NANO LETTERS

Open Access

pubs.acs.org/NanoLett

Letter

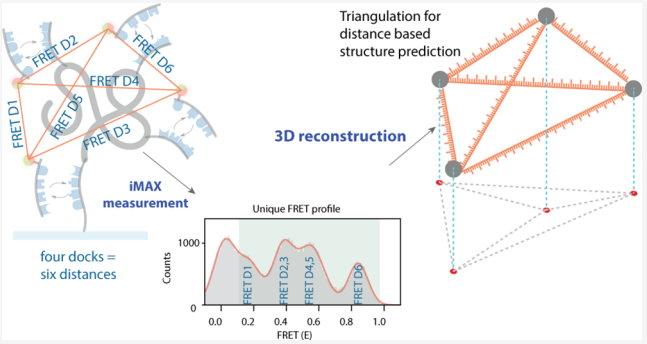
iMAX FRET (Information Maximized FRET) for Multipoint Single-Molecule Structural Analysis

Bhagyashree S. Joshi, Carlos de Lannoy, Mark R. Howarth, Sung Hyun Kim,* and Chirlmin Joo*

Cite This: https://doi.org/10.1021/acs.nanolett.4c00447

Read Online

ABSTRACT: Understanding the structure of biomolecules is vital for deciphering their roles in biological systems. Single-molecule techniques have emerged as alternatives to conventional ensemble structure analysis methods for uncovering new biology in molecular dynamics and interaction studies, yet only limited structural information could be obtained experimentally. Here, we address this challenge by introducing *iMAX* FRET, a one-pot method that allows *ab initio* 3D profiling of individual molecules using two-color FRET measurements. Through the stochastic exchange of fluorescent weak binders, *iMAX* FRET simultaneously assesses multiple distances on a biomolecule within a few minutes, which can then be used to reconstruct the coordinates of up to four points in each molecule, allowing structure-based inference. We demonstrate the 3D reconstruction of DNA nanostructures, protein quaternary structures, and conformational changes in proteins. With *iMAX* FRET, we provide a powerful approach to advance the understanding of biomolecular structure by expanding conventional FRET analysis to three dimensions.



KEYWORDS: *single-molecule FRET, single-molecule structural analysis, single-molecule conformational analysis, computational structure prediction, programmable DNA binding*

Three-dimensional structure dictates the functions of biomolecules.[1] Thus, their analysis is fundamental to understanding their biological functions. Seemingly small perturbations—such as an amino acid substitution, temperature changes, or intramolecular interaction—can lead to profound structural changes, potentially leading to diseased cellular states.[2−6] Analyzing the structures of individual single molecules and complexes is a prerequisite to understanding all cellular functions. However, traditional analysis techniques such as nuclear magnetic resonance and X-ray crystallography determine only the ensemble-averaged structure[7,8] and are unable to capture the structure variation of individual molecules that may underpin crucial biological processes. Furthermore, these methods often impose artificial conditions during measurements (such as crystallization)[9−11] requiring complex methodology.[12,13] Single-molecule techniques such as single-molecule fluorescence resonance energy transfer (smFRET) and single-particle cryoelectron microscopy have emerged as cutting-edge techniques for interrogating structures. While the complex workflow and reliance on specialized experts of single-particle cryoEM hamper its cross-domain adaptability, smFRET is arguably less complex in execution and more accessible. smFRET can measure distances between fluorescent dye pairs attached to a biomolecule in the 2−10 nm range and has been successfully used for conformational and kinetic analyses of biomolecules.[14,15] However, due to the signal complexity, only one or two dye-labeled points in a single molecule can be tracked at a time,[16] precluding a comprehensive understanding of the three-dimensional perspective without prior knowledge of the molecular structure.

In our previous study, we developed FRET X, an extension of conventional smFRET that allows multidistance observations between a single reference position and several monitoring positions within a molecule.[17] Although this multipoint analysis mitigated some limitations of the conventional smFRET, the necessity of a single reference point provides limited information, sufficient only to obtain structural fingerprint of a single molecule but inadequate for *de novo* structural reconstruction. Meanwhile, *de novo* structural reconstruction from smFRET data has been previously demonstrated by using triangulation of positions in three-dimensional space.[18−20] However, the authors separately prepared and measured a series of samples, each designed to

Received: January 26, 2024
Revised: July 2, 2024
Accepted: July 2, 2024

report a distance of different combinations of the points of interest. Consequently, 3D reconstruction was achievable only by combining data sets from ensembles of single molecules. To date, *de novo 3D reconstruction from* a single individual molecule remains unexplored.

We now present information MAXimized FRET (*iMAX* FRET), a one-pot experimental method that measures all possible mutual distance information between multiple points within a single molecule. Unlike hitherto reported methods that require prior structural knowledge to interpret data, *iMAX* FRET is the first method that enables *ab initio* structural analysis solely from smFRET data. The unique "one-pot measurement scheme" for stochastic multipoint sampling, i.e., no multiple repeated measurement with buffer exchange required, is realized by repurposing the probe exchange scheme, which has been utilized in recent studies to overcome photobleaching of organic dyes for long-term kinetics measurement.[21,22] Using our newly developed software pipeline, we show that *iMAX* FRET data can determine up to six distances from four positions in 3D space, from which the conformation of a molecule can be reconstructed through geometrical modeling. iMAX FRET provides a comprehensive understanding of structural heterogeneity within a biological sample.

The Principle of *iMAX* FRET. *iMAX* FRET employs weak binders to rapidly assess multiple points in native biomolecules and heteromeric complexes (Figure 1). In this work, we

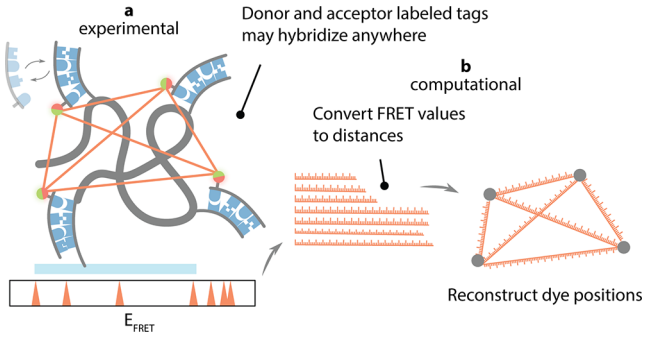


Figure 1. General concept of *iMAX* FRET. **a**, Experimental module. A biomolecule consists of 2−4 coordinates carrying weak binder targets, here DNA docking strands to which cognate imagers can reversibly bind. The imagers are labeled with either donor or acceptor (green as donor and red as acceptor), and they both compete for the binding sites. Each successful FRET event has a particular FRET efficiency (E_{FRET}) between two coordinates and, over time, all possible FRET efficiencies accumulate to give rise to the FRET histogram. **b**, Computational module. The apparent FRET efficiencies (E_{FRET}) for single molecules are converted into distances. They are run through geometrical reconstruction to predict the most optimal fit for the structure. This designates the predicted structure calculated based on the apparent FRET efficiencies.

utilized short single-stranded DNA (ssDNA) as weak binders, taking advantage of their programmable binding kinetics. Specific positions of interest within a protein, nucleic acid nanostructure, or multimeric complex were labeled with ssDNA molecules, referred to as docking strands. These docking strands transiently hybridize with complementary DNA oligos in solution, termed imagers, which are labeled with either a donor or acceptor fluorophore (Figure 1a). As these imager binding events occur stochastically and since each docking strand can serve as both the donor and acceptor binding site, all distances between the target positions can eventually be deduced from single-pair FRET events, where only one donor and acceptor imager pair is bound to the target biomolecule. The lengths and concentrations of the imagers are tuned to ensure that only one FRET pair is observed for a significant fraction of the recording time. The collected FRET values are subsequently translated to distances, which are then fit together in a three-dimensional construct (Figure 1b); all possible three-dimensional constructs using these lengths are generated, and the construct that violates the originally measured lengths the least is considered the correct fit. This method allows for per-molecule three-dimensional reconstruction without any prior knowledge of the identity or structure of the molecule; only basic geometry rules are applied.

One advantage of *iMAX* FRET is its relative ease of implementation. A single round of standard two-color FRET measurement is sufficient to obtain all the necessary structural information, whereas other methods for multiple distance observation often require the inclusion and observation of more dye colors, repeated measurements, or multiple sample preparations with different labelings.[16,23−27]

***iMAX* FRET Can Delineate Single-Stranded DNA Profile.** First, we aimed to assess the feasibility of the simultaneous multidistance measurement with the one-pot stochastic probe exchange scheme using ssDNA as a target molecule carrying multiple docking sequences. We prepared four ssDNA targets each of which contains two or three interspaced copies of an otherwise identical docking sequence at different positions, designated A, B, and C (Figure S1a, sequences in Table S1). Simultaneous binding to positions A and B—spaced 12 nt apart—was expected to yield high FRET, B and C were 16 nt apart which should generate an intermediate FRET, and the summed 28 nt distance between positions A and C should result in a low FRET signal (Figure S1a).

A mixture of donor and acceptor imagers of 8 nt was added to the sample chamber containing immobilized targets. The 8 nt imager with binding dwell time $\tau_{binding}$ = 1.0 ± 0.1 s was adapted from our previous study.[26] We added 10-fold excess of acceptor-labeled imagers over donor-labeled imagers to increase the probability of both fluorophores being present simultaneously. All events were collected from time traces of individual molecules (Figure S1b) and we built a histogram of the averaged FRET value/event (Figure S1c). All DNA samples showed the expected FRET efficiencies of 0.73 ± 0.01, 0.52 ± 0.01, and 0.21 ± 0.02 for positions A, B, and C, respectively (Figure S1c). Notably, the DNA sample carrying all three docking sites showed all three peaks, confirming the capability of our stochastic exchange scheme for simultaneous multipoint assessment.

We noted that the majority of binding events showed FRET efficiency of 0.0 (star, peak area of ∼69%), indicating that donor-only binding events were still dominant (Figure S1c). Thus, the acquisition of sufficient FRET events, i.e., simultaneous binding of a donor and an acceptor imager, required precise adjustment of the binding kinetics of imagers; event duration controlled by imager lengths, and event frequency controlled by imager concentrations. Using Monte Carlo simulations of our experiment (Supplementary Methods) across various concentrations and binding dwell times, we inferred that employing a 10-fold excess of acceptor combined with longer acceptor binding times produced the optimal number of single FRET-pair events (Figure S1d and e).

