UNIVERSITY OF CALIFORNIA SAN DIEGO

Conformational Dynamics in NF-κB Transcriptional Regulation

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

in

Chemistry

by

Wei Chen

Committee in charge:

Professor Elizabeth A. Komives, Chair
Professor Neal Krishna Devaraj
Professor James A. McCammon
Professor Johannes Schoeneberg
Professor Douglas E. Smith

2020

The Dissertation of Wei Chen is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____
Chair

University of California San Diego
2020

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I am fortunate to have an extremely supportive advisor who helped me reach my potential. I thank Professor Elizabeth Komives, Betsy, for leading me into the fantastic world of protein biophysics, for bringing me into exciting and challenging projects, for troubleshooting experiments and fixing instruments with me, for providing me every opportunity for career success, for willing to help with whatever I encountered in life, and for her inspiration and enthusiasm at every moment.

I am grateful to have many wonderful mentors throughout the PhD journey. I thank Clarisse Ricci for kindly teaching me how to perform all-atom molecular dynamics simulations. I enjoyed our meetings and the fun lunch at the iCafé. I thank Nick Schafer for his insightful guidance on the TAD structure prediction project and for being a genuine friend I can talk to during my visits to Rice University. I thank Professor Hajin Kim for both his overseas and on-site advice on the instrumentation of TIRF microscopy. I thank Professor Andy McCammon for his great supports, insights, and kindness.

I would never forget the time I shared with the Komives lab. Constanza Torres Páris, Connie, is a cool scientist with an international vision and a kind heart. I thank her for being a wonderful human being whom I, an introvert, enjoy talking to. I thank Riley Peacock for being a great ally with whom I start and finish PhD, and for bringing a lot of fun to the lab. I thank Shih-Ting Huang, Terry, for being a curious and hard-working undergrad to work with. I thank Jimmy Marion, Holly Dembinski, Kristen Ramsey, and Dominic Narang, for passing on their knowledge of NF-κB to me. I thank Maj Ghassemian for his knowledge in protein mass spectrometry and for being a friendly lab neighbor. I also thank Hannah Baughman, Calvin Lin, Philip Belzeski, Steven Silletti, Ryan Lumpkin, Francis Alipranti, Alla Ahmad, and other Komives group members for making the lab a supportive environment. I will miss the fun parties at Betsy's, the evening teatime,

the unfinished lab Pokémon championship, our annual lab cleanup on Betsy's birthday, and the world's cleanest and most organized lab space.

I thank my parents for teaching me the value of hard work and integrity while giving me a lot of freedom to explore the world. I thank my sister Lin for being one of my best friends. I thank my friends in Taiwan for constantly supporting me and sharing their lives with me.

Finally, I would like to express my gratitude to Chi-Yun Lin for being an extraordinary partner and an inspirational scientist to me. I was fortunate to grow together with you over the past seven years of our long-distance relationship.

Chapter 2, in part, appears in "RelA-Containing NFκB Dimers Have Strikingly Different DNA-Binding Cavities in the Absence of DNA" by Narang, D.*; Chen, W.*; Ricci C. G.; Komives, E. A., published in *Journal of Molecular Biology.* **2018,** *430*, 1510-1520. The dissertation author is one of the co-first authors denoted by asterisks.

Chapter 6, in part, appears in "Exclusivity and Compensation in NF-κB Dimer Distributions and IκB Inhibition." by Ramsey, K. M.*; Chen, W.*; Marion J. D.*; Bergqvist, S.; Komives, E. A., published in *Biochemistry* **2019**, *58* (21), 2555-2563. The dissertation author is one of the co-first authors denoted by asterisks.

I thank the Ministry of Education of Taiwan for the partial financial supports for two years.

VITA

2014 B.S. in Chemistry, National Taiwan University

2017 M.S. in Chemistry, University of California San Diego

2020 Ph. D. in Chemistry, University of California San Diego

PUBLICATIONS

Ramsey, K. M.*; **Chen, W.***; Marion, J. D.*; Bergqvist S; Komives, E. A. Exclusivity and Compensation in NFκB Dimer Distributions and IκB Inhibition. *Biochemistry.* **2019,** *58,* 2555-2563. *Equal contribution.

Narang, D.*; **Chen, W.***; Ricci, C. G.; Komives, E. A. RelA-containing NFκB Dimers Have Strikingly Different DNA-Binding Cavities in the Absence of DNA. *Journal of Molecular Biology.* **2018,** *430*, 1510-1520. *Equal contribution.

Ramsey, K. M.; Dembinski H. E.; **Chen, W.**; Ricci, C. G.; Komives, E. A. DNA and IκBα Both Induce Long Range Conformational Changes in NFκB. *Journal of Molecular Biology.* **2017,** *429*, 999-1008.

Hwang-Fu, Y.-H.; **Chen, W.**; Cheng, Y.-C. A Coherent Modified Redfield Theory for Excitation Energy Transfer in Molecular Aggregates. *Chemical Physics*. **2015**, *447*, 46–53.

ABSTRACT OF THE DISSERTATION

Conformational Dynamics in NF-κB Transcriptional Regulation

by

Wei Chen

Doctor of Philosophy in Chemistry

University of California San Diego, 2020

Professor Elizabeth A. Komives, Chair

The NF-κB family of transcription factors is a central mediator of immune and inflammatory responses. Dysregulation of NF-κB results in numerous diseases including cancers. Immediate activation and deactivation of NF-κB through macromolecular interactions is crucial. This dissertation investigated the role of conformational dynamics in these interactions by combining single-molecule FRET (smFRET) and molecular dynamics (MD) simulations. All-atom MD simulations showed that three different NF-κB dimers, RelA–p50, the RelA homodimer

and the p50 homodimer, display distinct relative domain motions leading to different degrees of binding cavity exposure that correlate with experimental DNA-binding rates. These motions were directly observed in real-time by smFRET, revealing a heterogeneous conformational ensemble. Interconversion between conformational states and DNA-binding events occur on the same time scale, suggesting they are coupled. Remarkably, the motions were also observed when NF-κB was bound to DNA, inferring a dynamic picture unexpected from crystal structures. These motions were allosterically altered by the inhibitor protein IκBα, which is known to accelerate DNA-dissociation and prevent DNA-binding. Coarse-grained MD simulations were also performed to predict the conformational ensemble of the disordered transactivation domain of NF-κB and captured the helix propensity of the motifs that become structured when bound to the transcription coactivator CBP domains. Subunit exchange between NF-κB dimers was also investigated with simple chemical kinetic equations. Together, this dissertation provides a molecular understanding of the conformational dynamics in NF-κB transcriptional regulation and may contribute to the ultimate deciphering of the dynamics-function relationship of proteins.

Chapter 1 Introduction

The NF-κB family of transcription factors responds to a large number of extracellular stimuli and regulates patterns of gene expression[1]. As a critical regulator, NF-κB is found in almost all animals and all types of cells and is implicated in various cellular functions including cell growth, proliferation, apoptosis, development, and stress responses. Dysregulation of NF-κB is found in numerous diseases, especially chronic inflammation and cancers[2].

In the absence of stimuli, NF-κB is held inactive in the cytoplasm by the inhibitory protein IκB. When a cell receives signals such as cytokines or viral attack, the IκB kinase (IKK) is activated to phosphorylate NF-κB-bound IκBα, which is then ubiquitylated and degraded by the proteasome[3] (**Figure 1.1**). Degradation of IκBα exposes the nuclear localization signal (NLS) of NF-κB and shifts the equilibrium of NF-κB from being in the cytoplasm to the nucleus, where NF-κB binds to DNA. NF-κB recognizes DNA with the consensus κB sequence 5'-GGGRNWYYCC-3', where N is any base, R is a purine, W is an adenine or thymine, and Y is a pyrimidine. There are over 10^4 of κB sites in the human genome, primarily in the promoter and enhancer regions[4].

As a rapidly inducible transcription factor, NF-κB activates transcription of hundreds of genes upon stimulation, including the one encoding its own inhibitor, IκBα. Activation of IκBα synthesis creates a negative feedback loop, in which the newly synthesized IκBα relocates to the nucleus and binds to NF-κB to block DNA binding[1]. Eventually the NF-κB:IκBα complex is then exported from the nucleus[5].

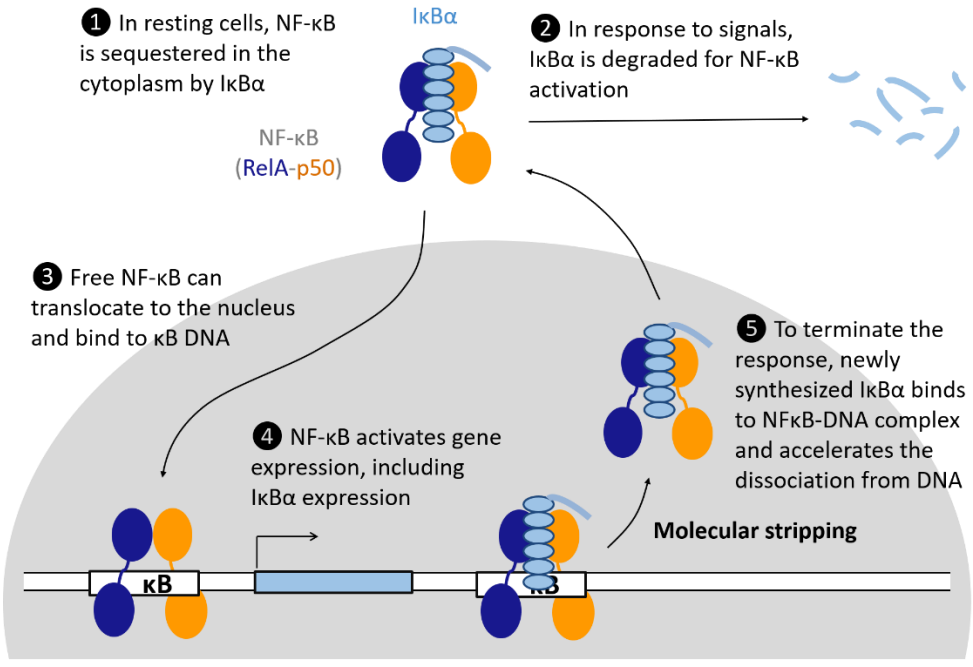


Figure 1.1 Schematic of NF-κB transcriptional regulation.

In mammals, there are five protein members in the NF-κB family—RelA (or p65), RelB, cRel, p50, and p52—which can form homodimers and heterodimers for DNA-binding. Among the 15 potential dimers that can be formed, 12 of them are capable of DNA-binding and 9 of them can activate transcription (**Figure 1.2 a**). Each family member has a conserved Rel homology domain (RHD), which contains two immunoglobulin (Ig)-like beta-barrel domains connected by a linker. The N-terminal Ig-like domains (NTDs) recognize target DNA with mostly irregular loops, and the C-terminal Ig-like domains are responsible for dimerization and therefore are referred to as dimerization domains (**Figure 1.2 b**). Following the RHD is the nuclear localization signal (NLS). RelA, RelB, and c-Rel have a C-terminal transactivation domain (TAD) which is absent in p50 and p52, and thus only dimers containing RelA, RelB, or c-Rel are capable of transcription activation.

The RelA–p50 heterodimer is the most abundant NF-κB dimer in general, while other NF-κB dimers have been found in different cell types and activate different sets of genes. Each NF-

κB dimer preferentially binds to members of the IκB family. RelA–p50 is primarily regulated by IκBα, while IκBβ stabilizes the RelA homodimer and IκBε binds to RelA- and cRel-containing dimers. IκB proteins share an ankyrin repeat (AR) domain which forms a large binding interface for NF-κB dimers (**Figure 1.2c**).

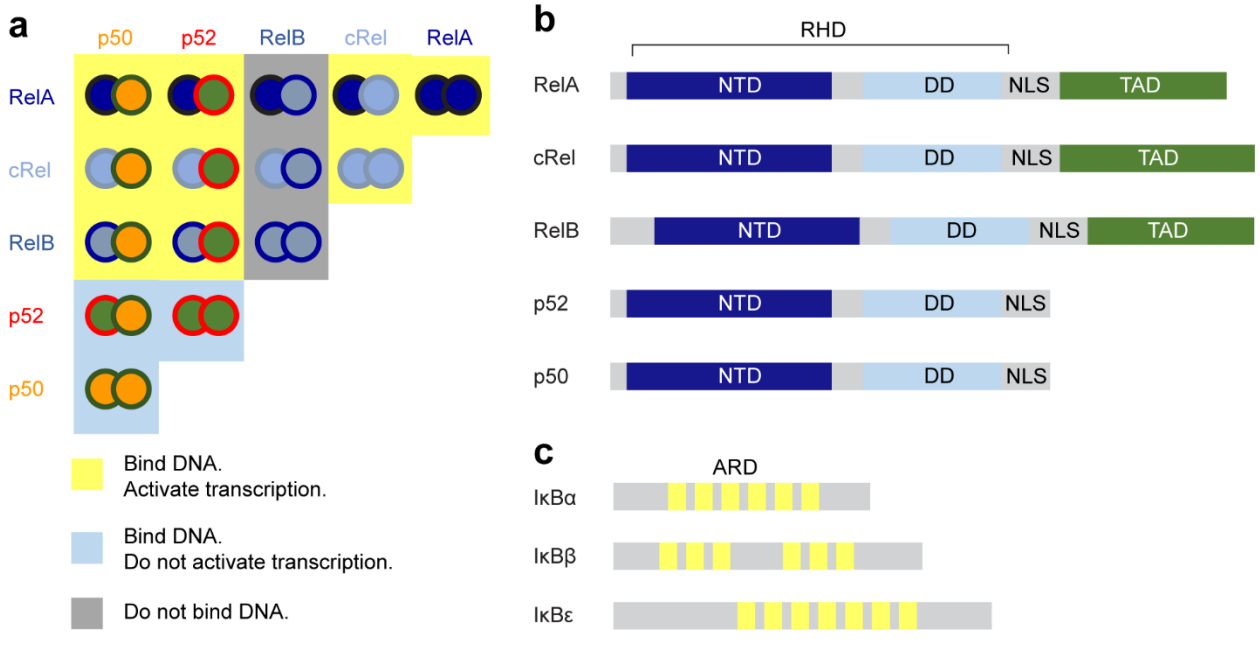


Figure 1.2 Components of the NF-κB signaling system. **(a)** The 5 NF-κB family members and form 15 potential dimers, of which 12 can bind DNA and 9 can activate transcription. **(b)** The NF-κB family of proteins share a Rel homology domain (RHD) containing an N-terminal domain (NTD) and a dimerization domain (DD). There is a nuclear localization signal (NSL) following the RHD. RelA, RelB, and c-Rel contain a transactivation domain (TAD). **(c)** The IκB family of proteins shares an ankyrin repeats domain (ARD).

Crystal structures of NF-κB bound to DNA reveal the conserved structures of the RHD and the NF-κB:DNA interactions among dimers (**Figure 1.3a**). Interestingly, NF-κB does not utilize classical DNA-binding motifs such as helix-turn-helix but instead contacts the DNA with loops in the NTDs, the dimerization domains, and the linkers. No crystal structures are available for free NF-κB or the TAD, possibly due to their dynamic nature in the absence of binding partners.

There are two crystal structures for IκBα-bound RelA–p50, each of which covers different regions of IκBα and NF-κB (**Figure 1.3b**). One of the structures shows the NLS of RelA bound to IκBα, explaining the cytoplasmic sequestration[6], while the other captures the PEST sequence in the C-terminus of IκBα[7], which is important for competing for binding with DNA. IκBα binds to the NF-κB dimerization domains with a large interface formed by the ARs. In both structures, the NTD of p50 is missing in the protein construct and therefore may create space for the RelA NTD to adopt an unnatural orientation. The orientation of the RelA NTD is different in the two NF-κB:IκBα structures and the NF-κB:DNA structure, suggesting rigid-body-like motions of the NTDs can exist. Since the two NTDs in a dimer are responsible for recognizing κB sites, positions and orientations of the NTDs in free and IκB-bound NF-κB are important to understand the molecular details of DNA-binding and inhibition.

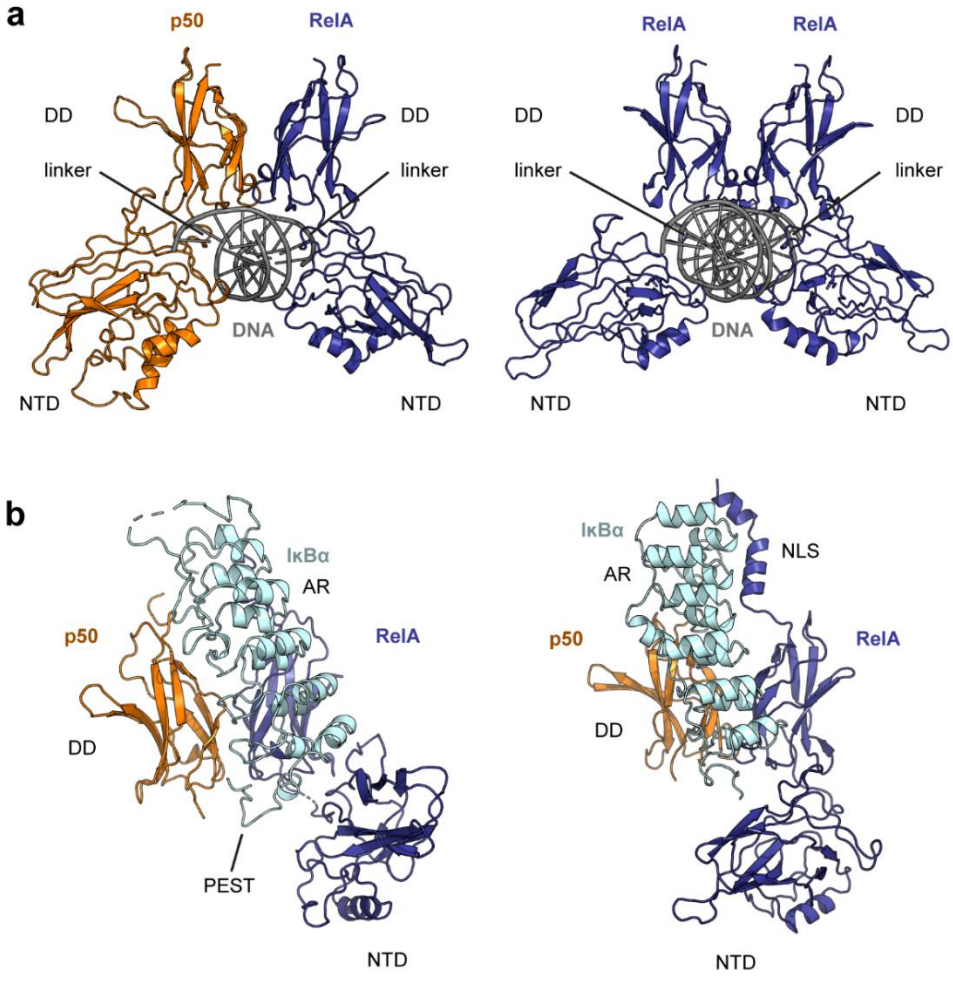


Figure 1.3 Structures of NF-κB bound to DNA and IκBα. **(a)** Crystal structures of NF-κB dimers RelA–p50 and the RelA homodimer show conserved RHD structures and protein-DNA interactions. Other NF-κB dimers adopt the same RHD structure and DNA-binding conformation and are not shown here. The NF-κB dimer binds to DNA through the loops in the NTDs and the dimerization domains (DD) and also through the linkers. **(b)** Two IκB-bound NF-κB structures with one showing the PEST sequence of IκBα (left) and the other showing the NLS of RelA (right). IκBα binds to the NF-κB dimerization domains with its ankyrin repeats (ARs). The NTD of p50 is missing from the protein construct that crystallized.

As a master regulator of hundreds of genes, the termination of NF-κB signaling must be efficient. Previous kinetic experiments revealed molecular stripping—a process in which IκBα actively removes NF-κB from DNA[8]. The discovery of molecular stripping suggests that the termination of NF-κB signaling is not just an equilibrium thermodynamic result of IκB

outcompeting DNA for binding to NF-κB but rather it is achieved through IκBα-accelerated dissociation of NF-κB from DNA in the nucleus. During molecular stripping, a transient ternary complex of NF-κB:IκBα:DNA is observed and the DNA dissociation rate is proportional to the IκBα concentration[9]. A stripping impaired IκBα mutant results in slower nuclear exportation of NF-κB in the cell[10], suggesting the termination of NF-κB transcriptional regulation is under kinetic control rather than at equilibrium.

Conformational dynamics plays a key role in the mechanism of molecular stripping as shown by coarse-grained molecular dynamics simulations, particularly the relative motions of NF-κB NTDs[11] and the dynamics of the disordered PEST sequence of IκBα[12]. In the simulations, IκBα twists the relative orientation of the two NTDs in the NF-κB dimer to a conformation where only one of the NTDs binds DNA. This conformational change in the NTDs induced by IκBα weakens the NF-κB:DNA interaction and therefore lowers the dissociation barrier. In the later stage of molecular stripping, the negatively charged PEST sequence of IκBα electrostatically repels DNA and prevents DNA rebinding[11,12].

It has been shown by hydrogen-deuterium exchange mass spectrometry (HDX-MS) that the NTDs in RelA–p50 become more protected from hydrogen-deuterium exchange when NF-κB is bound to IκBα, suggesting that even though IκBα binds to the NF-κB dimerization domains, it can induce conformational changes in the NTDs[13]. Neutralization of the negatively charged PEST sequence of IκBα results in a stripping impaired mutant both in test tubes and in cells[10].

The full picture of NF-κB transcriptional regulation by conformational dynamics is far from complete, particularly due to its highly heterogeneous nature that is difficult to resolve in bulk experiments. In this dissertation, I tackled the problem by observing the conformational dynamics of NF-κB one molecule at a time, either with single-molecule FRET experiments or molecular dynamics simulations, to understand its role in transcriptional regulation.

In Chapter 2, the conformational dynamics of three NF-κB dimers was investigated with all-atom molecular dynamics simulations. The results reveal distinct relative domain motions and conformational ensembles of the three dimers in the absence of DNA, which appear to be important in controlling the rate of DNA association.

Chapter 3 describes the establishment of single-molecule FRET technique to probe the relative domain motions predicted by molecular dynamics simulations.

In Chapter 4, the predicted relative domain motions of NF-κB are directly observed in real-time with single-molecule FRET experiments. The results reveal extremely slow and heterogeneous motions in the NTDs of NF-κB and their responses to DNA and IκBα, suggesting the importance of domain motions in both DNA-binding and molecular stripping.

In Chapter 5, coarse-grained molecular dynamics simulations are used to predict the possible structures of the disordered TAD of RelA.

Chapter 6 explores the subunit exchange between NF-κB dimers with chemical kinetic equations.

Together, this dissertation provides a molecular understanding of the conformational dynamics in NF- κB transcriptional regulation and may contribute to the ultimate deciphering of the dynamics-function relationship of proteins.

References

1. Hoffmann, A.; Levchenko, A.; Scott, M. L.; Baltimore, D., The IkappaB-NF-kappaB signaling module: temporal control and selective gene activation. *Science* **2002,** *298* (5596), 1241-5.

2. Kumar, A.; Takada, Y.; Boriek, A. M.; Aggarwal, B. B., Nuclear factor-kappaB: its role in health and disease. *J Mol Med (Berl)* **2004,** *82* (7), 434-48.

3. Traenckner, E. B.; Baeuerle, P. A., Appearance of apparently ubiquitin-conjugated I kappa B-alpha during its phosphorylation-induced degradation in intact cells. *J Cell Sci Suppl* **1995,** *19*, 79-84.

4. Martone, R.; Euskirchen, G.; Bertone, P.; Hartman, S.; Royce, T. E.; Luscombe, N. M.; Rinn, J. L.; Nelson, F. K.; Miller, P.; Gerstein, M.; Weissman, S.; Snyder, M., Distribution of NF-kappaB-binding sites across human chromosome 22. *Proc Natl Acad Sci U S A* **2003,** *100* (21), 12247-52.

5. Arenzana-Seisdedos, F.; Turpin, P.; Rodriguez, M.; Thomas, D.; Hay, R. T.; Virelizier, J.-L.; Dargemont, C., Nuclear localization of I kappa B alpha promotes active transport of NF-kappa B from the nucleus to the cytoplasm. *Journal of cell science* **1997,** *110* (3), 369-378.

6. Jacobs, M. D.; Harrison, S. C., Structure of an IκBα/NF-κB complex. *Cell* **1998,** *95* (6), 749-758.

7. Huxford, T.; Huang, D.-b.; Malek, S.; Ghosh, G., The crystal structure of the IκBα/NF-κB complex reveals mechanisms of NF-κB inactivation. *Cell* **1998,** *95* (6), 759-770.

8. Bergqvist, S.; Alverdi, V.; Mengel, B.; Hoffmann, A.; Ghosh, G.; Komives, E. A., Kinetic enhancement of NF-kappaBxDNA dissociation by IkappaBalpha. *Proc Natl Acad Sci U S A* **2009,** *106* (46), 19328-33.

9. Alverdi, V.; Hetrick, B.; Joseph, S.; Komives, E. A., Direct observation of a transient ternary complex during IkappaBalpha-mediated dissociation of NF-kappaB from DNA. *Proc Natl Acad Sci U S A* **2014,** *111* (1), 225-30.

10. Dembinski, H. E.; Wismer, K.; Vargas, J. D.; Suryawanshi, G. W.; Kern, N.; Kroon, G.; Dyson, H. J.; Hoffmann, A.; Komives, E. A., Functional importance of stripping in NFkappaB signaling revealed by a stripping-impaired IkappaBalpha mutant. *Proc Natl Acad Sci U S A* **2017,** *114* (8), 1916-1921.

11. Potoyan, D. A.; Zheng, W.; Komives, E. A.; Wolynes, P. G., Molecular stripping in the NF-kappaB/IkappaB/DNA genetic regulatory network. *Proc Natl Acad Sci U S A* **2016,** *113* (1), 110-5.

12. Potoyan, D. A.; Zheng, W.; Ferreiro, D. U.; Wolynes, P. G.; Komives, E. A., PEST Control of Molecular Stripping of NFkappaB from DNA Transcription Sites. *J Phys Chem B* **2016,** *120* (33), 8532-8.

13. Ramsey, K. M.; Dembinski, H. E.; Chen, W.; Ricci, C. G.; Komives, E. A., DNA and IkappaBalpha Both Induce Long-Range Conformational Changes in NFkappaB. *J Mol Biol* **2017,** *429* (7), 999-1008.

Chapter 2 Large-Scale Relative Domain Motions of NF-κB Revealed by Molecular Dynamics

Adopted and updated from Narang, D.*; Chen, W.*; Ricci C. G.; Komives, E. A., RelA-Containing NFκB Dimers Have Strikingly Different DNA-Binding Cavities in the Absence of DNA. *Journal of Molecular Biology.* **2018,** *430,* 1510-1520. * Equal contribution.

2.1 Introduction

The five protein members in the NF-κB family of transcription factors, RelA (or p65), RelB, c-Rel, p50, and p52, form homo- and heterodimers that recognize the consensus κB DNA sequence and regulate transcription. In each NF-κB protein, there are two immunoglobulin-like (Ig-like) β-barrel subdomains connected by a 10-amino-acid linker. The NF-κB proteins dimerize through the C-terminal Ig-like domains and recognize DNA through the irregular loops of the two N-terminal domains (NTDs) in the dimer. The C-terminal Ig-like domains are referred to as the dimerization domains, but they also contribute substantial DNA binding affinity through phosphate-ribose contacts by amino acids from five ordered loops[1].

This chapter focuses on three NF-κB dimers, the canonical RelA–p50 heterodimer, the RelA homodimer, and the p50 homodimer. Among these three dimers, RelA–p50 binds to single-κB-site-containing DNA sequences with the highest affinity in the 3-30 nM range[2-3]. The DNA-binding affinity of the p50 homodimer is of the same order of magnitude but slightly weaker[3], and the RelA homodimer generally binds much less tightly to DNA, with the difference being the most dramatic for the IgκB promoter sequence—K_D of 6.3 nM for RelA–p50 vs. 7.4 μM for the RelA homodimer[2-3].

It is not clear from the DNA-bound crystal structures of the three dimers[4-6] why they have different DNA-binding affinities. The order of affinity from RelA–p50 to the p50 homodimer to

the RelA homodimer indicates that affinity is not subunit additive and is unlikely to be determined by the number of residue-residue contacts in the NF-κB:DNA complex. Structures of free NF-κB dimers in the absence of DNA or inhibitory proteins are lacking, preventing a mechanistic understanding of the DNA-binding kinetics and regulation. This chapter describes the use of all-atom molecular dynamics (MD) simulations to explore the structure and conformational dynamics of three NF-κB dimers, RelA–p50, the RelA homodimer, and the p50 homodimer, in their free states, to understand the role of protein dynamics in the NF-κB:DNA interaction.

MD simulations showed that the three dimers underwent distinct conformational changes via large-scale relative domain motions, which lead to the formation of a novel interface between NTDs in the RelA homodimer not found in the other two dimers. The new interface in the RelA homodimer was observed in hydrogen-deuterium exchange mass spectrometry (HDX-MS) as a decrease in the hydrogen-deuterium exchange level in that region. A slower DNA association rate was obtained for the RelA homodimer by stopped-flow fluorescence experiments, indicating an interface between NTDs could impede DNA-binding. The results suggested that different NF-κB dimers utilize distinct relative motions in the NTDs to control their DNA-binding kinetics.

2.2 Methods

Molecular Dynamics Simulations. The atomic coordinates of *mus musculus* NFκB dimers, RelA–p50, the RelA homodimer, and the p50 homodimer, were obtained from crystal structures of NFκB-DNA complexes (PDB entries: 1LE9, 1RAM, and 1NFK), in which RelA ranges from residue 19 to 291, missing the NLS region (292-321), and p50 ranges from residue 39 to 350. DNA molecules were computationally removed for simulating free NFκB. I also simulated a mutant of the RelA–p50 heterodimer, the QQQQ mutant, in which four glutamic acids at positions 284 to 287 were mutated to glutamine with Pymol[7] to neutralize this negatively charged

region. Prior to simulations, the three NF-κB dimers were aligned. The RMSD of the aligned structures was 2.2 Å between RelA–p50 and the RelA homodimer, 3.4 Å between RelA–p50 and p50 homodimer, and 2.9 Å between RelA and p50 homodimers as calculated by Lovoalign[8]. The protonation states of ionic residues were assigned at pH 7 using the H++ server[9]. Proteins were solvated with TIP3P water[10] in a rectangular box with a 15 Å buffer region using tleap in the AMBER package[11]. Chloride ions were added to keep the systems electrically neutral. The total numbers of atoms are approximately 150000 for RelA–p50, 160000 for the RelA homodimer, and 180000 for the p50 homodimer. To allow the bonds and angles to relax and eliminate possible bad contacts, the complete and solvated model was submitted to the following refinement procedure using AMBER 14[11] and the ff14SB forcefield[12]: i) 1000 steps of energy minimization of the water molecules with the protein atoms restrained and ii) 2500 steps of energy minimization with no position restraints. Each stage was performed using the conjugate gradient method for 1000 and 2500 cycles.

Molecular dynamics simulations were performed with periodic boundary conditions and an integration time step of 2 fs. A Langevin thermostat was used for temperature control. The SHAKE algorithm was used to constrain hydrogen atoms during the simulation. The particle-mesh Ewald method[13] was used to handle long-range electrostatic interactions. A cutoff of 8 Å was used for short-range nonbonded interactions. Equilibration of the systems was performed in two stages. First, the system was heated up from 100 K to 300 K with position restraints on solutes with a force constant of 10 kcal/mol for 20 ps. Next, the entire system was equilibrated at 300 K and 1 bar for 1 ns to equilibrate the density. After equilibration, two independent 400-ns production runs were performed for each system. In the NVT regime, atomic coordinates were recorded every 2 ps and the resulting trajectories were visualized using Visual Molecular Dynamics (VMD)[14]. CPPTRAJ[15] in the AMBER package was used to perform RMSD and distance analysis. The

distance between the two NTDs in the NFκB dimers was defined as the distance between their centers of mass. VMD was used to performed salt bridge analysis. A salt bridge is defined as an acidic-basic residue pair with a distance between their O and N atoms less than 3.2 Å.

Electrostatics analysis with APBS. Electrostatic calculations were performed with the Adaptative Poisson-Boltzmann Solver (APBS)[16] plugin in Pymol[7], for the initial relaxed and minimized NFκB dimer structures described above. Calculations were performed at 298 K, with the traditional Linearized Poisson Boltzmann Equation (LPBE), with a grid size of 0.5 Å. using dielectric constants of 2 and 78 for the protein and water environments, respectively. We used an internal dielectric constant of 2.0 to represent the non-polar environment of the protein and an external constant of 78.0 to represent the polar aqueous environment of the solvent.

