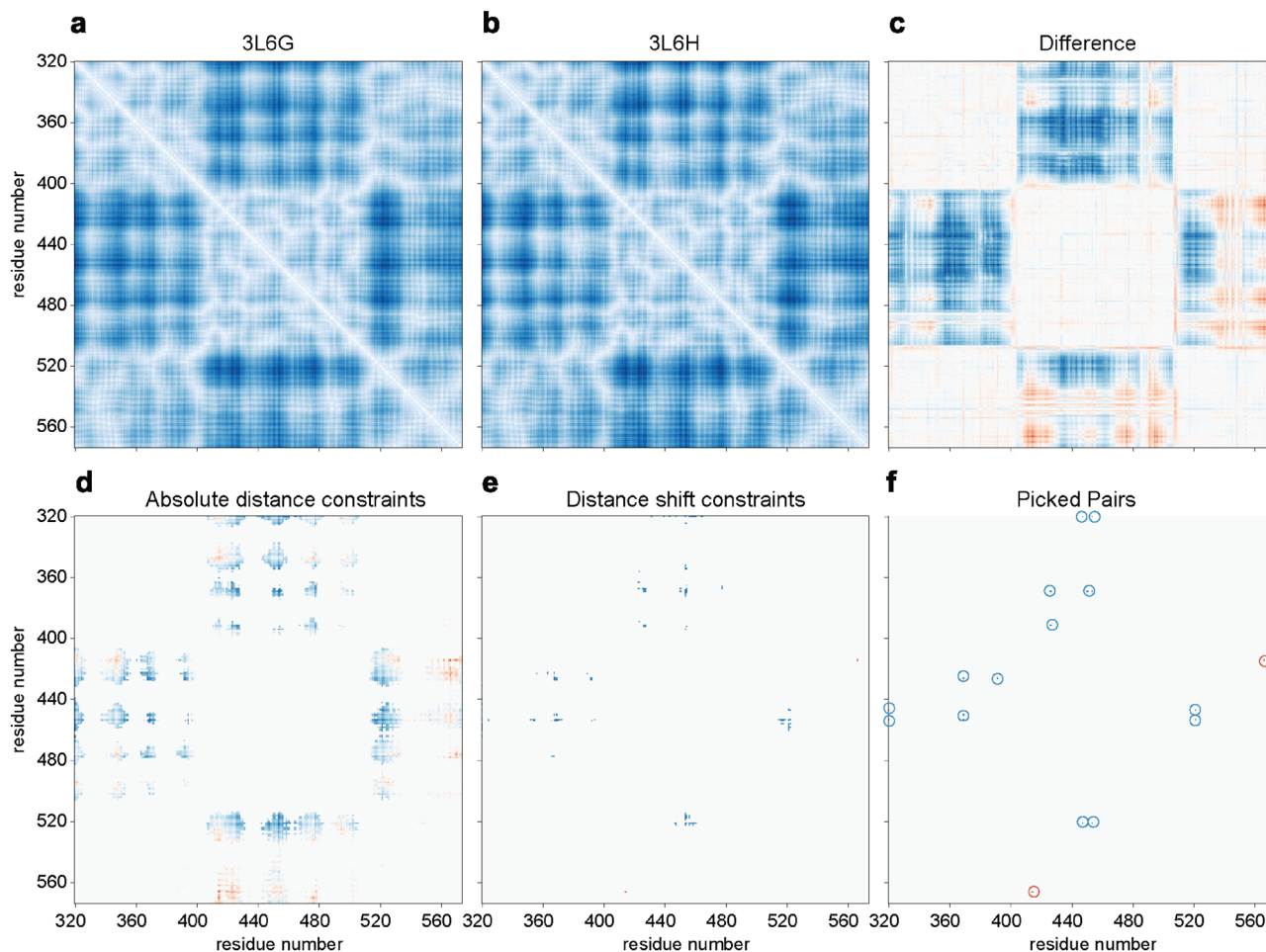
$$d_{\text{min}} < d_{\text{conformationB}} < d_{\text{max}} \quad (3)$$

We then establish a minimum threshold for the distance shift required for the smFRET measurements (Eq. 4). The resulting pairs are shown in Fig. 1e.

$$d_{\text{conformationAB}} > d_{\text{shift-threshold}} \quad (4)$$

Finally, the absolute accessible area is calculated using the DSSP program, which makes use of the structure of the protein to calculate properties as secondary structure, bond and torsion angles and water-exposed surface area<sup>21</sup>. Using solvent exposed residues is important to ensure accessibility for the probe to react but also to allow free rotation of the label. The total accessible area from the DSSP program is then divided by the theoretical total surface area for that residue (used values are the calculated surface area for the amino acid X in a Gly-X-Gly tripeptide from<sup>22</sup>, giving the relative surface accessibility (RSA). The amino acid pairs with a sufficient RSA in both conformations are kept (Fig. 1f), all others are discarded. All the defined thresholds can be adjusted to suit specific needs or to reduce the number of remaining pairs. Similarly, one could easily extend the filtering method based on secondary structure, as labelling of loop regions is typically favoured over structured areas. Secondary structure is also calculated by the DSSP program. A customizable script is available on GitHub: <https://doi.org/10.5281/zenodo.4446814> or <https://github.com/MembraneEnzymology/ResiduePairs>.