https://doi.org/10.1021/acs.nanolett.4c00447
Nano Lett. XXXX, XXX, XXX−XXX

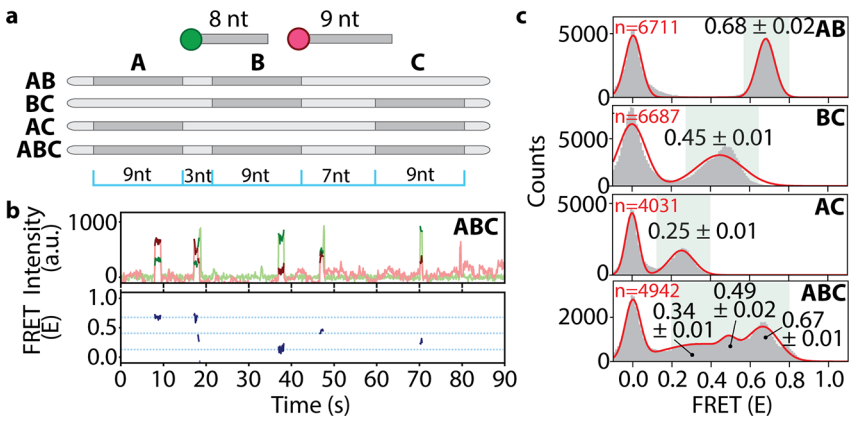


Figure 2. Resolution of three targets in linear DNA using *iMAX* FRET. **a**, Schematic representations of the linear DNA constructs. A, B, and C are the positions of identical docking sequences to which an 8 nt donor- and a 9 nt acceptor-labeled imager can bind. The molar ratio of the donor and acceptor strands were 1:10 (donor: acceptor). AB, BC, and AC are control constructs lacking either one of the three docking sequences, whereas ABC contains all three. The distances between A−B, B−C, and A−C are 12 nt, 16 nt, and 28 nt, respectively. The blue scale explains the distances between A, B, and C. **b**, Representative single-molecule intensity vs time trace (top panel) for the ABC construct (green for donor and red for acceptor intensities). Note that there are three different intensity peaks for the red i.e., acceptor intensity showing FRET events corresponding to successful donor−acceptor imager pair binding to A−C, B−C, and A−C docking sequences. The bottom panel shows the marked FRET efficiencies in blue. The highest blue line corresponds to A−B FRET, the middle line to B−C FRET while the lowest designates the A−C FRET event. **c**, Single-pair FRET event histograms from all molecules in a single field of view (gray bars). The mean FRET ± SEM is given for each peak in the histogram except the peak at 0.0 which corresponds to the donor-only binding events. Red solid lines are multi-Gaussian fit to the histograms. The FRET efficiency of each peak represents the distance between the designated docking sequences. Note the three peaks in the ABC construct corresponding to the three distances for A−C (0.34 ± 0.01), B−C (0.49 ± 0.01), and A−B (0.67 ± 0.01).

Indeed, utilizing longer 9 nt acceptor imager with 8 nt donor imager in the experiment led to a substantial increase in viable FRET events (Figure 2a−c, Figure S1f−g). This demonstrated that careful rational design of imager lengths, and hence dwell times, is pivotal in resolving multiple targets in *iMAX* FRET.

***iMAX* FRET Can Resolve DNA Nanostructures.** To demonstrate *iMAX* FRET's capability of *ab initio* three-dimensional structure determination, we analyzed a quadrangular DNA nanostructure outfitted with a docking strand at each corner (Figure 3a, left). This nanostructure featuring six distinct distances, referred to as D1 to D6 (Figure 3a, right), could be probed with four identical docking strands in *iMAX* FRET in a one-pot reaction. In this demonstration, however, we prepared each docking strand with a unique sequence for control purposes.

First, we probed each distance individually by adding two different imager strands, resulting in a single FRET peak (Figure 3b). We found that D3 and D5 were well-discernible from the other distances (FRET efficiency mean ± standard deviation of 0.83 ± 0.01 and 0.18 ± 0.01, respectively). D1, D2, D4, and D6 generated highly similar FRET values (0.35 ± 0.01 and 0.30 ± 0.01, 0.45 ± 0.01 and 0.37 ± 0.01, respectively). We increased the complexity by adding 3 imager strands for simultaneous analysis of three distances. Indeed, for each of the four possible triangles in this quadrangle, we could identify the expected number of FRET peaks (Figure 3c, panels i−iv). Triangle i (constructed from D1, D2, and D6) showed one major peak, whereas triangle ii (D3, D4, and D6) displayed two overlapping peaks, as expected based on single-distance analysis results. Similarly, triangles iii and iv showed three peaks for (D1, D4, D6), and (D2, D3, D5), respectively.

Next, we probed all six distances simultaneously by adding four different imagers together (Figure 3c, bottom plot) and observed four peaks. The highest at E = 0.84 and the lowest at E = 0.18 represented D3 and D5 respectively. However, the other two peaks were not straightforward to assign due to the overlapping FRET values of the other four distances D1, 2, 4, and 6. Nevertheless, the broad peak at 0.38 could be assigned as a degenerate peak of D1, D2, and D6, while the peak at E = 0.55 likely arose from D4.

Having acquired the distances, the reconstruction of triangle coordinates in Figure 3c is trivial as only one dissimilar triangle (i.e., ignoring rotation, translation, and reflection) can be constructed given the lengths of all three edges. Aligning and averaging triangle coordinates of all single molecules (Figure S2a), produced the shapes of the four triangle types (Figure S2b). To demonstrate that triangles reconstructed for single molecules contain sufficient information to allow recognition, we encoded the coordinates in a rotation-, translation-, and reflection-invariant embedding[28] and trained a tree-based machine learning algorithm to recognize each type. On held-out molecules, this classifier attained an accuracy of 74%, confirming that spatial reconstruction contains discriminative information (Figure 3d). Most errors were made between triangles that were expected to show more similarity due to the nanostructure's assumed symmetry (i+iii, ii+iv respectively). Even so, the classifier still correctly assigned classes to most molecules, indicating that the nanostructure is not truly symmetric.

Finally, we reconstructed the complete 3D quadrangle by using the FRET values obtained from the six-distances measurement in Figure 3c, bottom plot. In theory, 30 dissimilar quadrangles can be constructed given the lengths of all six edges. However, not all quadrangles can necessarily be built without violating the given lengths. An analysis pipeline was written (Supplementary Methods) that builds all possible dissimilar quadrangles and chooses the one for which the edge lengths are required to change the least to fit. We found that a 3D quadrangle could be constructed satisfying all distances without violating the FRET-derived lengths (Figure S3).

Similar to the triangle reconstructions, this 3D reconstruction indicated that the nanostructure has an asymmetric

https://doi.org/10.1021/acs.nanolett.4c00447
Nano Lett. XXXX, XXX, XXX−XXX

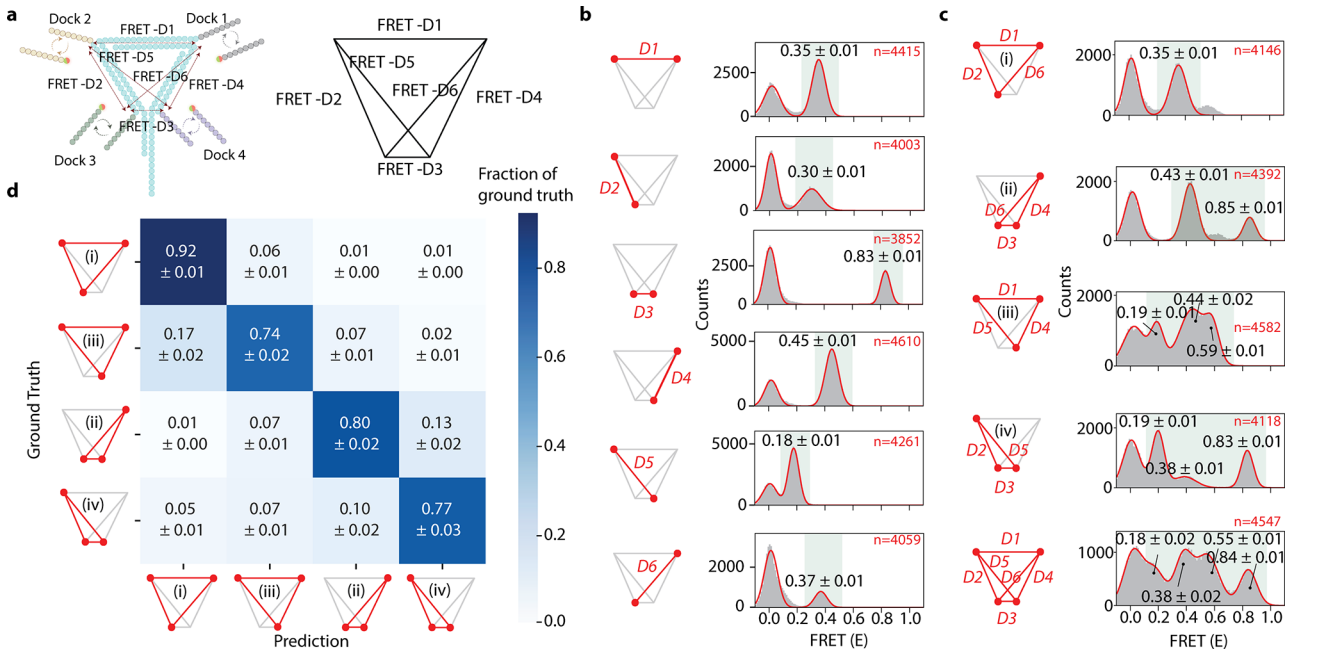


Figure 3. *iMAX* FRET provides structural analysis of a complex DNA nanostructure. **a**, Schematic representation of DNA nanostructure containing 4 overhangs of different DNA sequences which act as docking segments (Docks) for the imagers. As cognate imagers labeled with a donor or acceptor dye bind transiently to the Docks, FRET events occur in proportion to the distances between the Docks. Fifteen bp DNA length separates each pair of Docks 1−2 and 2−3, whereas the 13 bp segment separates the Docks 1−4 giving rise to FRET distances of FRET-D1, D2, and D4, respectively. As a result, Docks 3 and 4 are situated very close to each other giving rise to FRET-D3. The Docks 2−4 and 1−3 also make a pair culminating in FRET-D5 and -D6, respectively. The right panel is the line representation of all the distances generated from the DNA nanostructure and will be used henceforth as a model figure. **b**, *iMAX* FRET histograms of each FRET distance D1 to D6, separately. The red lines signify the FRET distance, red dots represent the Docks. Note that D1, D2, D4, and D6 are similar while FRET D2 and D3 mark the extremes in either direction. The shorter length of DNA (FRET D4, 13 bp) is reflected in slightly higher FRET efficiency (0.45 ± 0.01) as opposed to 2 bp longer FRET D1 and D2 (0.35 ± 0.01 and 0.30 ± 0.01), respectively. This hints at the distorted nanostructure due to differential side lengths. **c**, *iMAX* FRET histograms for combinations of all 3 spatial points forming triangles (i−iv). The red lines signify the FRET distances, red dots represent the Docks. Note that triangle (i) has one mid-FRET degenerate peak due to the three overlapping distances of D1, D2, and D6. Triangle (ii) has one mid-FRET degenerate peak from D4 and D6, and a high-FRET peak arising from D3. The bottom structure contains four peaks with 2 degenerate peaks and 2 single peaks as a result of all the FRET distances D1−D6. **d**, Confusion matrix showing classification accuracy and error modes of a tree-based machine learning classifier trained to identify the four triangles (i−iv) on a single molecule level and tested on held-out molecules. Each row denotes which fraction of total molecules for a given ground truth class are ascribed to which class, where the diagonal denotes correct classifications (i.e., the per-class accuracies).

conformation, rather than a planar symmetric conformation. We surmised that this reflected, in part, the slight out-of-plane attachment positions of the docking strands due to the helical structure of the double-stranded DNA, as well as the flexible carbon linker between the dyes and the DNA. To investigate whether these factors could truly contribute to the observed deviations, we prepared three quadrangles with one of the docking strands positioned at varying positions along the long edge of the structure, and reconstructed triangle shapes for each (Figure S4a−h). The helical nature of the dsDNA should be evident from the respective shapes and sizes of these triangles, as the variably positioned dye moves in and out of the quadrangle plane depending on its distance from the apex. We thus fitted the obtained triangles into the helical structure of dsDNA and calibrated the FRET radius and linker length (Supplementary methods). Satisfactorily, the reconstructed triangles fit the expected dye positions around the DNA helix shape with subangstrom accuracy (Figure S4g, Table S2), indicating that these factors were indeed contributing to the measured quadrangle shape.

