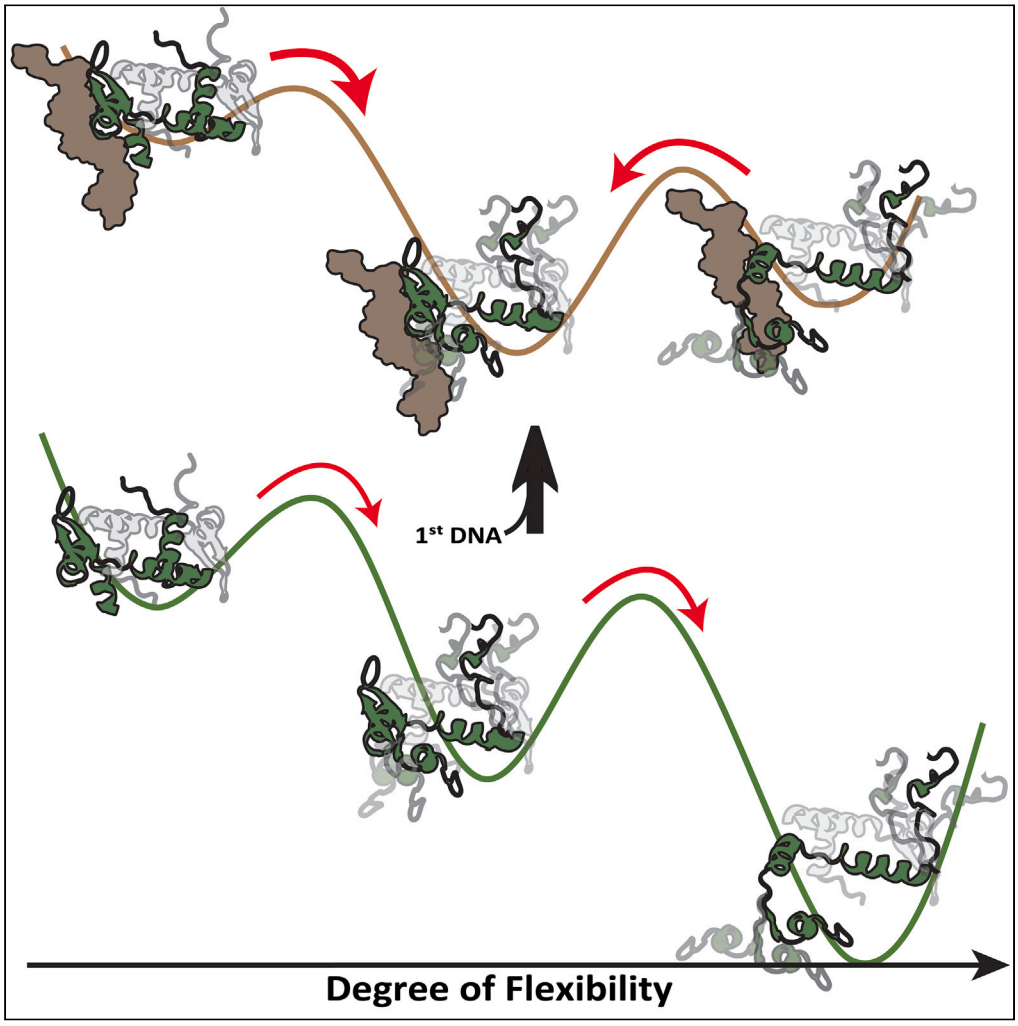
iScience

CellPress
OPEN ACCESS

Article

DNA facilitates heterodimerization between human transcription factors FoxP1 and FoxP2 by increasing their conformational flexibility



Ricardo Coñuecar, Isabel Asela, Maira Rivera, ..., Jorge Babul, Hugo Sanabria, Exequiel Medina

jbabul@uchile.cl (J.B.)
hsanabr@clemson.edu (H.S.)
exequiel.medinago@uchile.cl (E.M.)

Highlights

The forkhead (FKH) domains of FoxP1 and FoxP2 heterodimerize *in vitro*

The FKH heterodimer exhibits compact and extended conformations

The FKH heterodimer promotes extended configurations when bound to DNA

Binding to DNA increases the heterodimerization rate

Coñuecar et al., iScience *26*, 107228
July 21, 2023
https://doi.org/10.1016/j.isci.2023.107228

iScience

CellPress
OPEN ACCESS

Article

DNA facilitates heterodimerization between human transcription factors FoxP1 and FoxP2 by increasing their conformational flexibility

Ricardo Coñuecar,[1] Isabel Asela,[1] Maira Rivera,[2,3,8] Pablo Galaz-Davison,[2,3,8] Jorge González-Higueras,[2,3] George L. Hamilton,[4,8] Felipe Engelberger,[5] César A. Ramírez-Sarmiento,[2,3] Jorge Babul,[1,7,*] Hugo Sanabria,[6,*] and Exequiel Medina[1,6,9,*]

SUMMARY

Transcription factors regulate gene expression by binding to DNA. They have disordered regions and specific DNA-binding domains. Binding to DNA causes structural changes, including folding and interactions with other molecules. The FoxP subfamily of transcription factors in humans is unique because they can form heterotypic interactions without DNA. However, it is unclear how they form heterodimers and how DNA binding affects their function. We used computational and experimental methods to study the structural changes in FoxP1's DNA-binding domain when it forms a heterodimer with FoxP2. We found that FoxP1 has complex and diverse conformational dynamics, transitioning between compact and extended states. Surprisingly, DNA binding increases the flexibility of FoxP1, contrary to the typical folding-upon-binding mechanism. In addition, we observed a 3-fold increase in the rate of heterodimerization after FoxP1 binds to DNA. These findings emphasize the importance of structural flexibility in promoting heterodimerization to form transcriptional complexes.

INTRODUCTION

Transcription factors (TFs) are proteins with large disordered regions[1,2] Click or tap here to enter text.with specific domains that bind to DNA. Interactions with DNA often promote folding and subsequent homo- or heterotypic interactions that occur through different association domains.[3,4] Thus, DNA binding is crucial for TFs to assemble as functional quaternary complexes responsible for gene regulation. However, little is known about the ability of TFs to adopt homo- and heterotypic complexes in the absence of DNA and how DNA binding influences their structural dynamics.

One TF subfamily that adopts homo- and heterotypic complexes without interacting with DNA is the human FoxP, from the Forkhead box of TFs.[5–9] These proteins are crucial in immune, pulmonary, and neurological development. In addition, different mutations in the FoxP's DNA-binding domain (FKH) (Figure 1A) result in severe diseases.[10–13] In addition, brain[14–17] and esophagus[18] development require FoxP1:FoxP2 heterodimerization, whereas their absence or dysregulation results in autism spectrum disorders.[19–23] These antecedents suggest that human development and homeostasis require heterodimerization among FoxP TFs.

A unique property of the FKH domain of the FoxP subfamily of TFs (FoxP1-4) is that they dimerize via the three-dimensional domain swapping (3D-DS)[24] (Figure 1A). In this association process, two monomers exchange portions of their secondary structure content for adopting an intertwined dimer (or oligomer).[24–26] However, it is unclear if FoxP heterodimers use the 3D-DS mechanism. Nonetheless, all FoxP members share a high sequence identity (above 75%) in their FKH domains,[13] suggesting a high possibility of 3D-DS heterodimerization, because this mechanism relies on the conservation of intramolecular contacts at intermolecular fashion (Figure 1A). Moreover, FoxP proteins form heterodimers *in vivo* and *in vitro*, even when they exhibit significant differences in their dimerization equilibria.[5–8,27]

For example, the K_D value of the FoxP2 homodimer is 1,000-fold higher than FoxP1 homodimer,[7,9,27] indicating, at first sight, that the formation of FoxP1-FoxP2 heterodimers must rely on structural features

[1]Departamento de Biología, Facultad de Ciencias, Universidad de Chile, Santiago 7800003, Chile

[2]Institute for Biological and Medical Engineering, Schools of Engineering, Medicine and Biological Sciences, Pontificia Universidad Católica de Chile, Santiago 7820436, Chile

[3]ANID – Millennium Science Initiative Program – Millennium Institute for Integrative Biology (iBio), Santiago 8331150, Chile

[4]Department of Biochemistry and Molecular Pharmacology, New York University School of Medicine, New York, NY, USA

[5]Institute for Drug Discovery, Leipzig University Medical School, 04107 Leipzig, Germany

[6]Department of Physics & Astronomy, Clemson University, Clemson, SC 29634, USA

[7]Senior author

[8]These authors contributed equally

[9]Lead contact

*Correspondence: jbabul@uchile.cl (J.B.), hsanabr@clemson.edu (H.S.), exequiel.medinago@uchile.cl (E.M.)

https://doi.org/10.1016/j.isci.2023.107228

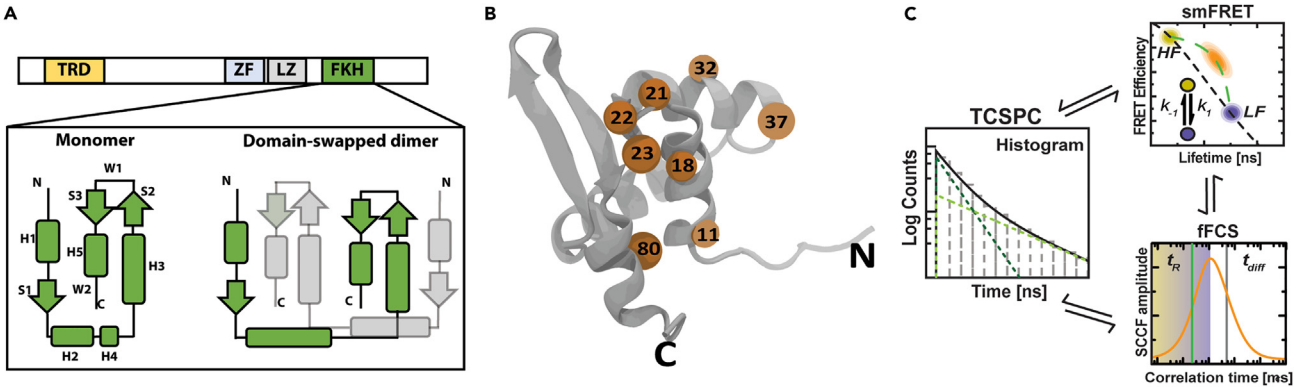


Figure 1. Structure of FoxP1 and the high-resolution experimental approach to determine its dynamics

(A) The complete protein's domains (Trans-Repression -TRD-, Zinc-Finger (ZF), Leucine Zipper (LZ), and Forkhead -FKH-) and the secondary structure of the FKH (in green) are shown.

(B) Structure of the FKH domain (PDB: 2KIU), showing in orange spheres the non-conserved residues between human FoxP1 and FoxP2 (11, 18, 21, 22, 23, 32, 37, and 80). As shown in the topology in (A), most of the FoxP1 and FoxP2 specific residues are located between helices *H*1 and *H*2.

(C) Single-molecule Multiparameter Fluorescence Spectroscopy approach to determine the structural dynamics from nano- to milli-seconds using FRET. Time-correlated Single Photon-Counting (TCSPC) histogram represents all donor-related photons recorded during the experiment to first analyze the FRET population distribution according to a model at nanoseconds timescales. These photons are also averaged into bursts (milliseconds timescales) and then represented as two-dimensional FRET efficiency versus donor lifetime in the presence of acceptor ($\langle\tau_{D(A)}\rangle_f$) (smFRET), to finally analyze their fluctuations (fFCS) to discriminate microseconds exchange. All strategies inform with each other to increase the robustness of the analysis.[24]

present in FoxP1 that favor its association with FoxP2. This is relevant considering that most evolutionary substitutions between human FoxP proteins are in the N-terminal region (helices *H*1-*H*2, Figure 1B), suggesting that localized changes modify the dissociation properties in this subfamily. However, the heterodimerization mechanisms and properties among FoxP TFs, the role of DNA binding on FoxP heterodimerization, and the inherent dynamics required to assemble transcriptional complexes remain elusive.