In the case of OpuAC we used the following thresholds:  $d_{\text{min}} = 40 \text{ \AA}$ ,  $d_{\text{max}} = 80 \text{ \AA}$ ,  $d_{\text{shift-threshold}} = 8 \text{ \AA}$ ,  $\text{RSA} = 60\%$  to obtain 9 pairs, shown in Fig. 2a,b. These pairs were exported to PyMOL for manual inspection, where we aimed for pairs located on the sides of OpuAC that do not interfere with the docking of the substrate-binding domain in the full-transporter complex (Fig. 2c). We selected two pairs, one with a positive FRET change upon glycine betaine binding (D320C/K453C) and one with a negative FRET change upon binding (N414C/K566C). We also include a pair (T504C/K521C) with a positive FRET signal and a low relative surface accessibility ( $\text{RSA} = 8\text{--}11\%$  for Thr504). The parameters of all three mutants plus the original mutant (V360C/N423C) are shown in Table 1.

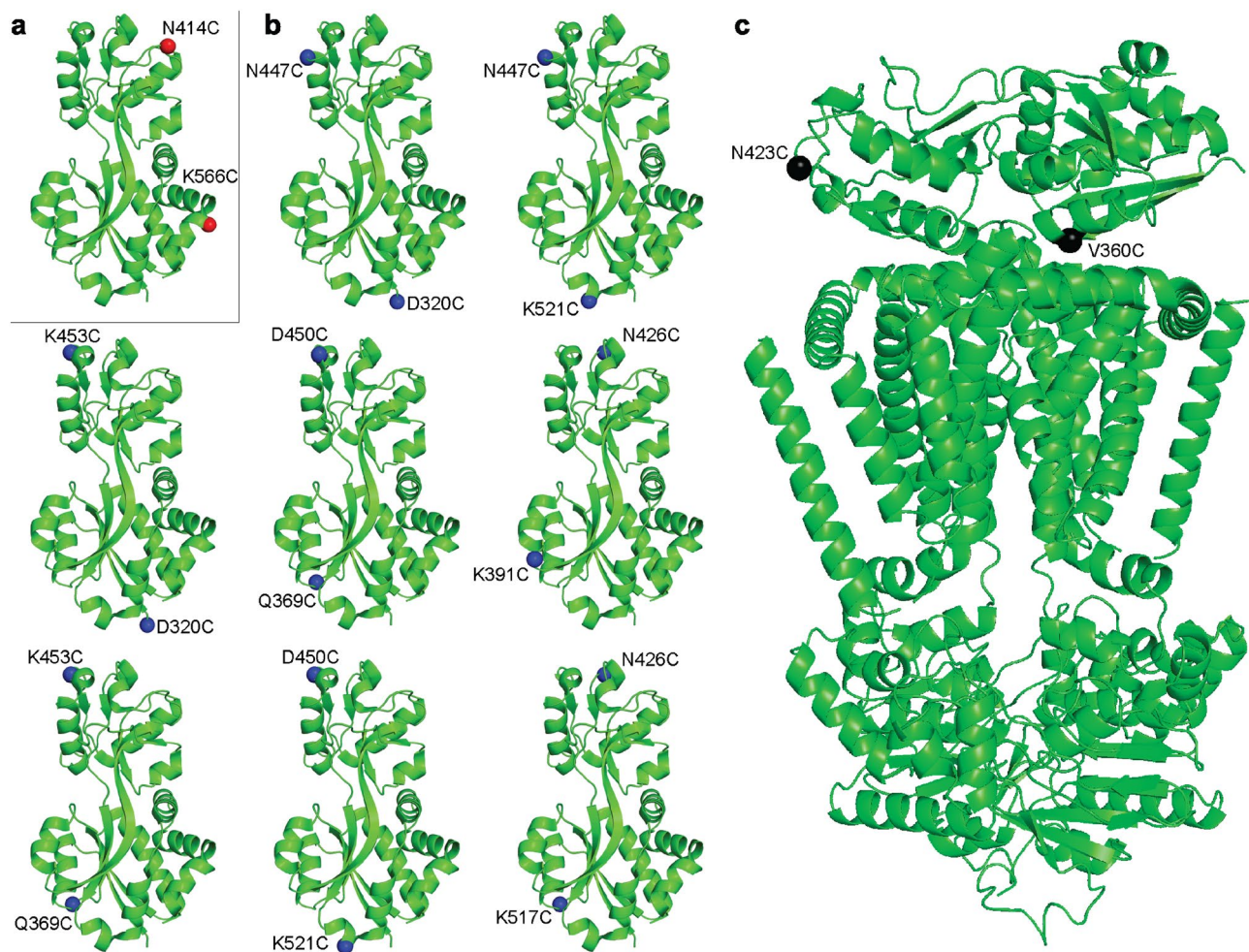


**Figure 1.** Maps, showing the distance between the centers of mass of the sidechains ( $C_{\alpha}$  for glycine) of all possible residue pairs in the (a) Open conformation (PDB: 3L6G) and (b) Closed liganded conformation (PDB: 3L6H) of OpuAC. Panel (c) shows the distance change when transitioning between the two conformations; the orange color indicates a decreasing distance upon binding of glycine betaine to the open conformation; the blue color shows an increasing distance upon ligand binding. Panel (d) The same as panel c but now the pairs with an absolute distance larger than 80 Å and smaller than 40 Å are filtered out. Panel (e), the same as panel d but now pairs with a distance change smaller than 8 Å are filtered out. Panel (f), the same as panel e but now all residues that are less than 60% surface-exposed are filtered out; circles are drawn around pairs/clusters to increase visibility.