In summary, *iMAX* FRET could successfully demonstrate the structural analysis of up to 4 points in a complex DNA nanostructure, and we could predict and retrieve these structural identities with high accuracy based on FRET fingerprints and computational modeling. This demonstrates that we could expand the signal space to 6 peaks (considering degenerate peaks) in a one-pot reaction requiring less than 2 min without using solution exchanges.[26] While the degeneracy of the FRET values may complicate data interpretation, it is worth noting that the number of observed distances provides a clue on the number of degenerate peaks—if the number of detected distances is not a triangular number (1, 3, 6, etc.), this suggests the requirement for the addition of degenerate peaks until the next triangular number. Systematically trying to add multiples of each distance will result in one or more best structures.

***iMAX* FRET Locates the Biotin Pockets in Tetravalent and Divalent Streptavidin Structures.** As *iMAX* FRET is well-suited to determine the relative position of three or more points in space, we set out to study multimeric structures, which are difficult to analyze with traditional FRET due to the inability to control labeling with donor and acceptor fluorophores of subunits within a multimeric protein.[29] Structural analysis of multimeric proteins by other techniques, including mass spectrometry, often requires complex stabilization using chemical linkers or cross-linking.[30−33] In contrast, *iMAX* FRET can be applied to native complexes. Moreover, ligand-binding multimers present a unique possibility for *iMAX*

https://doi.org/10.1021/acs.nanolett.4c00447
Nano Lett. XXXX, XXX, XXX−XXX

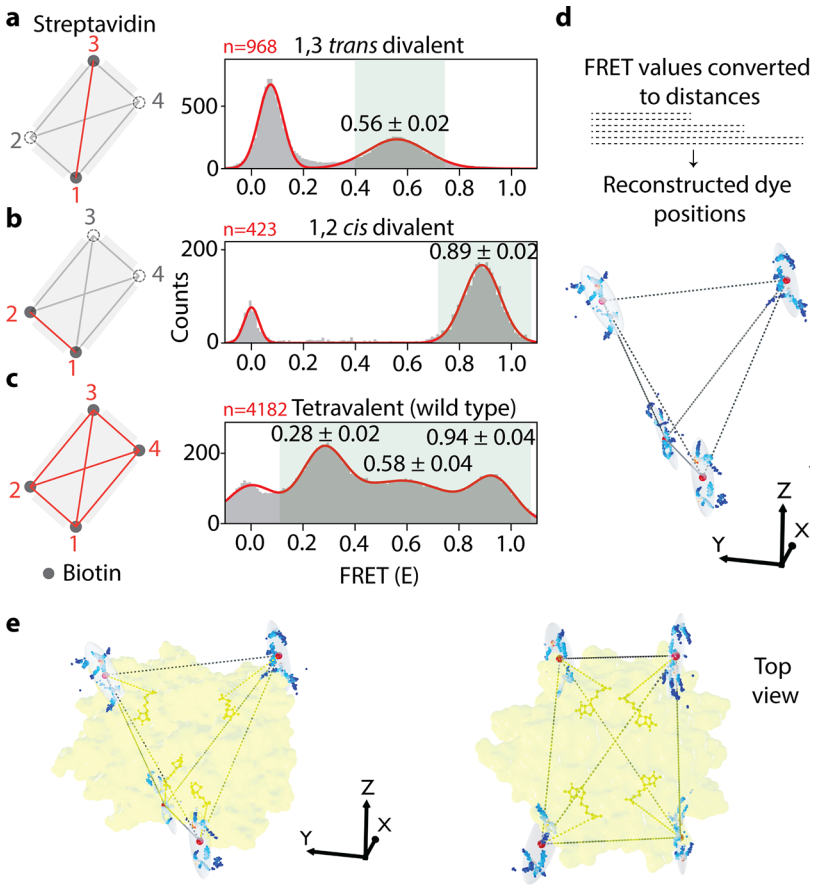


Figure 4. *iMAX* FRET-based structural analysis of streptavidin complexes. **a**, The mutant *1,3 trans* divalent streptavidin (PDB ID: 4BX6) can bind biotin (gray dots) to only two binding pockets, whereas the other two are mutated to abrogate the biotin-binding (dashed circles). Upon binding of an imager, the dye is facing toward the binding pocket. The distance between the bound biotins (red line) shows the FRET efficiency of 0.56 ± 0.02 in the histogram. **b**, The mutant 1,2 *cis* divalent streptavidin (PDB ID: 4BX5) can bind biotin as shown. The distance between the bound biotins (red line) shows a high FRET efficiency of 0.89 ± 0.02. **c**, The wild-type tetravalent with four active biotin-binding pockets. Hence, it can give rise to six distance possibilities $[n(n-1)/2]$. However, streptavidin is a symmetrical molecule, hence shows three degenerate peaks, each peak corresponding to two overlapping peaks. **d**, The FRET values are converted to six distances, and a structure is reconstructed for four biotin-binding pockets. Average positions for 1000 bootstrap iterations over all molecules are shown as dots (colored by density), the mean position is shown as a large red sphere and ovals report one standard error intervals (on average, 2.75 Å). **e**, The reconstructed structure is fitted into the reported (PDB ID: 2IZF) crystallographic streptavidin:biotin structure (yellow). Note that all four biotins can be fitted into the biotin-binding pockets with high accuracy.

FRET. For example, we can use docking strand-conjugated ligands to probe the positions of their binding pockets. We chose streptavidin as our model protein, as it contains four pockets for biotin. This also allowed us to indirectly immobilize streptavidin to a surface, by occupying one of its pockets with an immobilized biotinylated docking strand (Figure S5a). The other pockets were occupied by docking strands added to the solution.

Streptavidin is a tetramer organized in a tetrahedral (D2) symmetry with four biotin-binding pockets (Figure S5b). To derive single distances from four pockets, we measured two divalent streptavidin mutants—1,3 *trans* and 1,2 *cis* having only two active biotin binding pockets[34] (Figure 4a and b). As expected, a high FRET peak (0.89 ± 0.04) was observed for 1,2 *cis* and a mid-FRET peak (0.56 ± 0.02) for 1,3 *trans* (Figure 4a and b). Changing the dye positions from one end to another of the imagers proportionately reflected the changes in the FRET values, showing the ability of *iMAX* FRET to pinpoint the biotin-binding pockets accurately (Figure S5c and d). Subsequent *iMAX* FRET analysis on the wild-type streptavidin with four active binding pockets showed three different FRET efficiencies of 0.28 ± 0.02, 0.58 ± 0.04, and 0.94 ± 0.04 seen for pockets 1, 2, and 3, respectively (Figure 4c). Although, six distances were expected, the symmetric tetramer structure of streptavidin could exhibit only three peaks due to degeneracy. Nevertheless, by using these FRET values, we were able to reconstruct the relative positions of the binding pockets (Figure 4d). The reconstructed 3D spatial coordinates fit the known streptavidin structure gratifyingly well, accounting for a realistic average linker length of 1.8 nm, and showed limited variability over 1000 bootstrap iterations (SD of 2.75 Å averaged over all positions, Figure 4e). This confirms that *iMAX* FRET is capable of extracting three-dimensional features from multimeric proteins without the aid of complementary methods or additional information.

***iMAX* FRET Can Be Potentially Used to Analyze Protein Conformational Changes.** Next, we explore the compatibility of *iMAX* FRET with studying conformational changes of proteins without disturbing their activity. Many proteins undergo profound conformational changes upon binding to a ligand. A well-known example is substrate binding domain (SBD)[35] which captures extracellular substrates and

https://doi.org/10.1021/acs.nanolett.4c00447
Nano Lett. XXXX, XXX, XXX−XXX

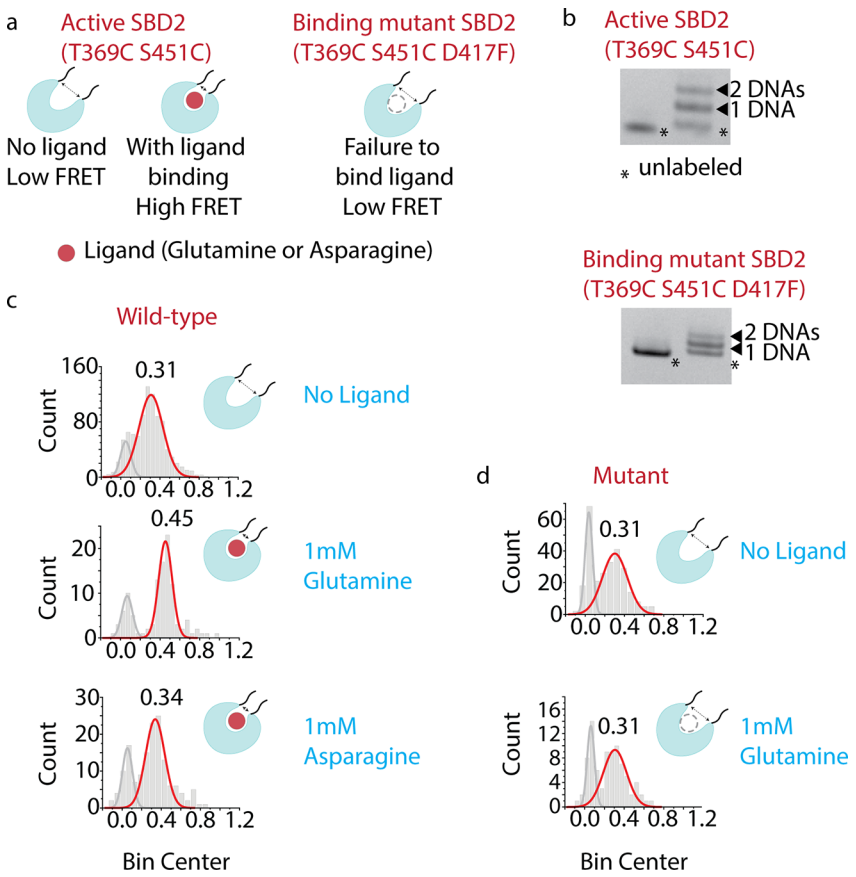


Figure 5. Structural analysis of conformational changes in SBD2-ligand complexes. **a**, 2 mutant SBD2 proteins—active (T369C S451C) and null (T369C S451C D417F). The cysteines are strategically added for DNA labeling. When a cognate ligand is bound, the conformation change results in higher FRET, whereas the null mutant retains the low-FRET value due to a lack of ligand binding. **b**, The SBD2 proteins are labeled with DNA using click chemistry. The ladder pattern suggests the weight shift due to the addition of one or both DNAs attached to the protein. **c**, The SBD2 protein changes its 0.31 FRET value (no ligand) to 0.45 upon its preferred glutamine ligand binding. When Asparagine is added, it stabilizes at 0.34 FRET. **d**, The mutant SBD2, due to the inability of ligand binding, remains at 0.31 FRET after the application of glutamine.

delivers them to transporters. We focused on Gln *PQ* from *Lactococcus lactis*, involved in amino acid sensing and import of asparagine and glutamine.[36,37] Here, we attempted to detect the open-to-closed conformational switch after ligand binding to the SBD2 protein.[38,39]

A wild-type protein that can bind glutamine and asparagine and a null mutant that does not bind any ligand as control were prepared[39] (Figure 5a). For DNA labeling, two cysteines were inserted into both proteins at strategic positions with no known adverse effects,[39] each located at one of the two lobes in SBD2 (Figure 5b). The distance between these two positions undergoes a significant change after a ligand binding according to the crystal structures.[38] Indeed, observed FRET increases from 0.31 to 0.45 and from 0.31 to 0.34 upon binding of Glutamine and Asparagine, respectively (Figure 5c). The smaller FRET shift with Asparagine reflected the fact that SBD2 undergoes a higher conformational change when bound to Glutamine compared to Asparagine.[39] In contrast, we did not observe a FRET shift from the mutant, confirming the FRET shift is indeed induced by ligand binding (Figure 5d). We conclude that *iMAX* FRET with the stochastic DNA probe exchange method can apply to dynamic structural analysis of proteins as a response to stimuli.