Previously, we characterized the dynamics responsible for the FKH homodimerization in FoxP1 via high-spatial and temporal resolution toolbox[28,29] (Figure 1C). In this approach that combines molecular dynamics (MD) and single-molecule multiparameter fluorescence spectroscopy (smMFS), we showed the presence of complex and heterogeneous structural changes in the FKH domain that promote the homotypic association.[28] We hypothesize that to promote the heterotypic association, this structural feature should be required. Therefore, the intramolecular behavior of FoxP1 in the heterodimeric complex and the impact of the DNA must be further explored.

In this work, we characterized the structural and functional properties of the heterotypic association between the FKH domain of FoxP1 and FoxP2 and the impact of DNA binding, employing the high-spatial and temporal resolution toolbox[29] (Figure 1C). We first determined computationally that the heterodimerization is favored by FoxP1 by the presence of specific amino acids, supporting their role as an evolutionary gate to modulate heterodimerization. smMFS and biochemical assays helped us identify transitions between a folded state and extended configurations that are predominant in the FoxP1:FoxP2 heterodimer, favoring it 100-fold over FoxP2 homodimerization. Surprisingly, the DNA not only does not promote the folding but increases the extended state, predicting, computationally, that the binding of DNA with FoxP1 in the heterodimer mainly evokes structural fluctuations at short timescales in FoxP1 while favoring heterodimerization.

These findings highlight the importance of structurally flexible complexes and the role of DNA in promoting heterotypic interactions by dynamically modulating local flexibility. This trait could enable FoxP TFs to act as versatile master gene-expression.

RESULTS

FoxP1:FoxP2 heterodimerization propensity prediction at residue-level

A striking aspect of the FoxP subfamily of TFs is its ability to form intertwined dimers in the absence of DNA. Although several studies have demonstrated that the favoring of the dimerization significantly varies for

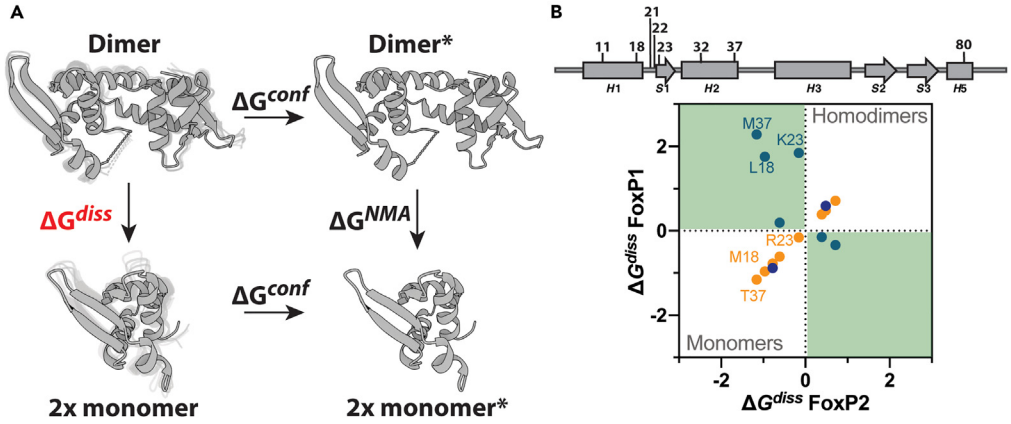


Figure 2. CCR simulations to dissect the per-residue dissociation free energy changes (ΔG^{diss}) in FoxP homodimers

(A) The thermodynamic cycle of the CCR simulations, where the dissociation free energy change (ΔG^{diss}) is obtained by first confining both monomers and dimers (ΔG^{conf}), followed by determining the difference in the free energy between the confined dimer and monomer (ΔG^{NMA}). The figure shows the general process either for FoxP1 or FoxP2.

(B) Dissociation free energy change (ΔG^{diss}) per residue only in residues that differentiate FoxP1 (in blue) from FoxP2 (orange), showing those that strongly stabilize the monomers ($\Delta G^{diss} < 0$) or homodimers ($\Delta G^{diss} > 0$) in both proteins and those stabilizing the dimer only in one protein and thus promote heterodimerization (green quadrants). The secondary structure topology of the DNA-binding domain shows the residue positions in which FoxP1 and FoxP2 have different amino acids.

each member of the subfamily,[5–8,27] these studies do not explain how differences in the sequence identity of FoxP members could promote homodimerization over heterodimerization or vice versa.

To predict how the residues that differ between FoxP1 and FoxP2 in their FKH could promote heterodimerization, we performed CCR (Confine-Convert-Release) simulations[30] and calculated the dissociation energy for both FoxP1 and FoxP2 homodimers (Figure 2A). In these simulations, the confinement of a structural ensemble (monomer and dimer) to its lowest energy state and the calculation of the confined monomer-dimer transition allows us to extract the dissociation free energy change per residue.[30,31] First, we built the homodimer of FoxP1 using the crystal structure of FoxP2 (PDB: 2A07, see STAR Methods) as a template. As a result, for each state (homodimers and monomers), we performed the confinement and then computed the per-residue free energy difference between the confined dimeric and monomeric states in the dissociation reaction (ΔG^{diss}) (Figure 2A). Hence, we use the results from these simulations to predict the relative residue-specific contribution to the dissociation, categorizing them as dimer or monomer stabilizing residues and defining which residue(s) from each protein could selectively stabilize the heterodimer over the homodimer (Figure 2B). Under this assumption, we defined what residues that stabilize the homodimer in one protein but stabilize the monomer in the other can stabilize the heterodimer (Figure 2B).

Our results indicate that the most significant ΔG^{diss} values for both proteins fall in the range of −5 to 5 kcal mol^{-1} (Table S1), which we interpret as stabilizing either the monomeric ($\Delta G^{diss} < 0$) or dimeric state ($\Delta G^{diss} > 0$), respectively. We determined the contribution of specific FoxP1 and FoxP2 residues toward heterodimerization by comparing the ΔG^{diss} of FoxP1 against FoxP2, identifying those residues that stabilize the dimer in FoxP1 but not in FoxP2 and vice versa (Figure 2B). Despite the scarce sequence differences between FoxP1 and FoxP2 in their DNA-binding domain (8%), we observe significant differences in the energetic contributions of these residues. Three of the eight FoxP1-specific residues stabilize the heterodimer, whereas their homologous residues in FoxP2 mainly stabilize the monomer or the homodimer (Figure 2B and Table S1). Based on these results, we rationalized that the L18, K23, and M37 are potential heterodimer-promoting residues from FoxP1, highlighting that the region of strand *S1* is an important secondary structure element that modulates the heterodimerization ability.

Disordered-like ensembles dominate the FoxP1:FoxP2 heterodimer

Previous reports have shown that FoxP1:FoxP2 heterodimerization occurs.[19,32] However, these studies did not characterize the biophysical properties of protein association by observing the isolated DNA-binding

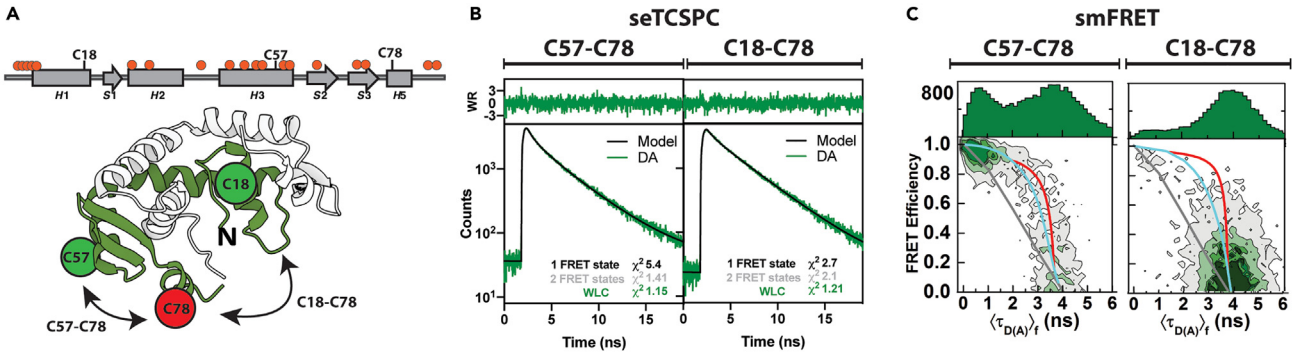


Figure 3. Single-molecule Multiparameter Fluorescence Spectroscopy to dissect the structural dynamics of the heterodimer

(A) Positions of the double-cysteine mutants of FoxP1 monomer to label with donor (green circle) or acceptor (red circle) dyes in the context of the heterodimer (green cartoon; white cartoon is shown only for visualization of each monomer). The DNA-binding regions are shown as orange circles.

(B) sub-ensemble fluorescence decay of both FRET variants (C57-C78 and C18-C78), showing the best fit (worm-like chain model, WLC) in green. In black and gray, 1 and 2 FRET states are shown as a comparison.

(C) smFRET pattern for both FRET variants (C57-C78 and C18-C78), where the lines explain the WLC behavior (cyan) and the dynamics between compact (High FRET, HF) and extended (Low FRET, LF) ensembles (red).

domains. We then wanted to describe the structural changes and dynamics of FoxP1 in the heterodimer with FoxP2, testing first the FoxP1:FoxP2 FKH association.

We used a fluorescent labeled FoxP1 FKH and an unlabeled version of FoxP2 FKH to monitor the association kinetics by fluorescence anisotropy. As controls, we performed the same approach but using either an unlabeled mutant of FoxP1 (A39P), which has been characterized as not forming dimers,[5,7] or a monomer of wild-type FoxP1 (wt) freshly isolated by size exclusion chromatography (SEC). Our results showed an increase in anisotropy when the labeled FoxP1 was combined with unlabeled FoxP1 or, to a lesser extent, with FoxP2 (Figure S1). These results indicate that this heterotypic association occurs *in vitro*.

We then wanted to describe the structural dynamics of FoxP1 in the context of the heterodimer, employing our single-molecule (sm) fluorescence toolbox[29] (Figure 1C). This approach enables the determination of structural dynamics from nanoseconds (ns) to milliseconds (ms) by combining time-resolved fluorescence (Time-Correlated single-photon counting, TCSPC), filtered fluorescence correlation spectroscopy (fFCS), and photon distribution analysis (PDA) (Figure 1C), thus monitoring dynamic processes spanning from fast and local chain motions to large-scale conformational changes.[28,33,34]

We used the Förster Resonance Energy Transfer (FRET) as a molecular ruler, using double-cysteine variants of FoxP1 FKH (FRET variants C18–C78 and C57–C78, hereafter) attached with Alexa 488 (A488, donor of FRET) and Alexa 647 (A647, acceptor of FRET) dyes (Figure 3A). We previously used these positions to characterize the dimeric FoxP1, showing at first that none of these mutations significantly alters the dimerization properties of the protein.[28] The labeling positions were used follow both the structural changes between regions that are not exchanged (C57–C78) and regions that are exchanged (C18–C78) in the 3DS-DS dimer[28] (Figure 3A). In that configuration, the dynamics between helices *H1*-*H3* and *H1*-*H5* can be analyzed.