**Mutations in OpuAC, the substrate-binding domain of OpuA.** The three newly selected mutants (Table 1) were first constructed in the SBD of OpuA, which were expressed as water-soluble proteins (named OpuAC) and purified to homogeneity. Glycine betaine titrations were performed to assure normal function of the mutant and fluorophore-labelled proteins. OpuAC with double cysteines were labelled with the fluorescence donor (Alexa555) and acceptor (Alexa647), using maleimide derivatives of the dyes. Glycine betaine titration of these labelled OpuAC mutants (Fig. 3a) was monitored by solution-based alternating laser excitation (ALEX) single-molecule FRET (Fig. 3b). Indeed, we see the FRET signal decreasing in OpuAC (N414C/K566C) and increasing in the other two mutants. The mutant (T504C/K521C) that was predicted to be least surface accessible (8–11%) also showed a glycine-betaine dependent conformational change, however, the apparent dissociation constant ( $K_D$ ) of 38  $\mu\text{M}$  is an order of magnitude higher than reported for the wildtype protein and the (D320C/K453C) and (N414C/K566C) mutants<sup>10,20</sup>. Moreover, a low surface accessibility may influence the rotational freedom of the labels. Although we cannot say with certainty that the increased  $K_D$  of OpuA (T504C/K521C) is due to the labelling of the buried Thr-504, we believe that the RSA is a valuable parameter to restrain in the initial selection of labelling sites. We propose to lower the restraints when the number of pairs is too low, but the labelled protein should always be tested for functionality. The (D320C/K453C) and (N414C/K566C) mutants show  $K_D$  values in the same range (1–4  $\mu\text{M}$ ) as reported for the wildtype protein and were used for further studies.

**Mutations in the full-length transporter OpuA.** Next, we verified that the labelling positions do not interfere with the activity of the full-length transporter. OpuA has two SBDs covalently linked to the transmembrane domain, and therefore four cysteines per complex. The three mutant pairs were constructed in the full-



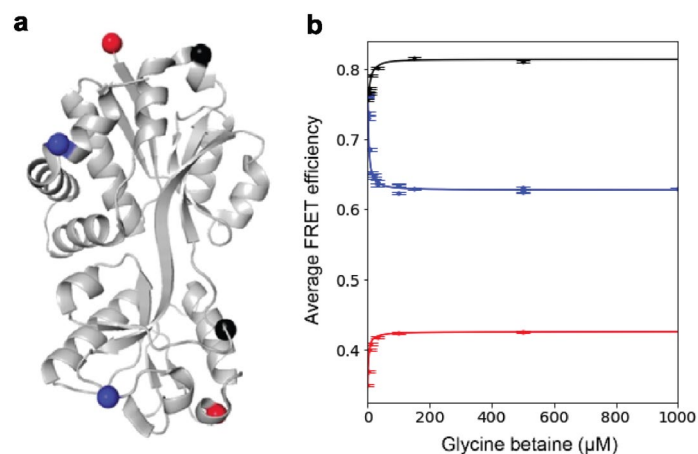


**Figure 2.** The 9 selected residue pairs shown in cartoon representation in OpuAC (PDB: 3L6G) (from left to right, top to bottom: N414C/K566C, D320C/N447C, N447C/K521C, D320C/K453C, Q369C/D450C, K391C/N426C, Q369C/K453C, D450C/K521C and N426C/K517C). The location of the amino acid pairs for dye labeling are shown as spheres. (a) Red spheres indicate a decreasing distance upon glycine betaine binding. (b) Blue spheres indicate an increasing distance upon binding. (c) Shows the positions of the Cys residues (V360C/N423C) previously used for labeling of OpuAC in the full-length protein (PDB: 7AHD), see Sikkema et al<sup>19</sup> for the full-length structure of OpuA; labeling of these residues affects the docking of the SBDs and therefore they cannot be used for smFRET studies of full-length OpuA.

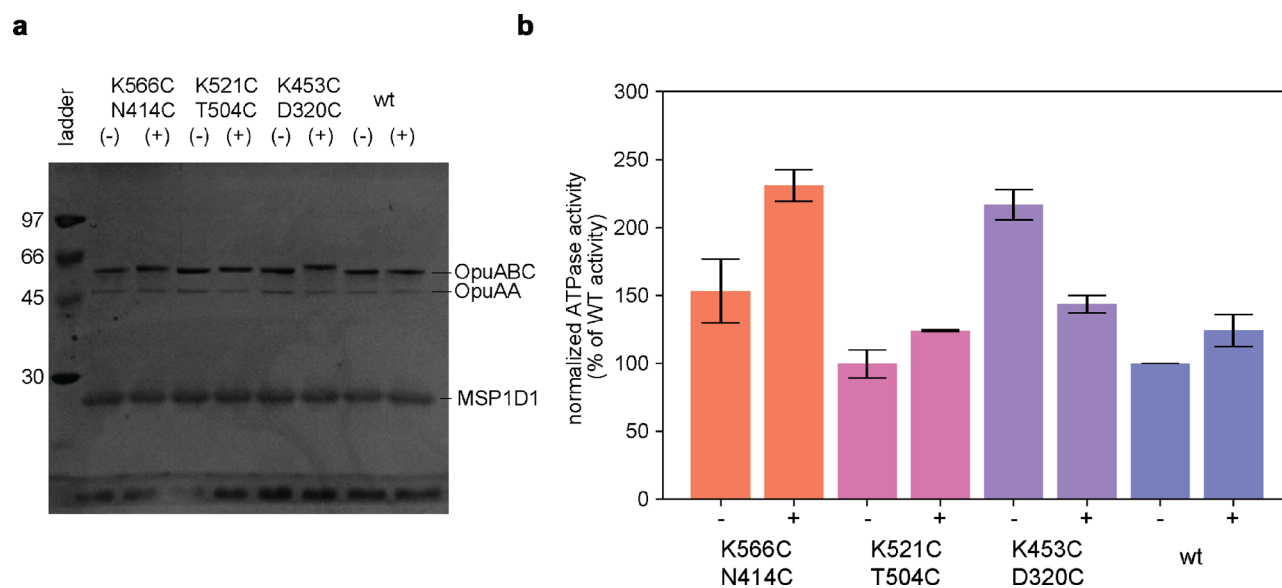
Residue pair	$\Delta d_{AB}$	$d_A$	$d_B$	$RSA_{1B}$	$RSA_{2B}$	$RSA_{1A}$	$RSA_{2A}$
D320C/K453C	9.4	67.3	57.9	91.7	90.7	105.7	88.6
T504C/K521C	6.4	47.8	41.4	8.8	88.1	11.4	91.7
N414C/K566C	-8.4	40.0	48.5	79.8	72.5	77.7	77.2
V360C/N423C	9.6	48.5	39.0	29.5	60.1	35.8	59.6