Stopped-Flow Fluorescence Experiments. Stopped flow experiments were carried out on an Applied Photophysics SX-20 apparatus at 25 °C. A hairpin DNA sequence corresponding to the ten-nucleotide IFN-κB site (underlined) 5'-AmMC6/GGGAAATTCCTCCCCCAGGAATTTCCC-3' (IDT Technologies) was labeled with pyrene succinimide (Sigma) and purified as described[17]. The mixing ratio of the apparatus is 1:1. The final concentration of the DNA was 0.1 μM. Different concentrations of NF-κB dimers after mixing were 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 1 μM. The pyrene-labeled DNA was excited by 343 nm wavelength light and the fluorescence signal was filtered with a 350 nm longpass filter. The data was collected for 0.5 seconds with 5000-time points. Multiple traces were collected and averaged for a better signal to noise ratio. The obtained data were fitted using a bi-exponential function.

2.3 Results

2.3.1 Distinct Domain Motions of NFκB Dimers Revealed by MD

To explore the conformational dynamics of NF-κB dimers, I set up all-atom MD simulations for the RelA–p50 heterodimer, the RelA homodimer, and the p50 homodimer in their free states. The DNA-bound conformations of NF-κB from NF-κB:DNA co-crystal structures (PDB entries: 1LE9, 1RAM, and 1NFK) were used as starting structures for the simulations. The three dimers adopted a similar DNA-bound conformation with a root mean square deviations (RMSD) of 2.2 Å between RelA–p50 and RelA homodimer, 3.4 Å between RelA–p50 and p50 homodimer, and 2.9 Å between RelA and p50 homodimers (**Figure 2.1 a**). For each dimer, two independent 400 ns simulations were performed starting with the same atomic coordinates but different random initial velocities.

The results showed that, within 100 ns, all three dimers relaxed from the starting conformations substantially via a large-scale conformational change caused by the relative movement of the two NTDs, as shown in the changes between their centers of mass over time (**Figure 2.1 b**). Despite the high similarity of their starting conformations derived from the DNA-bound structures, the three dimers underwent distinctive relative domain motions. The two NTDs of RelA–p50 twisted around the linkers while moving apart from each other, causing the inter-NTD distance to increase from 52 Å in the starting conformation to over 70 Å in the end. For the p50 homodimer, the "opening" of the two NTDs caused an even larger change in the inter-NTD distance from 54 Å to over 80 Å. The p65 homodimer, however, behaved completely differently. In the p65 homodimer, the two NTDs twisted around the linker and moved away from their original positions, but instead of moving apart, the NTDs came closed together and ended up with a shorter inter-NTD distance compared to the starting point.

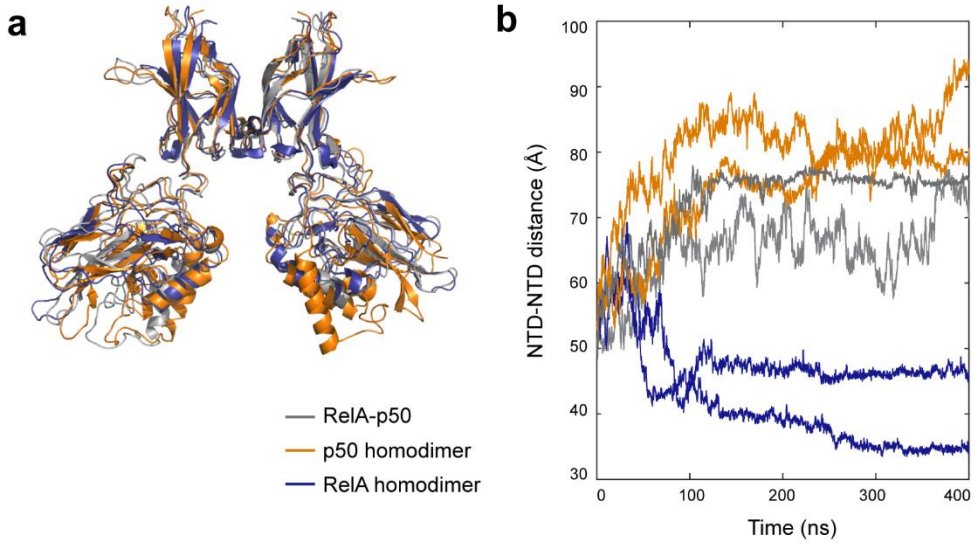


Figure 2.1 Three NF-κB dimers undergo distinctive relative domain motions in MD simulations. Grey: RelA–p50; Orange: the p50 homodimer; Blue: the RelA homodimer. **(a)** The three dimers, the RelA–p50 heterodimer, the RelA homodimer, and the p50 homodimer, started with similar conformations derived from their DNA-bound crystal structures. **(b)** Distance between the centers of mass of the two NTDs in the dimer as a function of time. The NTD-NTD distance increased in RelA–p50 and the p50 homodimer via an opening motion, while decreased in the RelA homodimer via a closing motion. Two independent simulations were performed for each dimer.

For each dimer, the two independent simulations were consistent in the general direction of the domain motions—opening in RelA–p50 and the p50 homodimer and closing in the RelA homodimer. However, each simulation visited a different set of conformations in its route as shown in the NTD-NTD distance change over time (**Figure 2.1 b**) and the representative structures at different time points (**Figure 2.2**). For RelA–p50, the two simulations reached the same degree of openness with different orientations of the NTDs, which was also the case for the p50 homodimer. For the RelA homodimer, the two simulations ended up with different inter-NTD distances. While a complete sampling of the domain positions and orientations is not possible with sub-microsecond simulations, the MD results revealed distinct relative domain motions for RelA–p50, the RelA homodimer, and the p50 homodimer that were unexpected from DNA-bound crystal structures.

Remarkably, in both RelA homodimer simulations, a new protein-protein interface appeared to form between the two NTDs, although the interfaces that formed in each trajectory were not the same due to different rotation directions.

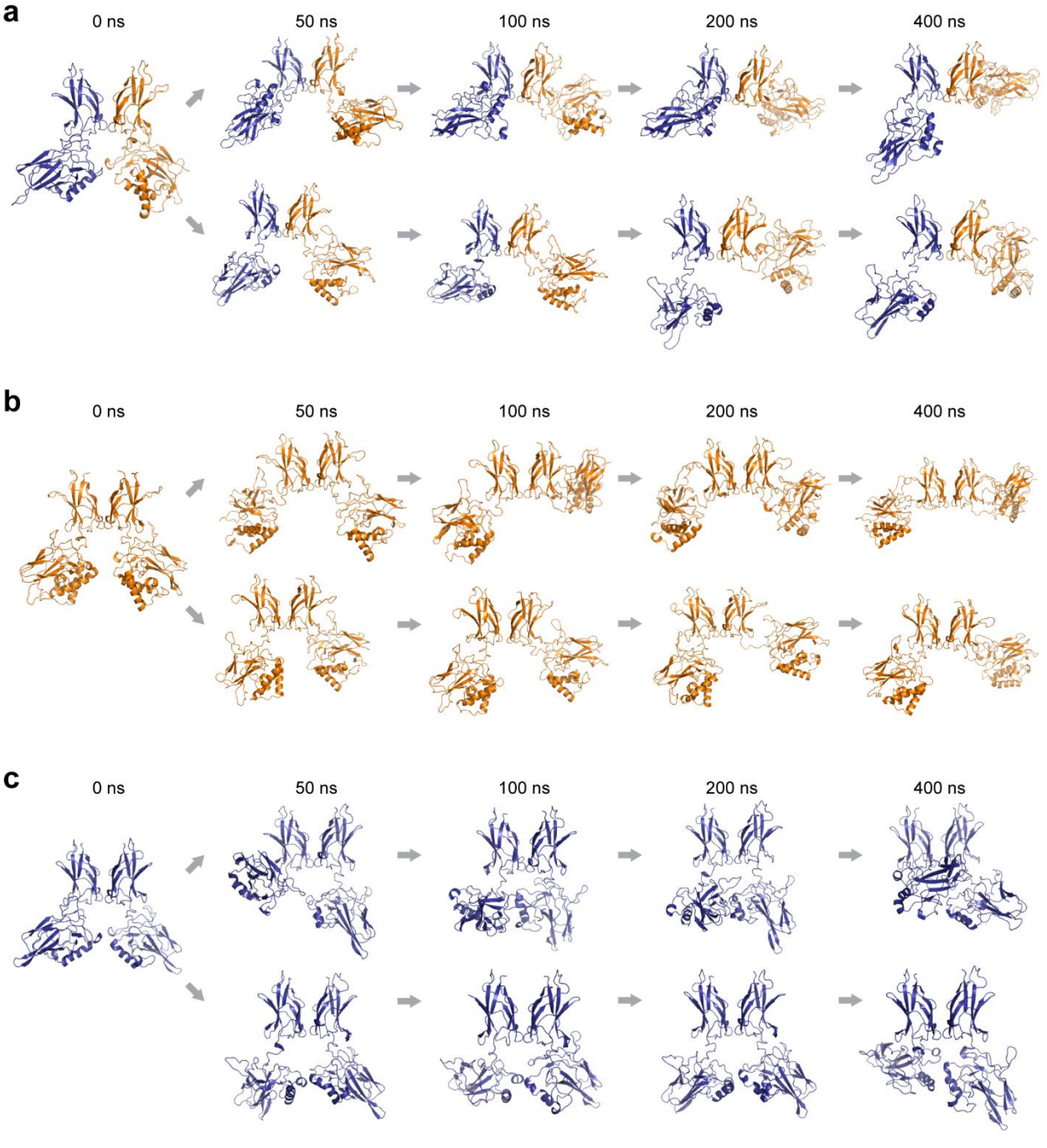


Figure 2.2 Representative structures at different time points in the simulations. Two independent simulations for each of the three dimers are consistent in the general direction of the opening/closing motion but visited different sets of conformations in their routes. **(a)** RelA–p50, **(b)** the p50 homodimer, and **(c)** the RelA homodimer. Orange: p50; Blue: RelA.

2.3.2 Relative Rotation of the NTDs

The relative motion of the NTDs was not simply hinge bending between an open and a closed state, but also involved relative rotations of the domains, as shown in representative conformations along trajectories (**Figure 2.2**). To quantify the relative rotation motions, two vectors, $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$, were each defined with the direction from the NTD center of mass to a specified residue in the NTD helix (**Figure 2.3**). The residues were chosen to be Q128 in RelA and F148 in p50. In a dimer, the angle between $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ would indicate the relative orientation of the two NTDs. If the relative domain motion is pure hinge bending, the angle would remain the same in the trajectory. But if there is a rotation component in the relative motion, the angle between $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ would change over time. In the DNA-bound structure, an obtuse angle of 125° between the NTDs is expected for RelA–p50 and the p50 homodimer, and an angle of 113° for the RelA homodimer. An acute angle indicates one of the NTDs is facing the opposite as in the crystal structure.

Calculation of the angle for the total six simulations showed relative rotation between the NTDs. The domain motions caused fluctuations in the angle with amplitudes over 50° for all trajectories. As was the case for the NTD-NTD distance, the angle also showed that the two simulations for a dimer sampled different sets of conformations. For RelA–p50, both simulations frequently visited conformations with angles from ~60° to over 150°, but one of the simulations reached an angle of 180° while the other did not. For the p50 homodimer, the two simulations took different routes in the beginning—one through a relative rotation that increased the angle between $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ and the other through one that decreased the angle. Both simulations ended up with fluctuations in the angle between ~50° and ~100°. The difference in conformational sampling was

more significant in the case of the RelA homodimer. Both simulations of the RelA homodimer started with the angle decreasing, with the angle in one of the simulations increased to ~150° while the angle in the other continued to decrease and reached ~40°. In the RelA homodimer, the two simulations underwent relative domain rotations in opposite directions as shown in the changes of the angle between v_1 and v_2, leading to two different closed conformations with distinct NTD-NTD interfaces (**Figure 2.2 c**).

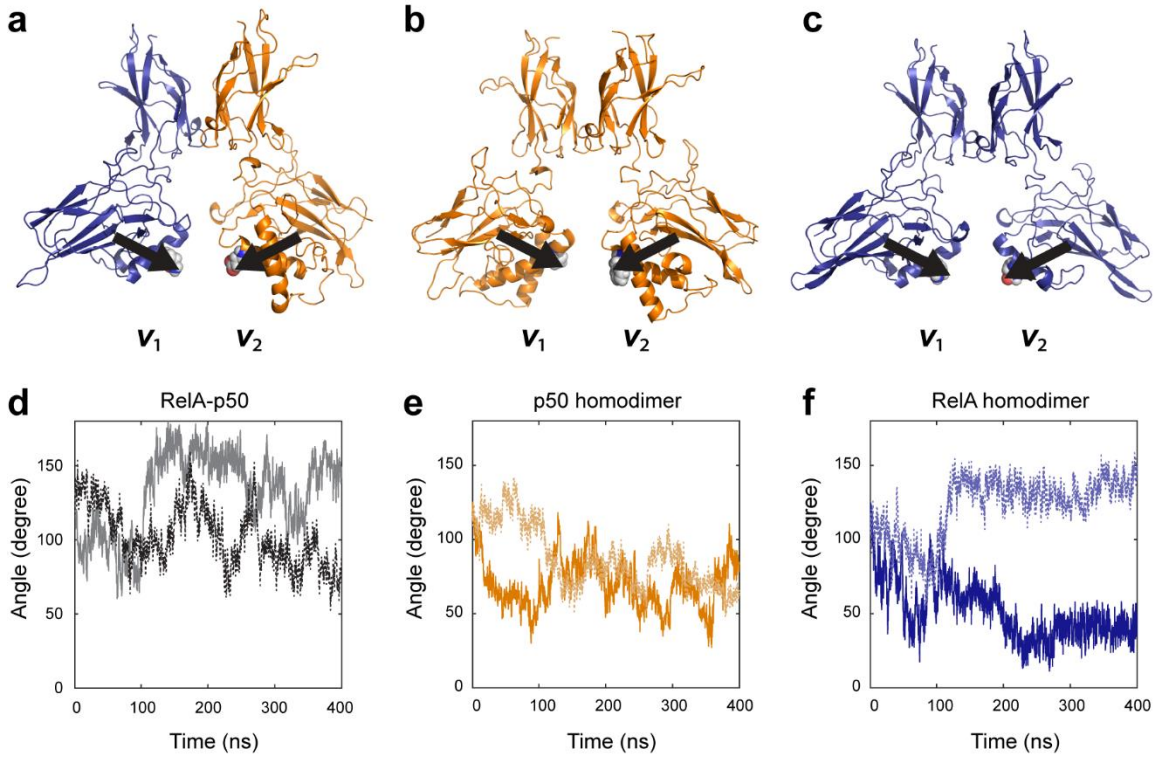


Figure 2.3 Relative rotation motions between the NTDs. Vectors, v_1 and v_2, defined with a direction from the NTD center of mass to a residue in the NTD helix (Q128 in RelA or F148 in p50, sphere) in **(a)** RelA–p50, **(b)** the p50 homodimer, and **(c)** the RelA homodimer. Changes in the angle between v_1 and v_2 indicated relative rotation between the two NTDs in **(d)** RelA–p50, **(e)** the p50 homodimer, and **(f)** the RelA homodimer. Two independent simulations for each of the three dimers sampled different sets of conformations. Solid line: simulation 1; Dotted line: simulation 2.

2.3.3 A Protein Interface Observed in the RelA Homodimer NTDs by Both MD and HDX-MS.

I examined the agreement between representative conformations of the major sub-states of RelA–p50 and the RelA homodimer observed by MD and the hydrogen-deuterium exchange patterns obtained from the hydrogen-deuterium exchange mass spectrometry (HDX-MS) experiments. HDX-MS on RelA in the context of RelA–p50 and the RelA homodimer showed that the entire NTD of RelA was overall more protected in the homodimer than in the heterodimer. A helix (residues 120-136) and the following loop (residues 137-156) exchanged 50% more in the RelA–p50 heterodimer than in the homodimer. This region was completely exposed in the RelA–p50 heterodimer simulations but was buried at the interface in both MD simulations of the RelA homodimer (**Figure 2.4**).

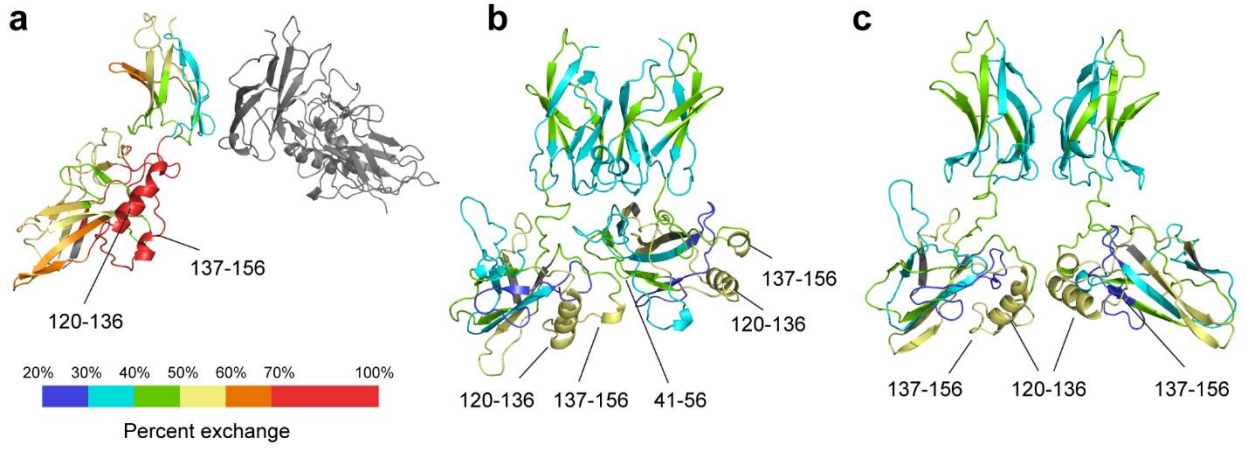


Figure 2.4 Representative conformations of RelA–p50 and the RelA homodimer observed in MD. The structures are colored according to the HDX-MS results on a rainbow scale representing percent uptake with red being the maximum and blue being the minimum. p50 is shown in gray. **(a)** RelA–p50 prefers an open conformation in which the only helix in the NTD (residues 120–136) and the following loop (residue 137–156), are exposed and show high deuterium exchange (red regions). **(b)** One of the two closed conformations from different simulations of the RelA homodimer. The loop containing residues 146–156 in one NTD interacts with the loop containing residues 41–56 in the other NTD. **(c)** The other closed conformation of the RelA homodimer, in which the 120-136 helices from the two NTDs interact with each other.

To understand the interactions across the NTD interface of RelA homodimer, I analyzed the salt bridges formed during the MD simulations (**Figure 2.5**). Five stable inter-NTD salt bridges (all listed with the residue from subunit 1 given first and the residue from subunit 2 given second); R50-E222, D53-R187, D53-K123, K56-D151, and R41-E147; were identified in simulation 1 and three; D151-K123, E127-R124, and R124-E146; in simulation 2. These salt bridges were formed in the later stage of the trajectories as the two NTDs came together (**Figure 2.6**). All these salt bridges at the NTD interface last for at least 100 ns with small fluctuations.

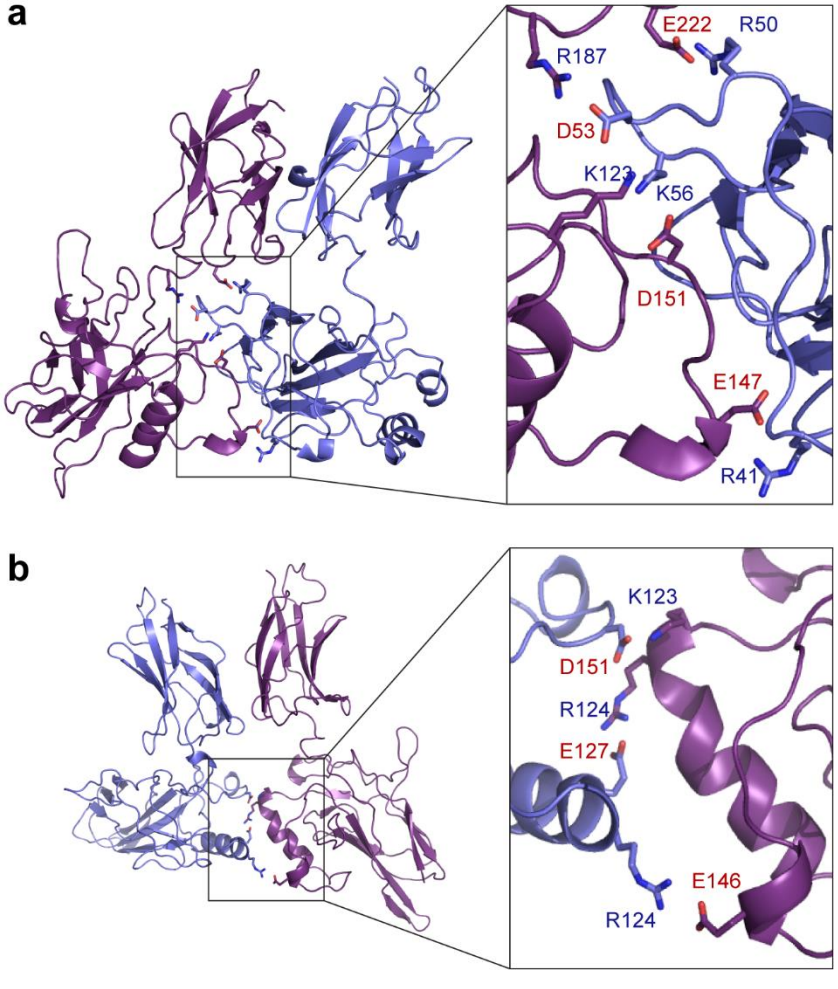


Figure 2.5 Salt bridges across the interface between the two NTDs in the RelA homodimer. The two RelA subunits are shown in blue and purple. Oxygen atoms of acidic side chains are shown in red and nitrogen atoms of basic side chains are shown in blue. **(a)** Five stable salt bridges (E222-R50, D53-R187, D53-K123, D151-K56, andE147-R41) were found at the interface in simulation 1. **(b)** Three stable salt bridges (D151-K123, E127-R124, and E146-R124) were found at the interface in simulation 2.

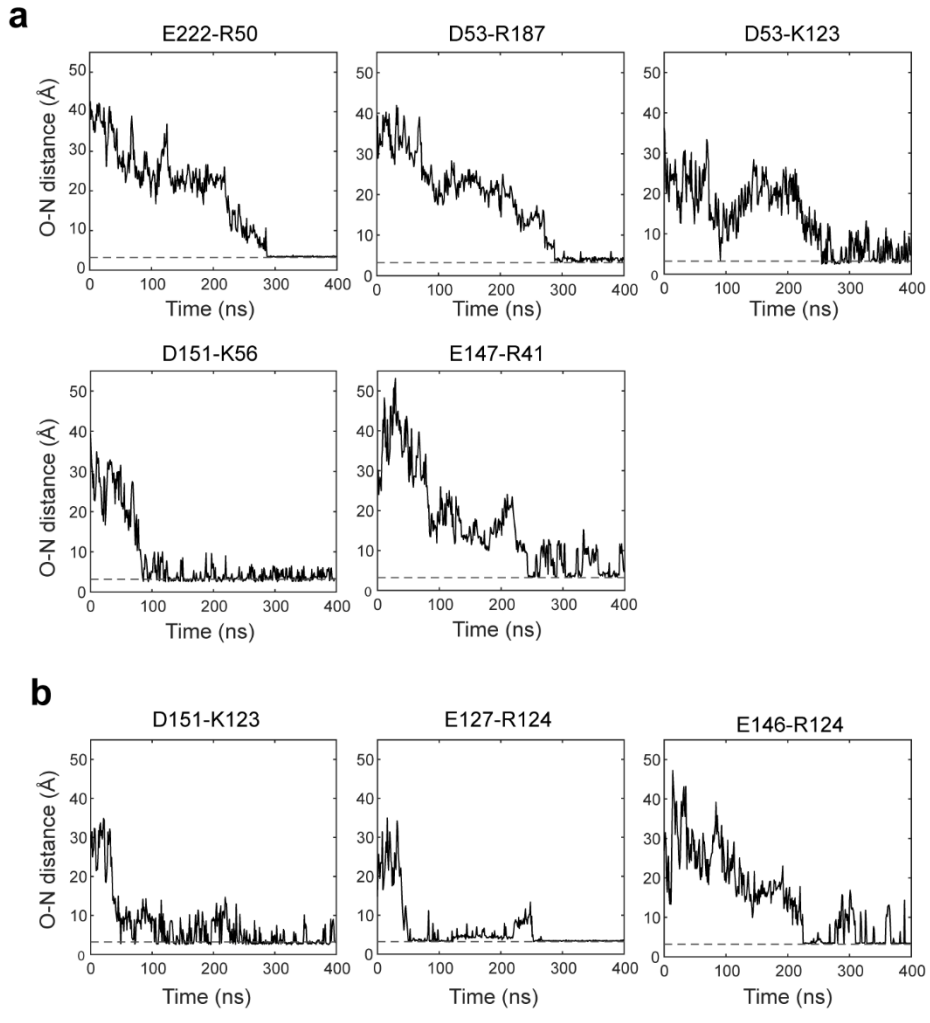


Figure 2.6 Stable salt bridges across the interface between the two NTDs in the RelA homodimer. All salt bridges lasted for at least 100 ns with small fluctuations in the O-N distance. **(a)** Simulation 1. **(b)** Simulation 2. Dash line: 3.2 Å threshold of O-N distance for a salt bridge.

2.3.4 Kinetics of DNA-Binding

The new interacting interface in the NTDs of the RelA homodimer suggested by both MD and HDX-MS could impede DNA binding. Binding kinetics of RelA–p50 and the RelA homodimer by stopped-flow fluorescence showed that the DNA association rate constant was 2.7 ± 0.06 x 10^8 M^{-1} s^{-1} for RelA–p50 and 0.75 ± 0.02 x 10^8 M^{-1} s^{-1} for RelA homodimer[18]. The

Association of DNA to the RelA homodimer was 3.6 times slower than the association to the RelA–p50, likely due to the NTD-NTD interface. To further examine whether the p50 homodimer, which also adopts an open conformational ensemble with an exposed binding cavity as in RelA–p50, also binds to DNA faster than the RelA homodimer does, I performed stopped-flow experiments on the p50 homodimer. DNA-binding kinetics was previously measured by monitoring the fluorescence intensity increase of a pyrene-labeled DNA hairpin corresponding to the IFN κB site upon rapid mixing with the NF-κB dimer[17]. As reported, fluorescence increased upon NF-κB binding to labeled DNA with bi-exponential kinetics (**Figure 2.7 a**), of which the second phase was slow with a weak dependence on protein concentration[17-18].

The concentration-dependent fast phase (k_1) was plotted against the p50 homodimer concentration. Linear fitting revealed the association rate constant for the p50 homodimer binding to DNA to be 2.5 ± 0.8 x 10^8 M^{-1} s^{-1} (**Figure 2.8 b**), which is similar to the rate constant for RelA–p50 and 3.3 times faster than that for the RelA homodimer. Together, the DNA association rates of RelA–p50, the p50 homodimer, and the RelA homodimer are consistent with their degrees of exposure of the DNA-binding cavity.

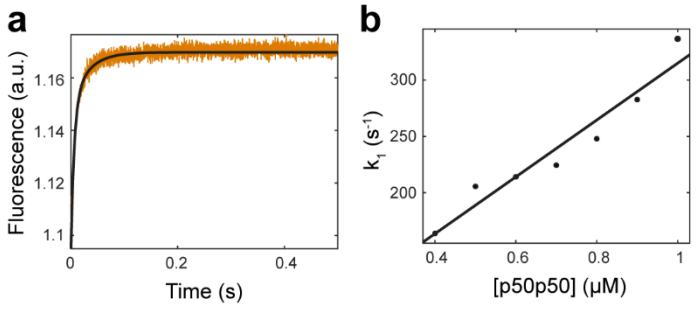


Figure 2.7 Stopped-flow fluorescence kinetics of the p50 homodimer binding to labeled DNA. **(a)** Representative kinetic traces (orange) of association of the p50 with DNA fitted with a biexponential function (black). Final concentrations of DNA and the p50 homodimer are 0.1 and 0.4 μM, respectively. **(b)** Concentration dependence of the p50 homodimer association with DNA. Slope = 2.5 ± 0.8 x 10^8 M^{-1} s^{-1}.

2.3.5 Electrostatic Potential Difference Drives Domain Rotations.

I next investigated whether electrostatic potential differences may be the underlying cause of the opening of the RelA–p50 and closing of the RelA homodimer. Adaptive Poisson-Boltzmann Solver (APBS) electrostatic evaluation of the initial structures after the DNA was computationally removed revealed a much more positive DNA-binding cavity in the RelA–p50 heterodimer and the p50 homodimer as compared to the RelA homodimer (**Figure 2.8**), which could lead to repulsion between the two NTDs. Computation of electrostatic field lines further suggested attractive electrostatic interactions within NF-κB[18]. An attractive interaction is observed between the positively charged region in the p50 NTD and a negatively charged region in its dimerization domain, which may explain why the p50 subunit shows a much larger displacement away from the DNA-bound conformation as compared to the RelA subunit and why the p50-containing dimers prefer an open conformational ensemble. In the RelA homodimer, there is an attractive interaction between parts of its NTDs, which may lead to the approaching of the two NTDs.

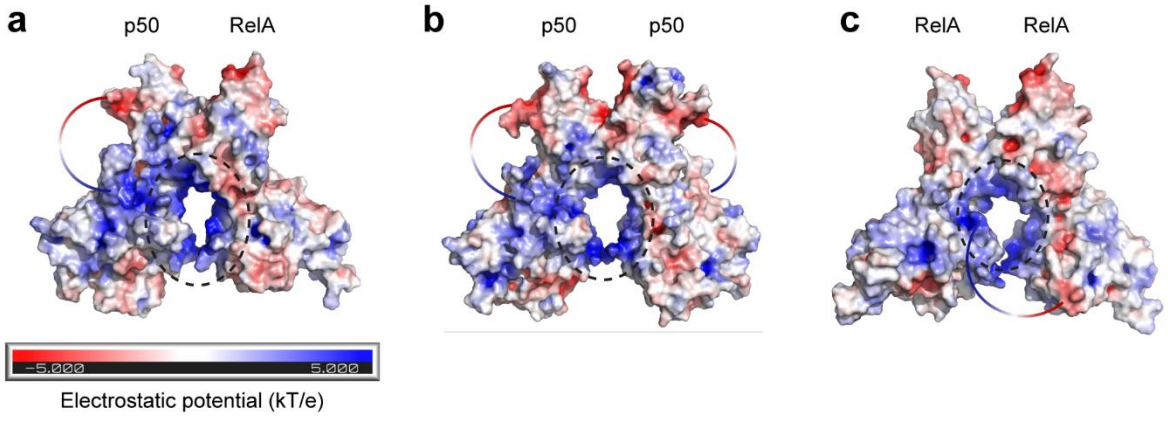


Figure 2.8 Electrostatic potential of RelA–p50, the RelA homodimer, and the p50 homodimer calculated with APBS. The electrostatic potential was mapped onto the 3D structures of the dimers after computational removal of the DNA and before they underwent any conformational changes in MD simulations. Positive regions are shown in blue and negative regions are shown in red, in a color scale ranging from −5 to +5 kT/e. **(a)** In RelA–p50, the DNA-binding cavity (dashed circle) is highly negatively charged. An electrostatic attraction (solid line) between the NTD and the dimerization domain of p50 is observed. **(b)** In the p50 homodimer, the DNA-binding cavity (dashed circle) is also highly negatively charged. An electrostatic attraction (solid line) between the NTD and the dimerization domain of p50 is observed as well. **(c)** In the RelA homodimer, the

DNA-binding cavity (dashed circle) appears less negatively charged as compared to the other two dimers. An electrostatic attraction (solid line) is observed between the two NTDs.

2.3.6 Relative Domain Motions of the QQQQ RelA–p50 Mutant

I next investigated the residues that contribute to the electrostatic attraction in p50 NTD and dimerization domain. Residues 284 to 287 are 4 consecutive glutamic acid residues at the tip of a loop in the dimerization domain of p50, while resides 74, 76, and 77 are three lysine residues in a loop in the p50 NTD. APBS electrostatic analysis suggested that these two oppositely charged regions could attract each other and drive the large-scale motion of the NTDs. Examination of the end conformations in the p50-containing simulations showed that these two regions make contacts with each other in RelA–p50 and in at least one of the subunits in the p50 homodimer (**Figure 9**).