Based on the preliminary heterotypic association assay results (Figure S1), all our smFRET measurements were performed using unlabeled FoxP2 at μM concentration under native conditions and at room temperature to ensure heterodimeric populations, whereas the labeled proteins were maintained at pM concentration (single-molecule conditions). In addition, we corroborated that at least the labeled FRET variants are in (hetero) dimeric conditions based on their diffusion times relative to monomers, as determined using FCS (Figure S2). As the total protein concentration of FoxP1 is ~10^6-fold lower than FoxP2, the heterodimer is the most likely state.

We first analyzed the time-resolved fluorescence data (Figure 3B), as the lifetime changes of the donor in the presence of the acceptor serve as a preliminary approach to define the most likely FRET model distribution of the respective FRET variants. In addition, photophysical artifacts such as PIFE[35] and dynamic quenching can be ruled out, as the lifetime is independent of the intensity-based changes.

As a previous report showed that the FoxP1 homodimer contains different disordered regions,[28] we considered two different FRET models to fit the lifetime of the donor: (1) the case of a flexible polymer behavior –Worm-like Chain (WLC)– and (2) the case of a defined Gaussian distribution (supplementary data analysis). We interpreted the WLC model as a highly heterogeneous conformational ensemble, whereas the single Gaussian distribution is more restricted and represented as a structured state. The dimers of both FRET variants were better fitted to the WLC model (Figure 3B), highlighting the need for a flexible and heterogeneous ensemble in the intramolecular dynamics of FoxP1.

Next, we analyzed our measurements as two-dimensional histograms based on the intensity-based FRET efficiency (E) and the average fluorescence lifetime of the donor in the presence of the acceptor ($\langle \tau_{D(A)} \rangle_f$) (Figure 3C). Both heterodimers showed High FRET (HF) and Low FRET (LF) populations, although the LF was significantly more pronounced in the C18–C78 FRET variant (Figure 3C). For both FRET variants, HF and LF are connected by a broad distribution, strongly suggesting the existence of at least dynamics in the observation time (milliseconds).[36]

To ascertain in our data the structural heterogeneity of FoxP1, we now analyzed the donor lifetimes in both HF and LF populations in the heterodimer variants (Table S2), finding that, in all cases, the LF population was better fitted using the WLC model. Of interest, only the HF population of the C18–C78 FRET variant was better fitted to a single Gaussian distribution with an average FRET distance of 39 $\pm$ 4 Å (Table S2), and close to the predicted (42 Å) in the domain-swapped structure of FoxP1.[28] The low fraction of molecules adopting such stable folded conformation (Figure 3C) indicates that this state is poorly accumulated in that region.

Of interest, both HF and LF populations in the C57–C78 FRET variant were better fitted to the WLC model, only differing in their apparent persistence lengths (Table S2). Furthermore, for the case of the LF population, the persistence length indicates a long average interdye distance (Table S2), confirming its extended nature. The determined persistence length calculated for the HF population in the case of the C57–C78 FRET variant suggests the highly flexible but compact nature of the region between helices *H*3 and *H*5. Although we do not have enough assurance to discriminate the correct structural behavior, these results align with those obtained for FoxP1 homodimers,[28] suggesting that an extended ensemble along the chain is dominant in the heterodimer.

Structural heterogeneity modulation in the heterodimer on DNA binding

Once we established that the FoxP1-FoxP2 heterodimer is dominated by high structural flexibility and heterogeneity, we reasoned that the interaction of FoxP1 with DNA could impact its structure and dynamics. As FoxP1 and FoxP2 show varying affinities to different DNA sequences, we used a specific double-stranded DNA (dsDNA) that is significantly better recognized by FoxP1 (Figure S3), so we can ensure that the DNA interacts only with FoxP1 in the heterodimer (Figure 4A).

We repeated the smFRET experiments with excess unlabeled FoxP2, whereas dsDNA was added at concentrations to maintain a 1:5 FoxP1: DNA stoichiometry (STAR Methods). The presence of DNA changed the diffusion time of the dye compared with the experiments in its absence (Figure S2), indicating the formation of the DNA:FoxP1:FoxP2 ternary complex.

When analyzed, both FRET variants maintain the behavior previously observed in the absence of the DNA, although the smFRET plots show an increased LF population (Figure 4B). Therefore, we repeated the analysis with our time-resolved fluorescence data on both FRET variants to determine and quantify the effect of DNA in the model that better represents the heterodimer. Of interest, none of the populations in both FRET variants changed their behavior compared with the condition in the absence of DNA (Table S2), suggesting that DNA does not influence the structural heterogeneity of the attributed to the heterodimer.

However, the presence of DNA increases the relative population of the LF ensemble as probed by both FRET variants, decreasing the persistence length of the extended ensemble for the C57–C78 FRET variant (Table S2). This result suggests compaction of the chain (between these labeling sites) at the local level, a transient folding of the C-terminal region, or a mix of these two. Considering that the DNA interacts with the helix *H*3 in the FKH domain,[6,8] this result suggests that the DNA is impacting mainly the binding region of FoxP1.

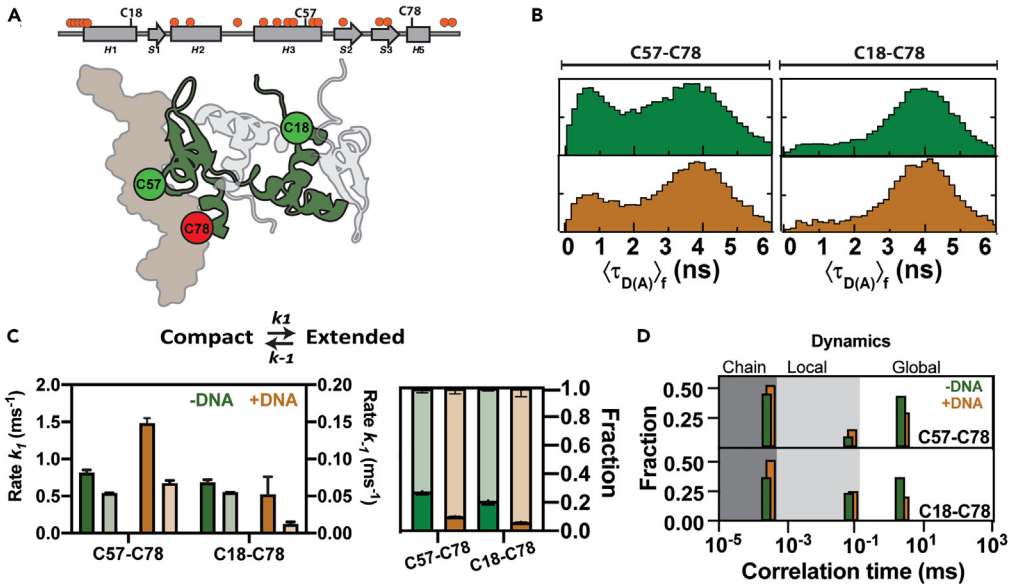


Figure 4. Effect of DNA in the structural dynamics of FoxP1 in the heterodimer

(A) The structural configuration of the heterodimer bound to DNA was analyzed by smMFS. The dye configuration is the same as in Figure 3. The DNA is represented in brown. The DNA-binding regions of the protein are shown as orange circles. The dimer structure was generated using the dimeric structure of FoxP2 (PDB: 2A07) as template.

(B) The average fluorescence lifetime of the donor in the presence of acceptor ($\langle\tau_{D(A)}\rangle_f$) patterns for both FRET pairs (C57-C78 and C18-C78), indicating the behavior in the absence (green) and presence of DNA (brown).

(C) PDA-derived rates of ensembles exchanging in the absence (green) and presence of DNA (brown). Dark color bars represent the rate transition from the compact to the extended ensemble (k_1), whereas light colors represent the reverse rate (k-1). The analysis was done in triplicate and it is shown with the respective standard deviation.

(D) Analysis of the fFCS data gives three FRET transitions (relaxation time t_R) from nano to milliseconds with their respective amplitude.

To analyze how DNA changed the kinetic connectivity of these HF and LF states of the heterodimer, we analyzed the equilibrium between the ensembles at fast and slow timescales, dissecting local and significant conformational changes in the complex. For slow timescales, we used photon distribution analysis (PDA),[36,37] assuming a simple two-state equilibrium between compact or native for the case of C18–C78- (HF) and extended (LF) populations (Figure 4C). PDA is useful to obtain the transition rates (k_1 and k_{-1}) in the heterodimer equilibrium in the ms timescale.[36,37]

In the absence of DNA, both FRET variants showed similar transition rates, favoring the extended ensemble in both cases but with differences between the rate constants. However, there are interesting local effects in the heterodimers in the presence of DNA. For example, for the C57-C78 FRET variant, DNA increases k_1 by ~2-fold while maintaining k_{-1} unaltered (Figure 4C). At the same time, DNA decreases the HF ensemble of the C18-C78 FRET variant (Figure 4C), lowering the k_{-1} value by ~3-fold (Figure 4B). These results indicate that DNA favors the accumulation of the LF ensemble by differentially impacting the dynamics of the heterodimer at a local level.

Next, we described the conformational dynamics in the absence and presence of DNA at smaller timescales using filtered FCS (fFCS).[34,38–40] The ability of fFCS to resolve dynamics at distinct timescales simultaneously allows us to characterize the conformational energy landscape with respect to distinct conformational exchange processes, capturing the exchange between the FRET states at sub-diffusion timescales (ns to ms regime).[34,38–40] We combined auto and cross-correlation (Figure S4) between the dyes to account for photophysical artifacts such as non-fluorescent dark states and determine then the effective relaxation times (t_R) because of changes in FRET.

Using fFCS, we found three relaxation times (t_R), from the hundreds of ns (t_{R1}), tens of microseconds, μs (t_{R2}), and low ms (t_{R3}), for both FRET variants in the absence and presence of DNA (Figure 4D, and Table S3). The observed transitions at the nano and microseconds timescale are hidden in the PDA analysis. In addition,

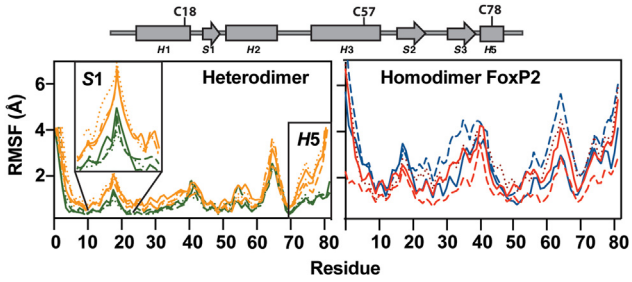


Figure 5. Local fluctuations of the heterodimer and homodimer FoxP2 alone or in the presence of DNA

Root mean squared distance fluctuations (RMSF) were calculated for each cluster from the backbone, and Cβ deviations of FoxP1 in the heterodimer and FoxP2 in the homodimer along the clusters were obtained from the tICA analysis. Green (FoxP1) and red (FoxP2) lines represent the RMSF values in free form, whereas yellow (FoxP1) and blue (FoxP2) represent the RMSF values in the presence of DNA. Each kind of line represents the analysis for a cluster. The secondary structure content of the FKH domain is shown to locate the positions labeled for our FRET experiments.

we determined the relative contributions of those t_R to the global dynamics exchange. The C57–C78 FRET variant shows transitions primarily at fast (t_{R1}) and slow timescales (t_{R3}), whereas the C18-C78 FRET variant showed transitions at a similar fraction across all timescales (Figure 4D). These results corroborate the heterogeneity in the structure changes in FoxP1 in the heterodimer, spanning from fast and localto slow and collective structural changes.[34,38–40]

Of interest, the presence of DNA decreases the relative contribution to the dynamics in both FRET variants. Specifically, in both FRET variants, we observed an increase in the contribution of the fast exchange (t_{R1}) and a decrease in the slowest (t_{R3}) relaxation time, indicating that the flexible and heterogeneous ensemble is accumulated on DNA binding, being able to rapidly sample different spatial conformations. We also noted an additional increase in the amplitude of the t_{R2} for the C57-C78 FRET variant. Considering the results obtained, we hypothesized that high conformational heterogeneity and flexibility dominate FoxP1 in the heterodimer, and binding to DNA accumulates the expanded and flexible ensemble.