**Table 1.** Distance and surface accessibility parameters for the original and newly selected mutants.  $\Delta d_{AB}$  is  $\Delta d_{\text{conformationAB}}$  in Å,  $\Delta d_X$  is  $\Delta d_{\text{conformationX}}$  in Å, and  $RSA_{nX}$  is the relative surface accessibility of residue n in conformation X and %.

length transporter and the proteins were purified and reconstituted in MSP1D1 nanodiscs. After reconstitution, half of the nanodiscs were labelled with 4-acetamido-4'-maleimidylstilbene-2,2'-disulfonic acid (AMdiS) and the other half of the sample was used as control. Like the fluorophores used for smFRET, AMdiS is a relatively bulky water-soluble maleimide but unlike the dyes it is affordable for large-scale protein labelling. We used SDS-PAGE gel electrophoresis to show that the proteins are quantitatively labelled with AMdiS, which is apparent from a significant shift in the migration of the OpuABC subunit of the OpuA complex (Fig. 4a). All fractions were then analysed for ATPase activity using a coupled enzyme assay (Fig. 4b). We do not want to interpret the



**Figure 3.** The three newly selected SBD mutation pairs that were constructed for OpuAC. **(a)** Structural representation (PDB:3L6G); red: (D320C/K453C), black: (T504C/K521C) and blue: (N414C/K566C). **(b)** FRET change upon binding of glycine betaine in the FRET pairs depicted in panel (a); the color coding is the same as in the structure; error bars represent the standard deviation obtained from the Gaussian fit in the smFRET data analysis.

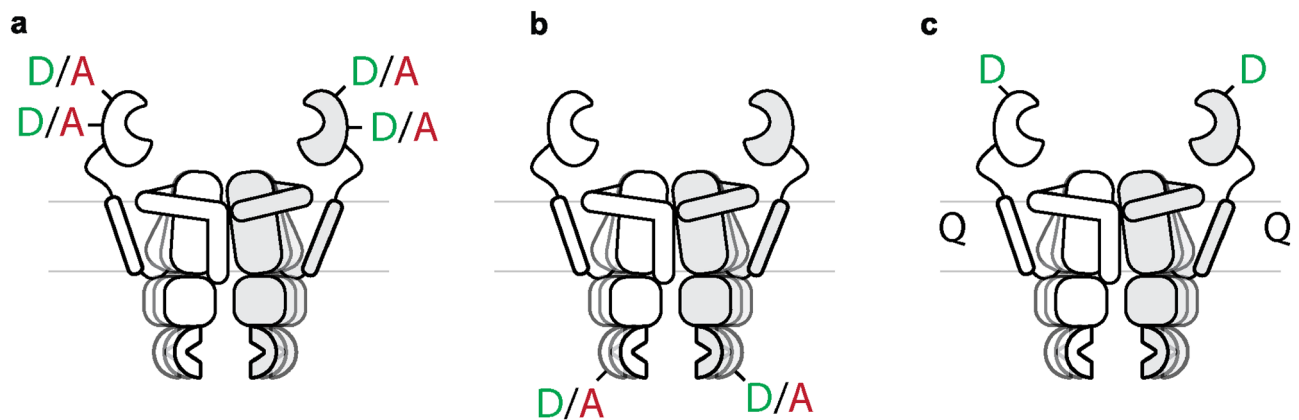


**Figure 4.** SBD mutations in the full-length transporter **(a)** SDS-PAGE gel showing OpuA nanodiscs of the three mutants and wild-type OpuA before (-) and after (+) labeling with AMdiS (4-acetamido-4'-maleimidylstilbene-2,2'-disulfonic acid) **(b)** Normalized ATPase activity in the presence of 10 mM MgATP, 62 μM glycine betaine and 300 mM KCl of the same nanodiscs as shown in (a). The ATPase activity was corrected for the  $A_{280}$  absorbance of the nanodiscs (contribution from the OpuA subunits and the MSP1D1 scaffolding protein) and normalized against the activity of unlabeled wild-type OpuA. Datapoints are shown as the average of two technical replicates. Error bars represent the standard deviation.

apparent increase (2 out of 3) or decrease (1 out of 3) in activity upon treatment of the OpuA nanodiscs observed with AMdiS because of uncertainties in the protein concentration. We infer that the glycine betaine-dependent ATP hydrolysis activity of the labelled mutants is comparable to that of the wildtype protein. Since OpuA labelling is better than 90% (unlabelled OpuABC is barely visible in Fig. 4a), we conclude that the OpuA mutants D320C/K453C, T504C/K521C and N414C/K566C are suitable candidates for future studies on the conformational dynamics of OpuA.