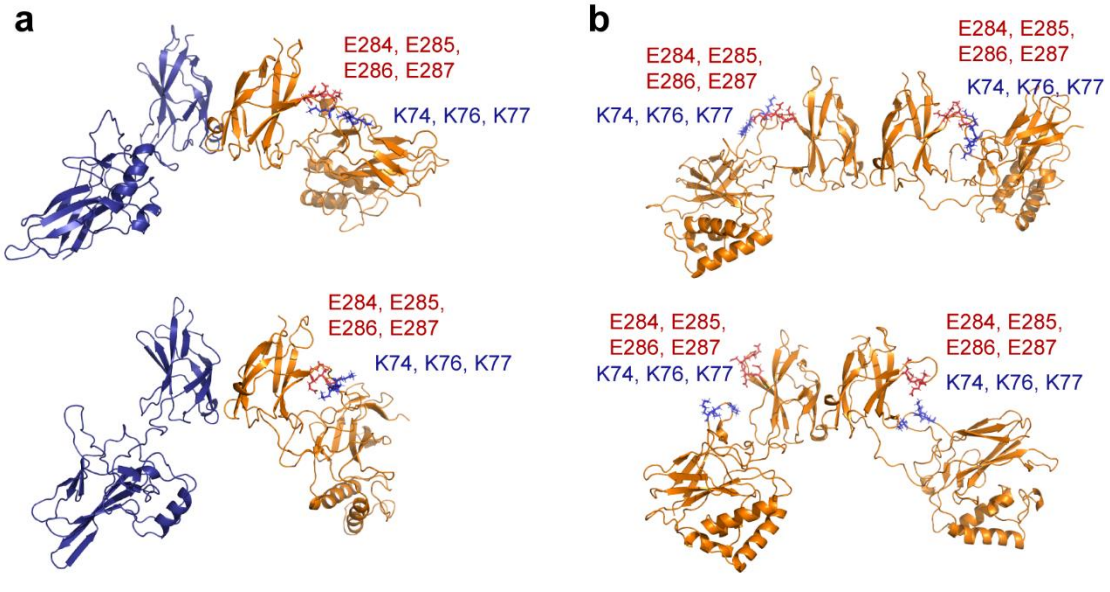


Figure 2.9 Residues that contribute to the electrostatic attraction in p50. Residues E284, E285, E286, and E287 (red sticks) reside at the tip of a loop in the dimerization domain of p50 form a negatively charged region that attracts positively charged residues K74, K76, and K77 (blue sticks) in the NTD. The interaction between the two oppositely charged regions leads to an open conformation in **(a)** RelA–p50 and **(b)** the p50 homodimer, where at least in one of the p50 subunits the two regions make contacts with each other. Top: simulation 1 at 400 ns; Bottom: simulation 2 at 400 ns.

To investigate whether the mutation of EEEE in p50 into four non-charged glutamine residues QQQQ would inhibit the opening motion of the NTDs, I performed a 150 ns all-atom MD simulation for the QQQQ RelA–p50 mutant. The distance between the centers of mass of the NTDs increases overtime as in wild type RelA–p50 and wild type p50 homodimer (**Figure 2.10 a**), suggesting that the electrostatic attraction between EEEE in the dimerization domain and three lysine residues in the NTD is not the only driving force of the opening motion, which agrees with the APBS analysis. In the QQQQ RelA–p50 mutant, the p50 NTD did not flip all the way up and make contacts with the dimerization domain as in the wildtype RelA–p50 did in 100 ns, suggesting that without the negatively charged loop, p50 could no longer stabilize the electrostatic interaction between its dimerization domain and NTD (**Figure 2.10 b**).

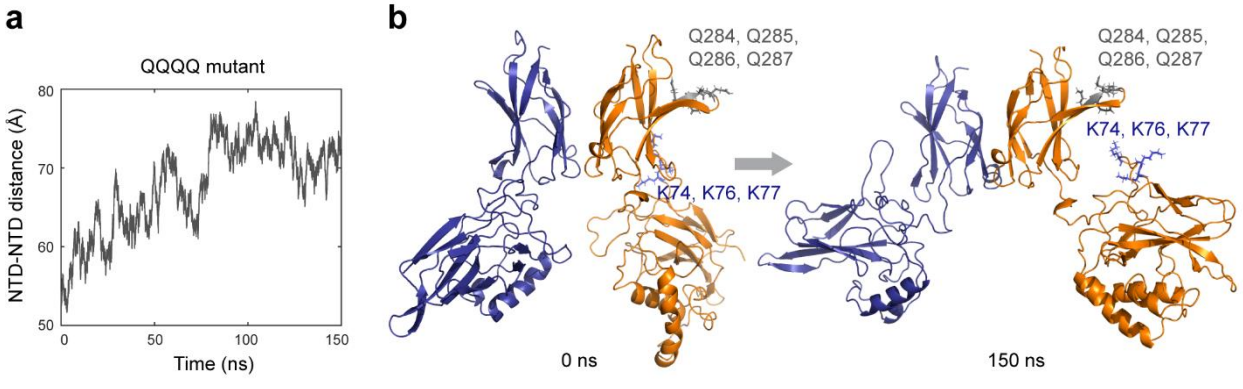


Figure 2.10 Simulations of the QQQQ mutant of RelA–p50. Mutation of negatively charged EEEE into non-charged QQQQ in p50 eliminated the electrostatic interaction between the p50 NTD and the p50 dimerization domain in the MD simulation of RelA–p50. **(a)** Distance between the centers of mass of the two NTDs as a function of time. The NTD-NTD distance increased over time, suggesting that there are other driving forces of the NTD opening motion. **(b)** Starting and end conformation of the QQQQ RelA–p50 mutant in the simulation. The two-NTD moved apart from each other as in the wild type, but the p50 NTD did not flip up to interact with the dimerization domain due to the mutation of negatively charged residues into non-charged ones. Blue: RelA; Orange: p50.

2.4 Discussion

All-atom MD simulations revealed the distinct relative domain motions in three NF-κB dimers: the RelA–p50 heterodimer, the RelA homodimer, and the p50 homodimer. Although their DNA-bound conformations are highly similar, the NTDs of p50-containing dimers rapidly moved farther apart from each other whereas the NTDs of RelA homodimer moved closer together. These conformational relaxation processes via rotation of the NTDs around the linkers lead to an open conformational ensemble for RelA–p50 and the p50 homodimer and a closed conformational ensemble for the RelA homodimer.

In all of the simulations, the NTDs reoriented relative to the dimerization domain in both RelA and p50, which preferred somewhat different relative domain orientations. The p50 NTD flipped up and appeared to form contacts between its positively charged loop (residues 74-77, KNKK) and a negatively charged loop in the dimerization domain (residues 284-287, EEEE). This reorientation of the p50 NTD opened the DNA-binding cavity exposing it to higher amide exchange. In all simulations, the RelA NTDs rotated around the DNA-binding position but did not move much away from their original position as compared to p50. In the RelA homodimer simulations, the rotation of the NTDs appeared to allow the formation of inter-domain contacts.

An interface formed between the NTDs in the simulations of the homodimer. In one simulation, the interface formed between loops and in the other simulation it was formed between helices. Both the loops and the helices were the regions of RelA that showed the most dramatic difference in exchange suggesting that the interfaces formed in the MD simulations may both be contributing to the lower observed exchange in the homodimer. These different interfaces suggested that instead of forming a unique strong inter-domain interaction, the NTDs could interact from different orientations in the homodimer. Both interfaces were stabilized by salt bridges, which could break and allow the NTD to rotate before they interact again and form a new

interface. Such breaking and reforming of salt bridges are not expected to be observed within the 400 ns NTD simulations. Taken together, the HDX-MS and MD simulations are most consistent with the formation of an interface between the RelA NTDs in the homodimer.

The RelA homodimer was previously reported to bind DNA more weakly than RelA–p50 and the p50 homodimer[3], and this is true for a number of different κB promoter sites including those for MIP2, RANTES, IgκB, IFN, and urokinase[2]. For most of the promoters, the RelA homodimer bound 3-5 fold more weakly than the heterodimer. The weaker binding observed in the RelA homodimer could be due to either a lower association rate, or a faster dissociation rate, or both. If indeed, the free RelA homodimer occluded the DNA binding cavity due to the formation of an interface between the NTDs, the association rate should be decreased. With stopped-flow fluorescence, it was shown that for the IFN promoter sequence, all of the difference between RelA–p50 and the RelA homodimer in binding affinity could be accounted for by a decreased association rate. DNA-binding rate measured for the p50 homodimer using stopped-flow fluorescence is as fast as that for RelA–p50, suggesting that a faster DNA-binding rate is the result of a more exposed DNA-binding cavity.

Taken together, the results strongly support the formation of an NTD interface in the homodimer that occludes DNA-binding. We considered testing the propensity for NTD domain reorientation and homodimer NTD interface formation using site-directed mutagenesis, but the electrostatic potential calculations and the simulation of QQQQ RelA–p50 mutant suggest that all three parts of the proteins, the dimerization domain, the linker, and the NTD contribute to the electrostatic interaction which apparently drives the RelA homodimers toward each other and the p50 subunit away. Disruption of one group of charged residues among all of them is not likely to have a very large effect. In the homodimer, two different NTD-NTD interfaces were observed in the simulations suggesting that while an interface can form, it is not particularly stable or unique,

and therefore disruption of one interaction would not likely disrupt this "fuzzy" interaction completely. The results reveal distinctive domain motions in the NF-κB dimers and provide a mechanistic explanation for why the RelA homodimer binds DNA weaker, an observation that was not understood from the analysis of static structures of the DNA-bound dimers.

Acknowledgments

Chapter 2, in part, appears in "RelA-Containing NFκB Dimers Have Strikingly Different DNA-Binding Cavities in the Absence of DNA" by Narang, D.*; Chen, W.*; Ricci C. G.; Komives, E. A., published in *Journal of Molecular Biology.* **2018,** *430,* 1510-1520. The dissertation author is one of the co-first authors denoted by asterisks.

References

1. Huang, D. B.; Huxford, T.; Chen, Y. Q.; Ghosh, G., The role of DNA in the mechanism of NF kappa B dimer formation: crystal structures of the dimerization domains of the p50 and p65 subunits. *Structure* **1997,** *5* (11), 1427-1436.

2. Bergqvist, S.; Alverdi, V.; Mengel, B.; Hoffmann, A.; Ghosh, G.; Komives, E. A., Kinetic enhancement of NF-kappaB•DNA dissociation by IkappaBalpha. *Proc. Nat. Acad. Sci. U.S.A.* **2009,** *106* (46), 19328-19333.

3. Phelps, C. B.; Sengchanthalangsy, L. L.; Malek, S.; Ghosh, G., Mechanism of kappa B DNA binding by Rel/NF-kappa B dimers. *J Biol Chem* **2000,** *275* (32), 24392-24399.

4. Huang, D. B.; Vu, D.; Cassiday, L. A.; Zimmerman, J. M.; Maher, L. J., 3rd; Ghosh, G., Crystal structure of NF-kappaB (p50)2 complexed to a high-affinity RNA aptamer. *Proc Natl Acad Sci U S A* **2003,** *100* (16), 9268-73.

5. Chen, Y. Q.; Ghosh, S.; Ghosh, G., A novel DNA recognition mode by the NF-kappa B p65 homodimer. *Nat Struct Biol* **1998,** *5* (1), 67-73.

6. Ghosh, G.; Vanduyne, G.; Ghosh, S.; Sigler, P. B., Structure of Nf-Kappa-B P50 Homodimer Bound to a Kappa-B Site. *Nature* **1995,** *373* (6512), 303-310.

7. *The PyMOL Molecular Graphics System*, 2.3.2; Schrödinger, LLC.

8. Martinez, L.; Andreani, R.; Martinez, J. M., Convergent algorithms for protein structural alignment. *BMC Bioinformatics* **2007,** *8*, 306.

9. Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A., H++: a server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic Acids Res* **2005,** *33* (Web Server issue), W368-71.

10. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D., Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **1983,** *79*, 926-935.

11. Case, D. A.; Babin, V.; Berryman, J. T.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Cheatham, I., T.E.; Darden, T. A.; Duke, R. E.; Gohlke, H.; Goetz, A. W.; Gusarov, S.; Homeyer, N.; Janowski, P.; Kaus, J.; Kolossváry, I.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Luchko, T.; Luo, R.; Madej, B.; Merz, K. M.; Paesani, F.; Roe, D. R.; Roitberg, A.; Sagui, C.; Salomon-Ferrer, R.; Seabra, G.; Simmerling, C. L.; Smith, W.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Kollman, P. A. *AMBER 14*, University of California, San Francisco. , 2014.

12. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C., ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput* **2015,** *11* (8), 3696-713.

13. Darden, T.; York, D.; Pedersen, L., Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. *J Chem Phys* **1993,** *98* (12), 10089-10092.

14. Humphrey, W.; Dalke, A.; Schulten, K., VMD: visual molecular dynamics. *J Mol Graph* **1996,** *14* (1), 33-8, 27-8.

15. Roe, D. R.; Cheatham, T. E., 3rd, PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J Chem Theory Comput* **2013,** *9* (7), 3084-95.

16. Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A., Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* **2001,** *98* (18), 10037-41.

17. Alverdi, V.; Hetrick, B.; Joseph, S.; Komives, E. A., Direct observation of a transient ternary complex during IkappaBalpha-mediated dissociation of NF-kappaB from DNA. *Proc Natl Acad Sci U S A* **2014,** *111* (1), 225-30.

18. Narang, D.; Chen, W.; Ricci, C. G.; Komives, E. A., RelA-Containing NFkappaB Dimers Have Strikingly Different DNA-Binding Cavities in the Absence of DNA. *J Mol Biol* **2018,** *430* (10), 1510-1520.

Chapter 3 Direct Observation of NF-κB Conformational Dynamics with Single-Molecule FRET: Instrumentation and Sample Preparation

In Chapter 2, molecular dynamics (MD) simulations revealed large-scale relative domain motions in NF-κB that can play an important role in its association with DNA[1] and the dissociation from DNA accelerated by IκBα[2]—the two regulation points that specify the time window of gene expression and need to be tightly controlled. Quantitative characterization of the relative domain motions and their responses to DNA and IκBα is fundamental to a mechanistic understanding of gene regulation by protein conformational dynamics.

Given the heterogeneous nature and the nanometer length scale of the domain motions, single-molecule FRET (smFRET) is the best and the only experimental tool to make direct observation and quantitative characterization. This chapter describes the principle of FRET, instrumentation of an in-house total-internal reflection fluorescence (TIRF) microscope, experimental designs to probe NF-κB motions, preparation of site-specific labeled NF-κB, the photostability of dyes, and surface passivation allowing direct observation of the predicted relative domain motions of the NF-κB RelA–p50 heterodimer.

3.1 Förster Resonance Energy Transfer

Förster Resonance Energy Transfer (FRET) or Fluorescence Resonance Energy Transfer (FRET) is a mechanism of excitation energy transfer between two fluorophores[3]. FRET has been widely applied in biological sciences as the "molecular ruler" to measure the distance between two

fluorophores for probing intermolecular interactions between biomolecules or intramolecular conformational dynamics at the nanoscale.

There are two requirements for FRET. First, the emission spectrum of the donor must overlap with the excitation spectrum of the acceptor to allow excitation energy transfer (**Figure 3.1 a**); Secondly—the key to its biological application—the two fluorophores need to be in proximity to each other (**Figure 3.1 b**). The distance between the donor and the acceptor governs the extent of energy transfer with the equation

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

where E is the energy transfer efficiency, a number from 0 to 1, R is the distance between fluorophores, and R_0, also known as the Förster distance, is the distance at which E = 0.5. The shorter the distance is, the higher FRET efficiency is. FRET is most sensitive to distance in the region of 30 to 70 Å, which makes it a great tool to measure biomolecular interactions and conformational dynamics.

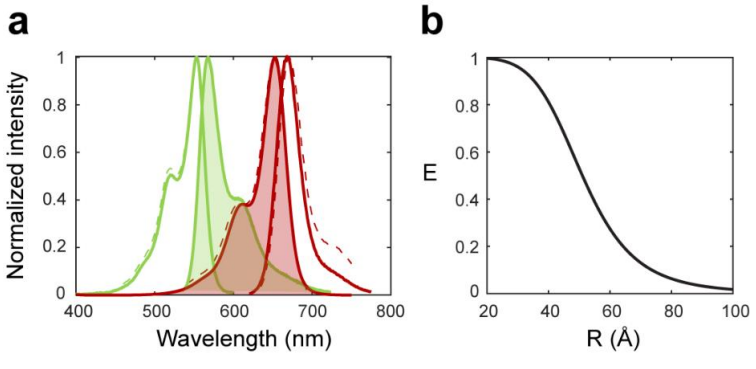


Figure 3.1 Two requirements for FRET: spectral overlap and physical proximity of the donor and the acceptor. **(a)** The excitation and emission spectra of the donor (green) and acceptor (red). The spectral overlap between the donor emission spectrum (green, filled) and the acceptor excitation spectrum (red, filled) is required for FRET. Dotted lines: conventional cyanine dye pair for FRET, Cy3 and Cy5. Solid lines: Alexa Fluor 555 and Alexa Fluor 647, FRET dye pair used in this study with better photostability and almost identical spectra as those of Cy3 and Cy5. **(b)** FRET efficiency (E) depends on the distance between the two fluorophores (R) and is most sensitive in the range from 30 to 70 Å.

The first measurement of single-molecule FRET was achieved in 1996[4]. This breakthrough allowed direct visualization of conformational substates and dynamic processes like interconversions between conformational states and macromolecular interactions. Currently, there are two approaches to smFRET depending on whether the molecules are freely diffusing or immobilized. In the diffusion-based approach, single molecules freely diffuse through a small observation volume and generate bursts of photons. Analysis of FRET efficiency fluctuations during single-molecule bursts, which last about 1 ms, can reveal the dynamics on the range from 10 μs to 0.1 s. On the other hand, with immobilized molecules, the observation time window can be as long as minutes, limited only by the photobleaching time of fluorophores, and thus allows characterization of slow dynamics on the time scale of 0.1 – 10 s.

To probe the relative domain motions of NF-κB, we chose to implement the immobilization-based approach with TIRF microscopy. In this way, single-molecule time traces could be obtained from movies containing hundreds of molecules in parallel. Small molecule dye pairs were covalently attached on the two N-terminal domains of the NF-κB dimer and relative domain motions will manifest as fluctuations in FRET efficiencies.

3.2 Total-Internal Reflection Fluorescence Microscopy

In TIRF microscopy, single-molecule resolution is achieved through exciting only the fluorescent molecules at the interface instead of in the bulk. At the interface between water and quartz, excitation light is totally reflected with only a small amount of light penetrating the interface 100-200 nm deep. This thin layer of light is called the evanescent wave and can excite only the molecules near the surface and effectively eliminate background signals.

We set up a prism-type TIRF microscope (**Figure 3.2**) for the purpose of smFRET. Two colors of lasers were installed: a 532 nm laser (SAPPHIRE 532-300 CW CDRH, Coherent) for exciting donors (green), and a 637 nm laser (OBIS™ 1196625 | 637nm LX 140mW Laser, Coherent) for exciting acceptors (red). The red laser was used to confirm the existence of acceptors in the sample and the green laser was used for exciting donors for FRET. Laser output powers were set to 50 mW. Excitation beams are controlled by shutters and guided by slits and mirrors. A polarizing beam splitter (PBS) was used for lining up the green and the red laser beams and a neutral density filter (ND filter) for attenuating brightness. Mirror 3 and 4 directed the excitation beam to a fused silica prism which contacts the quartz slide. Immersion oil (IMMOIL-F30CC, Olympus) was put in between the prism and the slide to match the refractive index. Total internal reflection occurred at the quartz (slide)-water (sample) interface.

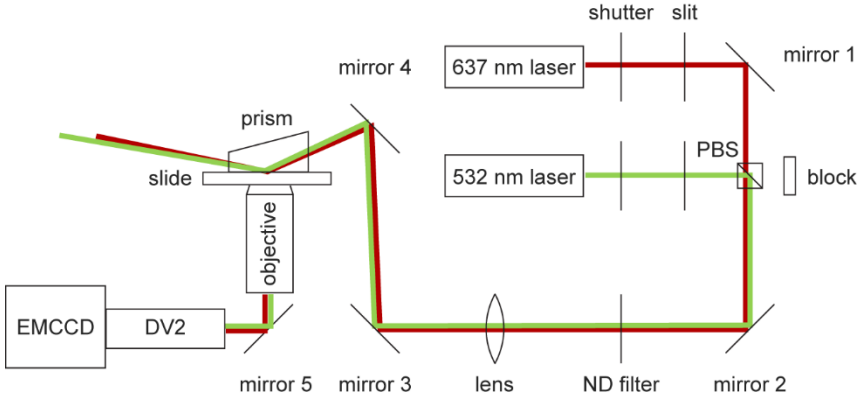


Figure 3.2 Schematic of a prism-type TIRF microscope. 532 nm (green) and 637 nm (red) lasers were set up to excite donor and acceptor respectively. Beam brightness and size could be controlled by the neutral density (ND) filter and the lens respectively. Excitation beams were guided by mirrors and the polarizing beam splitter (PBS) to focus on the prism. The Refractive index of the prism and slide was matched with that of the oil. Total internal reflection occurred at the quartz (slide)-water (sample) interface to excite only the molecules near the surface. FRET signals were collected by the objective, split into donor and acceptor channels with the DV2, and detected by EMCCD.

Fluorescent signals emitted by the dye pairs were collected by the objective (CFI Plan Apochromat VC 60XC WI, Nikon) inversely mounted on a Rapid Automated Modular Microscope System (Applied Scientific Instrumentation), and further split into donor and acceptor channels in a DV2 Two-Channel, Simultaneous-Imaging System (Photometrics). After passing bandpass filters (ET585/65m for donor channel and ET685/56m for acceptor channel, Chroma), donor and acceptor emission was detected by Electron Multiplying Charge-Coupled Device (EMCCD) camera (iXon^{EM}+ EMCCD camera, DU-897E-CS0-#BV, Andor Technology).

With a camera-based approach, about two hundred molecules can be viewed from a 256×512 field of view in a ~0.5 mm^2 excitation area. Movies were recorded with 100 ms time resolution. Donor and acceptor fluorescence intensities for individual molecules were further extracted as time traces using the data acquisition and analysis software obtained from the Ha Lab at Johns Hopkins University http://ha.med.jhmi.edu/resources/#1464200861600-0fad9996-bfd4.

3.3 Experimental Design for NF-κB Relative Domain Motions

To probe the distance between the two N-terminal domains of RelA–p50, the experiments were designed to have NF-κB dimers immobilized on a passivated surface through the 6XHis tag on the C-terminus of RelA interacting with penta-His antibody biotin conjugate (34440, Qiagen). (**Figure 3.3 a**). The surface was passivated with DDS-Tween 20[5] and biotinylated BSA, which binds to NeutrAvidin that interacts with the biotin-conjugated antibody. E128 in RelA and F148 in p50 were chosen as fluorophore labeling sites. The distance between the two labeling positions is 27 Å in the DNA-bound crystal structure, which would lead to a ~ 1.0 FRET efficiency. In MD simulations, NF-κB could adopt an open conformation with 70 Å inter-dye distance, leading to ~0.1 FRET efficiency (**Figure 3.3 b**), and the two N-terminal domains explored a wide range of

conformations, leading to a broad distribution of inter-dye distance and FRET efficiency (**Figure 3.3c**). The two residues for labeling reside in the helices in the N-terminal domains (**Figure 3.3 b**), and therefore the FRET efficiency changes were expected to reflect the rigid-body domain motions rather than local structural fluctuations caused by flexible loops.

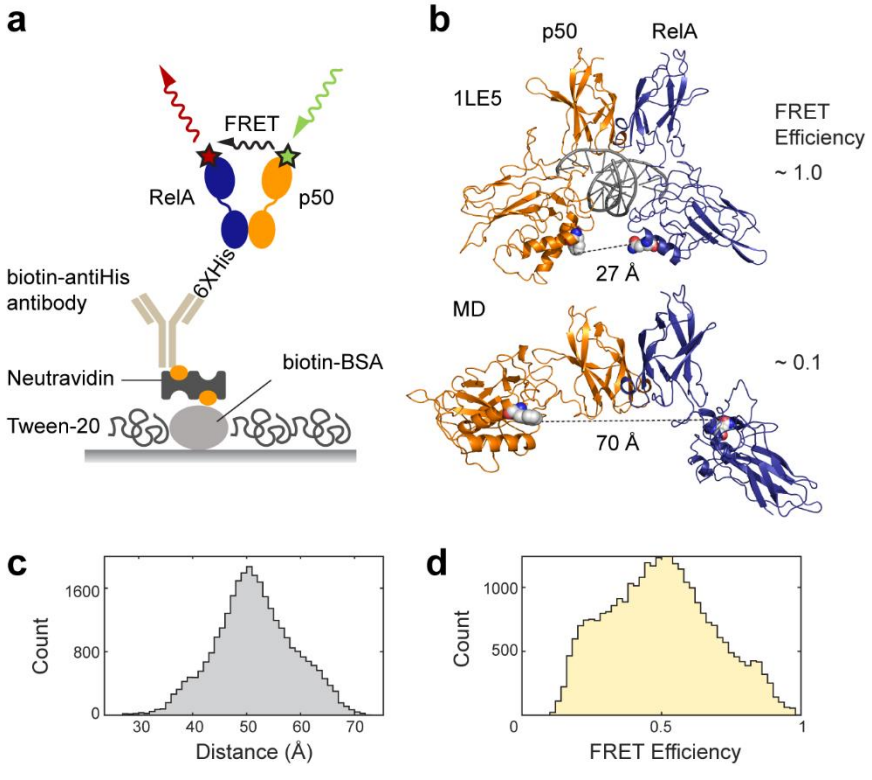


Figure 3.3 smFRET experimental design for probing the relative domain motions of NF-κB. **(a)** NF-κB (RelA–p50) immobilized on passivated surface. A series of interactions through the 6XHis-tag on RelA, biotin-conjugated anti-His antibody, NeutrAvidin, and biotin-conjugated bovine serum albumin (BSA) were designed for immobilizing NF-κB. Tween-20 was be used to reject non-specific binding. The two N-terminal domains were labeled with FRET donor (green star) and acceptor (red star) to probe the distance between them. **(b)** In the NF-κB:DNA crystal structure (PDB: 1LE5), the distance between the two labeling positions, E128 in RelA and F148 in p50 (spheres), would lead to a high FRET efficiency of 1. In MD simulations, free NF-κB could adopt an open conformation, leading to a low FRET efficiency of 0.1. **(c)** MD simulation revealed a broad distribution of the distance between theOE1 atom of Q128 in RelA and the CZ atom of F148 in p50. **(d)** MD predicted FRET efficiency distribution converted from the E128-F148 distance distribution.

Conventionally, the attachment of small molecule dyes to proteins is achieved by utilizing the reactive and relatively rare amino acid—cysteine. The target labeling site in the protein would be mutated into a cysteine and the thiol group of the cysteine would react with the maleimide conjugate on the dye to form covalent bonds. However, a problem with this approach is that in large proteins there are many native cysteines that could be nonspecifically labeled and would need to be mutated prior to labeling creating an unnatural protein that might not behave the same way as the wild type.

In the NF-κB RelA–p50 dimer, there are 8 native cysteines in RelA and 7 in p50 (**Figure 3.4**), most of which appear exposed to solvent in the crystal structure and could lead to problematic nonspecific labeling. To examine if these cysteines were more exposed than the targeted labeling site, RelA and p50 were separately mixed with N-ethylmaleimide (NEM), which carries a maleimide functional group that reacts with cysteines. Here the targeted sites were S131C for RelA and E152C for p50. C38 in RelA was known to be exposed and mutated to a serine. Proteolytic digestion following mass spectrometry were carried out to identify cysteines that were modified by NEM.

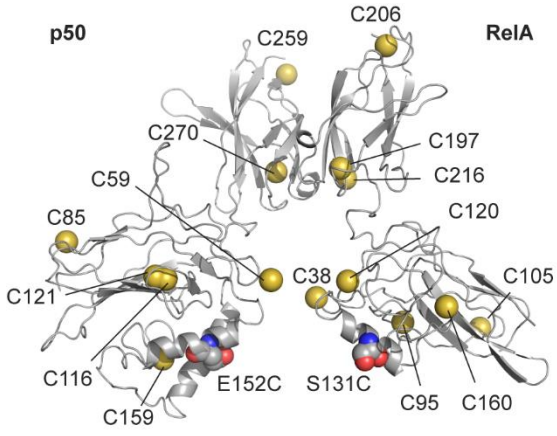


Figure 3.4 Cysteines in the NF-κB RelA–p50 dimer. There are 8 native cysteines in RelA and 7 in p50, most of which in the crystal structure (PDB: 1LE5) appear to be exposed to solvent and could lead to nonspecific labeling for smFRET. In the mass spectrometry experiment to probe thiol reactivity, E152C in p50 and S131C in RelA are the targeted labeling sites. RelA and p50 were separately reacted with N-ethylmaleimide and then digested by proteases. Modifications on cysteines would be identified by mass spectrometry to reveal thiol reactivity of native and target cysteines.

For C38S S131C RelA, pepsin digestion lead to perfect sequence coverage (**Figure 3.5**). Three native cysteines C95, C120, C160 were found not modified by NEM, while C105, C197, C206, and C216 were modified at various degrees from 8.3% (C216, 1 in 12 peptides) to 50% (C95, 1 in 2 peptides). The target site C131 was not modified, indicating native cysteines were more reactive than the target one.

Proteolytic digestion of C38S S131C RelA with the GluC protease also gave perfect sequence coverage and generated more peptides for quantification than with pepsin (**Figure 3.6**). Except for C120 and C160, which were not modified, all native cysteines were modified at different degrees—C95 (1 in 20 peptides, 5%), C105 (5 in 33 peptides, 15%), C197 (3 in 15 peptides, 20%), C206 (3 in 8 peptides, 38%), C216 (4 in 39 peptides, 10%). The target site C131 was not modified. Both pepsin and GluC digestions concluded that most native cysteines were

more reactive toward maleimide than the target site and would lead to undesired nonspecific labeling.

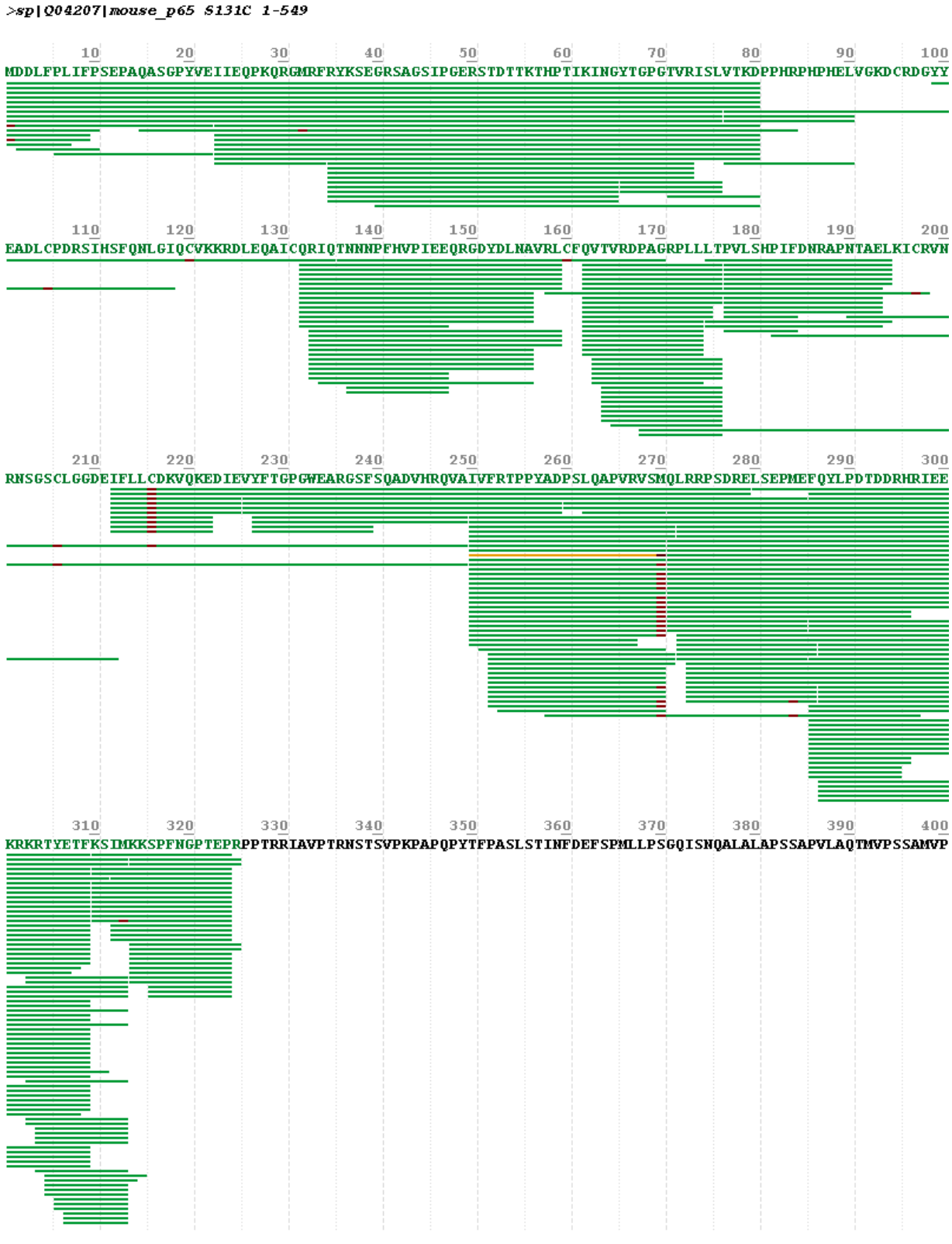


Figure 3.5 Mass spectrometry coverage map of C38S S131C RelA1-325 by pepsin digestion. Native cysteines C105, C197, C206, and C216 were modified by NEM while target site C131 was not. Green: peptides; Red: modified amino acids. Some methionines were oxidized and marked red.