In this study, we presented *iMAX* FRET, a pioneering structural analysis tool designed for probing multiple pairwise distances through the utilization of high-resolution smFRET and a weakly interacting probe seamlessly integrating *iMAX* FRET with geometric modeling for structural inference. Our approach facilitates the comprehensive assessment and prediction of molecular architectures leveraging their distinctive FRET signatures with the ultimate sensitivity of a single molecule measurement. This innovative methodology advances the frontier of *ab initio* structural prediction and unlocks new avenues for investigating conformational dynamics heterogeneity within molecular systems. *iMAX* FRET has many advantages over established techniques. First, as we use the stochastic exchange scheme for probing all possible points in a molecule with otherwise identical probes, the imaging time can be cut down considerably compared to other DNA hybridization-based imaging techniques.[26,27,40] Probe-labeled samples can be prepared within 24 h,[26] while weak-binder-based measurement can be completed in as little as 2 min. Second, *iMAX* FRET offers simplicity in sample preparation, circumventing the requirements associated with sample preparation required for other methods, such as crystallization. With only picomolar-range quantities needed, this approach facilitates the analysis of precious samples including patient-derived materials. Additionally, *iMAX* FRET allows for the interrogation of multiple distances in a nano object, including complex DNA nanostructures, proteins, and heteromeric complexes. It paves the way for studying the static and dynamic structural analysis of challenging multimeric proteins

https://doi.org/10.1021/acs.nanolett.4c00447
Nano Lett. XXXX, XXX, XXX−XXX

such as transcription factors and transmembrane proteins. Furthermore, *iMAX* FRET has the potential to provide quantitative insights into the species abundance of multimers and their characteristics within complex mixtures of homo and heteromers. Thus, it can replace cumbersome biochemical assays used to delineate the differential populations of homomers and heteromers present in a particular solution. Lastly, with recent developments on the incorporation of constraints into AlphaFold,[41] the information provided by *iMAX* FRET can be used to attain more accurate de novo structural predictions of any protein. This would be especially useful to resolve conformations of difficult-to-predict dynamic entities such as intrinsically disordered proteins.

Presently, *iMAX* FRET is optimally suited for structural determination within a restricted framework, exemplified by its application to DNA nanostructures and rigid proteins in this study. It is worth noting that DNA conjugation to proteins may induce structural alterations. To avoid this, a rational design strategy for positioning DNA strands on exposed surface of the biomolecules guided by the known structures may be necessary. Further, a functional assay should be performed to confirm that the structure and activity of DNA-conjugated biomolecules are preserved. Moreover, certain challenging scenarios, such as proteins with (1) excessively large or complicated structures, (2) a low or high number of possible labeling points, or (3) considerable degeneracy in the measured distances, may be analyzed over iterative rounds of measurements by utilizing the programmable nature of a probe. Despite these enhancements, *iMAX* FRET would still be limited to molecules where all spatial points under investigation fall within the FRET detectable distance range. Of note is that the range of distances can also be covered by using different FRET pairs, such as Cy2-Cy3 or Cy5-Cy7 pairs, which offers different FRET radii suitable for shorter or longer distances. Further developments should, therefore, include identifying and evaluating widely applicable methods for orthogonal labeling of docking strands or the use of alternative weak binders that do not require conjugation of labels to a protein.

■ ASSOCIATED CONTENT

Data Availability Statement

Data are available upon reasonable request. All codes are documented and freely available at https://github.com/cvdelannoy/iMAX-FRET.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.nanolett.4c00447.

Materials and methods, additional experimental details, bioinformatic analysis including Monte Carlo simulations, and structural prediction pipelines (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Sung Hyun Kim − *Kavli Institute of Nanoscience, Department of Bionanoscience, Delft University of Technology, Delft 2629HZ, The Netherlands; Department of Physics, Ewha Womans University, Seoul 03760, Republic of Korea; New and Renewable Energy Research Center, Ewha Womans University, Seoul 03760, Republic of Korea*; Email: ifolium@gmail.com

Chirlmin Joo − *Kavli Institute of Nanoscience, Department of Bionanoscience, Delft University of Technology, Delft 2629HZ, The Netherlands; Department of Physics, Ewha Womans University, Seoul 03760, Republic of Korea*; orcid.org/0000-0003-2803-0335; Email: c.joo@tudelft.nl

Authors

Bhagyashree S. Joshi − *Kavli Institute of Nanoscience, Department of Bionanoscience, Delft University of Technology, Delft 2629HZ, The Netherlands*

Carlos de Lannoy − *Kavli Institute of Nanoscience, Department of Bionanoscience, Delft University of Technology, Delft 2629HZ, The Netherlands*; orcid.org/0000-0002-3754-0508

Mark R. Howarth − *Department of Biochemistry, University of Oxford, Oxford OX1 3QU, U.K.*; Present Address: Department of Pharmacology, University of Cambridge, Tennis Court Road, Cambridge CB2 1PD, UK; orcid.org/0000-0001-8870-7147

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.nanolett.4c00447

Author Contributions

B.S.J. and C.J. initiated and designed the project. B.S.J. and S.H.K. designed and performed the experiments. S.H.K. designed and ran FRET efficiency extraction workflows and Monte Carlo simulations. C.L. designed and ran 3D modeling and structure classification pipelines. M.H. provided the streptavidin mutants. B.S.J., C.L., S.H.K., and C.J. wrote and edited the manuscript. All the authors read and approved the manuscript.

Notes

The authors declare the following competing financial interest(s): C.J. and B.S.J. hold a patent on single-molecule protein characterization and analysis.

■ ACKNOWLEDGMENTS

We thank Mike Filius for his help and scientific advice. C.J. is supported by The Netherlands Organization for Scientific Research (NWO) (VI.C.202.015), t, Basic Science Research Program (NRF-2023R1A2C2004745), and Frontier 10-10 (Ewha Womans University). S.H.K. was supported by Brain Pool program funded by the Ministry of Science and ICT through the National Research Foundation of Korea (RS-2023-00261876). M.H. acknowledges funding from the Biotechnology and Biological Sciences Research Council (BBSRC, BB/I006303/1).

■ REFERENCES

(1) Dill, K. A.; MacCallum, J. L. The protein-folding problem, 50 years on. *Science* **2012**, *338* (6110), 1042−1046.

(2) Redler, R. L.; Das, J.; Diaz, J. R.; Dokholyan, N. V. Protein Destabilization as a Common Factor in Diverse Inherited Disorders. *J. Mol. Evol* **2016**, *82* (1), 11−16.

(3) Niroula, A.; Vihinen, M. Harmful somatic amino acid substitutions affect key pathways in cancers. *BMC Med. Genomics* **2015**, *8*, 53.

(4) Teng, S.; Srivastava, A. K.; Schwartz, C. E.; Alexov, E.; Wang, L. Structural assessment of the effects of amino acid substitutions on protein stability and protein protein interaction. *Int. J. Comput. Biol. Drug Des* **2010**, *3* (4), 334−349.

(5) Juritz, E.; Fornasari, M. S.; Martelli, P. L.; Fariselli, P.; Casadio, R.; Parisi, G. On the effect of protein conformation diversity in