## Discussion

We describe a straightforward approach to select sites for labelling of proteins for smFRET or EPR measurements. One can use proteins similar to the one used to showcase the approach, for instance, the receptor or substrate binding domains associated with ABC transporters, tripartite tricarboxylate transporters (TTTs),



**Figure 5.** (a) Stochastic labeling of a homodimer with 2 cysteine residues per protomer, leading to  $2^4$  possible species. (b) Stochastic labeling of a homodimer with one cysteine residue per protomer to probe inter-protomer distances.  $2^2$  possible species (because the twofold symmetry, both DA and AD species, thus 50% of all species, are available for FRET) (c) Labeling of a single cysteine residue per protomer with a fluorescence donor, a fluorescence quencher is added to the lipid bilayer.

tripartite ATP-independent periplasmic transporters (TRAP), some ligand-gated ion channels (LGI), metabotropic receptors (GPCRs) or 2-component regulatory systems<sup>23</sup>. In these classes of proteins alone, already in 2016 (last structural classification<sup>23</sup>) there were over 500 structures available. However, in principle, the method is not limited to these proteins, but can be used for any system, provided at least two structures or homology models in different conformations are available.

Like OpuA, many of the above-mentioned proteins are homodimeric with more than one SBD per functional complex, hence multiple pairs of cysteine residues are present per complex, complicating the smFRET analysis (Fig. 5a). By introducing a single cysteine per domain and stochastic labeling, hence two cysteine residues per complex in case of a homodimeric complex, it will be possible to observe interdomain movements (Fig. 5b), as has been shown for the ABC-transporter BtuCD by labelling the transmembrane domains<sup>24</sup>, the ABC-transporter MRP1 by labelling the NBDs<sup>25</sup>, but also for the ABC-transporters MsbA<sup>6</sup> and McjD<sup>7</sup>. A similar approach has been used in smFRET studies on BetP, a homotrimeric protein with three fluorophores per complex<sup>8</sup>. Alternatively, one could label the protein with a fluorescence donor and introduce a fluorescence quencher in the ligand or membrane to probe conformational dynamics. In another study on the ABC transporter BtuCD, the cobalt ion in the ligand (vitamine B<sub>12</sub>) has been used for quenching to determine transfer of the substrate from the SBD through the TMD<sup>9</sup>. By inserting a quencher in for instance the vesicle or nanodisc membrane, and a fluorescence donor in the SBD, one could determine the conditions under which the SBD gets closer to or further away from the membrane (Fig. 5c).

To facilitate smFRET measurements in homodimeric proteins such as OpuA with multiple identical subunits, it should be possible to create apparent heterodimeric complexes with e.g. one protomer containing the double cysteine mutation and one protomer being cys-less. One can then probe the opening and closing of the SBD in the context of the full-length transporter and e.g. determine if the two SBDs of OpuA deliver substrates stochastically or that a receptor domain once bound can deliver multiple substrates. We aim to take this approach in future studies, building on the work described in this paper. In short, we describe a systematic method to find candidates for FRET, EPR or other double mutation-based distance-reporting methods that can be used to make a pre-selection of suitable pairs using relevant distance and solvent accessibility constraints.

## Material and methods

**Residue selection protocol.** For the residue selection protocol, we recommend to follow the instruction of the script (<https://doi.org/10.5281/zenodo.4446814>). In short: two protein structures (different conformations) are read using the ProDy Python library<sup>26</sup>. Distance maps, containing the distance between the center of mass of the side chain (C $\alpha$  for glycine) of each pair of residues for each protein structure. The difference map of these two distance maps is generated by subtraction. By selection of distances within a specified range, the amino acid pairs are filtered and only the pairs with a suitable distance (e.g. depending on the probes used for FRET or EPR) between the residues are kept. The DSSP software (version 3.0.0–2)<sup>21,27</sup> is used to assess the secondary structure and surface accessibility for each of the residues. By filtering based on surface accessibility the number of possible pairs is further reduced. The script returns a list of suitable residue pairs, as well as a script that can be imported in the PyMOL Molecular Graphics System and used for direct visualization of the obtained amino acid pairs.

**Construction of expression strains.** The cysteines were introduced sequentially using Quikchange Mutagenesis and the *Escherichia coli* pREOpuAHis vector. Using restriction cloning (*AlwNI* and *BamHI*), the OpuAC region of the gene, where the mutations were introduced, was transferred into the *Lactococcus lactis*



pNZOpuACHis vector, which was used for expression of OpuAC and derivatives. The resulting plasmids were transformed into the OpuA deletion strain *L. lactis* Opu401.

**Expression of genes.** *L. lactis* Opu401 carrying pNZOpuAHis or pNZOpuACHis were cultivated semi-anaerobically in a 2-L bioreactor at 30 °C, in a rich medium with 2% (w/v) gistex LS (Strik BV, Eemnes, NL), 65 mM Potassium phosphate pH 7.0 supplemented with 1% (w/v) glucose and 5 µg mL<sup>-1</sup> chloroamphenicol. The pH was kept constant at 6.5 by adjusting the medium with 4 M potassium hydroxide. The *nisA* promoter was activated at an OD<sub>600</sub> of 2 by adding 0.05% (v/v) of the culture supernatant of the nisin A producing *L. lactis* strain NZ9700<sup>28</sup>. The cells were harvested by centrifugation (15 min, 6000×g, 4 °C) after 2 h of induction, washed in 100 mM KPi pH 7.0, centrifuged again (15 min, 6000×g, 4 °C), resuspended in ice-cold 100 mM KPi pH 7.0 to a final OD of 100 and stored at – 80 °C until further use.