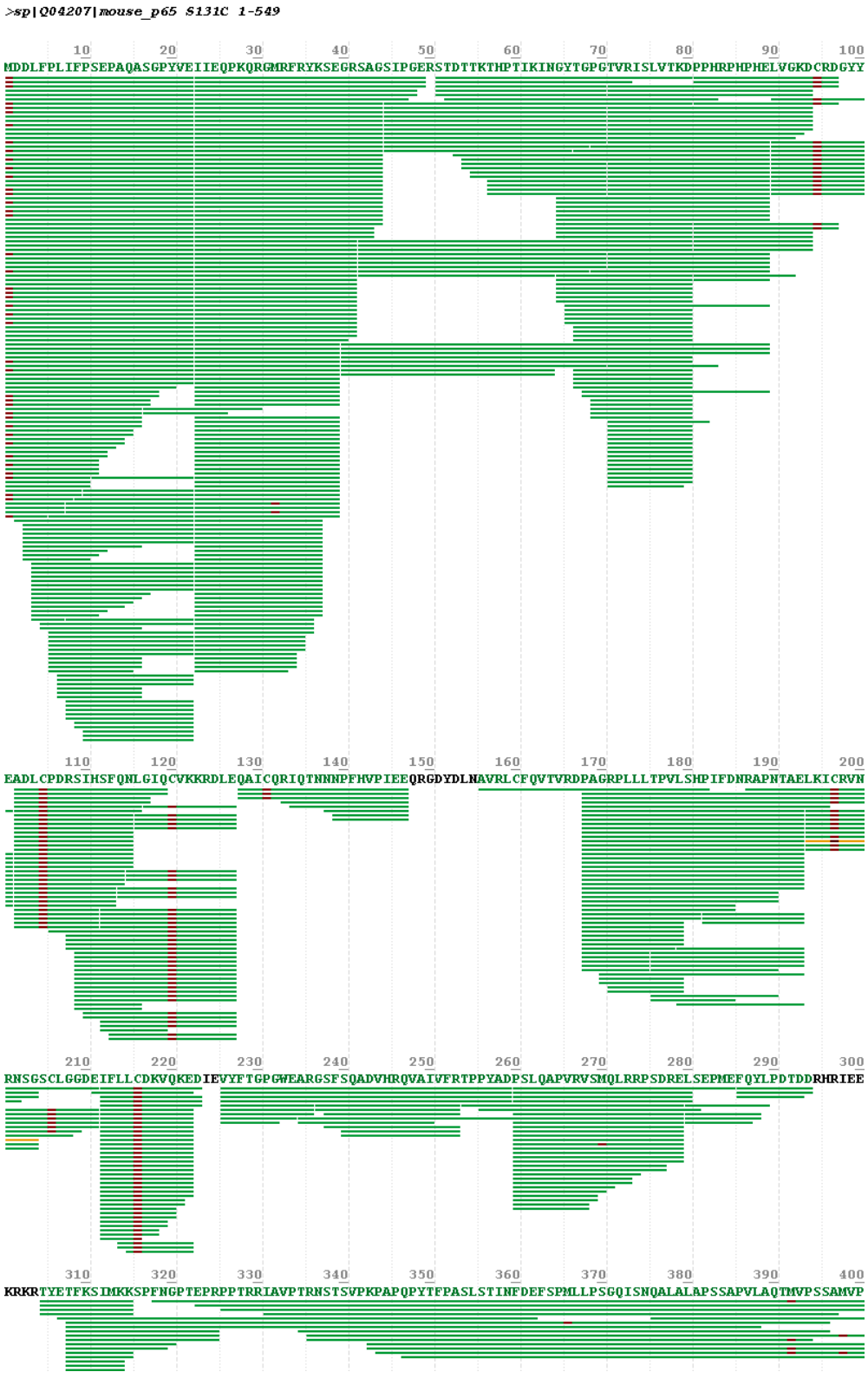


Figure 3.6 Mass spectrometry coverage map of C38S S131C RelA1-325 by GluC digestion. Native cysteines C95, C105, C197, C206, C216 were modified by NEM while target site C131 was not. Green: peptides; Red: modified amino acids. Some methionines were oxidized and marked red.

For E152C $p50_{39-363}$, pepsin digestion did not generate peptides that cover any of the cysteines (**Figure 3.7 a**), while GluC digestion generated peptides that cover C160, C259, and C270, all of which were modified by NEM (**Figure 3.7 b**). The target site C152 was not covered.

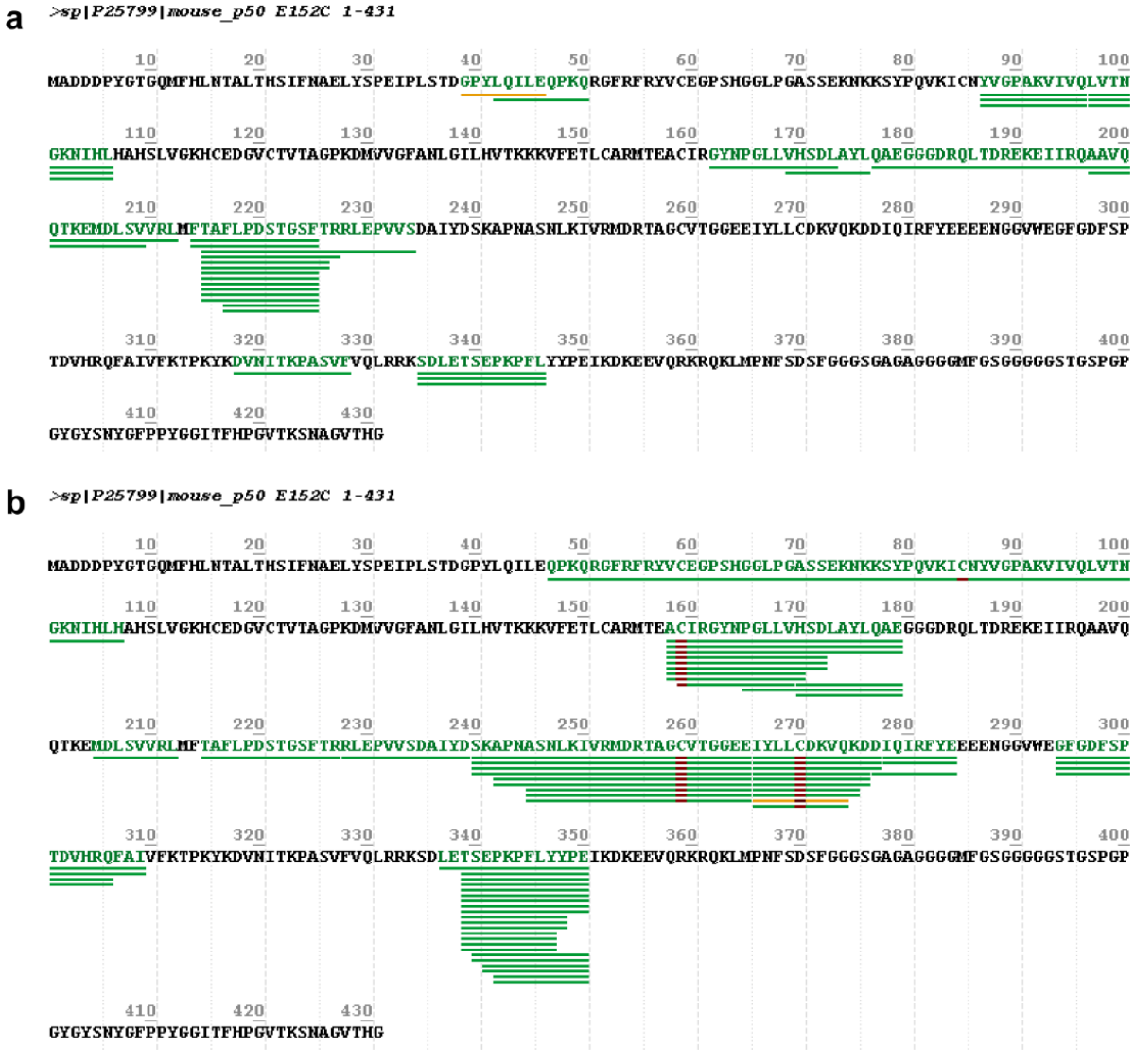


Figure 3.7 Mass spectrometry coverage map of E152C p5039-363. **(a)** Pepsin digestion. Cysteines were not covered. **(b)** GluC digestion. Native cysteines C160, C259, and C270 were modified by NEM. The target site C152 was not covered. Green: peptides; Red: modified amino acids.

NEM labeling and mass spectrometry showed that, in RelA, native cysteines were more reactive toward maleimide than the targeted cysteine was. Although peptides in p50 did not cover the target cysteine, given that it resides in the helix as the RelA target site does, it is very likely to be less reactive than the native cysteines in loops that could be easily labeled. From mass spectrometry, it was clear that for site-specific labeling, the labeling site must be a unique amino acid in the NF-κB subunit to avoid unwanted labeling. The two ways to achieve specific labeling are to mutate all native cysteines to other amino acids or to install an unnatural amino acid in the labeling position.

3.4 Site-Specific Fluorophore Labeling using Unnatural Amino Acids

To achieve site-specific labeling without having to mutate out all native cysteines, I decided to incorporate an unnatural amino acid (UAA) at position 128 in RelA and 148 in p50 using amber suppression.

In amber suppression, the amber codon UAG, the least used stop codon by nature, is repurposed to code for unnatural amino acids. A large number of tRNA/aminoacyl-tRNA synthetases (aaRS) that pair the amber codon and a specific UAA has been engineered to directly incorporate a variety of UAAs in *E. coli* and mammalian cells[6]. With amber suppression, site-specific labeling for smFRET can be achieved by genetically incorporating a UAA carrying a unique functional group that can be used for dye conjugation.

Several UAAs have been applied for smFRET fluorophore labeling (**Table 3.1**). The desired UAA to incorporate in NF-κB would be the one that only requires mild labeling conditions similar to physiological conditions. The unnatural amino acid, p-azidophenylalanine (pAzF) was

chosen because it does not require harsh labeling conditions such as low pH or copper ions. However, pAzF can be easily reduced to a nonreactive form, so reducing agents must be avoided in the protein purification process before fluorophore labeling. I have adapted the expression and purification protocol of NF-κB to avoid reducing agents prior to the labeling reaction (**Appendix I**).

Table 3.1. Unnatural amino acids for smFRET labeling.

Unnatural amino acid	**Reactive group**	**Reactive partner**	**Remark**	**Reference**
p-acetylphenylalanine (AcF)	Ketone	Hydroxylamine	Reacts at pH 4.	[7]
N-propagyllysine (PrK)	Alkyne	Azide	Requires Cu^{2+}.	[8]
p-azidophenylalanine (pAzF)	Azide	Alkyne	Reduced by DTT/βME/TCEP.	[9]
p-azidomethylphenylalanine	Azide	Alkyne	Patent protected, not accessible.	[10]

In protein preparation for pAzF incorporation, the plasmid pEVOL-pAzF (Plasmid #31186, Addgene) encoding the engineered tRNA/aaRS pair for pAzF was co-transformed with the protein expression plasmid into *E. coli*. RelA and p50 were expressed separately. pAzF was added to the media during growth.

To examine whether pAzF incorporation was successful, two growths of p50, with and without the addition of pAzF, were compared side-by-side (**Figure 3.8 a**). As expected, full-length p50 was present in the growth with pAzF. In the growth without pAzF, the amber codon was read as a stop codon and thus no full-length p50 was produced. Site-specific incorporation of pAzF for

both RelA and p50 was further confirmed by mass spectrometry following pepsin digestion. The peptide covering position 128 in RelA exhibited a + 60 amu mass shift, indicating the replacement of the glutamine in wildtype with pAzF. A + 34 amu mass shift was also observed for the p-aminopheylanaline (pAmF), the reduced form of pAzF due to the presence of DTT in the buffer for mass spectrometry. Likewise, the peptide covering position 148 in p50 exhibited a +41 amu mass shift, indicating the replacement of phenylalanine with pAzF, and a +15 amu shift for pAmF.

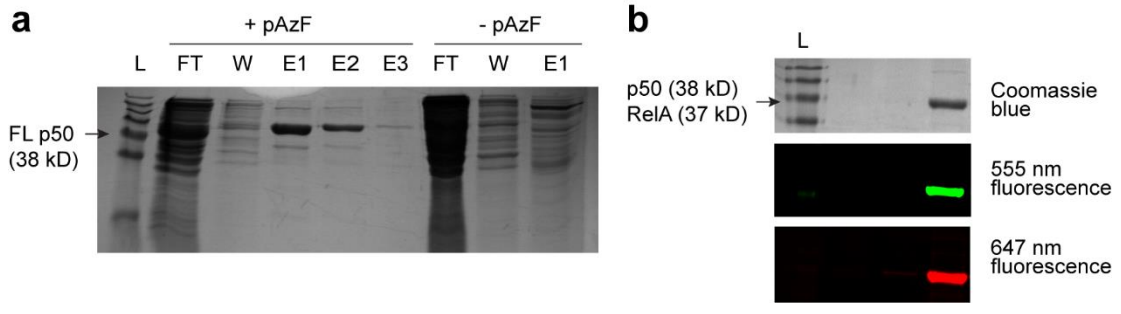


Figure 3.8 Site-specific fluorophore labeling with unnatural amino acid incorporation. **(a)** SDS-PAGE showing two *E. coli* growths with and without the addition of the unnatural amino acid pAzF showing successful expression of full-length (FL) p50 incorporating pAzF at position 148 which is an amber codon. In the growth without pAzF, the amber codon was read as a stop codon and thus FL p50 was not expressed. L: ladder; FT: flow-through of Ni-NTA purification; W: wash; E1-E3: eluted fractions. **(b)** SDS-PAGE showing donor-labeled p50 and acceptor labeled RelA were mixed to form dual-labeled RelA–p50 dimers. Donor (555 nm) and acceptor (647 nm) fluorescence were detected respectively.

The two subunits of the NF-κB dimer were further labeled with donor and acceptor separately. The azido group of pAzF undergoes the copper-free azide-alkyne cycloaddition reaction[11] with the cycloalkyne group of Alexa Fluor 555/647 sDIBO alkyne (Click-iT™ Alexa Fluor™ sDIBO Alkyne , Thermo Fisher Scientific). Labeling reactions were carried out at 4 °C in the dark for 2 to 7 days. The NF-κB monomer was mix with the dye with a 1-to-1 molar ratio with a concentration of 20 to 80 µM in 600 µL to 2 mL of 25 mM Tris pH 7.5, 150 mM NaCl, 0.5 mM EDTA. Labeled proteins were passed through PD-10 desalting columns to remove the access dye from each labeled protein. RelA and p50 were then combined to form RelA–p50 dimers.

Successful labeling was confirmed with SDS-PAGE and fluorescent gel imaging (**Figure 3.8 b**). Labeling efficiency was quantified by UV-Vis absorption using with the extinction coefficients (ε^{RelA}_{280} = 20,400 $M^{-1}cm^{-1}$, ε^{p50}_{280} = 23,380 $M^{-1}cm^{-1}$, $\varepsilon^{Alexa555}_{555}$ = 155,000 $M^{-1}cm^{-1}$, $\varepsilon^{Alexa647}_{650}$ = 239,000 $M^{-1}cm^{-1}$) and the correction factors accounting for UV absorption at 280 nm from the dyes ($CF^{Alexa555}$ = 0.08, $CF^{Alexa657}$ = 0.03). The labeling efficiency is 29% to 56 % for Alexa Fluor 555 to RelA, 20% to 35% for Alexa Fluor 647 to RelA, 52% to 83% for Alexa Fluor 647 to p50, and 18% to 50% for Alexa Fluor 647 to p50.

3.5 Photostability of Dyes

Accurate probing of protein conformational dynamics with smFRET depends on the photostability of dyes, and the two major problems are photobleaching and blinking. Photobleaching is the irreversible conversion to a permanently non-fluorescent state, mostly caused by molecular oxygen (O_2) in the buffer. A Short photobleaching lifetime limits the time length for observation, causing insufficient sampling of the conformational dynamics. Blinking is the undesired fluctuation in fluorescence intensity on the millisecond timescale caused by frequent visits of the dye to dark states. Blinking of the acceptor leads to a zero FRET efficiency that could be falsely assigned to a conformational state with a large inter-dye distance.

Photobleaching lifetime can be lengthened by enzymatic oxygen scavenging systems[12-13] which remove O_2 in the buffer. In this study, a combination of glucose oxidase and catalase was employed to remove O_2 by oxidizing glucose[12]. Blinking is often caused by transient visits to the electronic triplet state, which is naturally quenched by O_2. In the case of smFRET, in which O_2 is removed to lengthen the photobleaching lifetime, the duration of dark states can be prolonged to hundreds of milliseconds, resulting in complications of data interpretation. Trolox, a vitamin E

analog, was shown to effectively eliminate blinking and lengthen photobleaching lifetime[14] through quenching the triplet state[15] and would be used in this study.

The antiblinking capability of Trolox comes from the combination of Trolox (TX) and its oxidized species Trolox quinone (TQ) as an oxidizing and reducing system[15]. TX acts as the reductant that quenches the triplet state through electron transfer, resulting in a radical anion that is further returned to the ground state by TQ through a second electron transfer reaction.

In the preparation of TX/TQ reagent, a maturation process through a photochemical reaction of TX with oxygen is needed for TQ formation. Trolox has poor solubility in water as it acidifies the solution but can dissolve better with an increment of pH or in a constant pH buffer. The Trolox solution in this study was prepared by dissolving 10 mg Trolox in a constant pH 7.5 buffer (25 mM Tris, 150 mM NaCl, 0.5 mM EDTA) and stirring for 30 min so that all the powder dissolved. The solution was then incubated at room temperature with ambient light for TQ formation. UV absorption at 290 nm (A_{290}) was used to determine the concentration of TX at time 0 using the extinction coefficient $\varepsilon_{290}^{TX} = 2350 \pm 100\ \mathrm{M^{-1}cm^{-1}}$ [16]. TQ concentration can be determined using A_{255} using the equation

$$[TQ] = \frac{\frac{A_{255}}{d} - \varepsilon_{255}^{TX} \cdot [TX]_0}{\varepsilon_{255}^{TQ} - \varepsilon_{255}^{TX}},$$

in which $\varepsilon_{255}^{TX} \sim 400\ \mathrm{M^{-1}cm^{-1}}$ and $\varepsilon_{255}^{TQ} \sim 11600\ \mathrm{M^{-1}cm^{-1}}$.[15]

The time course of TQ formation was shown in **Figure 3.9**. Cordes *et al.* recommended to have 2 to 30% of TQ in the solution for antiblinking. Incubation at room temperature under ambient light for 5 hours leads to a sufficient amount of TQ for single-molecule experiments (**Figure 3.9**). In this study, Trolox was typically incubated overnight and then filtered through a 0.2 μm membrane.

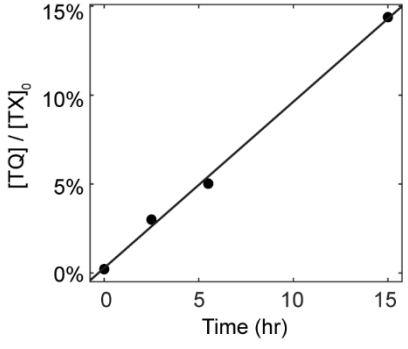


Figure 3.9 Time course of Trolox-quinone (TQ) formation. The combination of Trolox (TX) and TQ is required for effectively quenching the triplet state. After dissolving Trolox (TX), the solution was incubated at room temperature under ambient light for TQ formation. Incubation for 5 hours leads to 5% TQ, which is sufficient for antiblinking. Overnight incubation was carried out for all the smFRET experiments in this study.

Besides reagents, an improved version of the dye pair with better photostability was used. Alexa Fluor 555/647 pair is a superior alternative to the classic, most widely used FRET pair Cy3/Cy5 with better photostability and almost identical spectra (**Figure 3.1a**). The FRET distance for the Alexa Fluor 555/647 dye pair is 51 Å.

3.6 Surface Passivation

In TIRF-based smFRET, surface passivation of the quartz slide is crucial to reduce background fluorescence from impurities and nonspecifically bound molecules. Organic solvents are used to dissolve organic impurities and the surface can be hydroxylated and salinized to increase its hydrophobicity. A final coating with materials such as polyethylene glycol (PEG)[12] or proteins[17] blocks nonspecific binding of molecules.

In this study, the DDS-Tween-20 (DT20) surface passivation method[5] was used. Tween-20 is a surfactant that self-assembles on the DDS (dichlrodimethylsilane) surface to prevent nonspecific binding. Comparing to the conventional PEG surface, which can reject most nonspecific binding at low nanomolar concentrations leaving only 10 to 20 nonspecific spots, the DT20 surface can eliminate almost all nonspecific molecules for sample concentrations up to micromolar. A protocol for the DT20 surface passivation method is described in **Appendix II.**

3.7 Single-Molecule Detection of NF-κB Dynamics

With the instrumentation and sample preparation mentioned above, NF-κB incorporated the unnatural amino acid pAzF was site-specifically labeled and immobilized on a passivated surface for TIRF-based smFRET. Single NF-κB molecules were successfully observed (**Figure 3.10**). The donor and acceptor signals were anticorrelated while the total signal remained stable, indicating interconversions between conformational states with different FRET efficiencies. The sudden decreases in acceptor signal around 40 s and donor signal around 90 s are single photobleaching events. The large-scale fluctuations in FRET efficiency corresponded to the distance changes between the two N-terminal domains over time and provided the first direct evidence for the relative domain motions of NF-κB predicted by MD simulations in Chapter2. Detailed analyses of the motions are reported and discussed in Chapter 4.

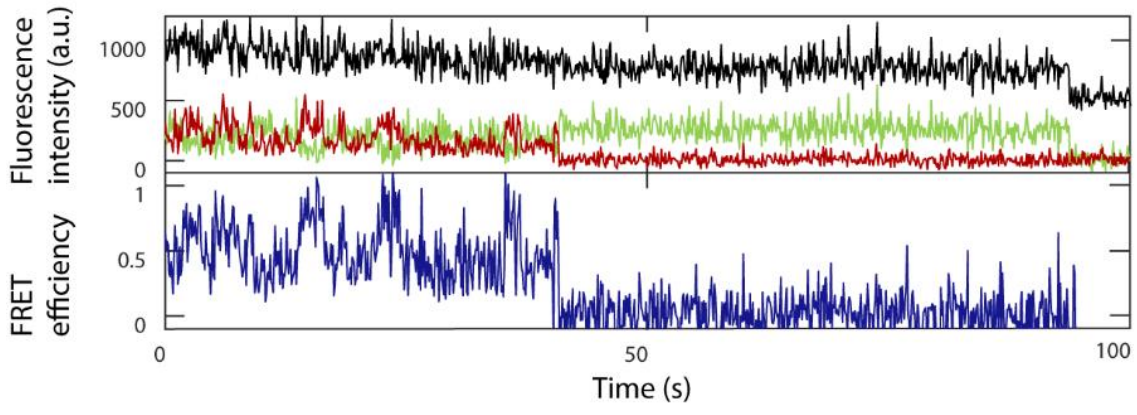


Figure 3.10 A representative smFRET trace of NF-κB. Anticorrelated donor (light green) and acceptor (dark red) signals and the stable total signal (black) indicated interconversions between different FRET states. Sudden drops of signals at 40 s and 90 s are single photobleaching events, showing that only one donor-acceptor pair was captured in this trace. Fluctuations in FRET efficiencies correspond to the distance changes between the two N-terminal domains and confirm the relative domain motions predicted by MD.

Acknowledgments

I thank Professor Hajin Kim at Ulsan National Institute of Science and Technology for his advice on setting up the TIRF microscope and Dr. Majid Ghassemian at the Biomolecular and Proteomics Mass Spectrometry Facility for carrying out the mass spectrometry experiments and analysis.

Appendix I

Protocol for site-specific labeling of NFκB with unnatural amino acids

Nonsense mutation

Mutate the target residue in the expression plasmid into an amber stop codon TAG using site-directed mutagenesis.

Co-transformation

1. Co-transform 50-100 ng of the protein expression plasmid encoding (RelA or p50) and 50-100 ng pEVOL-pAzF plasmid into 100 μL of BL21 (DE3) cells.
2. Plate the cells onto an ampicillin + chloramphenicol plate and let grow overnight.

Growth

3. Pick a colony and grow 10 mL overnight (about 16 hours) starter culture at 37˚C, 180 rpm.

<u>M9Zn starter culture (10 mL):</u>

- 8.6 mL Zn media
- 1.0 mL 15xM9 media
- 0.4 mL 20% Dextrose
- 10 μL MgSO4 (1M stock)
- 2 μL CaCl2 (1M stock)
- 10 μL ampicillin (200 mg/mL stock)
- 10 μL chloramphenicol (34 mg/mL stock)

4. Inoculate 1L M9 minimal media culture with the starter culture in next the morning.

M9 minimal media culture (1 L):

- 860 mL autoclaved DI water
- 100 mL 15xM9 media
- 40 mL 20% Dextrose
- 1 mL MgSO4 (1 M stock)
- 200 µL CaCl2 (1 M stock)
- 50uL Thiamine (20% stock)
- 1 mL ampicillin (200 mg/mL stock)
- 1 mL chloramphenicol (34 mg/mL stock)

5. Grow at 37˚C, 180 rpm until OD_{600} reaches 0.6-0.7. (Usually takes 6-8 hours) Place the flask on ice for 20 min to cool down.
6. Add 200 µL IPTG (1 M stock) and 1 mL arabinose (20% stock) to induce protein expression and aaRS/tRNA expression. Add 200 mg of pAzF (CAS 33173-53-4, Chem-Impex Int'l. Inc.).
7. Grow at 18˚C, 180 rpm overnight. (about 16 hours)

Harvest

8. Collect growth in the 1 L bottle for centrifugation. Spin down cells at 5K rpm (about 5K rcf) for 15-30 min

9. Re-suspend cell pellet in 35 mL of Buffer A (50 mM Na_3PO_4 pH 8, 150 mM NaCl, 10 mM Imidazole pH 8) with 150 μL protease inhibitor cocktail, 1 mM PMSF.
10. Sonicate 6*45 sec in a metal beaker on ice (2 min off in between) (power: 50 W). Because the ice meld due to the heat produced by sonication, adjust the height of the stage during the sonication break to make sure the sonicator tip is immersed. Save 100 uL for SDS-PAGE.
11. Collect lysate in the small centrifugation tube (about 35 mL). Spin down at 12K rpm (about 17K rcf) for 30-45 min

Ni-NTA Purification

Buffers

- Buffer A: 50 mM Na_3PO_4 pH 8, 150 mM NaCl, 10 mM Imidazole pH 8
- Buffer B: 50 mM Na_3PO_4 pH 8, 150 mM NaCl, 250 mM Imidazole pH 8
- Wash Buffer: 50 mM Na_3PO_4 pH 8, 150 mM NaCl, 20 mM Imidazole pH 8

12. Equilibrate 5-10 mL Ni-NTA resin with 5-10X column volume of Buffer A.
13. Load supernatant and collect flow-through.
14. Wash with 50 mL Wash Buffer. Collect wash.
15. Elute protein with 20 mL Buffer B. (Save 10 μL of the flow-through, wash, and elusion for SDS-PAGE)
16. Run SDS-PAGE to confirm protein expression.

Storage

17. Equilibrate a PD-10 desalting column with 25 mL SEC buffer (25 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA).
18. Load 2.5 mL protein sample onto PD-10. Discard flow-through.
19. Elute protein with 3.5 mL SEC buffer.
20. Repeat step 17-20 for the remaining sample.
21. Estimate protein concentration by A280 absorption (Higher 260/280 ratio due to the existence of pAzF. Actual concentration needs to be determined by BCA assay but a rough estimation is good enough for labeling purpose)
22. Separate into aliquots containing about 50 nmol protein (~0.6 mL, ~80 μM) and store at -80 ˚C.

Click reaction

23. Thaw an aliquot of protein and add 1 μL 50 mM sDIBO-Alexa dye (Alexa Fluor 555 or Alexa Fluor 647, in DMSO, stored at -80 ˚C). Keep the reaction in dark, at 4˚C, overnight (incubation for > 2 days gives the best labeling efficiency.)
24. Equilibrate a PD-10 column with SEC buffer containing reducing agent (25 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 1 mM DTT). Wrap the column with aluminum foil or do this in the dark.
25. Adjust sample volume to 2.5 mL and load onto PD-10. Discard flow-through.
26. Elute labeled protein with 3.5 mL SEC buffer containing a reducing agent.
27. Measure absorption at 280 nm, 550 nm, and 650 nm. Calculate labeling efficiency with the extinction coefficients (ε^{RelA}_{280} = 20,400 $M^{-1}cm^{-1}$, ε^{p50}_{280} = 23,380 $M^{-1}cm^{-1}$, $\varepsilon^{Alexa555}_{555}$

= 155,000 $M^{-1}cm^{-1}$, $\varepsilon^{Alexa647}_{650}$ = 239,000 $M^{-1}cm^{-1}$) and the correction factors accounting for A280 from the dyes ($CF^{Alexa555}$ = 0.08, $CF^{Alexa657}$ = 0.03).

Further purification

28. Combine RelA and p50 with a 3x to 5x excess of the acceptor labeled subunit.
29. Purify dimer with Mono S.
30. (optional) Purify dimer with S200 to remove aggregates.

Appendix II

Protocol for DT20 surface preparation

Cleaning

1. If reusing used slide/coverslip chambers, soak chambers in acetone to remove tapes and epoxy. Sonication can speed up the removal.
2. Rub quartz slides with ethanol using kimwipes or lens paper to remove residual tapes. Be careful and use wet kimwipes to avoid making scratches.

Removing Impurities

3. Place slides and new coverslips in glass jars and add acetone to the jar. Sonicate for 20 minutes.
4. Rinse slides and coverslips with methanol.
5. Sonicate slides and coverslips in methanol for 20 minutes.
6. Rinse slides and coverslips with MQ water.

Etching

7. Sonicate slides and coverslips in 5 M KOH for 1 hour to hydroxylate the surface.
8. Take out dichlorodimethylsilane (DDS, Aldrich, >99.5%) from 4 °C and warm it up to room temperature.
9. Rinse slides and coverslips with MQ water thoroughly (> 10 times).
10. Air-dry the slides and coverslips with nitrogen and place them in polypropylene (PP) slide holders.

Silanization

11. Rinse slides, coverslips, and a clean 150 mL glass beaker with hexanes (Fisher Chemical, Spectranalyzed™).
12. Add 75 mL hexanes and 50 µL DDS using a 100 µL glass syringe in the beaker. Pure the mixture in slide holders to cover slides and coverslips.

* *Put the needle tip under hexanes to avoid air contact and inject DDS quickly.*

13. Immediately cap the holders and gently shake them at room temperature for 1.5 hours.
14. Dump the solution into a waste bottle.
15. Sonicate slides and coverslips in hexanes for 1 minute. Repeat this step 3 times.
16. Air-dry slides and coverslips with nitrogen.
17. Put a pair of slide and coverslip in a 50 mL tube. Fill tubes with nitrogen and seal with parafilm. Store at -20 °C

Assemble sample chamber

On the day of experiment, warm up the slide and coverslip to room temperature and place them on a clean and dry surface. Assemble the slide and coverslip into a sample chamber as following (**Figure 3.11**).

18. Cut a piece of double-sided tape into 5 to 6 long strands.
19. Carefully place each strand on the slide to create channels.
20. Place the coverslip on top and use a pipette tip to scrub the surface to seal channel divisions.
21. Seal the edge with epoxy.

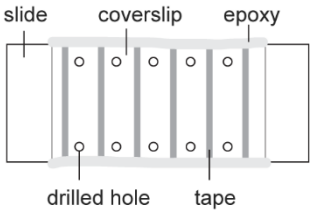


Figure 3.11 The sample chamber for smFRET. Channels are separated with double-sided tapes places between a slide with holes and a coverslip. Edges are sealed with epoxy.

Coating

22. Flow in 50 µL 0.2 mg/mL biotinylated BSA in SEC buffer (25 mM Tris, 150 mM NaCl, 0.5 mM EDTA) in a channel through the drilled hole. Incubate for 5 minutes.
23. Flow in 100 µL 0.2 Tween-20 in SEC buffer. Incubate for 10 minutes.
24. Flow in 50 µL 0.2 mg/mL NeutrAvidin in SEC buffer. Incubate for 5 minutes.
25. For biotinylated samples. Flow in 50 µL of the sample and incubate for 5 minutes. For His-tagged samples flow in 50 µL anti-His antibody biotin conjugate, incubate for 5 minutes, and then flow in 50 µL of the sample and incubate for 5 minutes.
26. Flow in 100 µL imaging buffer (5 M Trolox, 0.8 mg/mL glucose oxidase (G2133, Sigma-Aldrich), 0.2 mg/mL catalase (219001, CALBIOCHEM), 0.4% dextrose)

References

1. Narang, D.; Chen, W.; Ricci, C. G.; Komives, E. A., RelA-Containing NFkappaB Dimers Have Strikingly Different DNA-Binding Cavities in the Absence of DNA. *J Mol Biol* **2018,** *430* (10), 1510-1520.