discriminating among neutral and disease related single amino acid substitutions. *BMC Genomics* **2012**, *13*, S5.
(6) Liu, J. J.; Yu, C. S.; Wu, H. W.; Chang, Y. J.; Lin, C. P.; Lu, C. H. The structure-based cancer-related single amino acid variation prediction. *Sci. Rep* **2021**, *11* (1), 13599.
(7) Shi, Y. A glimpse of structural biology through X-ray crystallography. *Cell* **2014**, *159* (5), 995−1014.
(8) Nogales, E.; Scheres, S. H. Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity. *Mol. Cell* **2015**, *58* (4), 677−689.
(9) Lerner, E.; Cordes, T.; Ingargiola, A.; Alhadid, Y.; Chung, S.; Michalet, X.; Weiss, S. Toward dynamic structural biology: Two decades of single-molecule Forster resonance energy transfer. *Science* **2018**, DOI: 10.1126/science.aan1133.
(10) Henzler-Wildman, K. A.; Thai, V.; Lei, M.; Ott, M.; Wolf-Watz, M.; Fenn, T.; Pozharski, E.; Wilson, M. A.; Petsko, G. A.; Karplus, M.; et al. Intrinsic motions along an enzymatic reaction trajectory. *Nature* **2007**, *450* (7171), 838−844.
(11) Sikic, K.; Tomic, S.; Carugo, O. Systematic comparison of crystal and NMR protein structures deposited in the protein data bank. *Open Biochem J.* **2010**, *4*, 83−95.
(12) Doerr, A. Single-particle cryo-electron microscopy. *Nat. Methods* **2016**, *13* (1), 23.
(13) Nakane, T.; Kotecha, A.; Sente, A.; McMullan, G.; Masiulis, S.; Brown, P.; Grigoras, I. T.; Malinauskaite, L.; Malinauskas, T.; Miehling, J.; et al. Single-particle cryo-EM at atomic resolution. *Nature* **2020**, *587* (7832), 152−156.
(14) Lerner, E.; Barth, A.; Hendrix, J.; Ambrose, B.; Birkedal, V.; Blanchard, S. C.; Borner, R.; Sung Chung, H.; Cordes, T.; Craggs, T. D. FRET-based dynamic structural biology: Challenges, perspectives and an appeal for open-science practices. *eLife* **2021**, DOI: 10.7554/eLife.60416.
(15) Yao, Y.; Docter, M.; van Ginkel, J.; de Ridder, D.; Joo, C. Single-molecule protein sequencing through fingerprinting: computational assessment. *Phys. Biol.* **2015**, *12* (5), 055003.
(16) Clamme, J. P.; Deniz, A. A. Three-color single-molecule fluorescence resonance energy transfer. *ChemPhysChem* **2005**, *6* (1), 74−77.
(17) Filius, M.; van Wee, R.; de Lannoy, C.; Westerlaken, I.; Li, Z.; Kim, S. H.; de Agrela Pinto, C.; Wu, Y.; Boons, G. J.; Pabst, M.; et al. Full-length single-molecule protein fingerprinting. *Nat. Nanotechnol* **2024**, *19* (5), 652−659.
(18) Muschielok, A.; Andrecka, J.; Jawhari, A.; Bruckner, F.; Cramer, P.; Michaelis, J. A nano-positioning system for macromolecular structural analysis. *Nat. Methods* **2008**, *5* (11), 965−971.
(19) Andrecka, J.; Treutlein, B.; Arcusa, M. A.; Muschielok, A.; Lewis, R.; Cheung, A. C.; Cramer, P.; Michaelis, J. Nano positioning system reveals the course of upstream and nontemplate DNA within the RNA polymerase II elongation complex. *Nucleic Acids Res.* **2009**, *37* (17), 5803−5809.
(20) Andrecka, J.; Treutlein, B.; Arcusa, M. A. I.; Muschielok, A.; Lewis, R.; Cheung, A. C. M.; Cramer, P.; Michaelis, J. Nano positioning system reveals the course of upstream and nontemplate DNA within the RNA polymerase II elongation complex. *Nucleic Acids Res.* **2009**, *37* (17), 5803−5809.
(21) Vermeer, B.; Schmid, S. Can DyeCycling break the photobleaching limit in single-molecule FRET? *Nano Res.* **2022**, *15* (11), 9818−9830.
(22) Kummerlin, M.; Mazumder, A.; Kapanidis, A. N. Bleaching-resistant, Near-continuous Single-molecule Fluorescence and FRET Based on Fluorogenic and Transient DNA Binding. *ChemPhysChem* **2023**, *24* (12), No. e202300175.
(23) Hohng, S.; Joo, C.; Ha, T. Single-molecule three-color FRET. *Biophys. J.* **2004**, *87* (2), 1328−1337.
(24) Lee, N. K.; Kapanidis, A. N.; Koh, H. R.; Korlann, Y.; Ho, S. O.; Kim, Y.; Gassman, N.; Kim, S. K.; Weiss, S. Three-color alternating-laser excitation of single molecules: monitoring multiple interactions and distances. *Biophys. J.* **2007**, *92* (1), 303−312.
(25) Uphoff, S.; Holden, S. J.; Le Reste, L.; Periz, J.; van de Linde, S.; Heilemann, M.; Kapanidis, A. N. Monitoring multiple distances within a single molecule using switchable FRET. *Nat. Methods* **2010**, *7* (10), 831−836.
(26) Filius, M.; Kim, S. H.; Severins, I.; Joo, C. High-Resolution Single-Molecule FRET via DNA eXchange (FRET X). *Nano Lett.* **2021**, *21* (7), 3295−3301.
(27) Kim, S. H.; Kim, H.; Jeong, H.; Yoon, T. Y. Encoding Multiple Virtual Signals in DNA Barcodes with Single-Molecule FRET. *Nano Lett.* **2021**, *21* (4), 1694−1701.
(28) Durairaj, J.; Akdel, M.; de Ridder, D.; van Dijk, A. D. J. Geometricus represents protein structures as shape-mers derived from moment invariants. *Bioinformatics* **2020**, *36*, i718−i725.
(29) Sadler, E. E.; Kapanidis, A. N.; Tucker, S. J. Solution-Based Single-Molecule FRET Studies of K(+) Channel Gating in a Lipid Bilayer. *Biophys. J.* **2016**, *110* (12), 2663−2670.
(30) Mendoza, V. L.; Vachet, R. W. Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spectrom Rev.* **2009**, *28* (5), 785−815.
(31) Kiselar, J. G.; Chance, M. R. Future directions of structural mass spectrometry using hydroxyl radical footprinting. *J. Mass Spectrom* **2010**, *45* (12), 1373−1382.
(32) Schneider, M.; Belsom, A.; Rappsilber, J. Protein Tertiary Structure by Crosslinking/Mass Spectrometry. *Trends Biochem. Sci.* **2018**, *43* (3), 157−169.
(33) Yu, C.; Huang, L. Cross-Linking Mass Spectrometry: An Emerging Technology for Interactomics and Structural Biology. *Anal. Chem.* **2018**, *90* (1), 144−165.
(34) Fairhead, M.; Krndija, D.; Lowe, E. D.; Howarth, M. Plug-and-play pairing via defined divalent streptavidins. *J. Mol. Biol.* **2014**, *426* (1), 199−214.
(35) Berntsson, R. P.; Smits, S. H.; Schmitt, L.; Slotboom, D. J.; Poolman, B. A structural classification of substrate-binding proteins. *FEBS Lett.* **2010**, *584* (12), 2606−2617.
(36) Schuurman-Wolters, G. K.; Poolman, B. Substrate specificity and ionic regulation of GlnPQ from Lactococcus lactis. An ATP-binding cassette transporter with four extracytoplasmic substrate-binding domains. *J. Biol. Chem.* **2005**, *280* (25), 23785−23790.
(37) Fulyani, F.; Schuurman-Wolters, G. K.; Zagar, A. V.; Guskov, A.; Slotboom, D. J.; Poolman, B. Functional diversity of tandem substrate-binding domains in ABC transporters from pathogenic bacteria. *Structure* **2013**, *21* (10), 1879−1888.
(38) de Boer, M.; Gouridis, G.; Vietrov, R.; Begg, S. L.; Schuurman-Wolters, G. K.; Husada, F.; Eleftheriadis, N.; Poolman, B.; McDevitt, C. A.; Cordes, T. Conformational and dynamic plasticity in substrate-binding proteins underlies selective transport in ABC importers. *eLife* **2019**, DOI: 10.7554/eLife.44652.
(39) Gouridis, G.; Schuurman-Wolters, G. K.; Ploetz, E.; Husada, F.; Vietrov, R.; de Boer, M.; Cordes, T.; Poolman, B. Conformational dynamics in substrate-binding domains influences transport in the ABC importer GlnPQ. *Nat. Struct Mol. Biol.* **2015**, *22* (1), 57−64.
(40) Kummerlin, M.; Mazumder, A.; Kapanidis, A. N. Bleaching-resistant, Near-continuous Single-molecule Fluorescence and FRET Based on Fluorogenic and Transient DNA Binding. *ChemPhysChem* **2023**, *24*, No. e202300175.
(41) Yuanyuan, Z.; Zicong, Z.; Yuki, K.; Genki, T.; Bowen, Z.; Yi, X.; Daisuke, K. Distance-AF Modifying Predicted Protein Structure Models by Alphafold2 with User-Specified Distance Constraints. *bioRxiv*, December 4, **2023**. DOI: 10.1101/2023.12.01.569498.

Supplementary Materials for

iMAX FRET (Information Maximized FRET) for Multipoint Single-Molecule Structural Analysis

Bhagyashree S. Joshi[a], Carlos de Lannoy[a], Mark R. Howarth[b,c], Sung Hyun Kim[a,d,e*], Chirlmin Joo[a,d*]

[a]Kavli Institute of Nanoscience, Department of Bionanoscience, Delft University of Technology, Delft, 2629HZ, The Netherlands
[b]Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK
[c]Current address: Department of Pharmacology, University of Cambridge, Tennis Court Road, Cambridge, CB2 1PD, UK
[d]Department of Physics, Ewha Womans University, Seoul 03760, Republic of Korea
[e]New and Renewable Energy Research Center, Ewha Womans University, Seoul 03760, Republic of Korea

* Corresponding authors:
Sung Hyun Kim, ifolium@gmail.com;
Chirlmin Joo, c.joo@tudelft.nl

This PDF file includes:

Materials and Methods

Protein expression and purification

Divalent streptavidins were expressed in *Escherichia coli*, refolded from inclusion bodies, and purified by ammonium sulfate precipitation and ion-exchange chromatography, as reported in the original paper[1]. Tetravalent recombinant (wild-type) streptavidin was procured from Thermo Scientific. Plasmids encoding SBD2 (T369C/S451C) and SBD2 (T369C/S451C/D417C)[2] were a generous gift from Prof. Bert Poolman (Department of Biochemistry, University of Groningen, The Netherlands), and the proteins were expressed and purified using the reported protocol [2].

Protein labeling

Cysteine labeling was carried out as reported previously[2] with slight modifications as follows. Cysteine residues of purified proteins (25μM in the total volume of 50μl In PBS) were reduced with 50mM Tris(-2-carboethyl)phosphine (TCEP) at 40-fold molar excess for 30 minutes. Excess TCEP was removed with ZebaTM Spin desalting columns 7kDa MWCO (ThermoFisher) as it may interfere with the Maleimide reaction[3]. The proteins were then labeled with 25-fold molar excess monoreactive maleimide-Dibenzocyclooctyne (DBCO) (Sigma Aldrich) in Phosphate Buffered Saline (PBS) pH 7.4 overnight at room temperature. Excess maleimide-DBCO was removed with Zeba columns and reacted with 10-fold molar excess (ratio 1:10, cysteine to linker) of monoreactive Azidobenzoate-(5') functionalized DNA in PBS pH 7.4 and incubated overnight at room temperature.

Single-molecule Setup

All *iMAX* FRET measurements were performed on a custom-modified prism-type TIRF microscopy setup built around an inverted fluorescence microscope (Nikon, Ti2e)[4]. For illumination of samples immobilized on a quartz slide surface, a 532 nm diode-pumped solid-state laser and 640 nm diode laser (Oxxius, L6Cc) were directed to the surface with an incidence angle below the critical angle via a prism installed above the slide. Fluorescence signals of Cy3 and Cy5 dyes collected by an objective lens (Nikon, CFI Plan Apochromat VC 60X WI) placed below the quarts sample chamber were spectrally divided by a dichroic mirror (Chroma, T635lpxr) after removing scattered laser light by a laser blocking filter (Semrock, NF03-405/488/532/635E-25). The fluorescence signals were further cleared by bandpass filters (Chroma, ET585/65m for Cy3 and ET655LP for Cy5) and imaged on a sCMOS camera (Photometrics, PrimeBSI). The two lasers were operated with a trigger signal generated by the sCMOS camera for ALEX illumination scheme[5]. All the instruments were controlled by using commercial software (NIS elements, Nikon).

Single-molecule flow cell preparation and data acquisition

All single-molecule FRET experiments were performed at room temperature. The flow cells were prepared using our published protocol[6]. Briefly, quartz slides (G. Finkerbeiner Inc) were etched using acidic piranha and passivated with polyethylene glycol (PEG) to minimize any non-specific binding of molecules. mPEG-SVA and PEG-Biotin (Layson Bio) were used for the PEGylation. 50μl of 0.1mg/ml streptavidin (Thermofisher) was incubated into the flow channel for 5min. Excess was removed using 100 μL T50 (50mM Tris-HCl, pH 8.0, 50 mM NaCl). Next, 50 μL of 100pM biotinylated samples was introduced and incubated for 5min in the channel: linear DNA (Fig. 2), triangles (Fig. 3), or biotinylated Anti-His antibody (Fig. S5c and d). Unbound molecules were washed away with 100 μL T50. 100 μL of 10 nM donor labeled imager strands and 100 nM of acceptor labeled imager strands against the sequences under investigation were injected in imaging buffer (50 mM TrisHCl, pH 8.0, 500 mM NaCl, PCD (Merck), PCA (Merck) and 1 mM 6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid (Trolox) (Sigma)). See Table S1 for the full list of docking and imager strands.

Generally, for single-molecule studies, immobilization is carried out by biotin-streptavidin interactions[4]. However, it is highly difficult to precisely control the number of biotin molecules on the traditionally passivated surfaces (with Biotin-PEG). This raises the possibility of 2 or more binding pockets of streptavidin being occupied by biotins on the slide, leaving only one or two for actual fingerprinting. Thus, we modified the immobilization strategy for tetravalent and divalent streptavidin experiments (Fig. 4): Quartz slides were sonicated for >15min in Acetone, Methanol, and finally 1M KOH with washes with MilliQ in between. Next, the slides were flamed using a burner to remove organic residue if any, and immediately placed back in MilliQ. Finally, the slides were dried using a nitrogen blowgun and used for making the flow cell as explained above. The unused slides were stored at RT. 50µl of 1mg/ml of BSA-Azide (Click chemistry tools, 1535) was incubated with 15µl of 100uM (5') DBCO-DNA-Biotin (3') overnight in the dark at room temperature. 10nM of the resultant BSA-DNA-Biotin was added (50µl total volume) to the flow-cell and incubated for 10min. Excess BSA and free DNA were removed with 100µl T50. Next, 50µl of 1nM tetravalent or divalent Streptavidin was added to the channel and incubated for 5min. The excess was washed with 100µl of T50. Next, 100nM biotinylated docking strands were added to the flow cell and incubated for 30 min to ensure the labeling of all the streptavidin pockets. Unbound DNA was washed away with 100 µL T50. Following, 50 µL of 10 nM donor-labeled imager strands and 100 nM of acceptor-labeled imager strands prepared in the imaging buffer were injected into the flow cell.