**Isolation and purification of OpuAC.** The cells were lysed by passing them twice through a high-pressure device at 29 k psi (Constant Systems) in the presence of 100 µg mL<sup>-1</sup> deoxyribonuclease (DNase) and 2 mM MgSO<sub>4</sub>, followed by the immediate addition of 5 mM Na<sub>2</sub>-EDTA (pH 8.0), 1 mM PMSF plus 1 mM DTT for the cysteine mutants after breaking. Cell debris and membrane vesicles were removed by centrifugation (90 min; 125,000×g; 4 °C). The cell lysate was aliquoted in samples of 20 mL and flash-frozen in liquid nitrogen and kept at – 80 °C until further use. 1 mL column volume of Ni<sup>2+</sup>-Sephacryl resin was equilibrated with 12 column volumes of water and 2 column volumes of 50 mM potassium phosphate pH 7.0, 200 mM KCl, 15 mM imidazole and 1 mM DTT. 40 mL of cell lysate was thawed rapidly and supplemented with 10 mM imidazole, 1 mM DTT and the equilibrated resin. The column was poured, drained, and washed with 20 column volumes of 50 mM potassium phosphate pH 7.0, 200 mM KCl plus 1 mM DTT. The protein was eluted in 0.6 column volumes of elution buffer (50 mM potassium phosphate pH 7.0, 200 mM KCl, 500 mM imidazole plus 1 mM DTT) for the first fraction and 0.4 column volumes of the same buffer for the later fractions.

**Labeling of OpuAC for single-molecule FRET.** 100 µL column volume of Ni<sup>2+</sup>-Sephacryl resin was equilibrated with 1–2 column volumes of water and 1–2 column volumes of 50 mM potassium phosphate pH 8.0, 200 mM KCl and 1 mM DTT. 10 nmol of unlabelled OpuAC was added to the column and allowed to bind for 1 h at 4 °C. Then, the column was washed with 10 volumes of 50 mM potassium phosphate pH 8.0, 200 mM KCl to remove the DTT. The column was closed, and 1 mL of the same buffer was added. Next, 50 nmol of Alexa555 and Alexa647 were dissolved in 10 µL water free DMSO and added to the column. The labelling reaction was performed overnight at 4 °C. The next day, the column was washed with 10 column volumes of 50 mM potassium phosphate pH 8.0, 200 mM KCl and eluted in 700 µL of the same buffer supplemented with 500 mM imidazole.

**Isolation and purification of OpuA.** The cells were lysed by passing them twice through a high-pressure device at 29 k psi (Constant Systems) in the presence of 100 µg mL<sup>-1</sup> deoxyribonuclease (DNase) and 2 mM MgSO<sub>4</sub>, followed by the immediate addition of 5 mM Na<sub>2</sub>-EDTA (pH 8.0), 1 mM PMSF plus 1 mM DTT for the cysteine mutants after breaking. Cell debris was removed by centrifugation (15 min; 22,000×g; 4 °C), after which the membranes were harvested in an ultracentrifugation step (90 min; 125,000×g; 4 °C). The membranes were resuspended in 50 mM KPi pH 7.0 supplemented with 20% (v/v)w/v glycerol plus 1 mM DTT for the cysteine mutants to a total protein concentration of 10 mg mL<sup>-1</sup>, flash-frozen in liquid nitrogen and kept at – 80 °C until further use. 0.5 mL column volume of Ni<sup>2+</sup>-Sephacryl resin was equilibrated with 12 column volumes of water and 2 column volumes of 50 mM potassium phosphate pH 7.0, 200 mM KCl, 15 mM imidazole and 0.02% DDM. Membrane vesicles containing OpuA were thawed quickly and diluted to a total protein concentration of 3 mg mL<sup>-1</sup>. The solubilization of OpuA was performed at 4 °C with 0.5% DDM for 60 min, followed by a centrifugation step (20 min, 270,000×g, 4 °C) to remove the insoluble material. The supernatant was diluted 2.5 times and added to the equilibrated resin, after which they were incubated for 1 h at 4 °C. The column was poured, drained, and washed with 20 column volumes of 50 mM potassium phosphate pH 7.0, 200 mM KCl, 20% glycerol, 50 mM imidazole plus 0.02% DDM. OpuA was eluted in 0.6 column volumes of elution buffer [50 mM potassium phosphate pH 7.0, 200 mM KCl, 20% (v/v)w/v glycerol, 500 mM imidazole plus 0.02% DDM] for the first fraction and 0.4 column volumes of the same buffer for the later fractions. The obtained protein was used immediately for reconstitution in nanodiscs.

**Reconstitution of OpuA in MSP1D1 nanodiscs.** The reconstitution procedure was similar to<sup>29</sup>. In short: 1.4 µM of the purified OpuA was mixed with 14 µM purified MSP1D1 scaffold protein and 1.4 mM lipids (lipid composition: 50% DOPE, 12% DOPC, 38% DOPG) in 50 mM potassium phosphate pH 7.0, 4% (v/v) glycerol, 10 mM DDM plus 1 mM DTT to a total volume of 2 mL and was nutated for an hour at 4 °C. Then 2 g of SM2-Biobeads (Bio-rad) were added to adsorb the detergent and this mixture was allowed to incubate overnight. In the morning the supernatant was separated from the Biobeads with a syringe.