2. Potoyan, D. A.; Zheng, W.; Komives, E. A.; Wolynes, P. G., Molecular stripping in the NF-kappaB/IkappaB/DNA genetic regulatory network. *Proc Natl Acad Sci U S A* **2016,** *113* (1), 110-5.

3. Förster, T., Zwischenmolekulare Energiewanderung und Fluoreszenz. *Annalen der physik* **1948,** *437* (1-2), 55-75.

4. Ha, T.; Enderle, T.; Ogletree, D. F.; Chemla, D. S.; Selvin, P. R.; Weiss, S., Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc Natl Acad Sci U S A* **1996,** *93* (13), 6264-8.

5. Hua, B.; Han, K. Y.; Zhou, R.; Kim, H.; Shi, X.; Abeysirigunawardena, S. C.; Jain, A.; Singh, D.; Aggarwal, V.; Woodson, S. A.; Ha, T., An improved surface passivation method for single-molecule studies. *Nat Methods* **2014,** *11* (12), 1233-6.

6. Liu, C. C.; Schultz, P. G., Adding new chemistries to the genetic code. *Annu Rev Biochem* **2010,** *79*, 413-44.

7. Milles, S.; Lemke, E. A., Single molecule study of the intrinsically disordered FG-repeat nucleoporin 153. *Biophys J* **2011,** *101* (7), 1710-9.

8. Hartmann, A.; Krainer, G.; Keller, S.; Schlierf, M., Quantification of Millisecond Protein-Folding Dynamics in Membrane-Mimetic Environments by Single-Molecule Forster Resonance Energy Transfer Spectroscopy. *Anal Chem* **2015,** *87* (22), 11224-32.

9. McLoughlin, S. Y.; Kastantin, M.; Schwartz, D. K.; Kaar, J. L., Single-molecule resolution of protein structure and interfacial dynamics on biomaterial surfaces. *Proc Natl Acad Sci U S A* **2013,** *110* (48), 19396-401.

10. Lee, T. C.; Kang, M.; Kim, C. H.; Schultz, P. G.; Chapman, E.; Deniz, A. A., Dual Unnatural Amino Acid Incorporation and Click-Chemistry Labeling to Enable Single-Molecule FRET Studies of p97 Folding. *Chembiochem* **2016,** *17* (11), 981-4.

11. Jewett, J. C.; Bertozzi, C. R., Cu-free click cycloaddition reactions in chemical biology. *Chem Soc Rev* **2010,** *39* (4), 1272-9.

12. Roy, R.; Hohng, S.; Ha, T., A practical guide to single-molecule FRET. *Nat Methods* **2008,** *5* (6), 507-16.

13. Aitken, C. E.; Marshall, R. A.; Puglisi, J. D., An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophys J* **2008,** *94* (5), 1826-35.

14. Rasnik, I.; McKinney, S. A.; Ha, T., Nonblinking and long-lasting single-molecule fluorescence imaging. *Nat Methods* **2006,** *3* (11), 891-3.

15. Cordes, T.; Vogelsang, J.; Tinnefeld, P., On the mechanism of Trolox as antiblinking and antibleaching reagent. *J Am Chem Soc* **2009,** *131* (14), 5018-9.

16. Joo, C.; Ha, T., Imaging and identifying impurities in single-molecule FRET studies. *Cold Spring Harb Protoc* **2012,** *2012* (10), 1109-12.

17. Jeyachandran, Y. L.; Mielczarski, J. A.; Mielczarski, E.; Rai, B., Efficiency of blocking of non-specific interaction of different proteins by BSA adsorbed on hydrophobic and hydrophilic surfaces. *J Colloid Interface Sci* **2010,** *341* (1), 136-42.

Chapter 4 Extremely Slow and Heterogeneous Domain Motions in NF-κB/DNA/IκBα Interactions

4.1 Introduction

75% of eukaryotic proteins are multidomain proteins and their relative domain motions are important for macromolecular recognition, especially for efficient and reversible interactions required in cell signaling and transcription. In this chapter, I focus on the NF-κB RelA–p50 heterodimer, a multidomain transcription factor that recognizes target DNA with its two N-terminal domains (NTDs). In the NF-κB dimer, each NTD is connected to the C-terminal dimerization domain by a 10-amino-acid linker. In Chapter 2, Molecular dynamics (MD) simulations suggested that the flexible linkers allowed the two NTDs to undergo large-scale relative domain motions which lead to exposure of the DNA-binding cavity in RelA–p50.To experimentally probe these motions, single-molecule FRET (smFRET) was established as described in Chapter 3, where successful observation of the motions was also demonstrated.

In this chapter, smFRET results of the domain motions of NF-κB are further analyzed and discussed. Conformational heterogeneity of NF-κB suggested by MD simulations was confirmed experimentally. Extremely slow and heterogeneous domain motions occurring on the seconds- and minutes- timescales were discovered. The conformational ensemble and motions of NF-κB were altered by DNA and completely locked by the inhibitory protein IκBα, which is known to remove NF-κB from DNA and prevent DNA-binding. With smFRET, we obtained a quantitative characterization of the extremely slow and heterogeneous domain motions in the NF-κB/DNA/IκBα interactions that, to our knowledge, have not been reported in other protein systems.

4.2 Methods

Protein expression and purification. Plasmids of murine $RelA_{1-325}$ (UniProt entry Q04207), $p50_{39-363}$ (UniProt entry P25799), and human $I\kappa B\alpha_{67-287}$ (UniProt entry Q9UGJ8) in pET11a vectors (Plasmid #44744, Addgene) were provided by Gourisankar Ghosh at UCSD. A 6XHis-tag was introduced to the C-terminus of RelA and to the N-terminus of IκBα by site-directed mutagenesis. The DNA codons for residues Q128 of RelA and F148 of p50 were mutated into Amber stop codons by site-directed mutagenesis for unnatural amino acid incorporation and fluorophore labeling. The plasmid containing RelA or p50 was co-transformed with pEVOL-pAzF[1] (Plasmid #31186, Addgene), the plasmid encoding the aaRS/tRNA pair for the unnatural amino acid p-azidophenylalanine (pAzF, CAS 33173-53-4, Chem-Impex Int'l. Inc.), into BL21 (DE3) E. coli strain (New England Biolabs). Cells were first grown in M9 minimal media with 200 mg/L ampicillin and 34 mg/L chloramphenicol at 37 °C and 180 rpm. At 0.6 - 0.7 OD_{600}, the expression of RelA or p50 was induced with 0.2 mM IPTG, and the expression of the aaRS/tRNA was induced with 0.02% L-arabinose. pAzF was added to a 200 mg/L final concentration. Cells were then grown at 18 °C and 180 rpm for 16 hours and harvested by centrifugation at 5,000 rcf for 15 minutes.

The cell pellet containing the 6XHis-tagged protein was resuspended in 50 mM sodium phosphate pH 8, 150 mM NaCl, 10 mM imidazole, 0.5 mM PMSF with protease inhibitor cocktail (Sigma-Aldrich). Following sonication on ice, the cell lysate was centrifuged at 17,000 rcf for 30 minutes. The supernatant was passed through Ni-NTA resins (Thermo Fisher Scientific) and washed with 10 column volumes of the wash buffer (50 mM sodium phosphate pH 8, 150 mM NaCl, 20 mM imidazole). The 6XHis-tagged protein was eluted with 50 mM sodium phosphate, 150 mM NaCl, 250 mM imidazole. Elution was pass through a PD 10 desalting column (GE

Healthcare) to remove imidazole and eluted with 25 mM Tris pH 7.5, 150 mM NaCl, 0.5 mM EDTA. After added the protease inhibitor cocktail and glycerol (5% final concentration), the sample was aliquoted and stored at -80 °C for further fluorophore labeling and purification.

The cell pellet containing non-tagged proteins was resuspended in 25 mM Tris pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 0.5 mM PMSF. After sonication and centrifugation, the supernatant was passed through SP Sepharose Fast Flow resins (GE Healthcare). A NaCl gradient from 0 to 700 mM NaCl was used to elute the non-tagged protein. The elution fraction containing the non-tagged protein was purified with a PD10 desalting column and stored in 20 mM Tris pH7.5, 150 mM NaCl, 0.5 mM EDTA, 5% glycerol at -80 °C. Reducing agents was avoided to prevent reducing pAzF to pAmF (p-aminophenylalanine). Expression of full-length $RelA_{1-325}$ and $p50_{39-363}$ was confirmed by SDS-PAGE.

IκBα was expressed with BL21 (DE3) E. coli strain growing in M9 minimal media with 200 mg/L ampicillin at 37 °C and 180 rpm. At 0.6 - 0.7 OD_{600}, protein expression was induced with 0.2 mM IPTG. Cells were then grown at 18 °C and 180 rpm for 16 hours and harvested as the same way for NF-κB. The 6XHis-tagged IκBα was purified with Ni-NTA using the above-mentioned buffers and size-exclusion chromatography (Superdex 75, GE Healthcare) in Tris pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 1mM DTT.

Site-specific fluorophore labeling. Site-specific incorporation of pAzF was further confirmed by pepsin digestion and tandem mass spectrometry. Peptides DLE*AISQR (M+H 1093.567, residue 126 – 133 of RelA) and GILHVTKKKV*ETL (M+H 1653.984, residue 138 – 161 of p50) containing pAzF (denoted by an asterisk) were identified. Fluorophore labeling was achieved through the copper-free click reaction between the azido sidechain of the incorporated pAzF and the sDIBO functional group on the fluorophore. RelA and p50 were thawed from -80 °C and labeled separately with Click-iT™ Alexa Fluor™ 555 sDIBO Alkyne (FRET donor,

Thermo Fisher Scientific) or Click-iT™ Alexa Fluor™ 647 sDIBO Alkyne (FRET acceptor, Thermo Fisher Scientific) with a 1:1 protein-to-dye molar ratio at 4 °C for 2-7 days. The FRET radius of the donor/acceptor pair used in this study is 51 Å. Excess dyes were removed with PD10 desalting columns. Labeling efficiency was determined with the extinction coefficients ($\varepsilon^{RelA}{}_{280}$ = 20,400 $M^{-1}cm^{-1}$, $\varepsilon^{p50}{}_{280}$ = 23,380 $M^{-1}cm^{-1}$, $\varepsilon^{Alexa555}{}_{555}$ = 155,000 $M^{-1}cm^{-1}$, $\varepsilon^{Alexa647}{}_{650}$ = 239,000 $M^{-1}cm^{-1}$) and the correction factors accounting for A280 from the dyes ($CF^{Alexa555}$ = 0.08, $CF^{Alexa657}$ = 0.03). The labeling efficiencies for Alexa 647 RelA and Alexa 555 p50 were 35% and 52% respectively. Labeled RelA and p50 were mixed with a 1:5 donor-to-acceptor molar ratio to form RelA–p50 heterodimer. The dual-labeled heterodimer was then purified by cation exchange chromatography (Mono S 10/100, GE Healthcare) in 25 mM Tris pH 7.5, 0.5 mM EDTA, 1mM DTT with an NaCl gradient from 0 to 700 mM and size exclusion chromatography (Superdex 200, GE Healthcare) in Tris pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 1mM DTT as previously used for wildtype RelA–p50 heterodimer[2].

Single-molecule FRET data collection. A prism-type total internal reflection fluorescence (TIRF) microscopy was used. For Alexa 555/ Alexa 647 dye pair, 532 nm laser (SAPPHIRE 532-300 CW CDRH, Coherent) and 637 nm laser (OBIS™ 1196625 | 637nm LX 140mW Laser, Coherent) were set up to excite donor and acceptor respectively. Laser output powers were set to 50 mW for smFRET experiments. 532 nm and 637 nm laser beams were guided by mirrors and a beam splitter to colocalize on the excitation area on the prism. An inverted microscope was set up by mounting the objective (CFI Plan Apochromat VC 60XC WI, Nikon) on a Rapid Automated Modular Microscope System (Applied Scientific Instrumentation). Emission light was further split into donor and acceptor channels with a DV2 Two-Channel, Simultaneous-Imaging System (Photometrics) with bandpass filters (ET585/65m for donor channel and ET685/56m for acceptor channel, Chroma). Fluorescent signals were detected by

Electron Multiplying Charge-Coupled Device (EMCCD) (iXon^{EM}+ EMCCD camera, DU-897E-CS0-#BV, Andor Technology). DDS-Tween 20 passivated surface[3] with 0.2 mg/mL biotinylated bovine serum albumin (A8549, Sigma-Aldrich) and 0.2 mg/mL NeutrAvidin (31000, Thermo Fisher Scientific) were used to immobilize NF-κB. For free NF-κB experiments, NF-κB molecules were immobilized through the 6XHis-tag on the C-terminus of RelA by penta-His antibody biotin conjugate (34440, Qiagen). For DNA-bound NF-κB experiments, non-tagged RelA and p50 were used and DNA:NF-κB complexes were formed by mixing the proteins and DNA prior experiments and were immobilized through the biotinylated DNA:

Biotin-GCATGCGGGAAATTCCATGCATGCCCCCCATGCATGGAATTTCCCGCATGC

Experiments were performed at room temperature in the imaging buffer (25 mM Tris pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 0.8 mg/mL glucose oxidase (G2133, Sigma-Aldrich), 0.2 mg/mL catalase (219001, CALBIOCHEM), 0.4% dextrose, and 5 mM Trolox (CAS 53188-07-1, Sigma-Aldrich)). Trolox was prepared by dissolving 10 mg Trolox powder in 10 mL of 25 mM Tris, pH 7.5, 150 mM NaCl, 0.5 mM EDTA with vortexing for 30 minutes and then incubation at room temperature and ambient light overnight. The 637 nm laser was used to confirm the existence of acceptors and the 532 nm laser was used for collecting data. Data acquisition and analysis software was obtained from the Ha Lab at Johns Hopkins University http://ha.med.jhmi.edu/resources/#1464200861600-0fad9996-bfd4. Single-molecule movies were recorded with 100 ms time resolution.

Single-molecule FRET data analysis. Individual single-molecule traces were extracted from the acquired movies with the IDL scripts in the Ha Lab software package. The identified traces were further selected based on the criteria: (i) clear single photobleaching step for both donor and acceptor; (ii) anticorrelation pattern for donor and acceptor intensities; (iii) stable total fluorescence intensity before photobleaching; (iii) > 20 sec lifetime for both donor and acceptor.

For each trace, the background was defined as the mean values of the donor and acceptor fluorescence intensities after photobleaching and subtract from the data. To account for the donor signal leakage to the acceptor channel, we measured the correction factor by collecting data of Alexa 555 labeled RelA. After background subtraction for individual traces, the correction factor was defined as the mean value of the ratio of the acceptor intensity to the donor intensity and estimated to be 0.04. Therefore, for a single-molecule trace, the background and donor leakage are corrected by the equation, in which I is the fluorescent intensity:

$$I_{donor} = I_{donor}^{raw} - I_{donor}^{background}$$

$$I_{acceptor} = I_{acceptor}^{raw} - I_{acceptor}^{background} - 0.04 \times I_{donor}$$

The FRET efficiency (E) was calculated using the corrected fluorescent intensities. $E = I_{acceptor} / (I_{donor} + I_{acceptor})$.

Traces were then categorized as stable or fluctuating based on the fluctuations in their FRET efficiencies. If the amplitude of the fluctuation was larger than that caused by noise (shown in the signal after photobleaching), then the trace was categorized as fluctuating because the fluctuation was caused by the internal dynamics of NF-κB. If the fluctuations had comparable amplitude to that caused by noise, no internal dynamics could be inferred, and then the trace was categorized as stable.

Histograms were constructed by binning the FRET efficiency data points of each trajectory before photobleaching. The histogram for long-lived states was constructed by counting each visit to a long-lived state to avoid possible boarding caused by noise. First, a histogram for all the data points of FRET efficiency for one trace was built and fitted with a Gaussian distribution. Only those long-live states that can be fitted with Gaussian distribution are selected. Finally, the mean values from fitting are binned to construct the total histogram. respectively.

Cross-correlation analysis was carried out for fluctuating traces to estimate the fluctuation time scale. The cross-correlation function is defined as previously described[4]. Only traces longer than 30 seconds were included in the analysis and the correlation up to 15 seconds were calculated. Results were fitted with a bi-exponential function to compare the characteristic fluctuation time. Cross-correlation analysis of the defined stable traces showed zero correlation and indicated no fluctuations in the FRET efficiency during observation.

MD simulations and analysis. Distance between the OE1 atom of Q128 in RelA and the CZ atom of F148 in p50 was calculated for every 0.01 ns in the published all-atom MD simulations[5] with CPPTRAJ[6] in the AMBER package. Trajectory after 100 ns where the conformation relaxed to the open conformation was used. Two independent trajectories each with 100 to 400 ns were used, resulting in 600 ns of data.

Homology modeling. The homology model of RelA–p50 for the low-FRET state in the DNA-bound form was built with MODELLER [7] using the p50 homodimer bound to two RNAs as a template. Only the protein part is considered in the modeling process. To our knowledge, there is no openly available tool to homology model protein-nucleic acid complexes.

4.3 Results

4.3.1 NF-κB Display Conformational Heterogeneity as Predicted by MD

Previous MD simulations showed that the two NTDs in an NF-κB dimer can rotate and translate with respect to each other, leading to a broad distribution of conformations[5, 8]. FRET efficiency between the two labeling positions in the NTDs (**Figure 3.3**), residue 128 in RelA and residue 148 in p50, calculated from MD simulations[5] suggested that conformations sampled in hundreds of nanoseconds already can lead to a broad distribution in FRET efficiency ranging from

~0.1 to 1 (**Figure 4.1 a**). To experimentally probe the conformational heterogeneity and the relative motions in the NTDs, I carried out TIRF-based smFRET experiments as described in Chapter 3. For free NF-κB in the absence of DNA, 430 single-molecule traces were collected, each of which has a time length from 50 to 200 seconds depending on the photobleaching lifetime. The experimental FRET histogram, calculated from the first 10 seconds of the traces to avoid bias caused by long traces, showed a broad distribution from low to high FRET efficiency (**Figure 4.1 b**), as predicted by MD, which corresponds to distances spanning a range from 30 Å to 70 Å. Obtaining accurate distances from smFRET requires alternating two-color laser excitation (ALEX)[9] and is not the focus of this work. The distances we reported are close estimates based on rigorous background and donor leakage corrections and are comparable to crystal structures and MD simulataios. These smFRET results validated the predicted high conformational heterogeneity in the relative positions of the two NTDs.

4.3.2 Heterogeneous Long-Lived States in Free NF-κB

The smFRET traces were further characterized as stable (83% of the traces) or fluctuating (17% of the traces). A stable trace was defined as containing long-lived state(s) with lifetimes longer than the photobleaching lifetime, which is typically 90 sec obtained from fitting an exponential function to the distribution (**Figure 4.1 d**). In most stable traces, the single NF-κB molecules spent the entire observation time in one long-lived state until photobleaching. A broad range of long-lived states witj low- to mid- to high-FRET efficiencies were all observed (**Figure 4.1 e**). Occasionally, there were transitions between two states in a trace (**Figure 4.1 f**), indicating these long-lived states are interconvertible via slow dynamics likely in the minute timescale. To avoid broadening in the histogram caused by noise, a histogram for long-lived states was built by counting their mean FRET efficiency (**Figure 4.1 c**). Each count is one visit to a long-lived state.

The histogram constructed in this way showed that these long-lived states could not be separated into a few groups with hidden Markov modeling, but instead had a heterogeneous, continuous distribution spanning from low- to high-FRET efficiency, which corresponded to a broad range of distances of 30 Å and 70 Å between labeling positions. The long-lived states are likely deep wells of a rugged energy landscape that are interconvertible via slow dynamics in the minute timescale.

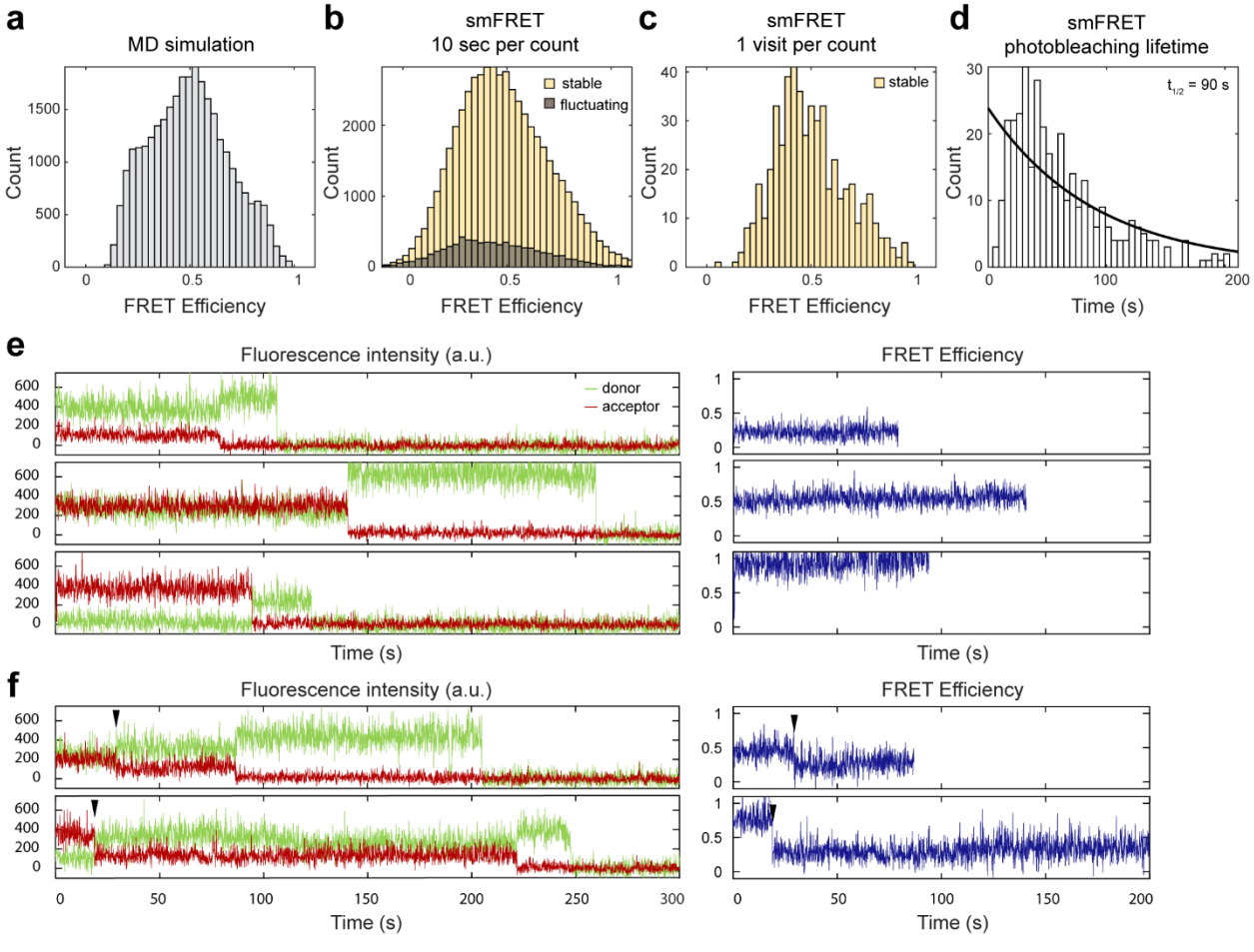


Figure 4.1 Conformational heterogeneity and long-lived states in the NF-κB NTDs. **(a)** All-atom MD simulations suggested that the conformations NTDs can adopt would lead to a broad distribution in FRET efficiency. **(b)** Experimental FRET histogram from smFRET validated the broad distribution of conformational ensemble. The histogram was constructed by counting the first 10 seconds of the traces. **(c)** Experimental FRET histogram constructed by counting each visit to a long-lived state suggested a continuous distribution of stable states that cannot be separated into a few groups. **(d)** Distribution of photobleaching lifetime. Exponential fitting revealed the typical lifetime to be 90 s. **(e)** Representative stable smFRET traces with low, mid, and high FRET efficiency. Other long-lived states with different FRET efficiencies are not shown here. **(f)** Transitions (arrowhead) between long-lived states were captured in smFRET traces.

4.3.3 Distribution Shift of Long-Lived States by DNA

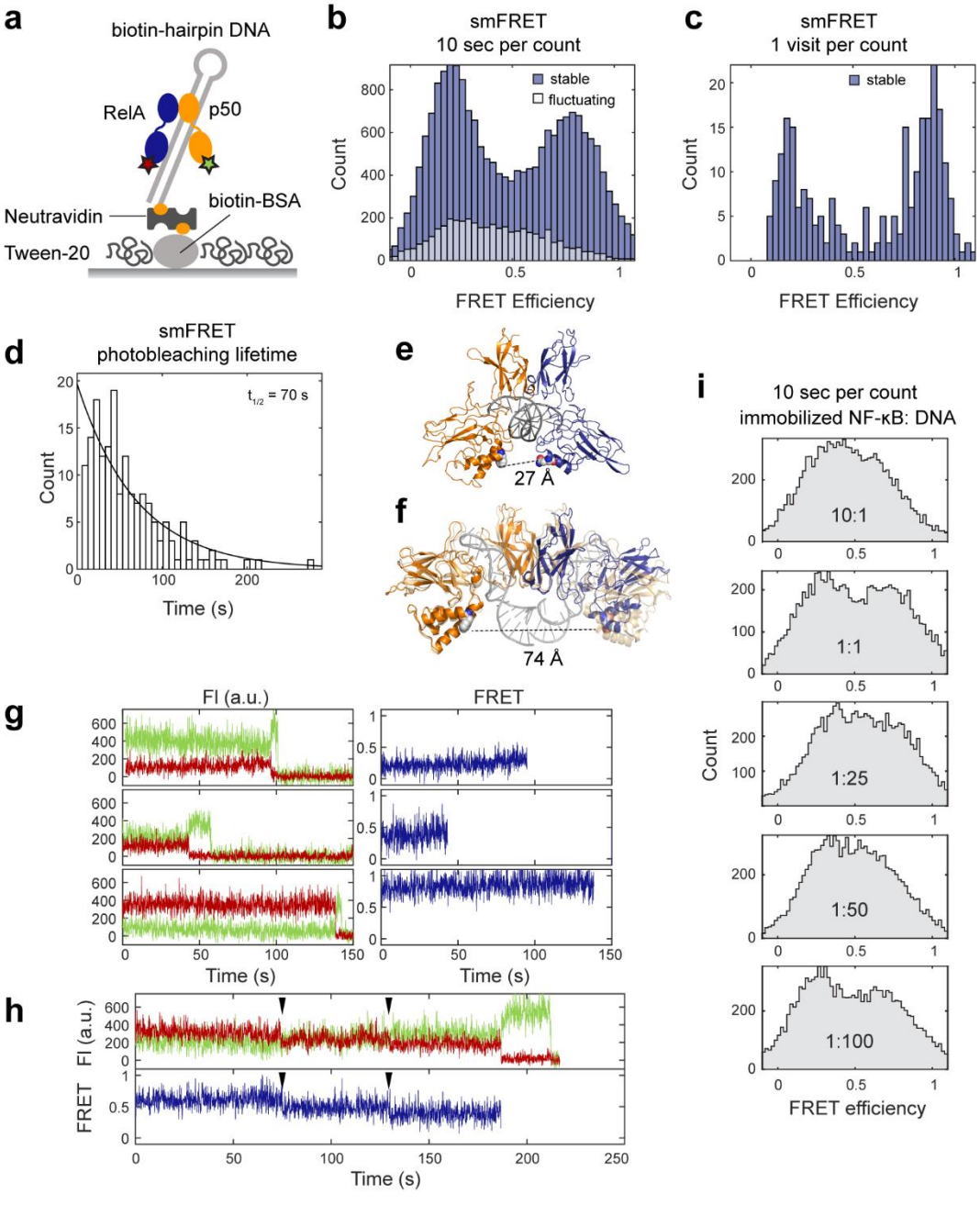
To investigate how DNA affects the relative positions and the motions of the two NTDs, I carried out smFRET experiments with NF-κB in complex with biotinylated hairpin DNA containing the IFN κB binding site. The NF-κB:DNA complex was then immobilized on the passivated surface through biotin-NeutrAvidin interactions (**Figure 4.2 a**). By immobilization through biotinylated DNA, it is ensured that every NF-κB molecule observed is DNA-bound. I collected 336 smFRET traces for DNA-bound NF-κB, of which 84% could be categorized as stable—containing long-lived state(s) with lifetimes longer than the photobleaching lifetime. (**Figure 4.2 b**). Surprisingly, DNA did not shift the conformational ensemble to a single high-FRET state as expected from the crystal structure (PDB entry: 1LE5) (**Figure 4.2 e**), but instead shifted the distribution of long-lived states to two major populations—a high-FRET population in good agreement with the crystal structure and an unexpected low-FRET population. The FRET histogram constructed by counting each visit to a long-lived state recapitulates the presence of two major populations (**Figure 4.2 c**). The time each molecule spent in the long-lived state is typically longer than the photobleaching time life, which is 70 s by fitting an exponential function to the distribution (**Figure 4.2 d**). It is important to note that traces in which molecules sampled a distribution of mid-FRET states were also observed (**Figure 4.2 g**). As was the case for free NF-κB, transitions between states were captured, indicating the long-lived states are also interconvertible in DNA-bound NF-κB likely via dynamics on the minutes timescale (**Figure 4.2 h**).

Since all crystal structures of DNA-bound NF-κB indicate a ~30 Å distance between the labeling positions (PDB entries: 1VKX, 1LE5, 1LE9, 1LEI, 2I9T, 2O61, and 3GUT), it is surprising to observe a major population at low FRET efficiency. However, a crystal structure of

NF-κB p50 homodimer bound to two double-stranded RNA molecules[10] (PDB entry 1OOA), in which the two NTDs are separated by two copies of nucleic acids instead of one as in DNA-bound structures, suggesting the observed low FRET state could be caused by binding to two DNA molecules on the surface of the microscope slide. The homology model of NF-κB RelA–p50 built from 1OOA suggested a 72 Å distance between the labeling positions, which would lead to a 0.1 FRET efficiency (**Figure 4.2 f**).

To further characterize the low-FRET state observed in DNA-bound NF-κB, I carried out smFRET titration experiments with immobilized NF-κB and a series of DNA concentrations. If the low-FRET state is a conformation caused by binding to two DNA molecules, the relative peak height for the low-FRET population would rise as the DNA concentration increases, since at higher DNA concentration it is more likely for NF-κB to bind to two DNA molecules. The results showed that a series of NF-κB:DNA molar ratios from 10:1 to 1:100 lead to almost identical FRET histograms built by counting the first 10 sec, each of which had a broad band caused by free NF-κB and two slightly higher peaks around high and low FRET efficiencies contributed by DNA-bound NF-κB (**Figure 4.2 i**). The independence of the FRET histogram on the DNA concentration suggested the low-FRET state is not a conformation resulting from binding to two copies of DNA.

Figure 4.2 smFRET experiments for DNA-bound NF-κB. **(a)** The experimental schematic. A biotinylated hairpin DNA with a κB site was used for immobilization through biotin-NreutrAvidin interactions. **(b)** FRET histogram showing two major populations at low and high FRET efficiencies contributed by stable traces. Only the first 10 seconds of each trace was included. **(c)** FRET histogram constructed by counting each visit to a long-lived state showing a continuous distribution with two dominating peaks. **(d)** Distribution of photobleaching lifetime. Exponential fitting revealed the typical lifetime to be 90s. **(e)** The crystal structure of DNA-bound NF-κB showing the expected distance between labeling positions (spheres) is 27 Å, which would lead to a FRET efficiency of 1. Blue: RelA; Orange: p50. **(f)** The crystal structure (PDB entry: 1OOA) of NF-κB p50 homodimer (light orange) bound to two copies of double-stranded RNA (grey) from which the homology model of RelA–p50 (blue: RelA and orange: p50) was built. With two copies of nucleic acids in between the two NTDs, the distance between the labeling positions is 74 Å, which corresponds to a FRET efficiency of 0.1. **(f)** Representative stable traces showing long-lived states with low, mid, and high FRET efficiencies. **(h)** A representative trace showing transitions between long-lived states from high to mid to low FRET efficiencies. **(i)** smFRET titration experiments with immobilized NF-κB and varying DNA concentrations. A series of NF-κB:DNA molar ratios from 10:1 to 1:100 showed similar FRET histograms with a broad band of free NF-κB and two slightly higher peaks at low and high FRET efficiencies. The independence of the conformational ensemble to DNA concentration suggested the low FRET state is not a result of binding to two copies of DNA.


a
biotin-hairpin DNA
RelA
p50
Neutravidin
biotin-BSA
Tween-20
b
smFRET
10 sec per count
stable
fluctuating
Count
FRET Efficiency
c
smFRET
1 visit per count
stable
Count
FRET Efficiency
d
smFRET
photobleaching lifetime
t1/2 = 70 s
Count
Time (s)
e
27 Å
f
74 Å
g
FI (a.u.)
FRET
Time (s)
Time (s)
h
FI (a.u.)
FRET
Time (s)
i
10 sec per count
immobilized NF-κB: DNA
10:1
1:1
1:25
1:50
1:100
Count
FRET efficiency

4.3.4 Relative Domain Motions on the Seconds Timescale

Besides the stable traces, 17% of the smFRET traces were categorized as fluctuating, defined as anticorrelated donor and acceptor signals resulting in fluctuations in FRET efficiency. The fluctuations reflected the relative motions between the two domains which caused the labeled residues to move closer and farther apart. In some traces the motions were large enough to connect low- and high-FRET states, corresponding to a 70 Å to 30 Å inter-dye distance change (**Figure 4.3 a, top**). The motions displayed a high level of heterogeneity and as was the case for the long-lived states, the motions could not be described by a few discrete states. In other traces (**Figure 4.3 a, middle**), the two domains fluctuated around a low-FRET state (FRET efficiency about 0.25, distance about 60 Å), and then transitioned to a high-FRET state (FRET efficiency about 0.8, distance about 40 Å) and fluctuated around there. This showed the hierarchy of the rugged energy landscape of NF-κB, which consists of states connected by larger and slower motions and substates interconvertible via smaller and faster motions. I also observed that NF-κB transitioned from fluctuating to a long-lived state and then was temporarily trapped there (**Figure 3 a, bottom**). This again confirmed the long-lived states are not static but are interconvertible with other conformational states.