Structural changes across the chain in the FoxP1:FoxP2 heterodimer

A question that remains unanswered is about the collective changes in FoxP1 in the heterodimer and the effect of binding to DNA. To get a detailed model of the structural changes in the heterodimer, we computationally characterized the heterodimer alongside the effect of the DNA on its dynamics, using the time-structure Independent Component Analysis (tICA),[41] a powerful method that explores more in detail the dynamics in proteins than PCA. We generated a model of the heterodimer, altogether with the complex with the FoxP1-binding DNA sequence, further used as input for explicit solvent MD simulations. Three 500 ns dynamics were generated using either the free model or its complex with DNA as an input. As a comparison, we performed FoxP2-FoxP2 free and in complex with its recognized DNA (STAR Methods). These simulations were analyzed through tICA, using the alpha carbon distances in the fluorophore-bound residues in FoxP1 (or FoxP2 for the homodimer) and its root mean squared deviation (RMSD) of the atomic positions to the initial structure as the input features for this method. The resulting clusters (Figure S5) determined from tICA were analyzed to capture the flexibility pattern of the free and DNA-bound ensembles by calculating the cluster RMSD (Figures 5, and S5).

After clustering the total production of 1.5 μs for each system into their representative members, three basins for both the free heterodimer and homodimer were obtained (Figure S5). In the presence of DNA, the number of clusters in the homodimer of FoxP2 decreased, whereas in the heterodimer increased to four clusters, suggesting that binding to DNA differently impacts the protein, increasing the number of structural ensembles in the heterodimer but not in the homodimer (Figure 5). Next, we calculated these clusters' backbone and sidechain fluctuations (Figures 5, and S5). This analysis shows that overall, the structural heterogeneity in the homodimer is higher compared with the heterodimer.

On adding DNA to the simulated systems, an increase in the per-residue fluctuation is observed only for the heterodimer and localized mostly in the strand *S*1 and helix *H*5 (Figure 5). Of interest, these regions are monitored by the FRET variant C18-C78 (Figures 3 and 4). Finally, it is worth noting that the residues

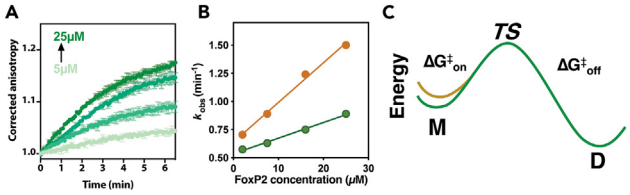


Figure 6. FoxP1-FoxP2 heterodimerization analysis via anisotropy kinetics

(A) Titration assays of the single-cysteine mutant FoxP1 (C32-OG488) with unlabeled FoxP2 show the corrected anisotropy change as a function of time. The unlabeled protein concentrations probed are indicated.

(B) The observed rates (k_{obs}) of heterodimerization in the absence (green) and presence (brown) of DNA were plotted against the concentration of unlabeled FoxP2, suggesting an increase in the association rate (k_{ass}) in the presence of DNA. The experiments were performed in triplicates.

(C) The energy diagram of heterodimerization in the absence of DNA (in green) shows the decrease in the association barrier in the presence of DNA (in brown).

belonging to helix *H3* are mostly unaffected by the presence of the DNA, suggesting that at least that region does not increase its flexibility. As this helix is mainly contacting the DNA[6,42] (Figures 3 and 4), the stabilization is probably because of the local stabilization of the DNA. Adding DNA increases the protein's overall motions, as it can be assessed from its local flexibility.

Impact of DNA on heterotypic association

All computational and experimental studies indicate that relevant structural rearrangements are adopted for the heterodimer when bound to DNA. Therefore, we reasoned if the interaction of FoxP1 with DNA could also impact the kinetic and thermodynamic parameters described for the heterotypic association with FoxP2, giving some details regarding the biological role of the heterodimer dynamics and the effect in the heterodimer lifetime.

To determine the functional properties of the heterodimer, we repeated the fluorescence anisotropy-based dimerization assay by titrating labeled FoxP1 with unlabeled FoxP2 (Figure 6A). Using this approach, we obtained an equilibrium dissociation constant (K_D) at 37°C and 38 ± 3 μM (Table 1). It is worth noting that although the FoxP1-FoxP2 heterodimer is thermodynamically unfavorable compared with the FoxP1 homodimer,[27,28] it is still 100-fold more favorable than the FoxP2 homodimer (K_D ~2,000 μM at 20°C).[9]

We repeated the association between FoxP1 and FoxP2 in the presence of DNA, using the same experimental setup as observed in the heterodimerization kinetics (Figure 6B). We fixed the concentration of the labeled monomer of FoxP1 and added dsDNA to generate a 1:1 FoxP1:DNA stoichiometry, thus finally adding different concentrations of unlabeled FoxP2. Compared with the kinetic in the free form, our results showed that the presence of DNA does not influence the dissociation rate (k_{diss}) but generated an increase of ~3-fold in k_{ass}, decreasing then the K_d from 38 μM to 18 μM and indicating that the presence of the DNA favors the association between FoxP1 and FoxP2 (Table 1). In this scenario, the interaction of one monomer with its DNA seems sufficient to promote heterotypic protein-protein association by decreasing the association energy barrier (Figure 6C), possibly because of the accumulation of the transiently folded but extended ensembles.

DISCUSSION

TFs gate access to genes, and many require the presence of DNA to adopt the necessary quaternary structure for their regulatory function.[4,43,44] Contrary to many bZIP TFs,[43,45] FoxP proteins do not require interacting with DNA to assemble into quaternary complexes. This property is relevant among FoxP proteins, as they are mutually expressed in different tissues and have been described as heterodimerizing inside cells.[18,32,46,47] Does DNA influence the function of these complexes by modulating their structural dynamics? Or does the adoption of the dimeric state before DNA interaction serve a biological purpose? These questions are complicated by the fact that, for FoxP proteins, the DNA-binding domain is also a dimerization domain.[5–8]

Table 1. Dissociation properties of the heterodimer in the free form and in the presence of DNA

Condition	K_d (μM)	k_{diss} (min^{-1})	k_{ass} ($\cdot 10^3$) ($M^{-1} \cdot min^{-1}$)
Free	38 ± 2.1	0.54 ± 0.01	14 ± 0.8
+DNA	18 ± 1.2	0.64 ± 0.03	35 ± 1.8

To answer these questions, we used a hybrid approach combining MD simulations, ensemble, and high-resolution approaches. First, we revealed the structural dynamics of the DNA-binding domain of the FoxP1 and FoxP2 heterodimer and the constraints imposed on the dimer by the DNA. Our experimental data show that the most likely physical model that explains the fluorescence information is the equilibrium between the folded and a highly heterogeneous extended ensemble that accumulates in the presence of DNA (Figures 3 and 4), similar to what we reported in the homodimer of FoxP1.[28] Furthermore, this model needs to be complemented with additional biophysical tools to unequivocally ascertain the structural behavior of the protein.

Of interest, the residue differences between FoxP1 and FoxP2 that could promote heterodimerization ($L^{FoxP1}18M^{FoxP2}$, $K^{FoxP1}23R^{FoxP2}$, and $M^{FoxP1}37T^{FoxP2}$) are mainly located in the region spanning secondary elements *H*1-*S*1-*H*2, and, considering the heterogeneity behavior described in our tICA analysis (Figure 5), we hypothesize that these evolutionary substitutions could at least partially stabilize the structural flexibility of the dimeric chain compared with the same region in FoxP2. For example, the C18–C78 FRET variant is almost exclusively represented by the extended ensemble (Figure 2), although we observed an HF population in the C18-C78 heterodimer with an interdye distance of 39 ± 4 Å, consistent with a well-structured region.

Moreover, because helix *H*1 should be surrounded by secondary structure elements of FoxP2 (Figure 2), we suggest that FoxP1 is locally stabilizing FoxP2 in the heterodimer, considering the high structural heterogeneity of the homodimer FoxP2 (Figure 5). However, the C57–C78 FRET variant reports mostly the C-terminal intrachain dynamics of FoxP1. This FRET variant reveals extended and compact ensembles (Figures 3 and 4), both consistent with high structural heterogeneity, in line with our previous observations with the homodimer of FoxP1,[28] and the tICA analysis in both hetero and homodimer (Figure 5).

In addition, both FRET variants exhibit structural dynamics that span fast and slow timescales, with significant conformational changes that could be associated with ensemble switching.[48,49] Considering the results in Figure 2, the broad distribution of the low FRET population is itself evidence of additional equilibrium between fully extended and other partially extended ensembles that are only observable under our fFCS measurements at μs timescales (Figure 4C).

The highly heterogeneous behavior of the heterodimer suggests that the conformational dynamics of FoxP1 are complex, and that the presence of multiple ensembles could be relevant in modulating the specificity and/or affinity of the heterodimer for other partners, allowing regulation of the complex's lifetime and, therefore, its repressor role in the cell.

Of interest, our results show that the association between FoxP1 and at least one dsDNA does not result in folding but instead still favors the extended ensemble (Figures 4B and 4C) and promotes heterodimerization by decreasing the energy barrier of association (Figure 4D). The fact that DNA increases the heterogeneity only of the heterodimer (as shown in our tICA analysis) highlights the critical role of flexibility in the dimerization properties among FoxP proteins.

In addition, DNA reduced the apparent persistence length of the C-terminal region, consistent with a possible shift of the equilibrium between fully and partially extended ensembles sampled in the μs timescale, as suggested by fFCS (Figure 4C). Structurally, helix *H*3 is crucial for DNA interaction, as observed in FoxP2[8] and FoxP3[6]; because helix *H*3 directly interacts with the DNA, the reduction of persistence length in this region suggests that the hypothesized partially extended intermediates are now accumulated when DNA is present. Altogether, the presence of one dsDNA does not promote the folding of the heterodimer but shifts the equilibrium to the partially extended ensembles (Figure 5), highlighting the modulation of the spectrum of protein disorder[50,51] to facilitate protein-protein association without sacrificing functional versatility.