Single-molecule fluorescence and FRET data analysis

The data collection and analysis were performed in multiple steps as reported previously[6]. A custom Python script was used to extract time traces of individual molecules from a sCMOS image collected at 0.1s exposure time per frame. Two-state K-means clustering algorithm were applied to the Cy3 and Cy5 fluorescence intensity traces to detect individual binding events of fluorescence imager strands. In order to ensure accurate results, binding events lasting for three or more consecutive frames were selected for further analysis. FRET efficiencies were calculated for each imager strand binding event and used to construct the FRET kymograph and histogram. From the events in which the acceptor probe dissociated or photobleached before the donor probe, we calculated the beta (leakage) and gamma correction factors for accurate FRET efficiency calculation following the method reported in a previous study[7]. Gaussian mixture modeling was applied to automatically classify populations in the FRET histogram. The Python-based automated analysis code can be freely accessed at the following link: https://github.com/kahutia/transient_FRET_analyzer2.

Supplementary Text

Monte Carlo simulations

In *iMAX* FRET, which has multiple identical docking sites, the chance of having single-pair FRET events, i.e. simultaneous binding of one Cy3- and one Cy5 probes, largely depends on the probe binding kinetics. Experimentally, the binding frequency and binding dwell time of a probe can be controlled by the concentration and the length of the DNA probe, respectively. To find the optimal condition that maximizes the chance of having FRET events, we carried out series of Monte Carlo simulations at various kinetic rates. We defined a system with three docking sites each of which had three states of 1) probe unbound, 2) Cy3 probe bound and 3) Cy5 probe bound states. Given transition rates of the two probes, each docking site of the system was allowed to freely transit between states 1 and 2 or states 1 and 3, but not between 2 and 3. Each simulation ran for 1-million-time steps from which we typically observed >5000 transitions. We then selected events in which the system entered into the single-pair FRET emitting state, in which only one Cy3 and Cy5 probe were bound among the three docking sites. After removing events that lasted shorter than three consecutive time steps, the number of the selected single-pair FRET events and the total time spent of the system in them were studied to understand the effect of probe binding kinetics. The simulation code was written in Matlab and freely available upon request.

Structure prediction and classification

A computational pipeline for the reconstruction of 3D-shapes and shape classification was implemented in Python 3.9. Briefly, the number of dyes is determined from the number of FRET efficiency values, which are translated to distances. Distances are used to construct all distinct distance matrices (*D*) using pre-computed index matrices. Each distance matrix is then converted to a coordinate matrix as follows[8]. We construct the Gramm matrix (*M*),

$$M_{ij} = \frac{D_{1j}^2 + D_{i1}^2 - D_{ij}^2}{2}$$

where *i,j* are row and column index respectively. After eigenvalue decomposition,

$$M = USU^T$$

the coordinate matrix *X* can be calculated by sorting U and S by descending order of eigenvalue size, taking the first 3 columns of *U* (*U*[:, :3]) and first 3 eigen values (*S*[:3]) and calculating:

$$X = U[:,:3]\sqrt{S[:3]}$$

Poorly fitting distance matrices generate negative eigenvalues and are excluded. Finally, the remaining coordinate matrices are calculated back to distance matrices, and the coordinate matrix for which distances are closest to the original FRET efficiency-derived distances is returned. The algorithm was implemented in numpy (v1.21.5)[9] with distance matrix calculation as implemented in scipy (v1.8.0)[10].

Numerical embedding of 3D shapes for classification was done using the Geometricus package (v0.3.0)[11]. Embedded coordinates were concatenated to the FRET fingerprint, after which a boosted tree classifier implemented using the XGBoost package (v.1.6.1)[12] was trained and tested on the data using a 10-fold cross validation scheme. The analysis code is freely available at https://github.com/cvdelannoy/iMAX-FRET.

Preparation of the custom DNA nanostructure for Förster radius fitting and classifier applicability analysis

The position of docking site 2 was changed to three different locations (Figure S4a) using click chemistry. To achieve this, alkyne handles were introduced into the DNA backbone at three different locations one at a time in separate constructs. The docking strand for site 2 was designed to contain an azide handle at its one end. The alkyne and azide-containing DNAs were reacted using copper-click chemistry. The clicked DNA products (cyan box, Figure S4b) were gel purified and then the triangles were assembled to generate three structurally similar nanostructures (Figure S4a, bottom left). The positional changes between the (variable) docking site 2 and the fixed docking site 3 were reflected in the FRET values (Figure S4c and d), while it remained constant for all the triangles for the undeviating distance between docking sites three and four (Figure S4e). We could similarly recapitulate the change in FRET values in three coordinates i.e. docking sites two, three and four (Figure S4f).

Conversion of FRET efficiency into the distance R

The following sixth-power relation between *R* and *E* was used to calculate the distance based on the experimentally acquired FRET efficiency.

$$E = \frac{1}{1 + (R/R_0)^6}$$

The Förster radius (R_0), a parameter that combines the influence of dye and medium properties, and relative dye orientations, was fitted using the above custom DNA construct with dyes positioned at known locations along one DNA arm (Figure S4g).

DNA structure modeling for Förster radius fitting

The Förster radius (R_0) denotes the dye distance at which the FRET efficiency is 0.5 and constitutes an essential parameter for the accurate calculation of distances from FRET efficiencies[13]. It factors in dye quantum yields and relative orientations, and the refractive index of the medium. In many applications, it suffices to approximate this value as a constant, however in structural biology, this may lead to unacceptable discrepancies with actual distances, as the effect of local environment and setup is ignored. Here we have used an elegant experiment to determine R_0, using our DNA nanostructure. Briefly, we measure FRET efficiencies for four triangles, created by click-chemistry (detailed in Figure S4).

To determine the Förster radius for our experiments, a single side of the DNA nanostructure was outfitted with clicked docking strands at positions 4, 7, and 15th base from a reference position. FRET efficiencies between clicked docking strands, the reference position, and a third position at one of the other angles of the nanostructure were then measured. We then used a parameter optimization approach with a tree-based Parzen estimator (TPE) implemented in the hyperopt package (v.0.2.7)[14] to estimate the Förster radius. Briefly, this algorithm generates randomized proposals for all one or more variable parameters within given ranges and chooses the combination that minimizes the objective function. The TPE constrains the parameter space based on objective values of previous rounds so that the next guess is more likely to return a lower objective value. Using this approach, we simultaneously fitted Förster radius, linker length, and two DNA geometry parameters (twist and axial rise) after 100 iterations. Here, DNA geometry parameters were allowed to vary slightly to account for unnatural stresses in the nanostructure. As an objective function, the squared sum of the difference between the modeled dye position after triangle construction using given FRET efficiencies (see above) and the expected position given the DNA geometry was used. Table S2 denotes ranges, step sizes, and fitted values for all parameters.

Supplementary Figures

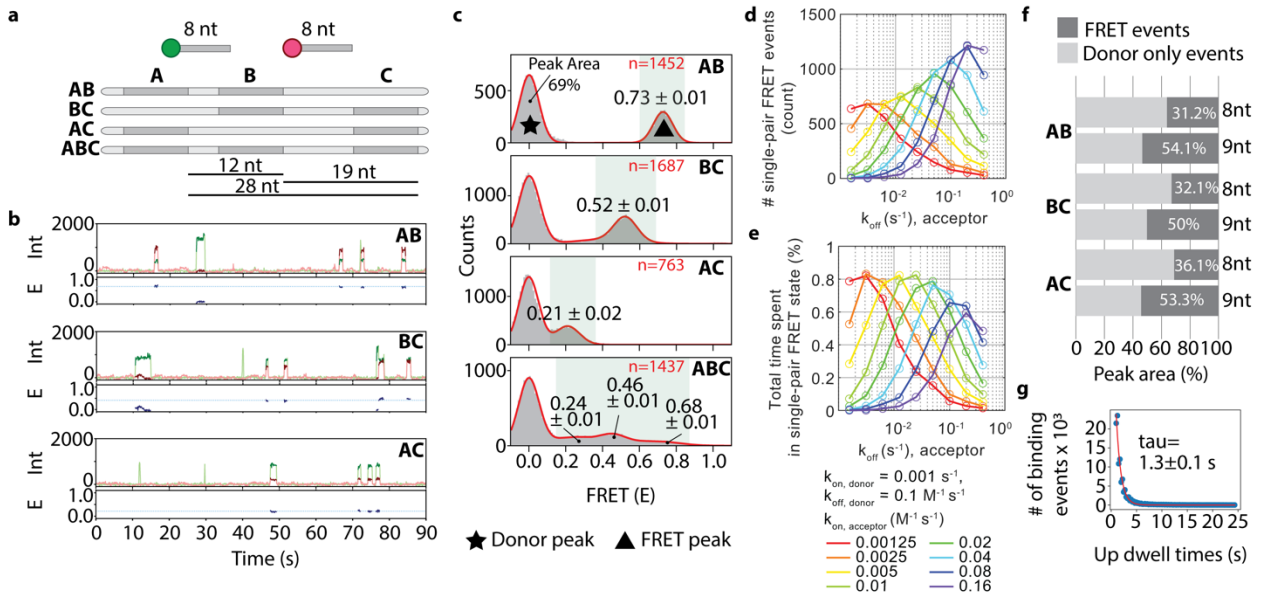


Figure S1: Rational design of the linear construct and imager characteristics for *iMAX* FRET.
a, Schematic representations of the linear DNA constructs. A, B, and C are the positions of identical docking sequences to which 8 nt donor- and acceptor-labeled imagers can bind. The donor to acceptor molar ratio was in a 1:10. The distances between the AB, BC, and AC segments are 12 nt, 19nt, and 28nt, respectively. **b**, Single-molecule intensity time traces for donor (green), acceptor (red) and FRET (blue) for the linear constructs AB, BC, and AC. **c**, Single-FRET event histograms from all molecules in a single field of view. Red solid lines are multi-Gaussian fit to the histograms. The three peaks in the ABC construct correspond to the three distances for A-C (0.24 ± 0.01), B-C (0.46 ± 0.01), and A-B (0.68 ± 0.01) (FRET ± SEM). Star designates the donor-only peak whereas the triangle reports the FRET events peak. **d-e**, The number of single-pair FRET events (d) and the total time spent (e) of a system with three docking strands, obtained from a series of Monte Carlo simulations with various probe binding kinetic rates. Given the donor binding ($k_{on, donor}$ = 0.001) and dissociation ($k_{off, donor}$ = 0.1) rates, the number of single-pair FRET events and total time spent within the state changed substantially with the acceptor binding ($k_{on, acceptor}$) and dissociation ($k_{off, acceptor}$) rates. While the maximum number of events was achieved with higher $k_{on, acceptor}$, the maximum time spent started decreasing when $k_{on, acceptor}$ was more than 10 times higher than that of the donor. At the optimal 10-times higher $k_{on, acceptor}$, $k_{off, acceptor}$ should be ~5-10 times lower than that of the donor to maximize the chance of observing single-pair FRET. **f**, Peak areas for donor only- and single-FRET events for each linear construct are plotted as percentages. Note the increase of 1.5-fold in the FRET events peak area when a longer acceptor imager (9nt) is used instead of an 8nt imager. **g**, Dwell time histogram of the 9nt acceptor imager binding events (Blue circles). The dwell time determined (tau) from single-exponential fit (red line) was 1.3±0.1 s under our experimental conditions (time ± SEM s)