To characterize the motions, I next carried out cross-correlation analysis for the donor and acceptor signals. 63 fluctuating traces with photobleaching time longer than 30 s were analyzed. The cross-correlation monotonically decayed as a function of time, suggesting no periodicity for the relative domain motion and reflecting the stochastic diffusion-like dynamics of single protein domains (**Figure 3 c**). The cross-correlation could be fitted with bi-exponential functions with the decaying rates for free NF-κB to be $2.0 \pm 0.3\ s^{-1}$ and $0.12 \pm 0.01\ s^{-1}$ and amplitudes to be -22 ± 2

and -38 ± 1. Characteristic fluctuation times, defined as the reciprocal of the rates, were 0.5 s and 8.4 s for free NF-κB.

4.3.5 Fluctuation Times for DNA-Bound NF-κB

The relative domain motions of NF-κB on the second timescale were also present when NF-κB was bound to immobilized DNA. However, large-scale fluctuations connecting low- and high-FRET states were no longer observed (**Figure 4.3 b, top and middle**). In the top panel of **Figure 3b**, the time period in the trace in which there were no donor or acceptor signals could be a donor dark state, an unbound state caused by dissociation of NF-κB from immobilized DNA, or a state in which an unlabeled NF-κB is bound to the immobilized DNA, and would not interfere with data interpretation. As in free NF-κB, transitions between fluctuating states and long-lived states were observed for DNA-bound NF-κB (**Figure 4.3 b, bottom**). Cross-correlation was calculated were carried out for 45 fluctuating traces and fitted with a bi-exponential function, with the decaying rates 1.1 ± 0.2 s^{-1} and 0.06 ± 0.01 s^{-1} and amplitudes -11 ± 1 and -8.4 ± 0.6, revealing the characteristic fluctuation times to be 0.9 s and 17 s. Both times are two-fold slower than those for free NF-κB. The smaller amplitudes reflect the lack of fluctuations connecting low- and high-FRET states. The decreased fluctuation amplitudes and rates for DNA-bound NF-κB are expected from the additional intermolecular interactions introduced by DNA.

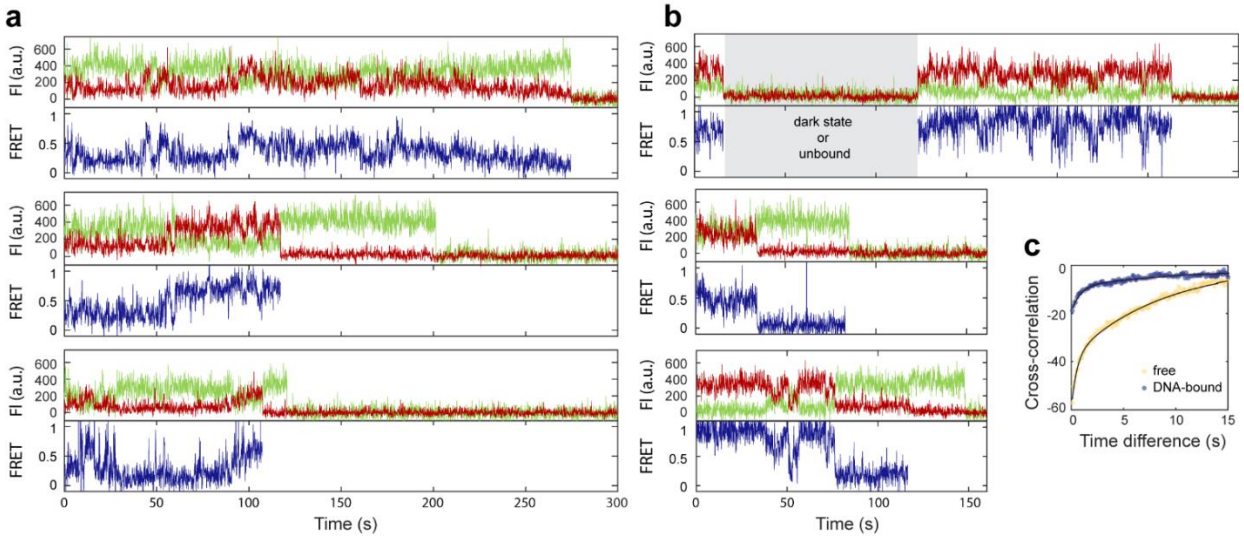


Figure 4.3 Fluctuating smFRET traces of NF-κB showing relative domain motions and the hierarchy of a rugged energy landscape. **(a)** Representative fluctuating traces for free NF-κB. Top: A trace with fluctuations connecting low- and high- FRET states, which corresponds to large scale domain motions connecting conformational states with 30 Å and 70 Å in the inter-dye distance. Middle: A trace with small fluctuations around a low-FRET state at first. At 60 seconds, the molecule transitioned to a high-FRET state and fluctuated around it. This trace demonstrates the transitions between conformational states in a hierarchical rugged energy landscape. Bottom: A trace with transitions between fluctuations and long-lived states, suggesting all the conformational states are interconvertible. **(b)** Representative fluctuating traces of DNA-bound NF-κB. Top and middle: small fluctuations in DNA-bound NF-κB. The time period with no signals (grey area) could be caused by a donor dark state or an unbound state, or a state in which an unlabeled NF-κB is bound to the immobilized DNA. Bottom: A trace with transitions between fluctuations and long-lived states. **(c)** Cross-correlation analyses for fluctuations of free and DNA-bound NF-κB showing that the relative domain motions have lower amplitudes and slower rates due to the interactions with DNA. Bi-exponential fitting revealed the characteristic fluctuation time to be 0.5 s and 8.4 s for free NF-κB and 0.9 s and 17 s for DNA-bound NF-κB. Dot: data; Line: fit.

4.3.6 IκBα Locks the NF-κB Domain Motions

To investigate how the inhibitory protein IκBα that strips NF-κB from DNA and prevents DNA rebinding affects the relative domain motions of NF-κB, I carried out smFRET experiments for IκBα-bound NF-κB. For these experiments, non-tagged RelA-p50 dimer was used and IκBα was immobilized via a 6XHis-tag (**Figure 4.4 a**). The FRET histogram constructed from 99 traces showed that every IκBα-bound NF-κB populated a stable low-FRET state (**Figure 4 b**), also shown in a representative trace (**Figure 4.4 e**), and the histogram built by counting each visit to a long-

lived state revealed a narrow conformational distribution in the range of 56 to 74 Å inter-dye distance (0.1 to 0.4 FRET efficiency) (**Figure 4.4 c**). All traces contained long-lived states with lifetime longer than the photobleaching lifetime, which is 34 s (**Figure 4.4 d**). No anticorrelation of donor and acceptor signals were observed for IκBα-bound NF-κB, suggesting the relative domain motions displayed in free and DNA-bound NF-κB were completely suppressed by IκBα. Transitions between long-lived states were not observed, either. The results suggested that IκB modified the rugged energy landscape of NF-κB to a funnel to a low-FRET state.

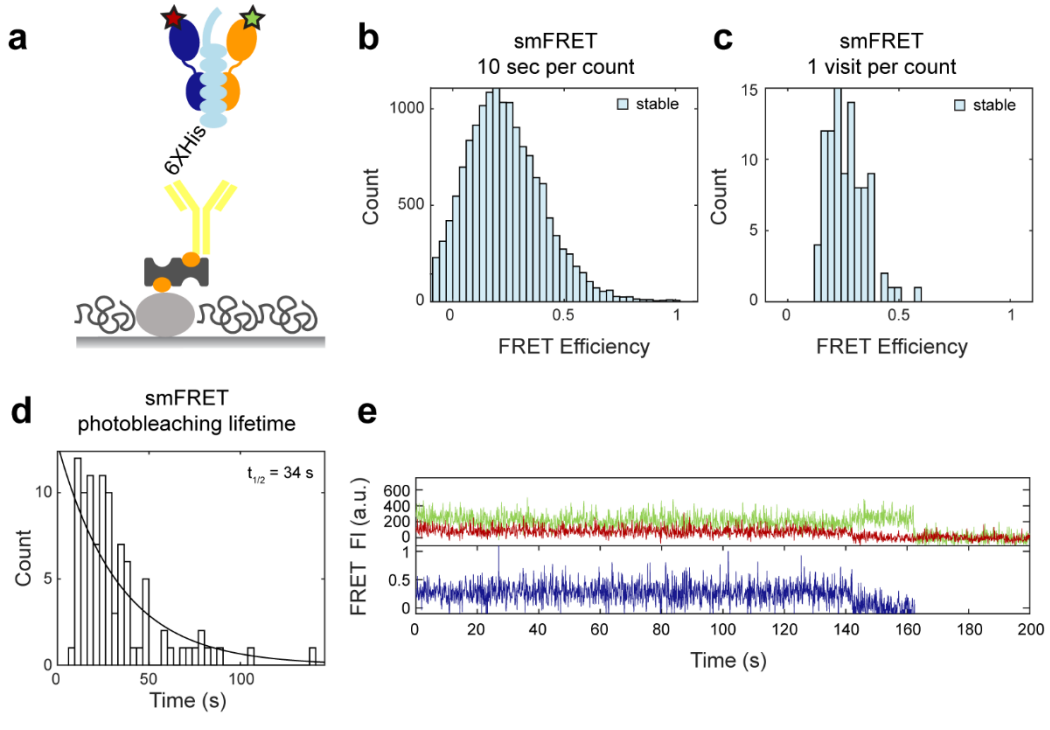


Figure 4.4 smFRET experiments for IκBα-bound NF-κB. **(a)** The experimental schematic. A 6XHis tag on IκBα was used to immobilize the NF-κB:IκBα complex through the interaction with anti-His antibody biotin conjugate. **(b)** FRET histogram showing the distribution of NF-κB molecules bound to IκBα all adopt a stable low FRET state. Only the first 10 seconds of the traces were used to avoid bias from long traces. **(c)** FRET histogram constructed by counting each visit to a long-lived state to reduce noise showing a narrow distribution at low FRET efficiency. **(d)** Distribution of photobleaching lifetime. Exponential fitting revealed the typical lifetime to be 90s. **(e)** A representative trace showing the stable low-FRET state. Every trace for the NF-kB molecules bound to IκBαs gave the same stable low FRET trace.

4.4 Discussion

4.4.1 The Rugged Energy Landscape and Slow Dynamics

Together, the smFRET results suggest a rugged energy landscape for NF-κB, with many deep wells leading to long-lived conformational states interconvertible on the minutes timescale superimposed on a rough surface of substates that interconvert with faster conformational dynamics on the seconds timescale. The hierarchy of the energy landscape manifests in the dynamics on different timescales visualized with smFRET. In contrast to a more accepted picture of conformational switching between just a few states, NF-κB appears to adopt a large number of conformations that constitute the continuous distribution observed by smFRET. Interconversions between these states occur on the second to minute timescale, which is too slow for any experimental technique except TIRF-based smFRET to capture.

To our knowledge, we are not aware of this kind of extremely slow and heterogenous dynamics being reported for other protein systems. The continuous conformational distribution of NF-κB spanning nanometers in distance resembles the continuous nature of conformational ensembles of intrinsically disordered proteins without any secondary or tertiary structures, but in the case of NF-κB, it is not an unstructured polypeptide chain that undergoes conformational dynamics on the ns to μs timescale, but a multidomain protein with well folded domains fluctuating unexpectedly slowly on the second to minute timescale.

The observed NF-κB populating a continuous density of states seems to contradict the NF-κB:DNA crystal structure. However, it has been reported that despite the compact conformation indicated by crystal structures, large-scale domain motions on the μs to ms timescale were observed for the 627-NLS domain of the PB2 segment of influenza polymerase by NMR, SAXS, and smFRET[11]. Another multidomain protein, the human U2AF65 protein for spliceosome

assembly, was also reported to sample a continuum of different conformations by NMR and SAXS[12].

4.4.2 Driving Forces of the Slow Motions

If one thinks about the individual domains in a multidomain protein as spheres chained together with flexible linkers, stochastic diffusion would lead to constant repositioning of the domains in different relative positions and orientations that manifest in changes in FRET efficiency. However, motions due to diffusion are typically fast and cannot explain the extremely slow conformational dynamics on the seconds to minutes timescale observed in NF-κB. Given the observed conformational heterogeneity, mechanisms that cause the slow motions seem to be nonspecific.

A probable mechanism that slows the motions is the electrostatic interaction between domains. The non-uniform charge distribution in NF-κB domains could drive relative motions between them as a result of balancing attractive and repulsive interactions. An analysis of the electrostatic surface of the RelA–p50 heterodimer revealed the electrostatic repulsion between the NTDs as well as the attraction between the NTD and the dimerization domain of p50, which could drive the large-scale conformational change from a closed to an open state[5]. Electrostatic interactions between domains could stabilize interfaces formed in their transient interactions, leading to the lived-lived states observed in smFRET. The frustration between these attractive and repulsive long-range electrostatic forces within the chained protein domains could lead to slow large-scale relative motions.

Besides electrostatics, linker rigidity might also cause the conformational dynamics to be slow. The highly conserved 10-amino-acid linker of NF-κB subunits contains a proline and several charged residues, with the sequence DNRAPNTAEL for RelA and DSKAPNASNL for p50. Prolines and charged residues are known to increase the rigidity of disordered linkers[13]. Therefore,

rather than undergo fast conformational fluctuations, the linkers of NF-κB might prefer certain configurations and undergo slow conformational changes leading to the slow domain motions observed.

4.4.3 Control Mechanism for DNA Association

A rugged energy landscape can be very useful in cell signaling and transcription, in which efficiency and reversibility of macromolecular interactions are important. Conformational flexibility could enable the functional versatility of a protein. In the case of NF-κB, the heterogenous long-lived states and the broad distribution of motional time scales could both be useful in controlling DNA association. To understand the relationship between the NF-kB conformational dynamics and DNA binding, we thought to first compare the timescale of binding to the timescale of conformational dynamics. The average time needed for a DNA-binding event can be estimated from the known rate constant (k_{on}) 2.7×10^{8} $M^{-1}s^{-1}$ [5] and the estimated nuclear concentration of κB sites (1-100 nM). By calculating $1/([\kappa B\ sites]\times k_{on})$ we obtain 0.1 – 10 seconds for a binding event. Recent *in vivo* single-molecule imaging showed that the binding time of RelA to DNA is in the same range[14].

The observed long-lived states typically have lifetimes longer than a minute and thus can be considered static from the viewpoint of DNA-binding. These long-lived states consist of a broad conformational distribution of the positioning of the NTDs with respect to each other. It is likely that only some of them are the “open conformations” with binding cavities accessible for DNA. The long-lived states could be a static control mechanism for NF-κB to regulate its DNA-binding rate. The domain motions observed in smFRET are characterized by fluctuation times 0.5 and 8.4 s, which are the same time scale as the DNA-binding event is in. The seconds motions of NF-κB are coupled to DNA-binding events and could be a dynamics control mechanism in which the re-

orientation of NTDs during the binding process, from first encountering DNA to forming a bound complex, affects the DNA-association kinetics.

4.4.4 Functions of Domain Motions in the NF-κB:DNA Complex

The amount of mRNA produced in transcription is controlled by the transcription time window defined by two time points—transcription factor binding to DNA and dissociation from DNA. Just as the motions in free NF-κB can be useful for DNA-binding, the motions in NF-κB:DNA can be important for the next event—facilitated DNA dissociation by IκBα, or "molecular stripping"[2]. Molecular stripping starts with IκBα encountering NF-κB:DNA, a transient IκBα:NF-κB:DNA ternary complex is then formed[15]. Finally, DNA dissociates and is unable to rebind the remaining NF-κB:IκBα complex. With molecular stripping, the DNA dissociation rate is proportional to the IκBα concentration and thus can be controlled.

The two crystal structures of IκBα:NF-κB showed that IκBα makes contacts with the NF-κB dimerization domains through its ankyrin repeats (PDB entries 1IKN and 1NFI). In both structures, the p50 subunit lacks the NTD and thus the relative positions of the two NTDs are unknown from crystallography. The C-terminal disordered PEST sequence of IκBα rich in proline (P), glutamic acid (E), serine (S), and threonine (T) residues is negatively charged and has been shown to bind to the DNA-binding cavity of NF-κB by NMR[16]. Coarse-grained MD simulations suggested that, during molecular stripping, IκBα induces a twist in the NTDs so that only one NTD can make contacts with the DNA, which lowers the barrier for DNA dissociation[8]. A previous smFRET study on IκBα showed the stretching motions in the ankyrin repeats were more pronounced in the presence of the NF-κB NTDs, suggesting a long-range communication between them[17]. The PEST sequence was thought to insert into the DNA binding cavity of NF-κB as a DNA mimic to exclude DNA-binding. The motions in DNA-bound NF-κB revealed by smFRET might

be utilized by IκBα for a series of conformational rearrangements during molecular stripping, especially for re-orientating the NTDs for DNA dissociation. In IκBα-bound NF-κB, the NTDs are locked in static positions which lead to long-lived low-FRET states. The smFRET studies show that IκBα prevents DNA binding by preventing the NTD motions required for DNA-binding.

4.5 Conclusion

Extremely slow and heterogenous domain motions were directly observed and characterized for NF-κB with smFRET, revealing a rugged energy landscape that can be reshaped by DNA and IκBα. A continuous distribution of conformations unexpected from crystal structure was discovered. We proposed the association and dissociation of NF-κB and DNA can be controlled with the domain motions in free and DNA-bound NF-κB respectively to achieve efficient transcriptional regulation.

References

1. Chin, J. W.; Santoro, S. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G., Addition of p-azido-L-phenylalanine to the genetic code of Escherichia coli. *J Am Chem Soc* **2002,** *124* (31), 9026-7.

2. Bergqvist, S.; Alverdi, V.; Mengel, B.; Hoffmann, A.; Ghosh, G.; Komives, E. A., Kinetic enhancement of NF-kappaBxDNA dissociation by IkappaBalpha. *Proc Natl Acad Sci U S A* **2009,** *106* (46), 19328-33.

3. Hua, B.; Han, K. Y.; Zhou, R.; Kim, H.; Shi, X.; Abeysirigunawardena, S. C.; Jain, A.; Singh, D.; Aggarwal, V.; Woodson, S. A.; Ha, T., An improved surface passivation method for single-molecule studies. *Nat Methods* **2014,** *11* (12), 1233-6.

4. Lamboy, J. A.; Kim, H.; Dembinski, H.; Ha, T.; Komives, E. A., Single-molecule FRET reveals the native-state dynamics of the IkappaBalpha ankyrin repeat domain. *J Mol Biol* **2013,** *425* (14), 2578-90.

5. Narang, D.; Chen, W.; Ricci, C. G.; Komives, E. A., RelA-Containing NFkappaB Dimers Have Strikingly Different DNA-Binding Cavities in the Absence of DNA. *J Mol Biol* **2018,** *430* (10), 1510-1520.

6. Roe, D. R.; Cheatham, T. E., 3rd, PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J Chem Theory Comput* **2013,** *9* (7), 3084-95.

7. Webb, B.; Sali, A., Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Protein Sci* **2016,** *86*, 2 9 1-2 9 37.

8. Potoyan, D. A.; Zheng, W.; Komives, E. A.; Wolynes, P. G., Molecular stripping in the NF-kappaB/IkappaB/DNA genetic regulatory network. *Proc Natl Acad Sci U S A* **2016,** *113* (1), 110-5.

9. Hellenkamp, B.; Schmid, S.; Doroshenko, O.; Opanasyuk, O.; Kuhnemuth, R.; Rezaei Adariani, S.; Ambrose, B.; Aznauryan, M.; Barth, A.; Birkedal, V.; Bowen, M. E.; Chen, H.; Cordes, T.; Eilert, T.; Fijen, C.; Gebhardt, C.; Gotz, M.; Gouridis, G.; Gratton, E.; Ha, T.; Hao, P.; Hanke, C. A.; Hartmann, A.; Hendrix, J.; Hildebrandt, L. L.; Hirschfeld, V.; Hohlbein, J.; Hua, B.; Hubner, C. G.; Kallis, E.; Kapanidis, A. N.; Kim, J. Y.; Krainer, G.; Lamb, D. C.; Lee, N. K.; Lemke, E. A.; Levesque, B.; Levitus, M.; McCann, J. J.; Naredi-Rainer, N.; Nettels, D.; Ngo, T.; Qiu, R.; Robb, N. C.; Rocker, C.; Sanabria, H.; Schlierf, M.; Schroder, T.; Schuler, B.; Seidel, H.; Streit, L.; Thurn, J.; Tinnefeld, P.; Tyagi, S.; Vandenberk, N.; Vera, A. M.; Weninger, K. R.; Wunsch, B.; Yanez-Orozco, I. S.; Michaelis, J.; Seidel, C. A. M.; Craggs, T. D.; Hugel, T., Precision and accuracy of single-molecule FRET measurements-a multi-laboratory benchmark study. *Nat Methods* **2018,** *15* (9), 669-676.

10. Huang, D. B.; Vu, D.; Cassiday, L. A.; Zimmerman, J. M.; Maher, L. J., 3rd; Ghosh, G., Crystal structure of NF-kappaB (p50)2 complexed to a high-affinity RNA aptamer. *Proc Natl Acad Sci U S A* **2003,** *100* (16), 9268-73.

11. Delaforge, E.; Milles, S.; Huang, J. R.; Bouvier, D.; Jensen, M. R.; Sattler, M.; Hart, D. J.; Blackledge, M., Investigating the Role of Large-Scale Domain Dynamics in Protein-Protein Interactions. *Front Mol Biosci* **2016,** *3*, 54.

12. Huang, J. R.; Warner, L. R.; Sanchez, C.; Gabel, F.; Madl, T.; Mackereth, C. D.; Sattler, M.; Blackledge, M., Transient electrostatic interactions dominate the conformational equilibrium sampled by multidomain splicing factor U2AF65: a combined NMR and SAXS study. *J Am Chem Soc* **2014,** *136* (19), 7068-76.

13. Sorensen, C. S.; Kjaergaard, M., Effective concentrations enforced by intrinsically disordered linkers are governed by polymer physics. *Proc Natl Acad Sci U S A* **2019,** *116* (46), 23124-23131.

14. Callegari, A.; Sieben, C.; Benke, A.; Suter, D. M.; Fierz, B.; Mazza, D.; Manley, S., Single-molecule dynamics and genome-wide transcriptomics reveal that NF-kB (p65)-DNA binding times can be decoupled from transcriptional activation. *PLoS Genet* **2019,** *15* (1), e1007891.

15. Alverdi, V.; Hetrick, B.; Joseph, S.; Komives, E. A., Direct observation of a transient ternary complex during IkappaBalpha-mediated dissociation of NF-kappaB from DNA. *Proc Natl Acad Sci U S A* **2014,** *111* (1), 225-30.

16. Sue, S. C.; Dyson, H. J., Interaction of the IkappaBalpha C-terminal PEST sequence with NF-kappaB: insights into the inhibition of NF-kappaB DNA binding by IkappaBalpha. *J Mol Biol* **2009,** *388* (4), 824-38.

17. Trelle, M. B.; Ramsey, K. M.; Lee, T. C.; Zheng, W.; Lamboy, J.; Wolynes, P. G.; Deniz, A.; Komives, E. A., Binding of NFkappaB Appears to Twist the Ankyrin Repeat Domain of IkappaBalpha. *Biophys J* **2016,** *110* (4), 887-95.

Chapter 5 Structural Prediction of the NF-κB Transactivation Domain

There are two essential domains in a transcription factor that defines its function—a DNA-binding domain that recognizes specific DNA sequences, and a transactivation domain (TAD) that binds to regulatory proteins to activate transcription. In the NF-κB family of proteins, only RelA, RelB, and c-Rel contain a TAD and thus only dimers with at least one of these subunits are capable of transcription activation.

The TAD of RelA contains 228 amino acids and occupies the C-terminus of the protein (**Figure 5.1 a**). Within the RelA TAD, TA1 and TA2 are the two regions identified to interact with the KIX and the TAZ1 domains of the regulatory protein CBP/P300 respectively[1-2]. The TAD is highly negatively charged with a theoretical pI value of 4.16 and enriched in prolines with 14.5% (**Figure 5.1 b**). The highly charged nature and the abundance of prolines make the TAD an intrinsically disordered region, as predicted by the protein disorder predictor server IUPred2A[3] (**Figure 5.1 c**).

It has been shown by NMR that TA1 and TA2 undergo disorder-to-order transitions upon binding to the corresponding CPB/P300 domains and assume helical structures in the bound state. Although NMR structures are available for these two regions when bound, the structural ensemble of the entire TAD in the free state remains unknown. In this chapter, *de novo* structure prediction with AWSEM (Associative memory, Water mediated, Structure and Energy Model) simulations were used to explore possible conformations of the TAD and the results are compared with hydrogen-deuterium exchange mass spectrometry (HDX-MS) experiments.

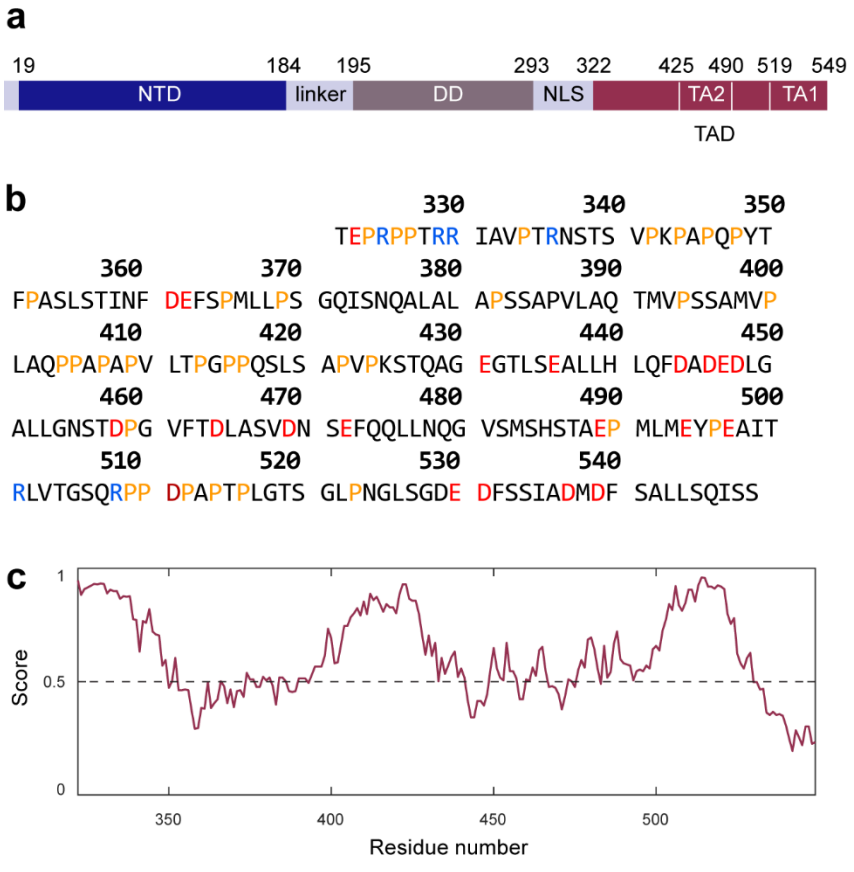


Figure 5.1 The transactivation domain (TAD) of NF-κB RelA. **(a)** Schematic of the domain organization of RelA. NTD: N-terminal domain; DD: dimerization domain; NLS: nuclear localization signal. TAD: transactivation domain. **(b)** The amino acid sequence of the TAD. Red: negatively charged residues; Blue: positively charged residues; Yellow: prolines. **(c)** Disorder prediction of the TAD with IUPred2A. A score above 0.5 indicates disorder.

5.1 Structure Prediction with AWSEM

The two main approaches to protein structure prediction are homology modeling and *de novo* prediction. Homology modeling predicts the structure of the target protein by making an analogy to solved structures with similar sequences, and thus is also called template-based modeling. *De novo* prediction methods directly predict the structure from the amino acid sequence without using templates. The RelA TAD has only a few sequence homologs according to HHblits[4],

none of which have complete structural templates for the TAD, and therefore AWSEM was used to perform *do novo* structure prediction for the RelA TAD.

AWSEM is a solvent-averaged free energy function of the protein's atomic coordinates[5]. For many globular proteins, the native functional ensemble of structures is the ensemble of conformations with the lowest free energy under physiological conditions. Prediction of native structures using AWSEM is done by searching for the global free energy minimum on the landscape, i.e., sampling conformations and finding the ensemble with the lowest free energy.

In practice, searching for the minimum of the AWSEM energy function is done by performing simulated annealing. Simulated annealing is a general algorithm widely used to efficiently find the global minima of functions in a high-dimensional search space. The word "annealing" shows the original inspiration from the technique of gradually cooling down materials to form crystals. In structure prediction with AWSEM, simulated annealing is performed by running a molecular dynamics simulation that starts with an unfolded conformation at a high temperature and is gradually cooled to a low temperature. A large number of conformations are visited during the simulated annealing, and a stable conformation is formed in the end. For practical purposes, multiple independent runs are performed to search for the best conformation. Each run is an independent search through the landscape, and, in a well-funneled landscape, many paths converge to the native structure.

Physical terms in the AWSEM energy function include backbone geometry, residue-residue contacts including water-mediated interactions, residue burial propensity, hydrogen bonding in secondary structures, and long-range electrostatics[6]. There is also a bioinformatically-based term, the associative memory term, which applies local structural constraints based on sequence-sequence alignments of the target sequence to fragments with known structures. A "memory" refers to a portion of a solved structure from the Protein Data Bank (PDB) that has a

similar sequence to a portion of the target protein. In standard structure prediction with AWSEM, each residue in the target is covered with 20 fragment memories with lengths less than or equal to 9 amino acids.

Adding the associative memory bioinformatic term is a simple way to complement the coarse-grained physical terms with local-in-sequence details. This approach utilizes the power of structural data—our current knowledge of sequence-structure correlation from the Protein Data Bank—to better sculpt energy landscapes and predict native structures. Unlike template-based methods, which predict the tertiary structure based on global sequence comparison between the target and templates, AWSEM uses bioinformatics locally and therefore can be used to predict the structure of proteins for which no global template exists in the PDB database. AWSEM need not rely on templates to fold a protein into its native tertiary structure. Instead, the fragment memories guide the formation of local structure, not tertiary structure. In AWSEM, the combination of physical terms and the bioinformatic terms guides the prediction of tertiary structures.

5.2 AWSEM Predicts a Diverse Set of Tertiary Structures for TAD

To investigate the conformational ensemble of TAD alone, I set up twenty AWSM *de novo* structure predictions. In each run, simulated annealing was performed starting with a temperature of 600 K and cooling down to 100 K over 4,000,000 of 2 fs time steps. The predictions resulted in twenty end structures with low similarity in tertiary structure with one another (**Figure 5.2 a**). To quantify the structure similarity, the mutual Q values between every pair of the predicted structures were calculated. The concept of Q comes from protein folding theories, in which it quantifies the degree of foldedness by counting the fraction of native contacts formed in folding intermediate.

The Q value between two protein structures has been used to indicate their similarity in tertiary structures[7] and is defined as:

$$Q^{\alpha\beta} = \frac{2}{(N-2)(N-3)} \sum_{j>j+2} \exp\left[-\left(r_{ij}^{\alpha}-r_{ij}^{\beta}\right)^2 / 2\delta_{ij}^2\right]$$

where α and β denote two protein structure, N is the total number of residues, r_{ij} is the distance between C_α atoms of residue i and j, and δ_{ij} is given as $(1 + |i - j|)^{0.15}$.[5, 7] A number of 0 indicates no similarity at all and 1 indicates identical structures. The mutual Q values between the twenty predicted structures were all below 0.3 indicating a very low similarity between any pairs of structures in the twenty predictions (**Figure 5.2 b**).

To test if the diverse set of structures were an artificial result of insufficient sampling from short simulated annealing, I set up another twenty AWSEM structure predictions with longer annealing processes from 600 K to 100 K with 8,000,000 of 2 fs time steps. Again, all the mutual Q values were below 0.3. Histograms of mutual Q values showed both short and long annealing sampled conformations with low similarity to one another (**Figure 5.2 c**), suggesting the TAD itself is intrinsically highly disordered and adopts a broad distribution of conformations.