**Labeling of OpuA for ATPase assay.** Two times 200 µL column volume of Ni<sup>2+</sup>-Sephacryl resin was equilibrated with 1 to 2 column volumes of water and 1–2 column volumes of 50 mM potassium phosphate pH 8.0, 200 mM KCl and 1 mM DTT. The reconstitution mixture was split in two samples of 2 mL. The mixture was let to bind to the column (1 hr 4 °C), which then was washed with 10–20 column volumes of buffer without DTT (50 mM potassium phosphate pH 8.0 plus 200 mM KCl). Then, 1 mL of the same buffer supplemented with 1 mM AMdiS (4-acetamido-4'-maleimidylstilbene-2,2'-disulfonic acid) was added and let to react for 1 h at

4 °C. The columns were washed with 10 to 20 column volumes of 50 mM potassium phosphate pH 8.0, 200 mM KCl plus 1 mM DTT. 500 µL of the same buffer supplemented with 500 mM imidazole was added, the columns were closed and left O/N. The next day the elution was collected and immediately used for the ATPase activity assay.

**ATPase activity assays.** As described in detail in<sup>29</sup>, we used a coupled enzyme assay with pyruvate kinase (PK) and lactate dehydrogenase (LDH) to determine the ATPase activity of OpuA, which is stoichiometrically coupled to the NADH absorbance decrease at 340 nm. The enzymes PK and LDH were present in excess over OpuA in terms of activity. The NADH absorbance was followed in 96-well plates using a Tecan Spark 10 m plate reader. Each well contains 50 mM KPi pH 7, 300 mM KCl, 4 mM phosphoenolpyruvate, 62 µM glycine betaine, 300 µM NADH, 2.1–3.5 units of pyruvate kinase and 3.2–4.9 units of lactate dehydrogenase. The wells were then supplemented with 100 µL of the elution fraction after labelling, corresponding to roughly 1–5 µM of OpuA. The reaction was started by the addition of 10 mM Mg ATP.

**Single-molecule FRET.** Solution-based smFRET and alternating laser excitation (ALEX) experiments were carried out at 5–25 pM of labelled protein at room temperature in 50 mM KPi pH 7.0 supplemented with 1 mM Trolox and 10 mM MEA for photo stabilization plus the reported glycine betaine concentrations. Microscope cover slides (no. 1.5H precision cover slides, VWR Marienfeld) were coated with 1 mg/mL of BSA for 30–60 s to prevent fluorophore and/or protein interactions with the glass material. The excess BSA was then removed by washing and exchanged with 50 mM KPi pH 7.0. All smFRET experiments were performed with a home-built confocal microscope. In brief, two laser-diodes (Coherent Obis) with emission wavelength of 532 and 637 nm were directly modulated for alternating periods of 50 µs and used for confocal excitation. The laser beams were coupled into a single-mode fibre (PM-S405-XP, Thorlabs) and collimated (MB06, Q-Optics/Linos) before entering an oil immersion objective (60X, NA 1.35, UPlanSAPO 60XO, Olympus). The fluorescence was collected by excitation at a depth of 20 µm. Average laser powers were 30 µW at 532 nm (30 kW/cm<sup>2</sup>) and 15 µW at 637 nm (15 kW/cm<sup>2</sup>). Excitation and emission light were separated by a dichroic beam splitter (zt532/642rpc, AHF Analysentechnik), which is mounted in an inverse microscope body (IX71, Olympus). Emitted light was focused onto a 50 µm pinhole and spectrally separated (640DCXR, AHF Analysentechnik) onto two single-photon avalanche diodes (TAU-SPADs-100, Picoquant) with appropriate spectral filtering (donor channel: HC582/75; acceptor channel: Edge Basic 647LP; AHF Analysentechnik). Registration of photon arrival times and alternation of the lasers was controlled by an NI-Card (PXI-6602, National Instruments). Analysis of the photon arrival times were done as described before<sup>10</sup>. In short, a ‘dual channel burst search’<sup>30</sup> was used to identify fluorescence bursts. The NDA (acceptor emission upon donor excitation), NDD (donor emission upon donor excitation) and NAA (acceptor emission upon acceptor excitation) photocounts were measured per burst and assignments are based on the excitation period and detection channel<sup>31</sup>. The photon counts were corrected for background, where the background counts were estimated by calculating the mean count rate over all bins with more than 20 counts. The apparent FRET efficiency was calculated as  $NDA/(NDA + NDD)$  and the Stoichiometry  $S$  by  $(NDA + NDD)/(NDA + NDD + NAA)$ <sup>24</sup>. Binning the detected bursts into 2D histograms with the apparent-FRET versus Stoichiometry allowed the selection of the donor and acceptor labelled molecules and reduce fluorophore bleaching artifacts<sup>31</sup>. The selected 1D apparent-FRET histograms were fitted, using the method of least squares, with a Gaussian distribution, yielding a 95% confidence interval for the mean.