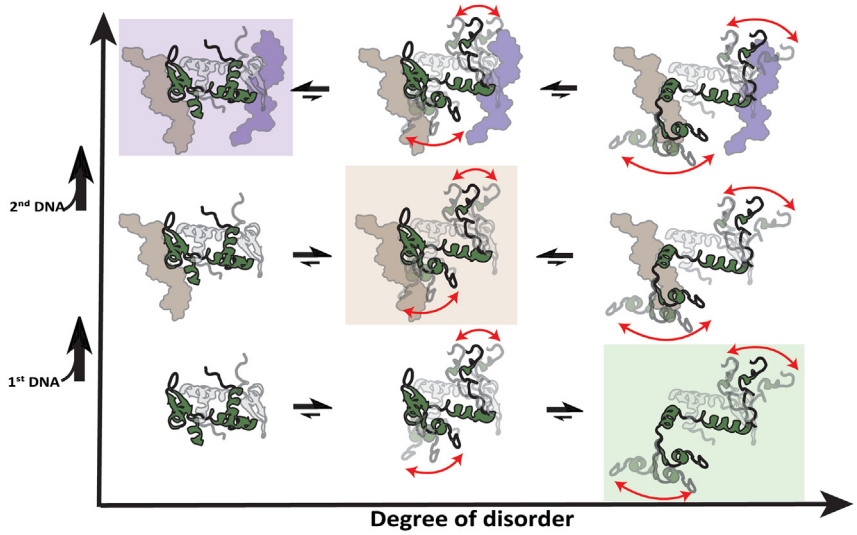


Figure 7. Structural dynamics of the heterodimer and effect of DNA

The characterized heterodimer of FoxP1 (in green) and FoxP2 (in gray) consists of a dynamic exchange between a compact ensemble by which the C-terminal (monitored by FRET variant C57-C78), and the N-terminal (monitored by the C18-C78 FRET variant) adopts the structured conformation (high FRET population in Figure 2) and an extended ensemble (green box). In addition, partially extended ensembles can be obtained, but their population is too low to be observed in the free form. The presence of one molecule of DNA that specifically binds to FoxP1 generates an accumulation of the partially extended ensembles, identified by the decrease of the apparent persistence length in the C-terminal region of FoxP1 (brown box). The purple box shows the effective compaction of the heterodimer once the quaternary complex is adopted, as observed in the crystal structure of dimeric FoxP2 and FoxP3.

The formation of the FoxP heterodimer-DNA ternary complex is notably different from the widely described folding-upon-binding mechanism by which many TFs assemble into protein-DNA complexes[3,52,53] (Figure 5). Moreover, the structural flexibility of this ternary complex implies an intrinsic ability to be available to bind another DNA molecule by dynamic anticipation,[54] continuously targeting different chromosomal elements without a kinetic "lock" related to protein folding. This dynamical behavior is an alternative to the observed monkey bar mechanism by which many proteins recognize DNA.[55]

Although we experimentally demonstrated the FoxP1:FoxP2 heterodimerization, the dissociation constant of the heterodimer is significantly lower than the homodimer of FoxP2, and the obtained value suggests that the homodimerization between monomers of FoxP1 is dominant over the heterodimerization. In this context, one question that must be answered from our results is how much heterodimers are expected inside cells and the effective role of the DNA.

The comparison of the heterodimerization in the absence and presence of DNA indicates that the impact is localized in the monomers' association (Figure 6), which depends on the relative FoxP1:FoxP2 concentration. For example, cortical regions in the adult brain show higher expression of FoxP2 with respect to FoxP1,[56] and therefore, heterodimerization could be the most likely event in those cells. In addition, the adoption of a quaternary complex with two DNA molecules could be an effective way to stabilize the heterodimeric structure (Figure 7). In this scenario, this dynamic complex could be perturbed by other known co-transcriptional factors such as NFAT, β-catenin, CtBP1, and others,[57–59] regulating the repressor effect on FoxP heterodimers and the formation of transcriptional complexes. For example, FoxP3 and FoxP2 have been crystallized bound to both DNA and NFAT,[8,57] suggesting the importance of FoxP scaffolding of transcriptional complexes.

Finally, we postulate that the ability to heterodimerize even in the presence of DNA could diversify the dynamic regulatory effect of FoxP TFs by interacting with non-identical DNA segments, as all FoxP members recognize specific consensus sequences.[6,8,28,60] However, one aspect still unclear corresponds to the structural dynamics of FoxP2 in the context of the heterodimers and how its properties are affected by the intrinsic dynamics of FoxP1. Although we demonstrated that the binding of DNA to one monomer

increases the heterogeneity in the complex, the effect of adopting the quaternary complex is still unclear and needs to be explored in both homotypic and heterotypic protein association scenarios.

Limitations of the study

This article is based on the *in vitro* analysis of two proteins that occurs at micromolar concentration. The main limitation is the contextualization in the biological environment, where the effective concentration can be significantly lower. We do not discard that additional processes such as multivalent interactions and possible molecular condensations can assist in overcoming the energy limitation.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Protein expression and purification
  - Fluorescence labeling of FoxP1
  - Dimerization kinetics
  - DNA binding measurements
  - Fluorescence measurements
  - Single-molecule multiparameter fluorescence spectroscopy (smMFS)
  - Fluorescence analysis of smMFS experiments
  - Filtered fluorescence correlation spectroscopy (fFCS)
  - Confinement molecular dynamics (CCR)
  - Explicit solvent molecular dynamics
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Dimerization kinetics
  - Time-resolved fluorescence analysis
  - Filtered FCS (fFCS)
  - Photon distribution analysis (PDA)
  - Molecular dynamics simulations
  - Time-structure based independent component analysis (TICA) and clustering
  - Statistical uncertainties and error analysis
  - FRET models

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.107228.

ACKNOWLEDGMENTS

This work was supported by ANID through Fondo de Desarrollo Científico y Tecnológico (FONDECYT grants 1170701 to J.B, 11200729 to E.M., 1201684 to C.A.R-S. and 3190731 to M.R.), ANID Millennium Science Initiative Program ICN17_022. HS acknowledges support from NIH (MH081923, 1P20GM130451) and NSF (CAREER MCB 1749778). P.G-D. was supported by ANID Doctoral Scholarship 21181705. This research was partially supported by the supercomputing infrastructure of the NLHPC (ECM-02). We also gratefully acknowledge the support of Nvidia Corporation with the donation of the GPUs that were used to perform some of the MD simulations presented herein.

AUTHOR CONTRIBUTIONS

Conceptualization, E.M., H.S., J.B., and C.A.R-S.; Validation, E.M., H.S., C.A.R-S., P.G-D., and M.R.; Formal analysis and investigation, R.C., I.A., G.H., P.G-D., M.R., E.M., J.G.H., and F.E.; Writing, E.M., H.S., J.B., C.A.R-S., P.G-D., and F.E.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: October 3, 2022
Revised: May 15, 2023
Accepted: June 20, 2023
Published: June 28, 2023

REFERENCES

1. Shammas, S.L. (2017). Mechanistic roles of protein disorder within transcription. Curr. Opin. Struct. Biol. *42*, 155–161.

2. Basu, S., Mackowiak, S.D., Niskanen, H., Knezevic, D., Asimi, V., Grosswendt, S., Geertsema, H., Ali, S., Jerković, I., Ewers, H., et al. (2020). Unblending of transcriptional condensates in human repeat expansion disease. Cell *181*, 1062–1079.e30.

3. Seldeen, K.L., McDonald, C.B., Deegan, B.J., and Farooq, A. (2008). Coupling of folding and DNA-binding in the bZIP domains of Jun-Fos heterodimeric transcription factor. Arch. Biochem. Biophys. *473*, 48–60.

4. Miller, M. (2009). The importance of being flexible: the case of basic region leucine zipper transcriptional regulators. Curr. Protein Pept. Sci. *10*, 244–269.

5. Chu, Y.P., Chang, C.H., Shiu, J.H., Chang, Y.T., Chen, C.Y., and Chuang, W.J. (2011). Solution structure and backbone dynamics of the DNA-binding domain of FOXP1: Insight into its domain swapping and DNA binding. Protein Sci. *20*, 908–924.

6. Bandukwala, H.S., Wu, Y., Feuerer, M., Chen, Y., Barboza, B., Ghosh, S., Stroud, J.C., Benoist, C., Mathis, D., Rao, A., and Chen, L. (2011). Structure of a domain-swapped FOXP3 dimer on DNA and its function in regulatory T cells. Immunity *34*, 479–491.

7. Medina, E., Córdova, C., Villalobos, P., Reyes, J., Komives, E.A., Ramírez-Sarmiento, C.A., and Babul, J. (2016). Three-Dimensional Domain Swapping Changes the Folding Mechanism of the Forkhead Domain of FoxP1. Biophys. J. *110*, 2349–2360.

8. Stroud, J.C., Wu, Y., Bates, D.L., Han, A., Nowick, K., Paabo, S., Tong, H., and Chen, L. (2006). Structure of the forkhead domain of FOXP2 bound to DNA. Structure *14*, 159–166.

9. Perumal, K., Dirr, H.W., and Fanucchi, S. (2015). A Single Amino Acid in the Hinge Loop Region of the FOXP Forkhead Domain is Significant for Dimerisation. Protein J. *34*, 111–121.

10. Barzaghi, F., Passerini, L., and Bacchetta, R. (2012). Immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome: A paradigm of immunodeficiency with autoimmunity. Front. Immunol. *3*, 211. https://doi.org/10.3389/fimmu.2012.00211.

11. Lai, C.S., Fisher, S.E., Hurst, J.A., Vargha-Khadem, F., and Monaco, A.P. (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. Nature *413*, 519–523.

12. Vernes, S.C., Nicod, J., Elahi, F.M., Coventry, J.a., Kenny, N., Coupe, A.-M., Bird, L.E., Davies, K.E., and Fisher, S.E. (2006). Functional genetic analysis of mutations implicated in a human speech and language disorder. Hum. Mol. Genet. *15*, 3154–3167.

13. Lam, E.W.F., Brosens, J.J., Gomes, A.R., and Koo, C.Y. (2013). Forkhead box proteins: Tuning forks for transcriptional harmony. Nat. Rev. Cancer *13*, 482–495. https://doi.org/10.1038/nrc3539.

14. Hisaoka, T., Nakamura, Y., Senba, E., and Morikawa, Y. (2010). The forkhead transcription factors, Foxp1 and Foxp2, identify different subpopulations of projection neurons in the mouse cerebral cortex. Neuroscience *166*, 551–563.

15. Estruch, S.B., Graham, S.A., Deriziotis, P., and Fisher, S.E. (2016). The language-related transcription factor FOXP2 is post-translationally modified with small ubiquitin-like modifiers. Sci. Rep. *6*, 20911.

16. Becker, M., Devanna, P., Fisher, S.E., and Vernes, S.C. (2018). Mapping of human FOXP2 enhancers reveals complex regulation. Front. Mol. Neurosci. *11*, 47.

17. den Hoed, J., Devaraju, K., and Fisher, S.E. (2021). Molecular networks of the FOXP2 transcription factor in the brain. EMBO Rep. *22*, e52803.

18. Shu, W., Lu, M.M., Zhang, Y., Tucker, P.W., Zhou, D., and Morrisey, E.E. (2007). Foxp2 and Foxp1 cooperatively regulate lung and esophagus development. Development *134*, 1991–2000.

19. Sin, C., Li, H., and Crawford, D.A. (2015). Transcriptional Regulation by FOXP1, FOXP2, and FOXP4 Dimerization. J. Mol. Neurosci. *55*, 437–448.

20. Hamdan, F.F., Daoud, H., Rochefort, D., Piton, A., Gauthier, J., Langlois, M., Foomani, G., Dobrzeniecka, S., Krebs, M.O., Joober, R., et al. (2010). De novo mutations in FOXP1 in cases with intellectual disability, autism, and language impairment. Am. J. Hum. Genet. *87*, 671–678.