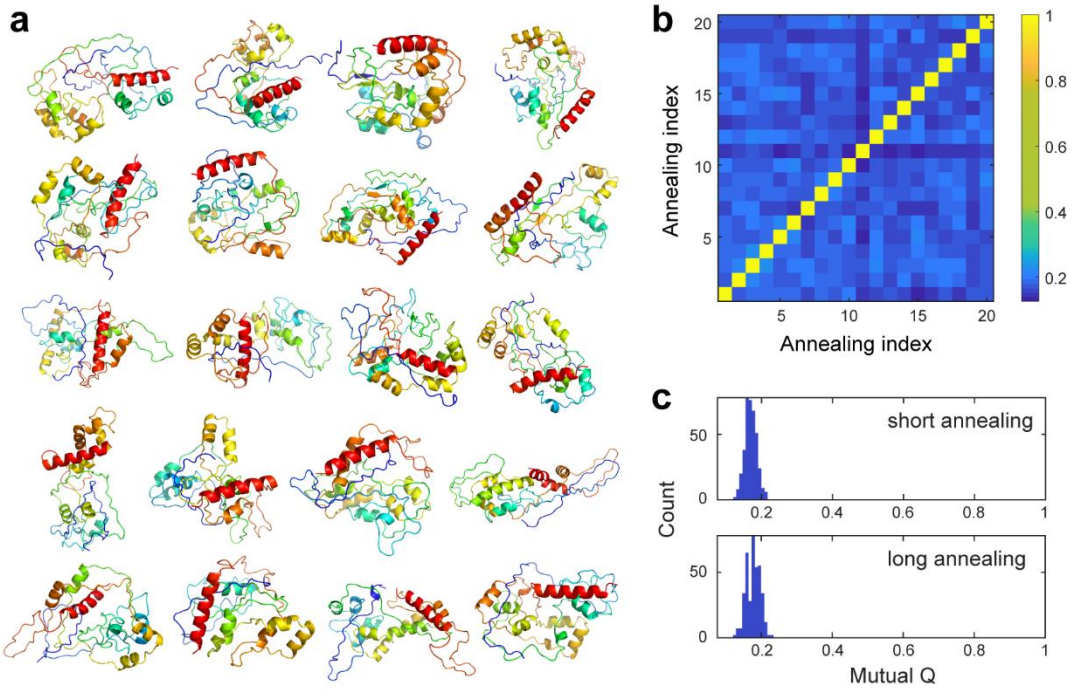


Figure 5.2 AWSEM structure predictions of the NF-κB TAD. **(a)** Predicted TAD structures from twenty AWSEM simulated annealing with different tertiary conformations. Structures are colored with a rainbow scheme with the N-terminus being blue C-terminus being red. **(b)** Mutual Q values between pairs of the predicted structures are all below 0.3, indicating very low similarity in tertiary structures. **(c)** Histograms of mutual Q values from short and long annealing simulations showing the low similarity between predicted structures is independent of sampling range.

5.3 Helical Propensity Predicted by AWSEM in Consistent with HDX-MS and NMR

Some regions in the TAD were consistently predicted to be helical despite the low similarity between the predicted tertiary structures. Analysis of the secondary structure with the Define Secondary Structure of Proteins (DSSP) algorithm in the 2Struc web server[8] showed a consistent pattern of helical formation among the twenty predicted structures (**Figure 5.3 a**). In all predictions, residues 382 to 290, 435 to 440, 460 to 466, 470 to 480, 496 to 502, and 528 to 549, were predicted to be helical. These regions include TA1 (residues 519 to 549) and TA2 (residue

425 to 490), which are two regions previously identified to undergo disorder-to-order transitions upon binding to CBP domains [1-2].

TA1 and TA2 were found by NMR to be disordered in the free state and to fold into helices when bound to the KIX and TAZ1 domains of CBP respectively. Secondary structure analysis on the NMR structures of TA1:KIX complex (PDB entry: 5U4K) and the TA2:TAZ1 complex (PDB entry: 2LWW) showed that the helical formation pattern is consistent with AWSEM prediction (**Figure 5.3 a**). Agreement between the secondary structural pattern can also be seen in the AWSM predicted structures and the CPB-bound NMR structures of the TAD, in all of which TA1 contains a C-terminal helix and TA2 contains several shorter helical segments (**Figure b-d**).

Hydrogen-Deuterium Exchange Mass Spectrometry (HDX-MS) on the full-length RelA–p50 with the TAD showed that the entire TAD exchanged all of its amide protons except for the regions containing TA1 and TA2 (**Figure 5.3 b**). Peptides with residues 440 to 444, 450 to 462, and 474 to 492 covered the TA2 region and exchanged 82 %, 82 %, and 90 % of its amide protons respectively, and the peptide with residues 503 to 540 covered the TA1 region has an exchange percentage of 87%. The observed protection from hydrogen-deuterium exchange for TA1 and TA2 is consistent with AWSEM predicted helical structures, suggesting these two regions are not completely random coils in the free states but instead biased to helical structures which moderately protected backbone amide hydrogen exchange.

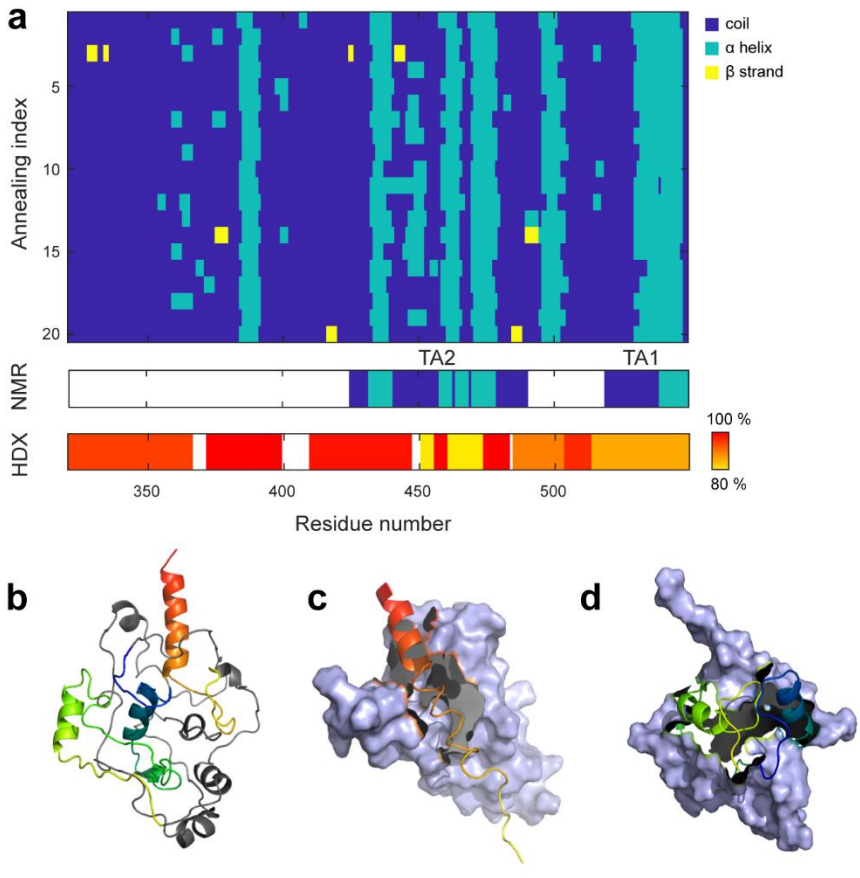


Figure 5.3 Secondary structure analysis of AWSEM predicted structures. **(a)** Agreement between AWSEM predicted secondary structures and experiments. Top: secondary structures of twenty AWSEM predicted structures form a consistent helical pattern. Middle: secondary structures from CBP-bound NMR structures of TA1 and TA2. Bottom: hydrogen-deuterium exchange percentage from HDX-MS. White region: no available data. **(b)** Representative AWSEM predicted structure showing helix formation in TA2 and (N- to C-terminus colored with a blue to green gradient) TA1 (N- to C-terminus colored with an orange to red gradient) **(c)** NMR structure of TA1 bound to CBP-KIX (grey surface). **(d)** NMR structure of TA2 bound to CBP-TAZ1 (grey surface).

References

1. Mukherjee, S. P.; Behar, M.; Birnbaum, H. A.; Hoffmann, A.; Wright, P. E.; Ghosh, G., Analysis of the RelA:CBP/p300 Interaction Reveals Its Involvement in NF-kappa B-Driven Transcription. *Plos Biol* **2013,** *11* (9).

2. Lecoq, L.; Raiola, L.; Chabot, P. R.; Cyr, N.; Arseneault, G.; Legault, P.; Omichinski, J. G., Structural characterization of interactions between transactivation domain 1 of the p65 subunit of NF-kappa B and transcription regulatory factors. *Nucleic Acids Res* **2017,** *45* (9), 5564-5576.

3. Meszaros, B.; Erdos, G.; Dosztanyi, Z., IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res* **2018,** *46* (W1), W329-W337.

4. Remmert, M.; Biegert, A.; Hauser, A.; Soding, J., HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat Methods* **2011,** *9* (2), 173-5.

5. Davtyan, A.; Schafer, N. P.; Zheng, W. H.; Clementi, C.; Wolynes, P. G.; Papoian, G. A., AWSEM-MD: Protein Structure Prediction Using Coarse-Grained Physical Potentials and Bioinformatically Based Local Structure Biasing. *J Phys Chem B* **2012,** *116* (29), 8494-8503.

6. Tsai, M. Y.; Zheng, W.; Balamurugan, D.; Schafer, N. P.; Kim, B. L.; Cheung, M. S.; Wolynes, P. G., Electrostatics, structure prediction, and the energy landscapes for protein folding and binding. *Protein Sci* **2016,** *25* (1), 255-69.

7. Chen, M. C.; Lin, X. C.; Lu, W.; Onuchic, J. N.; Wolynes, P. G., Protein Folding and Structure Prediction from the Ground Up II: AAWSEM for alpha/beta Proteins. *J Phys Chem B* **2017,** *121* (15), 3473-3482.

8. Klose, D. P.; Wallace, B. A.; Janes, R. W., 2Struc: the secondary structure server. *Bioinformatics* **2010,** *26* (20), 2624-2625.

Chapter 6 IκB Proteins Shift the Populations of NF-κB Dimers by Preferential Binding

Adapted and from Ramsey, K. M.*; Chen, W.*; Marion J. D.; Bergqvist, S.; Komives, E. A., Exclusivity and Compensation in NF-κB Dimer Distributions and IκB Inhibition. *Biochemistry* **2019**, *58* (21), 2555-2563. * Equal contribution.

6.1 Introduction

The population of various NF-κB dimers changes dramatically during cell differentiation. While RelA–p50 is most abundant in the early lineages of B-cell lines, p50-cRel dimers are present in later lineages[1]. In murine embryonic fibroblasts, RelA–p50 and the RelA homodimer are observed and appear to activate different sets of genes[2]. The p50-cRel heterodimer, rather than RelA, plays a key role in B-cell proliferation[3,4]. In B-cells, IκBε plays a specific role in limiting cRel and RelA-containing dimers[5].

NF-κB dimers respond to extracellular signals by turning on hundreds of genes and are highly regulated. Studies of the most abundant, canonical RelA–p50 dimer reveal that all dimers are kept in the cytoplasm by a bound inhibitor that requires signal-mediated degradation for removal[6,7]. For RelA–p50, once a stress signal is received, the inhibitory protein IκBα undergoes rapid degradation leading to exposure of the nuclear localization signal on NF-κB and promoting NF-κB nuclear translocation[8,9].

Our lab discovered that an equally stringent system is in place for signal termination. In a classical negative feedback loop, nuclear NF-κB activates transcription of new IκBα, which translocates to the nucleus and accelerates NF-κB dissociation from the DNA—a process termed molecular stripping[10,11]. In cells, the rate of molecular stripping has been shown to directly control the rate of export of NF-κB from the nucleus returning the system to the off state[12]. Without this

kinetic control, the termination of NF-κB transcription would be incomplete and leaky whereas with stripping every gene transcription event is terminated and all of the NF-κB is removed from the nucleus rapidly and completely[13,14].

But what about other dimers and inhibitors? RelA homodimer appears to be the most unstable dimer but is stabilized by IκBβ[15]. In fact, in IκBβ-/- cells, RelA homodimer is no longer observed. Thus, in addition to IκBα being able to remove NF-κB from the nucleus, it appears that IκBβ may stabilize the RelA homodimer[15].

A complete understanding of such a complex system, which is under kinetic control, requires rigorous measurement of binding affinities of all protein complexes. While some attempts have been made at such measurements, there is much disagreement in the field. Indeed, gel shift assays first estimated the affinity between NF-κB and IκBα to be 1 nM[16]. However, later isothermal titration calorimetry and surface plasmon resonance (SPR) estimated the affinity at 30 pM, and this much tighter binding was more consistent with the long intracellular half-life of the complex, which was estimated at 48 hrs in IKK-/- cells[9]. Previous reports estimated the RelA homodimer binding affinity at 800 nM[15]. Such a weak binding affinity would suggest a rapid dimer equilibrium resulting in a substantial monomer concentration at the cytoplasmic concentration of 350 nM. The observation of such a weak binding led to the suggestion that IκBβ may bind to monomeric RelA and perhaps facilitate dimerization[15].

Fluorescence anisotropy assays were developed as a more rigorous binding experiment to measure the equilibria for dimer exchange and IκB binding at concentrations in the nM range[17], which is more in line with intracellular concentrations. This chapter describes the fitting strategy for the complex assays and presents mathematical models for NF-κB subunit exchange and population shifts by IκBs.

6.2 Analysis of Fluorescence Anisotropy Assay for Homodimer Affinities

NF-κB dimer affinities were measured by fluorescence anisotropy assays developed by Ramsey *et al.*[17]. In the assays, a truncated NF-κB construct containing only the dimerization domain was fluorescently labeled and mixed with unlabeled full-length NF-κB containing both the dimerization domain and the N-terminal domain. The full-length NF-κB is in dynamic equilibrium between the monomer and the dimer. When a labeled dimerization domain binds to a full-length monomer, its fluorescence anisotropy increases due to the slower molecular tumbling rate of the dimer formed. Fluorescence anisotropy changes were measured for a series of mixtures with varying concentrations of the full-length NF-κB. A binding curve, which contains information of the dimer affinity, can be made by plotting the anisotropy change against full-length NF-κB concentrations (**Figure 6.1**).

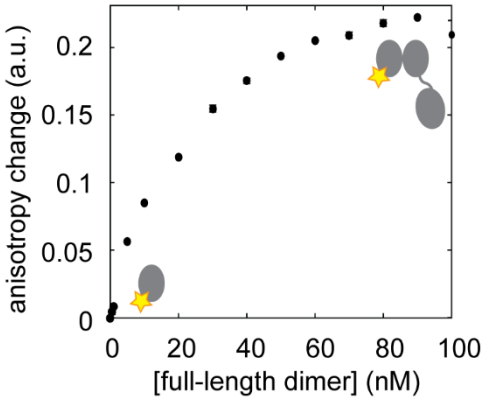


Figure 6.1 A binding curve obtained from fluorescence anisotropy assays of NF-κB. A fluorescently labeled truncated NF-κB construct with only the dimerization domain is fixed at a low concentration and titrated with the unlabeled full-length NF-κB dimer. NF-κB is at dynamic equilibrium between the monomer and the dimer. Once a full-length monomer binds to a labeled dimerization domain, the fluorescence anisotropy signal would increase. A binding curve of anisotropy change vs. full-length NF-κB dimer concentration contains the information of dimer binding affinities.

The affinities of two NF-κB homodimers (the RelA and p50 homodimers), and three heterodimers (RelA–p50, RelA-cRel, and p50-cRel) were measured. For the assays for heterodimers, the dimerization domain of RelA was used for RelA–p50 and RelA-cRel, and the dimerization domain of p50 was used for p50-cRel.

In each assay, there are multiple binding reactions and species at dynamic equilibrium (**Figure 6.2**). Since none of the assays contains only a simple bi-molecular reaction, the resulting binding curves cannot be conventionally fitted with hyperbola to obtain dissociation constants of the dimers[18]. To acquire the dissociation constants of target dimerization processes, I developed a numerical fitting strategy by simulating binding curves with a range of dissociation constants and compared them to the experimental binding curves.

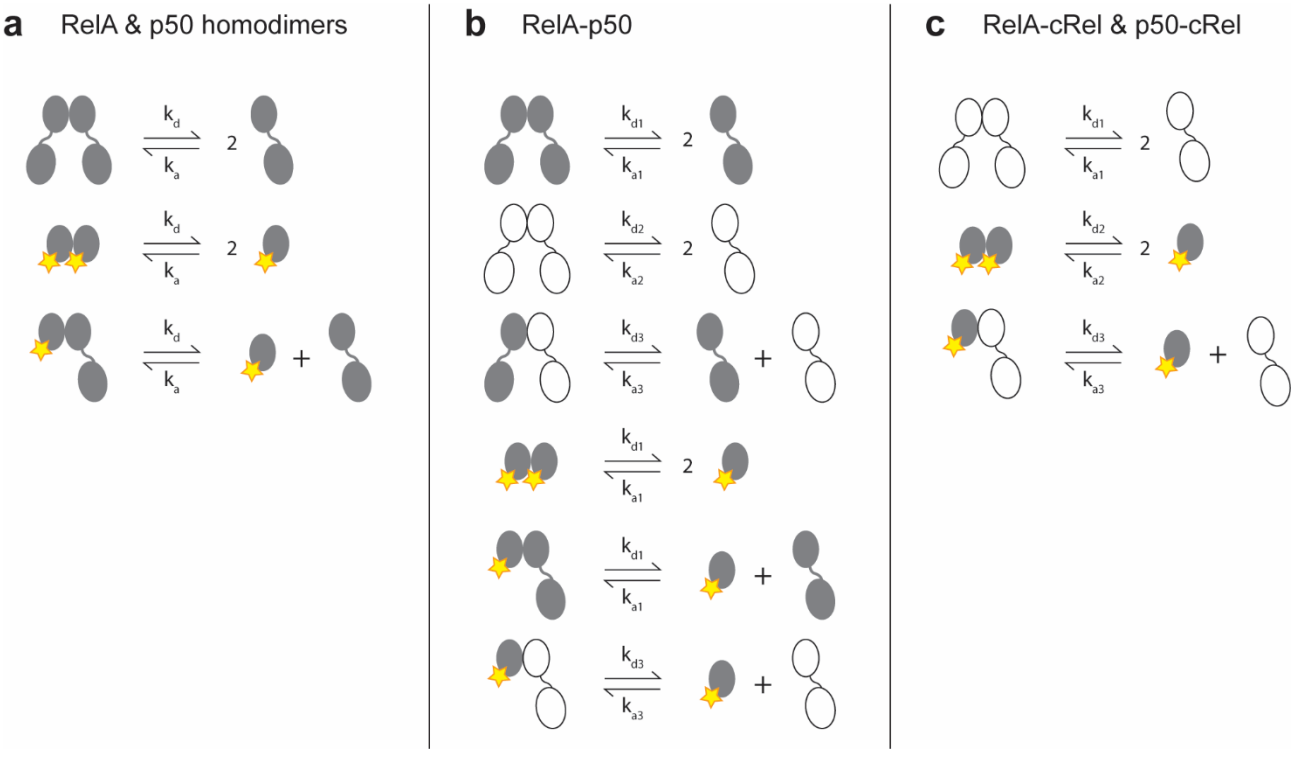


Figure 6.2 Binding reactions in the designed fluorescence anisotropy assays for **(a)** the RelA and p50 homodimer, **(b)** the RelA–p50 heterodimer, and **(c)** the RelA-cRel and p50-c-Rel heterodimer. In the RelA–p50 and the RelA-cRel assays, the dimerization domain of RelA was labeled. In the p50-cRel assay, the dimerization domain of p50 is labeled.

To simulate binding curves, the first step is to calculate equilibrium concentrations of the species, either by solving a system of equations connecting dissociation constants and concentrations or solving the ordinary differential equations (ODEs) describing the reaction kinetics. Here the latter approach is used. For homodimers, there are five ODEs, each of which describes the rate of concentration change of a species.

$$\frac{d}{dt}[N_{FL}] = -2k_a[N_{FL}]^2 + 2k_d[N_{FL}N_{FL}]$$

$$\frac{d}{dt}[N^*_{DD}] = -2k_a[N^*_{DD}]^2 + 2k_d[N^*_{DD}N^*_{DD}]$$

$$\frac{d}{dt}[N_{FL}N_{FL}] = k_a[N_{FL}]^2 - k_d[N_{FL}N_{FL}]$$

$$\frac{d}{dt}[N^*_{DD}N^*_{DD}] = k_a[N^*_{DD}]^2 - k_d[N^*_{DD}N^*_{DD}]$$

$$\frac{d}{dt}[N^*_{DD}N_{FL}] = k_a[N^*_{DD}][N_{FL}] - k_d[N^*_{DD}N_{FL}]$$

N_{FL} and N^*_{DD} denote the full-length monomer and the labeled dimerization domain respectively, whereas k_a and k_d denote the association and dissociation rate constants, which are connected by the dissociation constant K_D with the relationship $k_d = K_D \times k_a$. These five ODEs can be solved numerically to determine the concentration change of each species over time if the rate constants are given.

For protein-protein interactions, the association rate constant was often found in the range of 10^3 $M^{-1}s^{-1}$ to 10^9 $M^{-1}s^{-1}$ [19]. For our purpose here, the value of k_a does not affect equilibrium concentrations and was set to 10^9 $M^{-1}s^{-1}$. The equilibrium dissociation constant, K_D is set to be the same for all three reactions since only the dimerization domains are involved in the dimer interface

and the truncated and full-length NF-κB are expected to have the same dimerization thermodynamics.

This fitting strategy requires an initial guess of K_D. Usually, a good guess can be the concentration of the full-length species (x-axis of the binding curve) where the signal is half saturated. For example, a good initial guess of K_D of the binding curve in **Figure 6.1** is 30 nM. A wide range of K_D values from ten times lower to ten times higher than the initial guess is then generated. For each K_D value, the dissociation rate constant k_d can be converted from $k_d = K_D \times k_a$, and the corresponding five ODEs can be solved to obtain concentrations of the five species. It is important to make sure the time, t, is long enough to reach equilibrium.

Twenty K_D values spanning two orders of magnitude were generated to simulate twenty sets of equilibrium concentrations of the five species in the homodimer assay. For each set of concentrations, a binding curve was made by plotting the bound fraction of the labeled species, $[N^*_{DD}N^*_{FL}] / ([N^*_{DD}]+[N^*_{DD}N^*_{DD}]+[N^*_{DD}N^*_{FL}])$, against the initial concentration of the full-length dimer $[N_{FL}N_{FL}]_0$. The contribution of $N^*_{DD}N^*_{DD}$ was neglected due to its very low abundance.

The K_D value that fits the experimental data the best was determined by comparing the shapes of the twenty simulated binding curves to the experimental binding curve and choosing the one with the highest similarity, as quantified by the Pearson correlation coefficients (Rs). The constant connecting the bound fraction and anisotropy change was also determined by generating twenty values and choosing the one with the highest similarity. Deviation from the best fitting K_D would lead to significant changes in the curve shapes, which is also reflected in the R-squared values (**Figure 6.3)**, and therefore this numerical fitting strategy captures the dimerization affinities in the right order of magnitude. Using this strategy, the K_D values for the RelA homodimer and the p50 homodimer were determined to be 45 nM and 0.54 nM with R-squared values 0.9955 and 0.9621 respectively (**Figure 6.4 a, b**).

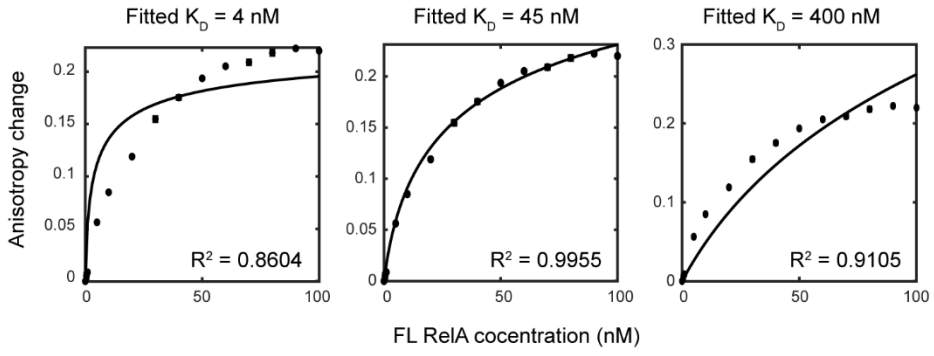


Figure 6.3 Simulated binding curves generated from different K_D values. The best-fitting K_D is determined by choosing the highest R-squared values between the experimental and simulated data.

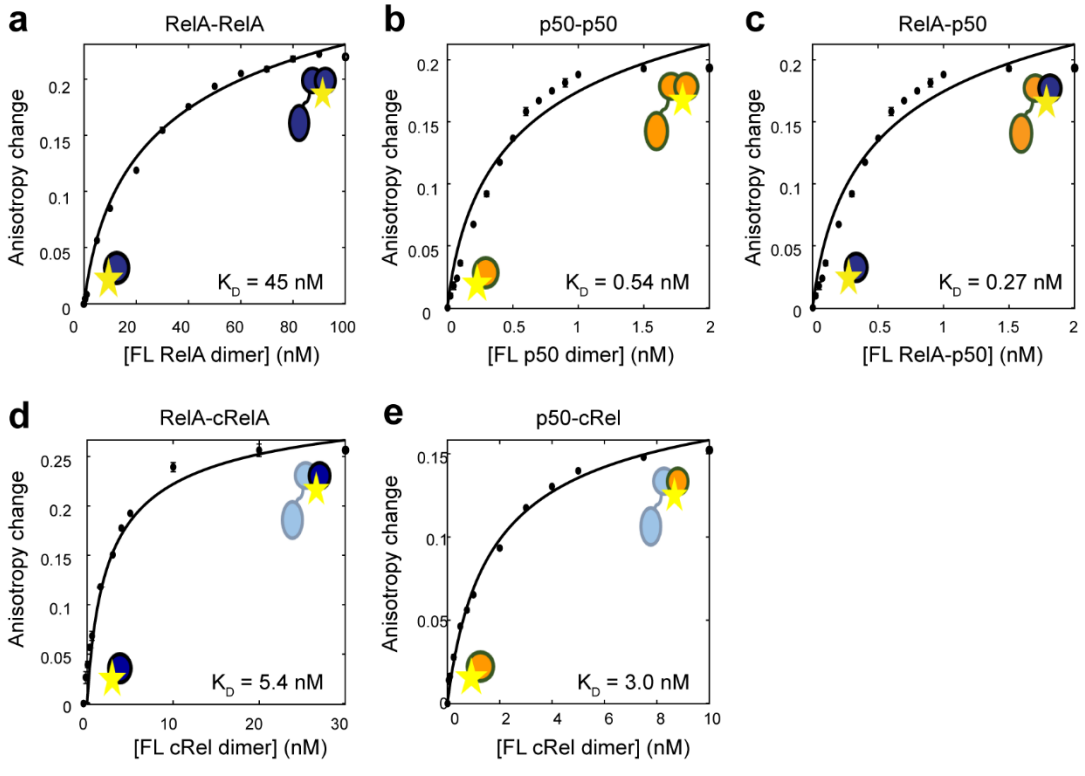


Figure 6.4 Numerical fitting of the fluorescence anisotropy data of NF-κB dimers. **(a)** The labeled RelA dimerization domain was titrated with the full-length RelA dimer. A K_D of 45 nM was determined for the RelA homodimer. **(b)** The labeled p50 dimerization domain was titrated with the full-length p50 dimer. A K_D of 0.54 nM was determined for the p50 homodimer. **(c)** The labeled RelA dimerization domain was titrated with the full-length RelA–p50 dimer. A K_D of 0.27 nM was determined for the RelA–p50 heterodimer. **(d)** The labeled RelA dimerization domain was titrated with the full-length cRel dimer. A K_D of 5.4 nM was determined for the RelA-cRel heterodimer. **(e)** The labeled p50 dimerization domain was titrated with the full-length p50-cRel heterodimer. A K_D of 3.0 nM was determined for the p50-cRel heterodimer.

6.3 Analysis of Fluorescent Anisotropy Assay for Heterodimer Affinities

The same numerical strategy was used to fit the binding data for NF-κB heterodimers RelA–p50, RelA-cRel, and p50-cRel. For the RelA–p50 heterodimer, there are nine species in the binding assay (**Figure 6.2 b**), each of which is related to an ODE.

$$\frac{d}{dt}[RelA_{FL}] = -2k_{a1}[RelA_{FL}]^2 + 2k_{d1}[RelA_{FL}:RelA_{FL}]$$

$$\frac{d}{dt}[RelA^*_{DD}] = -2k_{a1}[RelA^*_{DD}]^2 + 2k_{d1}[RelA^*_{DD}:RelA^*_{DD}]$$

$$\frac{d}{dt}[p50] = -2k_{a2}[p50]^2 + 2k_{d2}[p50:p50]$$

$$\frac{d}{dt}[RelA_{FL}:RelA_{FL}] = k_{a1}[RelA_{FL}]^2 - k_{d1}[RelA_{FL}:RelA_{FL}]$$

$$\frac{d}{dt}[RelA^*_{DD}:RelA^*_{DD}] = k_{a1}[RelA^*_{DD}]^2 - k_{d1}[RelA^*_{DD}:RelA^*_{DD}]$$

$$\frac{d}{dt}[RelA^*_{DD}:RelA_{FL}] = k_{a1}[RelA^*_{DD}][RelA_{FL}] - k_{d1}[RelA^*_{DD}:RelA_{FL}]$$

$$\frac{d}{dt}[p50:p50] = k_{a2}[p50]^2 - k_{d2}[p50:p50]$$

$$\frac{d}{dt}[RelA_{FL}:p50] = k_{a3}[RelA_{FL}][p50] - k_{d3}[RelA_{FL}:p50]$$

$$\frac{d}{dt}[RelA^*_{DD}:p50] = k_{a3}[RelA^*_{DD}][p50] - k_{d3}[RelA^*_{DD}:p50]$$

There were three sets of rate constants, k_{a1} and k_{d1}, k_{a2} and k_{d2}, k_{a3} and k_{d3}, for the dimerization kinetics of the RelA homodimer, the p50 homodimer, and the RelA–p50 heterodimer respectively. As was the case for the homodimers, the values of the association rate constants do not affect the equilibrium properties and were set to 10^9 $M^{-1}s^{-1}$ in the analysis. $RelA_{FL}$ and $RelA^*_{DD}$ denote the

full-length RelA monomer and the RelA dimerization domain respectively, and they have the same binding thermodynamics. There are three dissociation constants for the RelA homodimer, p50 homodimer, and the RelA–p50 heterodimer. The first two were determined by fitting the homodimer data in the previous section and the latter is to be determined here.

As was in the case of homodimers, twenty K_D values for RelA–p50 were generated ranging from ten times lower to ten times higher than the initial guess, which is the x value of the experimental binding curve corresponding to half saturation. For each K_D value, a binding curve was simulated and compared to the experimental data by calculating correlation coefficients, Rs. The best fitted R-squared value for the RelA–p50 heterodimer is 0.9700 and the determined K_D is 0.27 nM (**Figure 6.4 c**).

For the RelA-cRel and p50-cRel heterodimers, labeled RelA or p50 dimerization domains were titrated by the full-length cRel dimer respectively. As shown in **Figure 6.2 c**, five species and three reactions are involved. Five ODEs similar to the ODEs for homodimers can be written down. However, for these c-Rel heterodimers, there are three K_D values governing the reactions in the assay. The K_D values for RelA-cRel or p50-cRel are to be determined; the K_D for the p50 homodimer is known from previous results; the K_D for cRel homodimer is unknown due to lack of experimental data.

The same numerical fitting strategy was used to estimate the K_D for RelA:cRel, except that this time a wide range of K_D for the cRel homodimer was used to see which value gives the best correlation coefficient between the simulated and the experimental curves and to test the sensitivity of the fitting results to the cRel homodimer K_D value. Similarly, a wide range of the RelA homodimer K_D was also used to test if the previously determined K_D gives the highest similarity between simulated and experimental curves and how sensitive is the fitting results to this K_D value (**Table 6.1**). The same procedure was repeated for the p50-cRel heterodimer (**Table 6.2**).

Fortunately, the fitted K_D values for cRel-containing heterodimers were not sensitive to the K_D values of the homodimers, with about a five-fold change in response to 5 orders of magnitude change in the homodimer K_D. Using this strategy, the K_D values of RelA-cRel and p50-cRel were estimated to be 5.4 nM and 3.0 nM respectively (**Figure 6.4 d, e**).

Table 6.1. Fitted K_D and R^2 values for RelA-cRel dimerization with different K_D values for the RelA and cRel homodimers.