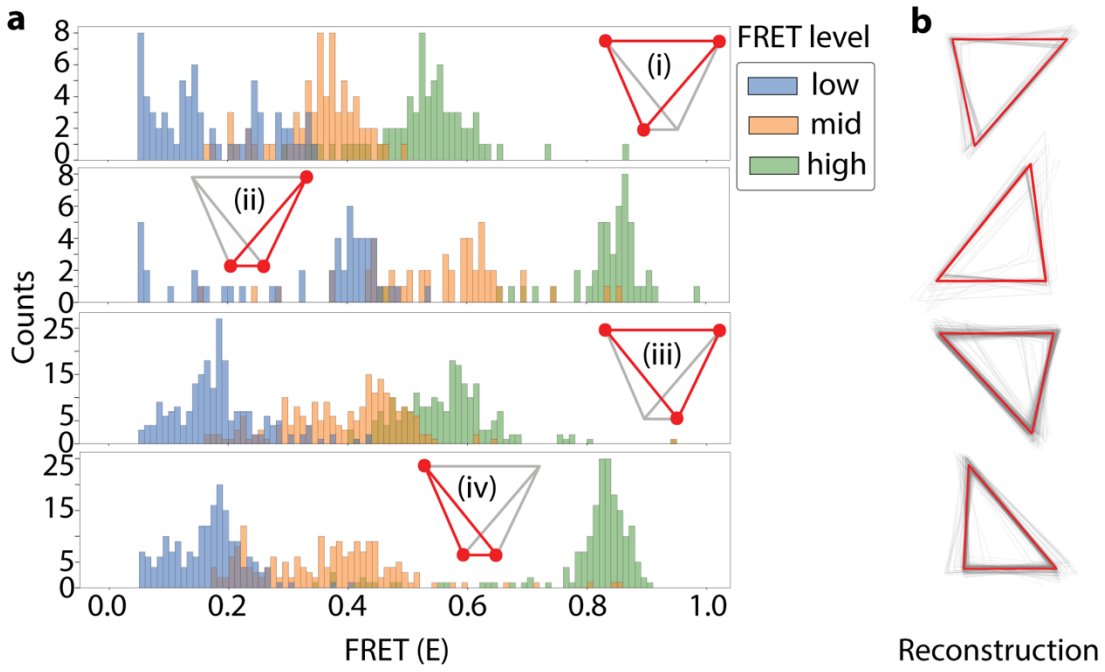


Figure S2: Reconstruction of triangles for single-molecule FRET histograms
a. Histograms displaying per-molecule FRET efficiencies separately for each of the three levels (low, mid, and high) per triangle type in the quadrangular DNA nanostructure (i to iv). Only molecules featuring all three values are shown. **b**. Aligned reconstructed triangles for all single molecules (gray) and the average triangle (red) per triangle type.

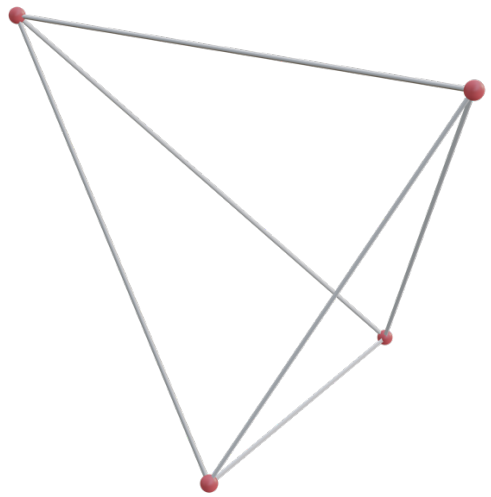


Figure S3: 3D reconstruction of the quadrangular DNA nanostructure
3D reconstruction of the relative dye positions in the nanostructure based on FRET values, revealing its asymmetric and staggered nature.

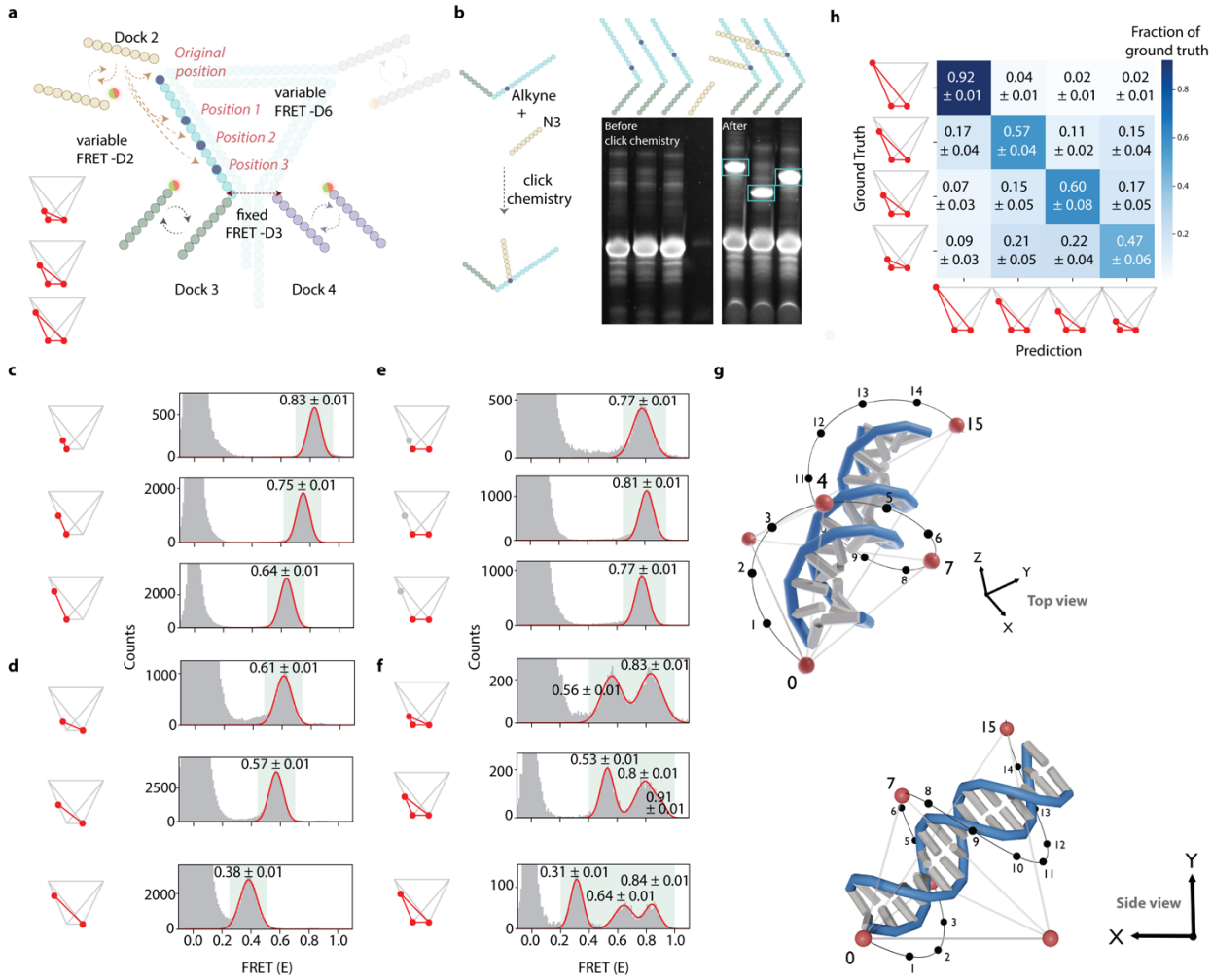


Figure S4: *iMAX* FRET-based analysis of closely related DNA nanostructures

a, In the complex DNA nanostructure, the position of Dock 2 is changed to three different positions giving rise to three FRET-D2 variations. **b**, Click-chemistry is used to attach an azide-linked Dock 3 to the backbone DNA with an alkyne handle. The clicked DNA products (cyan box) were gel extracted and then the nanostructures were reconstituted by hybridization. **c**, The FRET changes between the (variable) Dock 2 and fixed Dock 3 are reflected in the change in the differential position change of Dock 2. **d**, The changes in Dock 2 also changed the distance between sites 2 and 4, as confirmed by the changing FRET values. **e**, FRET values for the distances, between sites 3 and 4, as expected, remained largely unaffected. **f**, The change in FRET values in three points can be similarly recapitulated, for the triangles with imagers and Docks 2,3, and 4. Overall FRET values also shifted for triangles as well for all positions with respect to the original triangle (iv). **g**, 3D reconstruction of dsDNA strand (blue/white) with dye positions (red spheres) of three triangles with the same base reconstructed from FRET efficiencies. Triangles differed in the position of their third dye, which was located at nucleotides with indices 4, 7, or 15, counted from the base. Förster radius, DNA twist, DNA axial rise, and dye-DNA linker length were optimized using a tree-based Parzen estimator-based approach. Black numbers and dots denote expected dye positions and indices for linkers attached to different nucleotides, based on DNA geometry and linker length. Images rendered at two different view angles were generated in Blender (v3.6). **h**, our integrated computational approach can differentiate the 3D structures from each other on a single molecule level with up to 60% accuracy.