21. Araujo, D.J., Anderson, A.G., Berto, S., Runnels, W., Harper, M., Ammanuel, S., Rieger, M.A., Huang, H.-C., Rajkovich, K., Loerwald, K.W., et al. (2015). FoxP1 orchestration of ASD-relevant signaling pathways in the striatum. Genes Dev. *29*, 2081–2096.

22. Wang, B., Lin, D., Li, C., and Tucker, P. (2003). Multiple domains define the expression and regulatory properties of Foxp1 forkhead transcriptional repressors. J. Biol. Chem. *278*, 24259–24268.

23. Sollis, E., Graham, S.A., Vino, A., Froehlich, H., Vreeburg, M., Dimitropoulou, D., Gilissen, C., Pfundt, R., Rappold, G.A., Brunner, H.G., et al. (2016). Identification and functional characterization of de novo FOXP1 variants provides novel insights into the etiology of neurodevelopmental disorder. Hum. Mol. Genet. *25*, 546–557.

24. Liu, Y., and Eisenberg, D. (2002). 3D domain swapping: as domains continue to swap. Protein Sci. *11*, 1285–1299.

25. Bennett, M.J., Choe, S., and Eisenberg, D. (1994). Domain swapping: entangling alliances between proteins. Proc. Natl. Acad. Sci. USA *91*, 3127–3131.

26. Bennett, M.J., Schlunegger, M.P., and Eisenberg, D. (1995). 3D domain swapping: A mechanism for oligomer assembly. Protein Sci. *4*, 2455–2468. https://doi.org/10.1002/pro.5560041202.

27. Medina, E., Villalobos, P., Coñuecar, R., Ramírez-Sarmiento, C.A., and Babul, J. (2019). The protonation state of an evolutionarily conserved histidine modulates domain swapping stability of FoxP1. Sci. Rep. *9*, 5441.

28. Medina, E., Villalobos, P., Hamilton, G.L., Komives, E.A., Sanabria, H., Ramírez-Sarmiento, C.A., and Babul, J. (2020). Intrinsically Disordered Regions of the DNA-Binding Domain of Human FoxP1 Facilitate Domain Swapping. J. Mol. Biol. *432*, 5411–5429.

29. Medina, E., R Latham, D., and Sanabria, H. (2021). Unraveling protein's structural dynamics: from configurational dynamics to ensemble switching guides functional mesoscale assemblies. Curr. Opin. Struct. Biol. *66*, 129–138.

30. Cecchini, M., Krivov, S.V., Spichty, M., and Karplus, M. (2010). Calculation of free-energy differences by confinement simulations. Application to peptide conformers. J. Phys. Chem. B *114*, 6763.

31. Galaz-Davison, P., Molina, J.A., Silletti, S., Komives, E.A., Knauer, S.H., Artsimovitch, I., and Ramírez-Sarmiento, C.A. (2020). Differential local stability governs the

metamorphic fold switch of bacterial virulence factor RfaH. Biophys. J. *118*, 96–104.

32. Li, S., Weidenfeld, J., and Morrisey, E.E. (2004). Transcriptional and DNA binding activity of the Foxp1/2/4 family is modulated by heterotypic and homotypic protein interactions. Mol. Cell Biol. *24*, 809–822.

33. Yanez Orozco, I.S., Mindlin, F.A., Ma, J., Wang, B., Levesque, B., Spencer, M., Rezaei Adariani, S., Hamilton, G., Ding, F., Bowen, M.E., and Sanabria, H. (2018). Identifying weak interdomain interactions that stabilize the supertertiary structure of the N-terminal tandem PDZ domains of PSD-95. Nat. Commun. *9*, 3724.

34. Sanabria, H., Rodnin, D., Hemmen, K., Peulen, T.-O., Felekyan, S., Fleissner, M.R., Dimura, M., Koberling, F., Kühnemuth, R., Hubbell, W., et al. (2020). Resolving dynamics and function of transient states in single enzyme molecules. Nat. Commun. *11*, 1231.

35. Hwang, H., and Myong, S. (2014). Protein induced fluorescence enhancement (PIFE) for probing protein-nucleic acid interactions. Chem. Soc. Rev. *43*, 1221–1229.

36. Kalinin, S., Valeri, A., Antonik, M., Felekyan, S., and Seidel, C.A.M. (2010). Detection of structural dynamics by FRET: a photon distribution and fluorescence lifetime analysis of systems with multiple states. J. Phys. Chem. B *114*, 7983–7995.

37. Kalinin, S., Felekyan, S., Valeri, A., and Seidel, C.A.M. (2008). Characterizing Multiple Molecular States in Single-Molecule Multiparameter Fluorescence Detection by Probability Distribution Analysis. J. Phys. Chem. B *112*, 8361–8374.

38. Felekyan, S., Kalinin, S., Sanabria, H., Valeri, A., and Seidel, C.A.M. (2012). Filtered FCS: Species Auto- and Cross-Correlation Functions Highlight Binding and Dynamics in Biomolecules. ChemPhysChem *13*, 1036–1053.

39. Torres, T., and Levitus, M. (2007). Measuring conformational dynamics: a new FCS-FRET approach. J. Phys. Chem. B *111*, 7392–7400.

40. Price, E.S., Aleksiejew, M., and Johnson, C.K. (2011). FRET-FCS detection of intralobe dynamics in calmodulin. J. Phys. Chem. B *115*, 9320–9326.

41. Naritomi, Y., and Fuchigami, S. (2011). Slow dynamics in protein fluctuations revealed by time-structure based independent component analysis: the case of domain motions. J. Chem. Phys. *134*, 065101.

42. Wu, Y., Borde, M., Heissmeyer, V., Feuerer, M., Lapan, A.D., Stroud, J.C., Bates, D.L., Guo, L., Han, A., Ziegler, S.F., et al. (2006). FOXP3 Controls Regulatory T Cell Funcion through Cooperation with NFAT. Cell *126*, 375–387.

43. Levy, Y., Onuchic, J.N., and Wolynes, P.G. (2007). Fly-Casting in Protein—DNA Binding: Frustration between Protein Folding and Electrostatics Facilitates Target Recognition. J. Am. Chem. Soc. *129*, 738–739.

44. Shoemaker, B.a., Portman, J.J., and Wolynes, P.G. (2000). Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. Proc. Natl. Acad. Sci. USA *97*, 8868–8873.

45. Fujii, Y., Shimizu, T., Toda, T., Yanagida, M., and Hakoshima, T. (2000). Structural basis for the diversity of DNA recognition by bZIP transcription factors. Nat. Struct. Biol. *7*, 889–893.

46. Spaeth, J.M., Hunter, C.S., Bonatakis, L., Guo, M., French, C.A., Slack, I., Hara, M., Fisher, S.E., Ferrer, J., Morrisey, E.E., et al. (2015). The FOXP1, FOXP2 and FOXP4 transcription factors are required for islet alpha cell proliferation and function in mice. Diabetologia *58*, 1836–1844.

47. Konopacki, C., Pritykin, Y., Rubtsov, Y., Leslie, C.S., and Rudensky, A.Y. (2019). Transcription factor Foxp1 regulates Foxp3 chromatin binding and coordinates regulatory T cell function. Nat. Immunol. *20*, 232–242.

48. Saikia, N., Yanez-Orozco, I.S., Qiu, R., Hao, P., Milikisiyants, S., Ou, E., Hamilton, G.L., Weninger, K.R., Smirnova, T.I., Sanabria, H., and Ding, F. (2021). Integrative structural dynamics probing of the conformational heterogeneity in synaptosomal-associated protein 25. Cell Rep. Phys. Sci. *2*, 100616.

49. Choi, U.B., Sanabria, H., Smirnova, T., Bowen, M.E., and Weninger, K.R. (2019). Spontaneous switching among conformational ensembles in intrinsically disordered proteins. Biomolecules *9*, 114.

50. Schuler, B., Borgia, A., Borgia, M.B., Heidarsson, P.O., Holmstrom, E.D., Nettels, D., and Sottini, A. (2020). Binding without folding - the biomolecular function of disordered polyelectrolyte complexes. Curr. Opin. Struct. Biol. *60*, 66–76.

51. Olsen, J.G., Teilum, K., and Kragelund, B.B. (2017). Behaviour of intrinsically disordered proteins in protein–protein complexes with an emphasis on fuzziness. Cell. Mol. Life Sci. *74*, 3175–3183. https://doi.org/10.1007/s00018-017-2560-7.

52. Creamer, T.P. (2013). Transient disorder. Intrinsically Disord. Proteins *1*, e26412.

53. Rentzeperis, D., Jonsson, T., and Sauer, R.T. (1999). Acceleration of the refolding of Arc repressor by nucleic acids and other polyanions. Nat. Struct. Biol. *6*, 569–573.

54. Tsytlonok, M., Sanabria, H., Wang, Y., Felekyan, S., Hemmen, K., Phillips, A.H., Yun, M.K., Waddell, M.B., Park, C.G., Vaithiyalingam, S., et al. (2019). Dynamic anticipation by Cdk2/Cyclin A-bound p27 mediates signal integration in cell cycle regulation. Nat. Commun. *10*, 1676. https://doi.org/10.1038/s41467-019-09446-w.

55. Vuzman, D., Azia, A., and Levy, Y. (2010). Searching DNA via a "Monkey Bar" Mechanism: The Significance of Disordered Tails. J. Mol. Biol. *396*, 674–684.

56. Ferland, R.J., Cherry, T.J., Preware, P.O., Morrisey, E.E., and Walsh, C.A. (2003). Characterization of Foxp2 and Foxp1 mRNA and protein in the developing and mature brain. J. Comp. Neurol. *460*, 266–279.

57. Wu, Y., Borde, M., Heissmeyer, V., Feuerer, M., Lapan, A.D., Stroud, J.C., Bates, D.L., Guo, L., Han, A., Ziegler, S.F., et al. (2006). FOXP3 controls regulatory T cell function through cooperation with NFAT. Cell *126*, 375–387.

58. Rocca, D.L., Wilkinson, K.A., and Henley, J.M. (2017). SUMOylation of FOXP1 regulates transcriptional repression via CtBP1 to drive dendritic morphogenesis. Sci. Rep. *7*, 877.

59. Kim, J.-H., Hwang, J., Jung, J.H., Lee, H.-J., Lee, D.Y., and Kim, S.-H. (2019). Molecular networks of FOXP family: dual biologic functions, interplay with other molecules and clinical implications in cancer progression. Mol. Cancer *18*, 180.

60. Blane, A., and Fanucchi, S. (2015). Effect of pH on the Structure and DNA Binding of the FOXP2 Forkhead Domain. Biochemistry *54*, 4001–4007.

61. Kudryavtsev, V., Sikor, M., Kalinin, S., Mokranjac, D., Seidel, C.A.M., and Lamb, D.C. (2012). Combining MFD and PIE for Accurate Single-Pair Förster Resonance Energy Transfer Measurements. ChemPhysChem *13*, 1060–1078.

62. Kühnemuth, R., and Seidel, C.A.M. (2001). Principles of Single Molecule Multiparameter Fluorescence Spectroscopy. Single Mol. *2*, 251–254.

63. Webb, B., and Sali, A. (2014). Comparative Protein Structure Modeling Using MODELLER. Curr. Protoc. Bioinformatics *47*, 5.6.1–5.632.

64. Pearlman, D.A., Case, D.A., Caldwell, J.W., Ross, W.S., Cheatham, T.E., III, DeBolt, S., Ferguson, D., Seibel, G., and Kollman, P. (1995). AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. Comput. Phys. Commun. *91*, 1–41.

65. Hawkins, G.D., Cramer, C.J., and Truhlar, D.G. (1996). Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium. J. Phys. Chem. *100*, 19824–19839.