K_D^{AA} \ K_D^{CC}	**0.1 nM**	**1 nM**	**10 nM**	**100 nM**	**1000 nM**
0.5 nM	0.71 nM $R^2 = 0.9500$	1.72 nM $R^2 = 0.9613$	3.53 nM $R^2 = 0.9858$	4.54 nM $R^2 = 0.9960$	4.54 nM $R^2 = 0.9971$
5 nM	0.91 nM $R^2 = 0.9471$	2.32 nM $R^2 = 0.9614$	4.34 nM $R^2 = 0.9832$	5.35 nM $R^2 = 0.9943$	5.15 nM $R^2 = 0.9954$
45 nM best fitted	0.91 nM $R^2 = 0.9448$	2.32 nM $R^2 = 0.9594$	4.54 nM $R^2 = 0.9832$	5.35 nM $R^2 = 0.9938$	5.35 nM $R^2 = 0.9953$
500 nM	0.91 nM $R^2 = 0.9445$	2.32 nM $R^2 = 0.9591$	4.54 nM $R^2 = 0.9830$	5.35 nM $R^2 = 0.9937$	5.35 nM $R^2 = 0.9953$

Table 6.2. Fitted K_D and R^2 values for RelA–p50 dimerization with different K_D values for the RelA and p50 homodimers.

K_D^{p50p50} \ K_D^{CC}	**0.1 nM**	**1 nM**	**10 nM**	**100 nM**	**1000 nM**
0.05 nM	0.30 nM $R^2 = 0.9834$	0.91 nM $R^2 = 0.9925$	1.31 nM $R^2 = 0.9942$	1.51 nM $R^2 = 0.9907$	1.51 nM $R^2 = 0.9899$
0.54 nM best fitted	0.71 nM $R^2 = 0.9849$	1.92 nM $R^2 = 0.9919$	2.93 nM $R^2 = 0.9933$	2.93 nM $R^2 = 0.9925$	2.93 nM $R^2 = 0.9917$
5 nM	1.11 nM $R^2 = 0.9862$	2.32 nM $R^2 = 0.9913$	3.53 nM $R^2 = 0.9937$	3.74 nM $R^2 = 0.9920$	3.53 nM $R^2 = 0.9924$
50 nM	1.11 nM $R^2 = 0.9857$	2.52 nM $R^2 = 0.9913$	3.74 nM $R^2 = 0.9935$	3.74 nM $R^2 = 0.9927$	3.53 nM $R^2 = 0.9930$

6.4 Analysis of Fluorescent Anisotropy Assay for IκB Binding Affinities

The binding affinities between IκB proteins and NF-κB dimers were measured using the same fluorescent anisotropy assay[17]. To measure the affinities of IκB proteins binding to NF-κB, labeled NF-κB dimers are held at fixed concentrations and IκB proteins are varied in concentrations. Experimental binding curves for IκBα binding to RelA- and cRel-containing dimers, IκBβ binding to RelA- and cRel-containing dimers, and IκBε binding to RelA-, p50- and cRel-containing dimers were obtained. However, the above-mentioned numerical fitting strategy lead to sub-picomolar binding affinities, which did not seem physiologically correct. Therefore, K_D values for IκBs binding to NF-κB dimers were estimated by fitting the binding curve with the equation $\text{anisotropy} = \frac{\text{anisotropy}_{max} \times [\text{unlabled dimer}]}{K_D + [\text{unlabeled dimer}]}$, which is for fitting conventional bi-

molecular reactions[18]. The estimated K_D values were compared with previous results of K_D values measured by SPR[20].

Binding of the RelA homodimer to each IκB showed that IκBα bound with a K_D of 47 ± 6 pM. For comparison, the binding affinity measured by SPR was 54 ± 8 pM. IκBβ bound with a K_D of 25 ± 3 pM, and IκBε bound with a K_D of 170 ± 10 pM (**Figure 6.5 a**). Thus, RelA homodimer slightly prefers IκBβ over IκBα, and is strongly inhibited by all three IκBs as indicated in their pM range affinities.

IκBα bound to RelA homodimer dimerization domain (residues 190 to 325), $RelA_{dd}$-$RelA_{dd}$, with a K_D of 280 ± 31 pM, showing that the N-terminal domains contribute about 6-fold to the IκBα binding affinity similar to what was previously reported[20]. IκBβ bound to the dimerization domain construct, $RelA_{dd}$-$RelA_{dd}$, with a K_D of 72 ± 6 pM, a difference of three-fold (**Figure 6.5 b**).

RelA19-325-p50 heterodimer gave a K_D for IκBα of 42 ± 9 pM, very similar to the value of 39 pM measure by SPR and reported previously[20], a K_D for IκBβ of 890 ± 160 pM, again similar to that measured by SPR of 1100 ± 50 pM. RelA–p50 heterodimer binds to IκBε with a K_D of 800 ± 110 pM (**Figure 6.5 c, Table 6.3**). Thus, the RelA–p50 heterodimer dramatically prefers IκBα. The binding affinity of IκBα was within error the same for RelA homodimers (47 pM) as for RelA–p50 heterodimers (42 pM). IκBε did not bind particularly tightly to either the RelA homodimer or the RelA–p50 heterodimer. Binding of the p50 homodimer to IκBs was not measured because previous reports indicated that p50 homodimer binds weakly, if at all, to IκBα[21].

Two NF-κB dimers containing cRel were also characterized for their binding preferences among the IκBs. Dimers containing cRel bound most tightly to IκBε. RelA-cRel, p50-cRel, and cRel homodimer bound to IκBε with affinities of 98 ± 5 pM, 60 ± 10 pM, and 100 ± 8 pM, respectively (**Figure 6.5 d, e, Table 6.3**).

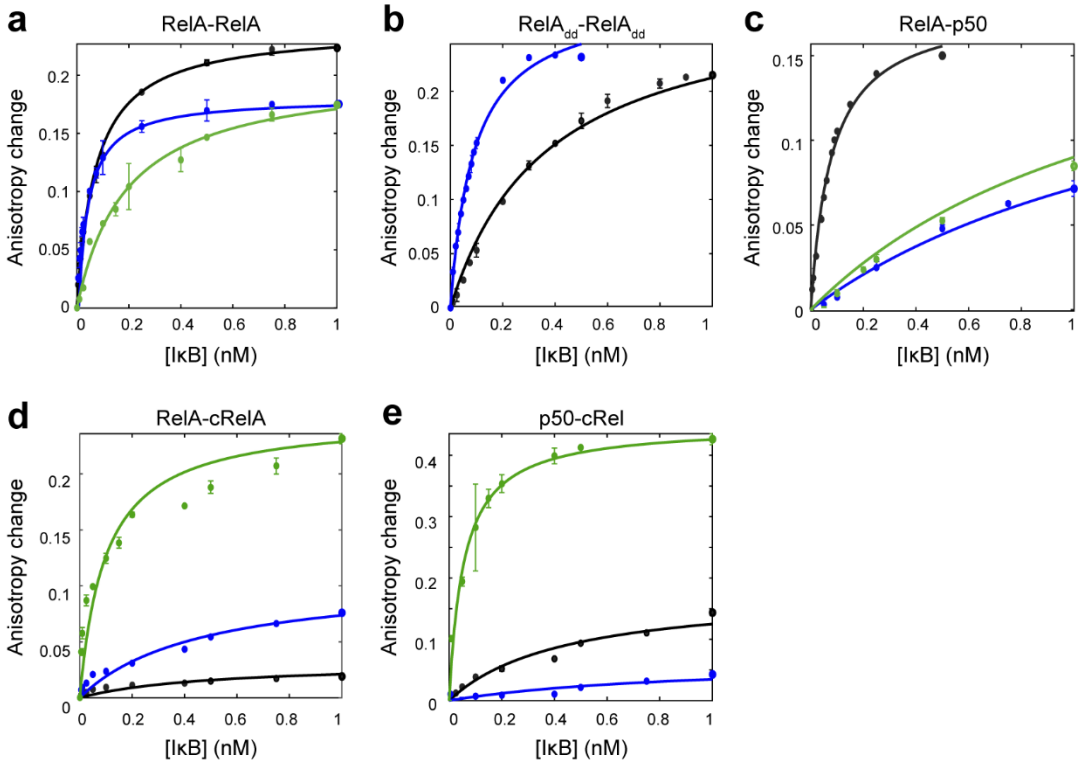


Figure 6.5 Binding of IκBs to labeled NFκB dimers by fluorescence anisotropy assays. Dots: experimental data; Lines: fitting curves. **(a)** Binding of the RelA homodimer to each IκB showed that IκBα (black) bound with a K_D of 47 ± 6 pM, IκBβ (blue) bound with a K_D of 25 ± 3 pM, and IκBε (green) bound with a K_D of 170 ± 10 pM. **(b)** Binding of $RelA_{dd}$ to each IκB showed that IκBα bound to the dimerization domain construct with a K_D of 280 ± 31 pM and IκBβ bound with a K_D of 72 ± 6 pM. **(c)** Binding of the RelA–p50 heterodimer (500 pM) gave a K_D for IκBα of 42 ± 9 pM, a K_D for IκBβ of 890 ± 160 pM, and a K_D for IκBε of 800 ± 110 pM. **(d)** Binding of the RelA-cRel heterodimer (5 nM) gave a K_D for IκBα of 500 ± 90 pM, a K_D for IκBβ of 460 ± 50 pM, and a K_D for IκBε of 98 ± 5 pM. **(e)** p50-cRel (5 nM) bound to IκBα with a K_D of 440 ± 50 pM, to IκBβ with a K_D of 920 ± 150 pM, and to IκBε with a K_D of 60 ± 10 pM.

Table 6.3. Binding affinities of NF-κB dimers as dissociation constants.

NF-κB	Dimer partner	K_D (pM)	IκBα (pM)	IκBβ (pM)	IκBε (pM)
RelA	**RelA**	45,000 +/- 6000	47 +/- 6 54 +/- 8 (SPR)	25 +/- 3 70 +/- 4 (SPR)	170 +/- 10
	p50	270 +/- 20	42 +/- 9 39 +/- 3 (SPR)	890 +/- 160 1100 +/- 50 (SPR)	800 +/- 110
	cRel	5400 +/-900	500 +/- 90	460 +/- 50	98 +/- 5
p50	**p50**	540 +/- 60	ND	ND	810 +/- 30
	cRel	3000 +/- 300	440 +/- 50	920 +/- 150	60 +/- 10
cRel	**cRel**	ND	ND	ND	100 +/- 8

6.5 Subunit Exchange of NF-κB Dimers

The fluorescence anisotropy results that showed both RelA–p50 and IκBα binding affinities in the pM range suggested subunit exchange between dimers may be limited in the presence of IκBα. To investigate the rates of subunit exchange of RelA–p50 and how IκBα may affect it, I carried out simulations by numerically solving kinetics equations.

RelA–p50 can re-equilibrate to form two corresponding homodimers. The faster the two subunits in a dimer dissociate and re-associate, the faster they can exchange and form other dimers. Association and dissociation rate constants are connected by the K_D we measured but their values are unknown. The typical association rate constant for protein-protein interactions was in the range of 10^3 $M^{-1}s^{-1}$ to 10^9 $M^{-1}s^{-1}$ [19]. Given rate constants and initial concentrations, the kinetic process of

RelA–p50 subunit exchange can be simulated by numerically solving the ODEs described in **Section 6.2**.

I first tested whether subunit exchange could occur in cells, where the estimated amounts of RelA, p50, and IκBα are equally 350 nM[15]. The maximal NF-κB subunit association rate was initially set to the diffusion limit $ka_1 = ka_2 = ka_3 = 10^9\ M^{-1}s^{-1}$ as had been done previously[15]. The resulting dissociation rates from the measured K_D values of each dimer were calculated; $k_{d1} = 18.6\ s^{-1}$; $k_{d2} = 0.3\ s^{-1}$; $k_{d3} = 0.4\ s^{-1}$. The numerical result showed that under these conditions, RelA–p50 can re-equilibrate and form the two homodimers within only 5 seconds (**Figure 6.6 a**). If the association rates were 100-fold slower, the dimers would still re-equilibrate in 5 min. Even if the association rates were only $10^5\ M^{-1}s^{-1}$, the dimers would re-equilibrate in 5 hr.

I next tested whether such fast subunit exchange would be affected by IκBα, which is always present in resting cells to keep NF-κB in the cytoplasm and inactive. Rate equations including IκBα binding to RelA–p50 were added to the simulation. The result showed that, starting with RelA–p50-IκBα, subunits never exchange over periods of at least 24 hrs. (**Figure 6.6 b**). Because of the high affinity of IκBα binding to RelA–p50, even when maximal association rates are assumed, the dimer exchange of RelA–p50 is totally blocked by IκBα.

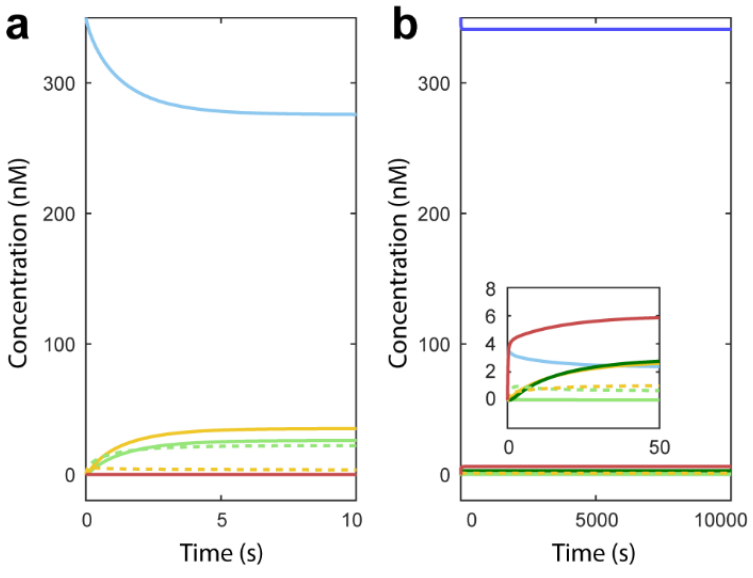


Figure 6.6 Numerical solutions of the kinetic equations show that RelA–p50 can only undergo subunit exchange when free from IκBα. **(a)** In the absence of IκBα, RelA–p50 can re-equilibrate to form RelA-RelA and p50-p50 homodimers within 5 seconds when assuming diffusion-limited association rate constants. At equilibrium, the dominant species is RelA–p50 due to its high dimerization affinity. **(b)** When bound to IκBα, RelA–p50 does not undergo subunit exchange. Light blue: ReA-p50; Dar blue: RelA–p50-IκBα; Solid light green: the RelA homodimer; Solid yellow: the p50 homodimer; Dark green: RelA-RelA-IκBα; Dotted green: the RelA monomer; Dotted yellow: the p50 monomer; Red: IκBα.

6.6 *In vivo* Half-Life of NF-κB RelA–p50

Given the widely varying reported affinities of IκBα binding to RelA–p50 from 39 pM determined by SPR[20] and 42 pM in this work to 1 nM determined by gel shift[1,15,16], I tested whether the ~40 pM affinity could be consistent with the intracellular half-life of RelA–p50-IκBα. Free IκBα degrades rapidly *in vivo* with a short half-life of 5 to 10 minutes, while NF-κB-bound IκBα is stable with a half-life longer than 48 hours[9]. The concentration change of RelA–p50-IκBα was expressed with the simple rate equations:

$$\frac{d}{dt}[A:50] = -k_a[A:50][\alpha] + k_d[A:50:\alpha]$$

$$\frac{d}{dt}[\alpha] = -k_a[A:50][\alpha] + k_d[A:50:\alpha] - k[\alpha]$$

$$\frac{d}{dt}[A:50:\alpha] = k_a[A:50][\alpha] - k_d[A:50:\alpha]$$

The degradation rate constant of free IκBα, k, is 0.002 s^{-1}.[9] Both limiting cases of fast and slow association between RelA–p50 and IκBα were tested. The association rate constant, k_a, was set to be 10^9 M^{-1} s^{-1} for the fast case and 10^4 M^{-1} s^{-1} for the slow case; the dissociation rate constant k_d was calculated from $K_D \times k_a$. The results showed that in both cases the 40 pM affinity greatly stabilizes the protein complex and explains its > 48 hours *in vivo* half-life. If IκBα bound to RelA–p50 with an affinity of only 1 nM, a substantial portion of free IκBα was present, and its rapid degradation would drive a decrease in the RelA–p50-IκBα concentration (**Figure 6.7**).

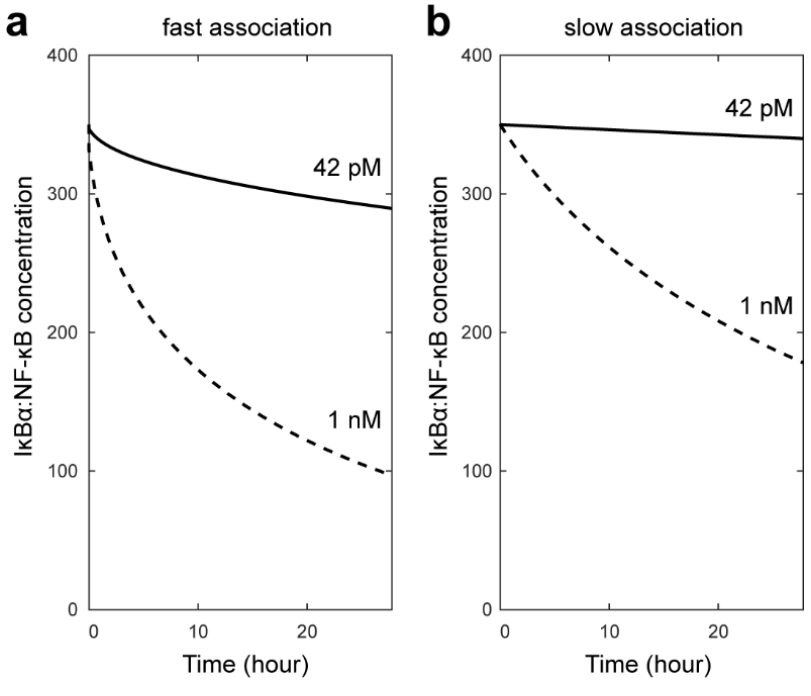


Figure 6.7 Numerical solutions of kinetic equations show that the long *in vivo* half-life of NFκB bound IκBα can only be explained by tight binding (42 pM). **(a)** If the NFκB-IκBα association rate were diffusion-controlled (10^9 $M^{-1}s^{-1}$), the 42 pM binding affinity shows that the half-life of bound IκBα is longer than 24 hrs, consistent with in vivo observations. **(b)** If the NFκB-IκBα association rate were slower (10^4 $M^{-1}s^{-1}$), the half-life of bound IκBα is even slower, but in both cases the 1nM binding affinity results in dissociation and subsequent degradation of IκBα.

6.7 Population Shift of NF-κB Dimers by IκB

The fluorescence anisotropy experiments showed that each IκB has a binding affinity in the pM range for at least one NF-κB dimer. IκBα binds preferentially to RelA–p50 and RelA homodimer; IκBβ binds preferentially to RelA-RelA; IκBε binds preferentially to c-Rel containing dimers. To investigate how IκB preferential binding can shift the equilibrium populations of NF-κB dimers, I carried out numerical simulations by solving the kinetic equations. The simplest case was considered, where the dominant molecules are RelA, p50, IκBα, and IκBβ, as in mouse embryonic fibroblasts. Our results showed that, in the absence of IκBα, IκBβ can shift the equilibrium population from RelA–p50 to RelA-RelA-IκBβ (**Figure 6.8**), consistent with previous experimental data[15]. The preference of IκBβ toward RelA-RelA stabilizes this dimer despite the weaker dimerization affinity of RelA-RelA (45 nM) compared to that of RelA–p50 (270 pM). Interestingly, in the presence of IκBα, IκBβ is not able to shift the dimer equilibrium population, because all RelA–p50 is bound tightly to IκBα and does not undergo dimer exchange.

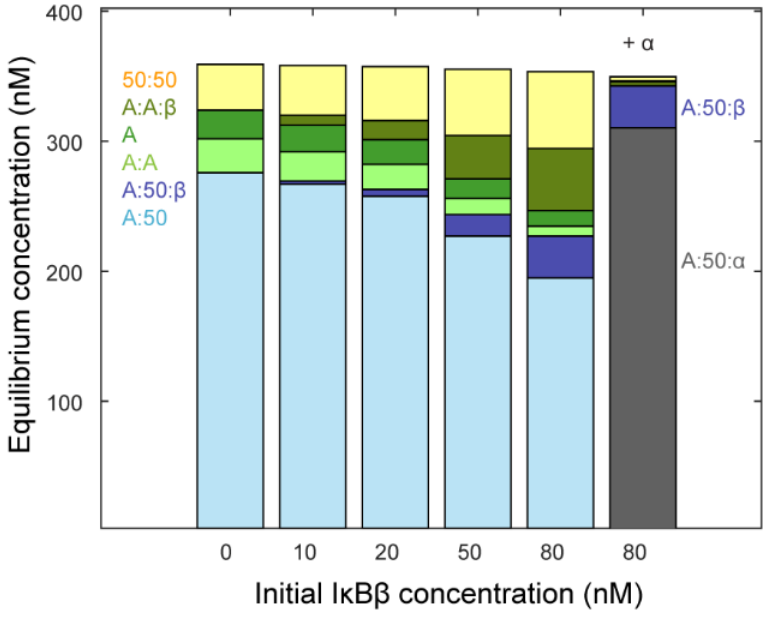


Figure 6.8 Numerical simulations of the dimer population shift upon increasing IκBβ concentration. In the absence of IκBα and of IκBβ, most of the RelA is bound to p50 (light blue) although some is also present as RelA homodimer (light green). As the concentration of IκBβ increases from 0 to 80 nM, the equilibrium population of free RelA–p50 (light blue), shifts towards IκBβ-bound RelA–p50, denoted A:50:β (dark blue). As the concentration of IκBβ increases, the population of free RelA homodimer is almost completely depleted and The IκBβ bound RelA homodimer (denoted A:A:β) increases over that of A:50:β due to the preferential binding of IκBβ to the RelA homodimer. However, in the presence of IκBα, (right-most bar) RelA–p50 is tightly bound to IκBα (grey) and thus the population shift is blocked.

6.8 Discussion

Most *in vivo* studies of NF-κB monitor fluorescently labeled RelA, but RelA can exist in homodimers or RelA–p50 or RelA-cRel heterodimers, and p50-cRel would not be observed. Each NF-κB dimer has been shown to have distinct functions, with, for example, RelA–p50 being important in inflammation and p50-cRel being important in cancer[22]. It therefore seemed important to establish the dimer binding affinities so that complete systems models of all dimers and their inhibitors could be generated, and predictions could be made about when dimer exchange might occur allowing an understanding of how different dimers prevail in different cell types.

A fluorescence anisotropy method was developed to measure the dimer equilibrium binding affinities under the lower concentration expected in cells. By mixing and matching labeled and unlabeled dimers with and without the DNA-binding domains, we measured the K_D of each dimer formation. Our results showed that the RelA–p50 dimer has the highest affinity, with a K_D of 0.27 nM. The p50 homodimer binds nearly as tightly with a K_D of 5.4 nM. Dimers containing cRel have affinities in the low nM range, and the RelA homodimer is weakest, with an affinity of 45 nM. These dimer binding affinities are much stronger than previously reported based on gel-shift and analytical ultracentrifugation experiments, although the order of affinities is the same as previously reported. Why does it matter that the dimerization affinities are stronger than have been used in previous models? First, because strong dimer affinities imply that essentially no NF-κB monomers will be present in the cell. Even the weakest RelA homodimer would have only 2.2% of the molecules in the monomeric form at cellular concentrations[15]. This result should put to rest any further speculation that monomeric NF-κB may be present and functional in cells.

Earlier reports of the binding affinity of IκBα for RelA–p50 used gel-shift assays and measured a K_D of 1nM. We reported a binding affinity of 39 pM[20], which was more consistent with the long intracellular half-life of this complex, but many in the field still deferred to the earlier report. Here, we not only re-confirm the picomolar binding affinity of IκBα for RelA–p50, but we also report picomolar binding affinities for other IκB-NF-κB complexes. This means that in the cytoplasm of resting cells, all NF-κB dimers will be bound to inhibitors, and essentially no dimer exchange is expected to occur unless one of the IκBs is knocked-out or degraded. Importantly, dimer stabilization by its preferred IκB provides a simple and precise explanation for the short half-lives of both IκBs and NF-κBs in the absence of their binding partner, but long half-lives when bound. Indeed, the melting temperature of the RelA–p50-IκBα complex was at least 30 °C higher than that of NF-κB alone[23] and 40 °C higher than IκBα alone[24].

We found that in most cases, each dimer has a preferred inhibitor; RelA homodimer prefers IκBβ, cRel dimers prefer IκBε, and RelA–p50 prefers IκBα. Interestingly, this preferential binding correlates with which NF-κB induces transcription of which inhibitor genes. RelA is required for induction of the IκBα gene whereas cRel induces only the IκBε gene[5].

The only exception to exclusivity is that IκBα binds almost as tightly to RelA homodimers as to RelA–p50 heterodimers. This is a very interesting result in light of our observation that during stimulation, IκBα is degraded releasing the inhibition of RelA–p50, which then enters the nucleus and activates transcription of new IκBα. This IκBα then enters the nucleus and "strips" NF-κB off the DNA and the NF-κB-IκBα complex is rapidly exported from the nucleus[12]. Single-cell imaging revealed that nearly all the RelA was rapidly exported from the nucleus, and now we know that is because when IκBα enters the nucleus, it will remove both RelA–p50 heterodimers and RelA homodimers. If IκBβ is already present in the nucleus, as has been shown previously[25], our results support the idea that it would preferentially capture RelA homodimers upon stimulation, but only until the newly synthesized IκBα enters the nucleus, at which point all of the DNA-bound RelA-containing dimers would be stripped off the DNA by IκBα and exported out of the nucleus. Finally, it was reported recently that cRel-containing dimers sometimes replace RelA containing dimers at κB site-containing promoters and promote transcription at later times[26]. Our results suggest that this is due to the preferential molecular stripping of RelA-containing dimers by IκBα leaving the cRel-containing dimers behind in the nucleus to continue transcription of late-onset genes.

Acknowledgments

Chapter 6, in part, appears in "Exclusivity and Compensation in NF-κB Dimer Distributions and IκB Inhibition." by Ramsey, K. M.*; Chen, W.*; Marion J. D.*; Bergqvist, S.;

Komives, E. A., published in *Biochemistry* **2019**, *58* (21), 2555-2563. The dissertation author is one of the co-first authors denoted by asterisks.

References

1. Liou, H. C.; Sha, W. C.; Scott, M. L.; Baltimore, D., Sequential induction of NF-kappa B/Rel family proteins during B-cell terminal differentiation. *Mol Cell Biol* **1994,** *14* (8), 5349-59.

2. Hoffmann, A.; Leung, T. H.; Baltimore, D., Genetic analysis of NF-kappaB/Rel transcription factors defines functional specificities. *EMBO J* **2003,** *22* (20), 5530-9.

3. Pohl, T.; Gugasyan, R.; Grumont, R. J.; Strasser, A.; Metcalf, D.; Tarlinton, D.; Sha, W.; Baltimore, D.; Gerondakis, S., The combined absence of NF-kappa B1 and c-Rel reveals that overlapping roles for these transcription factors in the B cell lineage are restricted to the activation and function of mature cells. *Proc Natl Acad Sci U S A* **2002,** *99* (7), 4514-9.

4. Gilmore, T. D.; Kalaitzidis, D.; Liang, M. C.; Starczynowski, D. T., The c-Rel transcription factor and B-cell proliferation: a deal with the devil. *Oncogene* **2004,** *23* (13), 2275-86.

5. Alves, B. N.; Tsui, R.; Almaden, J.; Shokhirev, M. N.; Davis-Turak, J.; Fujimoto, J.; Birnbaum, H.; Ponomarenko, J.; Hoffmann, A., IkappaBepsilon is a key regulator of B cell expansion by providing negative feedback on cRel and RelA in a stimulus-specific manner. *J Immunol* **2014,** *192* (7), 3121-32.

6. Baeuerle, P. A.; Baltimore, D., A 65-kappaD subunit of active NF-kappaB is required for inhibition of NF-kappaB by I kappaB. *Genes Dev* **1989,** *3* (11), 1689-98.

7. Hoffmann, A.; Levchenko, A.; Scott, M. L.; Baltimore, D., The IkappaB-NF-kappaB signaling module: temporal control and selective gene activation. *Science* **2002,** *298* (5596), 1241-5.

8. Hoffmann, A.; Baltimore, D., Circuitry of nuclear factor kappaB signaling. *Immunol Rev* **2006,** *210*, 171-86.

9. O'Dea, E. L.; Barken, D.; Peralta, R. Q.; Tran, K. T.; Werner, S. L.; Kearns, J. D.; Levchenko, A.; Hoffmann, A., A homeostatic model of IkappaB metabolism to control constitutive NF-kappaB activity. *Mol Syst Biol* **2007,** *3*, 111.

10. Alverdi, V.; Hetrick, B.; Joseph, S.; Komives, E. A., Direct observation of a transient ternary complex during IkappaBalpha-mediated dissociation of NF-kappaB from DNA. *Proc Natl Acad Sci U S A* **2014,** *111* (1), 225-30.

11. Bergqvist, S.; Alverdi, V.; Mengel, B.; Hoffmann, A.; Ghosh, G.; Komives, E. A., Kinetic enhancement of NF-kappaBxDNA dissociation by IkappaBalpha. *Proc Natl Acad Sci U S A* **2009,** *106* (46), 19328-33.

12. Dembinski, H. E.; Wismer, K.; Vargas, J. D.; Suryawanshi, G. W.; Kern, N.; Kroon, G.; Dyson, H. J.; Hoffmann, A.; Komives, E. A., Functional importance of stripping in NFkappaB signaling revealed by a stripping-impaired IkappaBalpha mutant. *Proc Natl Acad Sci U S A* **2017,** *114* (8), 1916-1921.

13. Potoyan, D. A.; Wolynes, P. G., Stochastic dynamics of genetic broadcasting networks. *Phys Rev E* **2017,** *96* (5-1), 052305.

14. Potoyan, D. A.; Zheng, W.; Komives, E. A.; Wolynes, P. G., Molecular stripping in the NF-kappaB/IkappaB/DNA genetic regulatory network. *Proc Natl Acad Sci U S A* **2016,** *113* (1), 110-5.

15. Tsui, R.; Kearns, J. D.; Lynch, C.; Vu, D.; Ngo, K. A.; Basak, S.; Ghosh, G.; Hoffmann, A., IkappaBbeta enhances the generation of the low-affinity NFkappaB/RelA homodimer. *Nat Commun* **2015,** *6*, 7068.

16. Malek, S.; Huxford, T.; Ghosh, G., Ikappa Balpha functions through direct contacts with the nuclear localization signals and the DNA binding sequences of NF-kappaB. *J Biol Chem* **1998,** *273* (39), 25427-35.

17. Ramsey, K. M.; Chen, W.; Marion, J. D.; Bergqvist, S.; Komives, E. A., Exclusivity and Compensation in NFkappaB Dimer Distributions and IkappaB Inhibition. *Biochemistry* **2019,** *58* (21), 2555-2563.

18. Pollard, T. D., A guide to simple and informative binding assays. *Mol Biol Cell* **2010,** *21* (23), 4061-7.

19. Schreiber, G.; Haran, G.; Zhou, H. X., Fundamental aspects of protein-protein association kinetics. *Chem Rev* **2009,** *109* (3), 839-60.

20. Bergqvist, S.; Croy, C. H.; Kjaergaard, M.; Huxford, T.; Ghosh, G.; Komives, E. A., Thermodynamics reveal that helix four in the NLS of NF-kappaB p65 anchors IkappaBalpha, forming a very stable complex. *J Mol Biol* **2006,** *360* (2), 421-34.

21. Beg, A. A.; Ruben, S. M.; Scheinman, R. I.; Haskill, S.; Rosen, C. A.; Baldwin, A. S., Jr., I kappa B interacts with the nuclear localization sequences of the subunits of NF-kappa B: a mechanism for cytoplasmic retention. *Genes Dev* **1992,** *6* (10), 1899-913.

22. Gilmore, T. D.; Gerondakis, S., The c-Rel Transcription Factor in Development and Disease. *Genes Cancer* **2011,** *2* (7), 695-711.

23. Ramsey, K. M.; Dembinski, H. E.; Chen, W.; Ricci, C. G.; Komives, E. A., DNA and IkappaBalpha Both Induce Long-Range Conformational Changes in NFkappaB. *J Mol Biol* **2017,** *429* (7), 999-1008.

24. Croy, C. H.; Bergqvist, S.; Huxford, T.; Ghosh, G.; Komives, E. A., Biophysical characterization of the free IkappaBalpha ankyrin repeat domain in solution. *Protein Sci* **2004,** *13* (7), 1767-77.

25. Rao, P.; Hayden, M. S.; Long, M.; Scott, M. L.; West, A. P.; Zhang, D.; Oeckinghaus, A.; Lynch, C.; Hoffmann, A.; Baltimore, D.; Ghosh, S., IkappaBbeta acts to inhibit and activate gene expression during the inflammatory response. *Nature* **2010,** *466* (7310), 1115-9.

26. Smale, S. T., Hierarchies of NF-kappaB target-gene regulation. *Nat Immunol* **2011,** *12* (8), 689-94.