Received: 4 December 2020; Accepted: 17 February 2021

Published online: 11 March 2021

## References

- Kühlbrandt, W. The resolution revolution. *Science* **343**, 1443–2144 (2014).
- Nakane, T. *et al.* Single-particle cryo-EM at atomic resolution. *Biorxiv* <https://doi.org/10.1101/2020.05.22.110189> (2020).
- Pandey, S. *et al.* Time-resolved serial femtosecond crystallography at the European XFEL. *Nat. Methods* **17**, 73–78 (2020).
- Hofmann, S. *et al.* Conformation space of a heterodimeric ABC exporter under turnover conditions. *Nature* **571**, 580–583 (2019).
- Reginsson, G. W. & Schiemann, O. Pulsed electron-electron double resonance: Beyond nanometre distance measurements on biomacromolecules. *Biochem. J.* **434**, 353–363 (2011).
- Liu, Y., Liu, Y., He, L., Zhao, Y. & Zhang, X. C. Single-molecule fluorescence studies on the conformational change of the ABC transporter MsbA. *Biophys. Rep.* **4**, 153–165 (2018).
- Husada, F. *et al.* Conformational dynamics of the ABC transporter McjD seen by single-molecule FRET. *EMBO J.* **37**, e100056 (2018).
- Jazi, A. A. *et al.* Caging and photoactivation in single-molecule Förster resonance energy transfer experiments. *Biochemistry* **56**, 2031–2041 (2017).
- Goudsmits, J. M. H., Slotboom, D. J. & Van Oijen, A. M. Single-molecule visualization of conformational changes and substrate transport in the Vitamin B12 ABC importer BtuCD-F. *Nat. Commun.* **8**, 1–10 (2017).
- de Boer, M. *et al.* Conformational and dynamic plasticity in substrate-binding proteins underlies selective transport in ABC importers. *eLife* **8**, e44652 (2019).
- Nanda, J. S. & Lorsch, J. R. Labeling of a protein with fluorophores using maleimide derivitization. *Methods Enzymol.* **536**, 79–86 (2014).
- Joseph, B. *et al.* In situ observation of conformational dynamics and protein ligand–substrate interactions in outer-membrane proteins with DEER/PELDOR spectroscopy. *Nat. Protoc.* **14**, 2344–2369 (2019).
- Halbmair, K., Wegner, J., Diederichsen, U. & Bennati, M. Pulse EPR measurements of intramolecular distances in a TOPP-labeled transmembrane peptide in lipids. *Biophys. J.* **111**, 2345–2348 (2016).
- Wood, E. Molecular probes: Handbook of fluorescent probes and research chemicals. *Biochem. Educ.* **22**, 83 (1994).
- Tsvetkov, Y. D., Bowman, M. K. & Grishin, Y. A. *Pulsed Electron-Electron Double Resonance. Pulsed Electron-Electron Double Resonance* (Springer International Publishing, Berlin, 2019).



16. Margeat, E. *et al.* Direct observation of abortive initiation and promoter escape within single immobilized transcription complexes. *Biophys. J.* **90**, 1419–1431 (2006).
17. Lee, N. K. *et al.* Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. *Biophys. J.* **88**, 2939–2953 (2005).
18. Müller, B. K., Zaychikov, E., Bräuchle, C. & Lamb, D. C. Pulsed interleaved excitation. *Biophys. J.* **89**, 3508–3522 (2005).
19. Sikkema, H. R. *et al.* Gating by ionic strength and safety check by cyclic-di-AMP in the ABC transporter OpuA. *Sci. Adv.* **6**, eabd7697 (2020).
20. Wolters, J. C. *et al.* Ligand binding and crystal structures of the substrate-binding domain of the ABC transporter OpuA. *PLoS ONE* **5**, e10651 (2010).
21. Kabsch, W. & Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637 (1983).
22. Tien, M. Z., Meyer, A. G., Sydykova, D. K., Spielman, S. J. & Wilke, C. O. Maximum allowed solvent accessibilities of residues in proteins. *PLoS ONE* **8**, e80635 (2013).
23. Scheepers, G. H., Lycklama a Nijeholt, J. A. & Poolman, B. An updated structural classification of substrate-binding proteins. *FEBS Lett.* **590**, 4393–4401 (2016).
24. Yang, M. *et al.* Single-molecule probing of the conformational homogeneity of the ABC transporter BtuCD. *Nat. Chem. Biol.* **14**, 715–722 (2018).
25. Wang, L. *et al.* Characterization of the kinetic cycle of an ABC transporter by single-molecule and cryo-EM analyses. *eLife* **9**, e56451 (2020).
26. Bakan, A., Meireles, L. M. & Bahar, I. ProDy: Protein dynamics inferred from theory and experiments. *Bioinformatics* **27**, 1575–1577 (2011).
27. Joosten, R. P. *et al.* A series of PDB related databases for everyday needs. *Nucleic Acids Res.* **39**, D411–D419 (2011).
28. Kuipers, O. P., de Ruyter, P. G. G. A., Kleerebezem, M. & de Vos, W. M. Quorum sensing-controlled gene expression in lactic acid bacteria. *J. Biotechnol.* **64**, 15–21 (1998).
29. Karasawa, A. *et al.* Physicochemical factors controlling the activity and energy coupling of an ionic strength-gated ATP-binding cassette (ABC) transporter. *J. Biol. Chem.* **288**, 29862–29871 (2013).
30. Nir, E. *et al.* Shot-noise limited single-molecule FRET histograms: Comparison between theory and experiments. *J. Phys. Chem. B* **110**, 22103–22124 (2006).
31. Kapanidis, A. N. *et al.* Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. *Proc. Natl. Acad. Sci. USA* **101**, 8936–8941 (2004).

## Acknowledgements

We thank Marijn de Boer and Ruslan Vietrov for help with the smFRET measurements, Marco van den Noort for fruitful discussions and Michiel Punter for critical assessment of the programming of the code for the residue selection protocol. The work was funded by an ERC Advanced Grant (ABCvolume; #670578).

## Author contributions

B.P. and H.S. designed the research; H.S. performed the research; B.P. and H.S. analysed data; B.P. and H.S. wrote the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to B.P.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021