66. Eyring, H. (1935). The Activated Complex and the Absolute Rate of Chemical Reactions. Chem. Rev. *17*, 65–77.

67. Antonik, M., Felekyan, S., Gaiduk, A., and Seidel, C.A.M. (2006). Separating structural heterogeneities from stochastic variations in fluorescence resonance energy transfer distributions via photon distribution analysis. J. Phys. Chem. B *110*, 6970–6978.

68. Tyka, M.D., Sessions, R.B., and Clarke, A.R. (2007). Absolute free-energy calculations of liquids using a harmonic reference state. J. Phys. Chem. B *111*, 9571–9580.

69. Hoffmann, M., Scherer, M., Hempel, T., Mardt, A., de Silva, B., Husic, B.E., Klus, S., Wu, H., Kutz, N., Brunton, S.L., and Noé, F. (2022). Deeptime: a Python library for machine learning dynamical models from time series data. Mach. Learn, Sci. Technol. *3*, 015009.

70. Peulen, T.O., Opanasyuk, O., and Seidel, C.A.M. (2017). Combining Graphical and Analytical Methods with Molecular Simulations to Analyze Time-Resolved FRET Measurements of Labeled Macromolecules Accurately. J. Phys. Chem. B *121*, 8211–8241.

71. Tsytlonok, M., Hemmen, K., Hamilton, G., Kolimi, N., Felekyan, S., Seidel, C.A.M., Tompa, P., and Sanabria, H. (2020). Specific conformational dynamics and expansion underpin a multi-step mechanism for specific binding of p27 with Cdk2/cyclin A. J. Mol. Biol. *432*, 2998–3017.

72. Barth, A., Opanasyuk, O., Peulen, T.-O., Felekyan, S., Kalinin, S., Sanabria, H., and Seidel, C.A.M. (2022). Unraveling multi-state molecular dynamics in single-molecule FRET experiments. I. Theory of FRET-lines. J. Chem. Phys. *156*, 141501.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
E. coli (DE3)Bl21-C41	Our laboratory	
Chemicals, peptides, and recombinant proteins		
HEPES	Sigma-Aldrich	Cat#H3375
NaCl	Merck	Cat#S1679
2-mercaptoethanol	Sigma-Aldrich	Cat#63689
Alexa488-maleimide	ThermoFisher Scientific	Cat#A10254
Alexa647-maleimide	ThermoFisher Scientific	Cat#A203347
Oregon Green488-maleimide	ThermoFisher Scientific	Cat#O6034
Oligonucleotides		
CAAGGTAAACAAGACAACGTAAACCAA	IDT-DNA	
Recombinant DNA		
pET-TEV-FoxP1(FKH)	A modified version of pET-28a	N/A
pET-TEV-FoxP2(FKH)	A modified version of pET-28a	N/A
pET-TEV-FoxP1(FKH)L18C/V78C	A modified version of pET-28a	N/A
pET-TEV-FoxP1(FKH)S57C/V78C	A modified version of pET-28a	N/A
pET-TEV-FoxP1(FKH)S57C	A modified version of pET-28a	N/A
Software and algorithms		
Multiparameter fluorescence detection analysis	Claus Seidel, HHU	

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Exequiel Medina (exequiel.medinago@uchile.cl).

Materials availability

The recombinant constructs developed in this work will be available upon request.

Data and code availability

- This article does not contain any new code.
- All the raw and analyzed data in this work will be shared by the lead contact upon request. This work did not generate any new accession code for the data.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

This work used the bacterial strain E. coli Bl21(DE3)-C41, which was grown in Luria-Bertani culture media at 37°C.

METHOD DETAILS

Protein expression and purification

All FoxP proteins (wild-type and mutants) were expressed in *Escherichia coli* Bl21(DE3)-C41 previously transformed with the respective plasmids. The purification steps were done and purified as described.[28] Briefly, for all proteins, a saturated culture of transformed bacteria supplemented with kanamicyn was used to inoculate a 1L of terrific broth to grow at 37°C, supplemented with kanamycin, under reaching

an optical density of 0.9. After that, an overnight induction of the gene expression was inducted with IPTG (isopropyl-β-D-thiogalactopyranoside), incubating at 25°C.

For the protein purification, the grown media was centrifugated at 6000 rpm for 10 min at 4°C, homogenizing the pellet with binding buffer (20 mM HEPES pH 7.8, 500 mM NaCl, 30 mM imidazole and 2 mM β–mercaptoethanol). The cells' lysis was done by sonication, centrifugated at 18000 rpm for 20 min and the clarified was loaded onto a Niquel affinity column, eluting then the protein using a 30–500 mM imidazole gradient. Standard buffer (20 mM HEPES pH 7.8, 20 mM NaCl, 2 mM β–mercaptoethanol) was used prior to each experiment unless otherwise indicated.

Fluorescence labeling of FoxP1

We prepared FoxP1 constructs where the native cysteine residue (at position 61) was replaced by serine (C61S). Single-cysteine (N32C) or double-cysteine (C57-C78 and C18/C78) mutants were then generated. Before labeling, all buffers were filtered and degassed. Then, all proteins were prepared for labeling as described.[28] The proteins were first dialyzed to remove the reducing agent and concentrated (~80–100 μM), and the labeling was performed as follow: for the single labeling, an overnight incubation with a 5-fold excess of Oregon Green 488 (OG488), whereas all double labeling reactions were performed in a two-step process: first, a two-hour incubation at room temperature with A488 followed by purification via cationic exchange chromatography using the Capto HiRes column (Cytiva, Marlborough, MA, USA); and second, overnight labeling at 4°C with A647 using the freshly purified labeled protein as a template. Finally, in all cases, excess fluorophore was removed via SEC.

Dimerization kinetics

The N32C mutant of FoxP1 labeled with OG488 (FoxP1 C32-OG488) was used in dimerization kinetics. First, 200 nM of protein was incubated at 37°C to promote the dissociation of all dimeric populations in the solution. This event was followed by a decrease in anisotropy (data not shown). Then, an aliquot of a freshly purified monomer of FoxP1, FoxP2, or the A39P mutant of FoxP1 was used for mixing with the equilibrated labeled protein, followed by anisotropy changes upon dimerization. Different FoxP1 or FoxP2 concentrations were used to determine the kinetic parameters, combined with a standard buffer to maintain the same final volume in all measurements. Each dimerization reaction was performed in triplicate at 37°C.

DNA binding measurements

A dsDNA (5′-CAAGGTAAACAAGACAACGTAAACAA-3′) was labeled in its 3′ terminal with FAM dye. For each titration experiment, 400 nM of this dsDNA was allowed to bind with different concentrations of unlabeled FoxP1 or FoxP2, ranging from 0-800 nM. All measurements were performed at 37°C and in triplicate.

Fluorescence measurements

Anisotropy measurements were performed in a Jasco FP-8300 spectrofluorometer (Jasco Corp, Tokyo, Japan), employing 1 cm path length cells. In dimerization and titration experiments, parallel and perpendicular fluorescent emissions at 525 nm, after excitation at 480 nm, were recorded using labeled protein or DNA, depending on the experiment. The G factor was determined using free dyes at 1 μM in a standard buffer to calculate the observed anisotropy in all experiments.

Single-molecule multiparameter fluorescence spectroscopy (smMFS)

All the measurements were done in a home-built confocal system containing a 485-nm and a 640-nm diode laser (LDH-D-C 485 and LDH-D-C 640 PicoQuant, Germany, operating at 20 MHz in a pulsed interleaved excitation scheme) exciting freely diffusing labeled molecules that passed through a detection volume of the 60×, 1.2 NA collar (0.17) corrected Olympus objective. The emitted fluorescence signal was collected through the same objective and spatially filtered using a 70-μm pinhole to define the confocal detection volume. The signal was divided into parallel and perpendicular components at two different colors (''green'' and ''red'') through bandpass filters, HQ 520/35, and HQ 720/150, for green and red, respectively. A total of four photon-detectors are used, two for green (PMA 40 Hybrid, PicoQuant, Germany) and two for red channels (PMA 50 Hybrid, PicoQuant). A time-correlated single-photon counting (TCSPC) module 5 (HydraHarp 400, PicoQuant, Germany) with Time-Tagged Time-Resolved (TTTR) mode and four synchronized input channels were used for data registration.

For smMFS measurements, homo or heterodimers using the double-labeled mutants were diluted in the standard buffer to pM concentration in the presence of 10 μM of unlabeled FoxP2, assuring ~1 burst per second. Those conditions allowed the obtaining of dimers, as corroborated via single-molecule fluorescence correlation measurements. The effect of DNA on heterodimers was analyzed by adding an aliquot of DNA in all steps to maintain the stoichiometry of 1:1 FoxP1: DNA. For all measurements, we used an oil immersion liquid with the refraction index of water (Immersol, Carl Zeiss Inc., Thornwood, NY, USA). NUNC chambers (Lab-Tek, Thermo Scientific) were used with a 500 μL sample volume. Standard controls (Rhodamine 110, Rhodamine 101 and Alexa 647) to determine detection efficiencies were measured.

Fluorescence analysis of smMFS experiments

smMFS experiments were analyzed using Multiparameter Fluorescence Detection (MFD).[61] Sub-ensemble TCSPC[62] of selected FRET bursts was fitted with the specified exponential decay model to determine the best one that represents the data (see FRET models). In-house written software used to perform the analysis can be downloaded from http://www.mpc.hhu.de/software.html and https://github.com/Fluorescence-Tools.

Filtered fluorescence correlation spectroscopy (fFCS)

fFCS was accomplished by selecting the single-molecule burst to differentiate the fluorescence photons of the DA-labeled samples from background photons. Then the fluorescence species was auto-correlated based on the detection spectral window (Green, *G*, and Red *R*) to generate four correlation curves ($G_{GG}^{(DA)}(t_c)$, $G_{RR}^{(DA)}(t_c)$, $G_{GR}^{(DA)}(t_c)$, $G_{RG}^{(DA)}(t_c)$), where the subscripts correspond to the spectral window, and the superscripts in parenthesis are the labeled species that were observed. These correspond to the color auto- and cross-correlation function of the FRET labeled samples at single-molecule resolution.

Confinement molecular dynamics (CCR)

To set up these simulations, we used the monomeric structures of FoxP1 (PDB ID: 2KIU) and FoxP2 (PDB ID: 2A07), whereas the dimeric structure of FoxP2 (PDB ID: 2A07) was used as a template to build the dimeric form of FoxP1 via homology modeling via Modeller,[63] as previously described.[28] These structures were initially energy minimized using a gradient descent algorithm from the molecular dynamics suite Amber20,[64] and subsequently, energy minimized using the Newton-Raphson algorithm through the NAB module of Amber until reaching a Δrms of 10^{-12}. The objective of this minimization is to reach the deepest minimum in the conformational space of the proteins so that molecular dynamics biased toward that microstate are carried out, named Confinement-Conversion Release or CCR.[30,31] Through 25 different simulations of 30 ns each, in which a cartesian restraining potential with harmonic form is applied to all the protein atoms of the simulation so that the motions are increasingly restricted to the global minimum, and the strength of such harmonic restraining constant k_R doubles from 0.000025 to 419 kcal·mol^{-1}. These simulations were run in implicit solvation[65] in the Amber20 using the CUDA-accelerated PMEMD with the amber ff19SB force field, alongside the Langevin thermostat and SHAKE for Hydrogens, which allowed us to run a simulation with a timestep of 2 fs. No periodic boundary conditions were used; hence no cutoff was used for electrostatics.