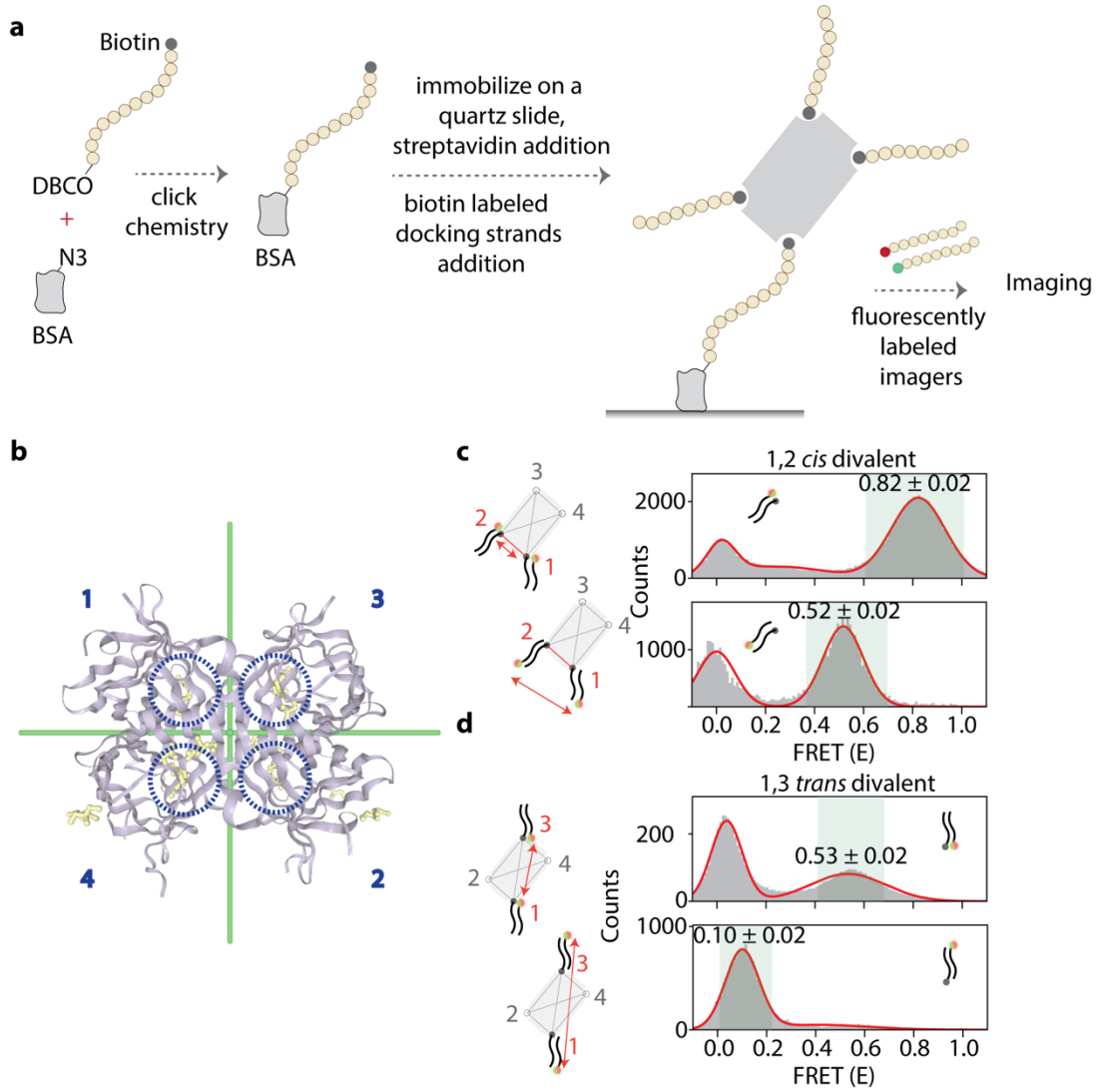


Figure S5: Immobilization scheme of streptavidins and their structural analysis

a, BSA-Azide was immobilized on a quartz slide, conjugated with DNA with a DBCO handle at one end and biotin at the other. The presence of only one Azide per BSA molecule allowed the attachment of one biotin, and thus one streptavidin molecule per BSA molecule. Using this newly developed immobilization scheme, we could ensure that only one pocket is filled with biotin for immobilization and that the remaining 3 pockets are available for binding biotinylated docking sequences for fingerprinting. **b**, D2 symmetry of the wild-type streptavidin tetramer (from PDB ID: 3RY2). 1,2,3 and 4 designate the numbering of subunits in the tetramer. Biotins (yellow space fills) are highlighted with dashed blue circles. **c and d**, With the use of imagers for probing as opposed to covalently conjugated dyes generally used in FRET assays, we could modify the location of dye to artificially change the distance between the 2 points. When 3' instead of 5' dye-labeled imagers were applied, we could see the relative FRET shift to lower efficiencies corresponding to the new distance, on the divalent structures. 1,2 cis divalent streptavidin shows a change of 0.30 FRET value, while it is 0.43 for the 1,3 trans divalent streptavidin mutants.

Supplementary Tables

Table S1: DNA constructs

Corresponding Figure	Description	Sequence (5'-3')	Modification	Supplier
Figure 2	Linear construct with POI-A and B docking sequences	TTTTTTTTTTTTTTTTTTTATAC ATCTATTTATACATCTA	5' Biotin	Ella Biotech (GmbH)
Figure 2	Linear construct with POI-B and C docking sequences	TTTTTATACATCTATTTTTTTA TACATCTATTTTTTTTTT	5' Biotin	Ella Biotech (GmbH)
Figure 2	Linear construct with POI-A and C docking sequences	TTTTTATACATCTATTTTTTTTT TTTTTTTTATACATCTA	5' Biotin	Ella Biotech (GmbH)
Figure 2	Linear construct with POI-A, B and C docking sequences	TTTTTATACATCTATTTTTTTA TACATCTATTTATACATCTA	5' Biotin	Ella Biotech (GmbH)
Figure 2	Donor imager strand	AGATGTAT	3' Cy3	Ella Biotech (GmbH)
Figure 2	Acceptor imager strand	AGATGTAT	3' Cy5	Ella Biotech (GmbH)
Figure 2	Longer acceptor imager strand	TAGATGTAT	3' Cy5	Ella Biotech (GmbH)
Figure 3	DNA Nanostructure backbone + Dock 1	ATTCATTCTCATCCTCTGTCG GGTGTACCGTAAGGTGAAT AGTGACTTTATACATCTA	-	Ella Biotech (GmbH)
Figure 3	DNA Nanostructure Left arm + Dock 2 and Dock 3	AGAGGAGGATTTCGGTACAC CCGACAGTTTTCAATGTA	-	Ella Biotech (GmbH)
Figure 3	DNA Nanostructure right arm + Dock 4	TCTTCATTACTTTTCGATAA CAATCAGGTCACTATTCACC TTA	-	Ella Biotech (GmbH)
Figure 3	DNA Nanostructure biotin strand	CTGATTGTTATCGAGGATGA GAATGAATTTTTTTTTTTTT TTT	Biotin – 3'end labeled	Ella Biotech (GmbH)
Figure 3	DNA Nanostructure donor imager strand Dock 1	AGATGTAT	3' Cy3	Ella Biotech (GmbH)
Figure 3	DNA Nanostructure acceptor imager strand Dock 1	TAGATGTAT	3' Cy5	Ella Biotech (GmbH)
Figure 3	DNA Nanostructure donor imager strand Dock 2	TCCTCCT	5' Cy3	Ella Biotech (GmbH)

Figure 3	DNA Nanostructure acceptor imager strand Dock 2	TCCTCCTC	5’ Cy5	Ella Biotech (GmbH)
Figure 3	DNA Nanostructure donor imager strand Dock 3	TACATTGA	3’ Cy3	Ella Biotech (GmbH)
Figure 3	DNA Nanostructure donor imager strand Dock 3	TACATTGAA	3’ Cy5	Ella Biotech (GmbH)
Figure 3	DNA Nanostructure donor imager strand Dock 4	AGTAATGA	5’ Cy3	Ella Biotech (GmbH)
Figure 3	DNA Nanostructure acceptor imager strand Dock 4	AGTAATGAAG	5’ Cy5	Ella Biotech (GmbH)
Figure S4	Clickable DNA Nanostructure Left arm + Dock 2 Position 2	CGGTA7ACCCGACAGTTTTCA ATGTA	7= C8-Alkyne-dC	Biomers.net (GmbH)
Figure S4	Clickable DNA Nanostructure Left arm + Dock 2 Position 3	CGGTACACC7GACAGTTTTCA ATGTA	7= C8-Alkyne-dC	Biomers.net (GmbH)
Figure S4	Clickable DNA Nanostructure Left arm + Dock 2 Position 4	CGGTACACCCGA7AGTTTTCA ATGTA	7= C8-Alkyne-dC	Biomers.net (GmbH)
Figure S4	Clickable Dock 2	AGAGGAGGATTT	5’ Azide-pro	Biomers.net (GmbH)
Figure 4, Figure S5	Docking strand for streptavidin WT and mutants	ATACATCTA	3’ Biotin	Ella Biotech (GmbH)
Figure 4, Figure S5	Immobilization strand for streptavidin WT and mutants	AAAAGAAAAGAAATACATCT AT	5’ DBCO, 3’ Biotin	Ella Biotech (GmbH)
Figure 5	Docking strand for proteins – SBD2 WT and mutants	TATACATCTAT	5’ Azide-pro	Ella Biotech (GmbH)

Table S2: **Förster radius fitting parameters**

Ranges, step sizes, and fitted values for all parameters fitted by the tree-based Parzen estimator optimization algorithm. Here, the structure diameter spans the DNA strand diameter and two times the linker length.

	Min	Max	Step size	Fitted value
Förster radius (Å)	50	60	0.1	53.8
DNA twist (°/bp)	32	40	1	39
Axial rise (Å/bp)	2.3	5.0	0.1	4.1
Structure diameter (Å)	25	70	0.5	38

References

(1) Fairhead, M.; Krndija, D.; Lowe, E. D.; Howarth, M. Plug-and-play pairing via defined divalent streptavidins. *J Mol Biol* **2014**, *426* (1), 199-214. DOI: 10.1016/j.jmb.2013.09.016.
(2) de Boer, M.; Gouridis, G.; Vietrov, R.; Begg, S. L.; Schuurman-Wolters, G. K.; Husada, F.; Eleftheriadis, N.; Poolman, B.; McDevitt, C. A.; Cordes, T. Conformational and dynamic plasticity in substrate-binding proteins underlies selective transport in ABC importers. *eLife* **2019**, *8*. DOI: 10.7554/eLife.44652.
(3) Kantner, T.; Alkhawaja, B.; Watts, A. G. In Situ Quenching of Trialkylphosphine Reducing Agents Using Water-Soluble PEG-Azides Improves Maleimide Conjugation to Proteins. *ACS Omega* **2017**, *2* (9), 5785-5791. DOI: 10.1021/acsomega.7b01094.
(4) Roy, R.; Hohng, S.; Ha, T. A practical guide to single-molecule FRET. *Nat Methods* **2008**, *5* (6), 507-516. DOI: 10.1038/nmeth.1208.
(5) Kapanidis, A. N.; Laurence, T. A.; Lee, N. K.; Margeat, E.; Kong, X.; Weiss, S. Alternating-laser excitation of single molecules. *Acc Chem Res* **2005**, *38* (7), 523-533. DOI: 10.1021/ar0401348.
(6) Filius, M.; Kim, S. H.; Severins, I.; Joo, C. High-Resolution Single-Molecule FRET via DNA eXchange (FRET X). *Nano Lett* **2021**, *21* (7), 3295-3301. DOI: 10.1021/acs.nanolett.1c00725.
(7) McCann, J. J.; Choi, U. B.; Zheng, L.; Weninger, K.; Bowen, M. E. Optimizing methods to recover absolute FRET efficiency from immobilized single molecules. *Biophys J* **2010**, *99* (3), 961-970. DOI: 10.1016/j.bpj.2010.04.063.
(8) Crippen, G. M., Havel, T. F., . Stable calculation of coordinates from distance information. *Acta Crystallographica Section A* **1978**, *2* (34), 282-284. DOI: 10.1107/S0567739478000522.
(9) Harris, C. R.; Millman, K. J.; van der Walt, S. J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N. J.; et al. Array programming with NumPy. *Nature* **2020**, *585* (7825), 357-362. DOI: 10.1038/s41586-020-2649-2.
(10) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* **2020**, *17* (3), 261-272. DOI: 10.1038/s41592-019-0686-2.
(11) Durairaj, J.; Akdel, M.; de Ridder, D.; van Dijk, A. D. J. Geometricus represents protein structures as shape-mers derived from moment invariants. *Bioinformatics* **2020**, *36* (Suppl_2), i718-i725. DOI: 10.1093/bioinformatics/btaa839.
(12) Guestrin, T. C. a. C. XGBoost: A Scalable Tree Boosting System. *arXiv* **2016**. DOI: 10.1145/2939672.2939785. (accessed 2024.01.26).
(13) Hohlbein, J.; Craggs, T. D.; Cordes, T. Alternating-laser excitation: single-molecule FRET and beyond. *Chem Soc Rev* **2014**, *43* (4), 1156-1171. DOI: 10.1039/c3cs60233h.
(14) Bergstra, J.; Yamins, D.; Cox, D. Making a Science of Model Search: Hyperparameter Optimization in Hundreds of Dimensions for Vision Architectures. In Proceedings of the 30th International Conference on Machine Learning, Proceedings of Machine Learning Research; 2013.