Explicit solvent molecular dynamics

Two systems were generated for the conventional molecular dynamics protocols, corresponding to a free version of the FoxP1-FoxP2 heterodimer and homodimer FoxP2 or their complexes with the respective DNA ligand. To generate the heterodimer, ColabFold (colabfold.com), an implementation of AF2 in Google Colaboratory, was used, considering as input the forkhead (FKH) domains of the human FoxP1 and FoxP2 proteins and using as a template the PDB70 with the rest as default settings. To generate the protein:DNA complex, the sequence 5′-CAAGGTAAACAAGACAACGTAAACAA-3′ for the heterodimer and 5′-AACTATGAAACAAATTTTCCT-3′ for the homodimer FoxP2 were modeled as a duplex using NAB from the Ambre20 suite, and both DNA and the respective dimer were merged through structural comparison with the FoxP2-DNA complex from the PDB ID: 2A07.

To avoid steric clashes resulting from the superposition of DNA and protein, we used RosettaRelax module from the Rosetta suite, which performs monte-carlo fragment explorations to minimize the sidechain and slightly readjust the backbone of both protein and DNA. We performed 500 relax protocols using the

option Relax:thorough, which generates a deep minimization of the system. The best configuration evaluated from the Rosetta score was further used for molecular dynamics.

The systems were solvated in 12 Å TIP3P water padding using the force fields ff19SB for protein and OL15 for DNA. Further charges were neutralized by adding counter ions away from the protein using the addIons2 module from the Amber20 suite. The systems were minimized, and temperature and pressure equilibrated before a production simulation of 100 ns under NPT ensemble. This relaxation MD was clustered using cpptraj from the Amber20 suite, and the main cluster was used for further production of 3 independent replicas of 500 ns each.

QUANTIFICATION AND STATISTICAL ANALYSIS

Dimerization kinetics

All the dimerization reactions were fitted into a single exponential curve, where the observed rate represents a pseudo-first-order reaction, considering that both unlabeled proteins (FoxP1 and FoxP2) are in excess with respect to the labeled N32C mutant. In that scenario, the true association rate (k_{ass}) is obtained from Equation 1 and depends on the concentration of the unlabeled protein used in homo and heterodimerization:

$$k_{obs} = (k_{ass} \cdot [FoxP]) + k_{diss} \quad \text{(Equation 1)}$$

This equation also obtains k_{diss} plotting k_{obs} vs. [FoxP]. Finally, assuming a simple two-state model, activation free energy changes ($\Delta G^{\ddagger}$) were calculated using the transition state theory approach.[66] In this approach, the kinetic rate (association and dissociation) can be related to an equilibrium between the respective state (monomer or dimer) and the transition state. Thermodynamically, the activation free energy changes ($\Delta G^{\ddagger}$) can be calculated from Equation 2:

$$\Delta G^{\ddagger}_{diss\ or\ ass} = -R \cdot T \cdot ln\,(K^{\ddagger}), \quad \text{(Equation 2)}$$

where $K^{\ddagger}$ is the equilibrium constant between native states (monomer or dimer) and the transition state. Fitting procedures were performed using the software GraphPad Prism 8.0 (www.graphpad.com).

Time-resolved fluorescence analysis

Time-resolved fluorescence decays ($F(t)$) were described using a multi-exponential model (Equation 3)

$$F_{norm}(t) = \sum_i^n x_i \exp\left(t \Big/ \tau_{DA}^{(i)}\right) \quad \text{(Equation 3)}$$

Where χ_i is the i-th population fraction, and $\tau_{DA}^{(i)}$ is the fluorescence lifetime of that population. The relationships between lifetime and species averaged lifetime, and with Donor-Acceptor interdye distances have been reported previously.[33] Fluorescence decays from heterodimers were analyzed using a single or double exponential model, obtaining the respective χ^2_r to compare both models and to choose, using the F-test criteria, the statistically more robust behavior.

Filtered FCS (fFCS)

To analyze the hydrodynamic properties of heterodimers, each $G_{GG}^{(DA)}(t_c)$ was fitted with the model function (Equation 4) that considers a 3-dimensional Gaussian confocal volume to identify the characteristic time of diffusion t_D.

$$G(t_c) = \frac{1}{N}\left(\frac{1}{1+\frac{t}{\tau_D}}\right)\sqrt{\left(\frac{1}{1+\left(\frac{\omega}{z}\right)^2\frac{t}{\tau_D z}}\right)}\left(1 - |d| + |d| exp\left(-\frac{t}{t_{pho}}\right)\right) + B \quad \text{(Equation 4)}$$

Where N is the average number of particles in the confocal volume, ω and z are the axes for the geometrical volume, and $\tau_D = \frac{\omega^2}{4D}$, where D is the characteristic diffusion constant.

Fitting of correlation curves was performed using Chisurf software. Error analyses of parameters for cross-correlation fits was performed by evaluating the chi-squared surface corresponding to the variation of each fit parameter. The error range was determined by using the F-test to compare the chi-squared values

sampled from the chi-squared distribution to the presented fit and identifying the parameter range corresponding to a confidence interval of 2σ.

The auto and cross-correlated signals were fitted to Equations 5 and 6:

$$AC_{(HF-HF,LF-LF)}(t) = B + \frac{1}{N\left(1+\frac{t}{\tau_D}\right)\sqrt{1+\frac{t}{s^2\tau_D}}}\left(1 + A_L \exp\left(-\frac{t}{\tau_L}\right) - A_L + \sum_{i=1}^{3} A_i \exp\left(-\frac{t}{\tau_i}\right) - A_i\right)\left(1 + A_P \exp\left(-\frac{t}{\tau_P}\right) - A_P\right) \quad \text{(Equation 5)}$$

$$CC_{(HF-LF,LF-HF)}(t) = B + \frac{1}{N\left(1+\frac{t}{\tau_D}\right)\sqrt{1+\frac{t}{s^2\tau_D}}}\left(1 - A_{CC}\sum_{i=1}^{3} A_i \exp\left(-\frac{t}{\tau_i}\right)\right)\left(1 - A_L \exp\left(-\frac{t}{\tau_L}\right)\right)\left(1 + A_P \exp\left(-\frac{t}{\tau_P}\right) - A_P\right) \quad \text{(Equation 6)}$$

Where B is the baseline value of each curve, S is the geometrical volume, τ_i and A_i are the anticorrelation time and its amplitude, and A_{CC} corresponds to the cross-correlation amplitude term. Photophysical terms for are described by τ_P and A_P, which correspond to times and its amplitude, respectively. Any longer exchange and/or bleaching term is contained in τ_L and its related amplitude A_L. Fitting of correlation curves was performed using Chisurf software.

Photon distribution analysis (PDA)

We used Probability Distribution Analysis (PDA)[36,37,67] to model the Efficiency (E) distributions and identify mean FRET distances ($\langle R_{DA}\rangle_E$) and their corresponding uncertainties. To properly account for the heterogeneity in the duration of bursts, bursts are split into equal time windows per burst with multiple time window sizes (Δt = 1, 3 and 5 ms), and the FRET indicator E histogram is obtained for each time window size. We globally fit all time windows with different models that vary in increasing level of complexity and the best model is selected based on the global figure of merit χ_r^2 and statistical uncertainties (see "statistical uncertainties and error analysis" section). A static model considering one and two states was not enough to fully describe the E histograms at these time windows. For details on the models used in each data analysis see "FRET models" section.

Molecular dynamics simulations

To analyze CCR simulations, we used a mathematical approach already described by us[31] and initially by others.[30,68] Briefly, this approach solves thermodynamic integration for the Confinement and Release step of the free-energy calculation, where the Release step is simply the negative of the Confinement step (Equation 7):

$$\Delta G_{conf} = \sum_{k_i}^{k_{f-1}} \frac{(\chi_{i+1}k_{i+1} - \chi_i k_i)}{\frac{\ln(\chi_{i+1}) - \ln(\chi_i)}{\ln(k_{i+1}) - \ln(k_i)}+1}; \text{where } \chi = \langle N \cdot RMSD^2 \rangle \quad \text{(Equation 7)}$$

In this equation, k_i is the *i-th* restraining constant starting from k_1 = 0.000025 kcal mol^{-1}, and k_f is the last restraining constant 419 kcal·mol^{-1}, and N corresponds to the total number of atoms in the system. This analysis only requires calculating the RMSD from the simulations, which decreases non monotonically for increasing restraining constraints. This calculation is done for each system, namely FoxP1 dimer, FoxP1 monomer, FoxP2 dimer and FoxP2 monomer. Then each corresponding system is summed correspondingly. Meanwhile, the conversion or transformation step is also calculated using the deeply energy-minimized structure obtained previously for each system and calculating its atomic frequencies using normal mode analysis (NMA) through NAB in Amber. Then each frequency is used as input for solving the Equation 8:

$$\Delta G_{HO} = \Delta E - kT\sum_{i=1}^{3N-6} \ln\left(\frac{\nu_1}{\nu_2}\right) \quad \text{(Equation 8)}$$

With v being the *i-th* non-zero frequency calculated from NMA, ΔE corresponding to the potential energy difference between configurations 2 and 1, and kT is the Boltzmann factor at 298.15 K. This calculation allows to determine the free energy of Conversion from state A to B by only comparing their relative frequencies, derived from the Harmonic Oscillator or HO approach. The same approach was used to estimate the per-residue free-energies, considering the fluctuations of each residue instead of the whole structure.

Finally, to estimate the errors in the free-energy calculations, we divided each simulation into 5 equal blocks, and calculated ΔG as per Equation 8, but using the *i-th* block instead of the whole simulation for each k_i value. Then, the standard deviation of the calculated ΔG from each block was propagated through the whole range of k_i.

Time-structure based independent component analysis (TICA) and clustering

Markov models were generated with the PyEMMA and DeepTime Python Library version 2.5.11 and 0.4.3 respectively.[69] For each simulated system, we combined the three five hundred nanosecond classical MD trajectories. As features for the subsequent analysis, we chose the internal coordinates of the residues used in the FRET pairs. Then we performed time-lagged independent component analysis (TICA) of the featured trajectories with a lag time of 50 frames equivalent to 250 ps. For each system, with and without DNA, we constructed a three and four-state Markov chain model respectively, resampled structures probabilistically from each metastable state, computed the stationary distribution for each metastable state. Distance distributions for the FRET pairs were estimated using the samples drawn from each metastable state.

Statistical uncertainties and error analysis

To determine the statistical uncertainties of the different model functions and resulted fit parameters for PDA, time-resolved fluorescence and fFCS, we employed a search algorithm over the figure of merit χ_r^2 against all varying parameters considering the number of degrees of freedom (ν) and the number of free parameter for each model, as described by us in previous reports.[28,34,48] We use the statistical test of the F-distribution with a 95% ($P = 2\sigma$) confidence interval to define the upper and lower limits for each parameter with respect to the χ_r^2 surface.

FRET models

To better understand the properties of the heterodimer at nanoseconds timescale, we considered the Worm-Like Chain (WLC) model, that describes how a dynamic population that exhibits WLC dynamics would diverge from the static FRET line, considering the effects of fluorophore linker dynamics on the observed FRET efficiency.[70] In this model, the fluorophores are coupled to the protein by a flexible linker, and thus, show Gaussian chain like characteristics. We describe the spatial probability distributions of the dyes by 3-dimensional Gaussians $p(RL, R_{DA})$ between the donor and the acceptor for a single conformation with a separation distance RL, as nicely described by.[71,72]