



University of  
Massachusetts  
Amherst

## Combining Simulation and the MspA Nanopore to Study p53 Dynamics and Interactions

Item Type	Thesis (Open Access)
Authors	Schultz, Samantha A
DOI	<a href="https://doi.org/10.7275/35904425.0">10.7275/35904425.0</a>
Download date	2026-05-20 17:23:46
Link to Item	<a href="https://hdl.handle.net/20.500.14394/33034">https://hdl.handle.net/20.500.14394/33034</a>

**Combining Simulation and the MspA Nanopore to Study p53 Dynamics and Interactions**

A Thesis Presented

by

SAMANTHA A. SCHULTZ

Submitted to the Graduate School of the  
University of Massachusetts Amherst in partial fulfillment  
of the requirements for the degree of

MASTER OF SCIENCE

September 2023

Molecular and Cellular Biology Program

# **Combining Simulation and the MspA Nanopore to Study p53 Dynamics and Interactions**

A Thesis Presented

by

SAMANTHA A. SCHULTZ

Approved as to style and content by:

---

Dr. Jianhan Chen, Co-Chair

---

Dr. Min Chen, Co-Chair

---

Dr. Scott Garman, Member

---

Dr. Margaret Stratton, Graduate Program Director

Molecular and Cellular Biology

## ACKNOWLEDGMENTS

I feel lucky to have many people to thank as I conclude my time at UMass. First, to my friends, family, and partner who have been my support system, thank you for believing in me and encouraging me always. I couldn't have made it through this year without all of you by my side in my most exciting times, overwhelming moments, and everything in between.

To my lab mates in the Jianhan Chen and Min Chen labs, I am grateful for the time I got to spend getting to know all of you and I will always think back fondly on our fun memories in and outside of the lab. I feel fortunate to have worked with such knowledgeable and capable peers, and all of this work would not have been possible without your guidance and expertise. Specifically, I would like to thank Erik Nordquist, Yumeng Zhang, Joshua Foster, and Kaitlyn Gilliam for being incredible mentors to me throughout my time as an undergraduate and graduate student.

Lastly, to my advisors Dr. Jianhan Chen and Dr. Min Chen, I cannot thank you enough for all you have done for me over the past few years. Both of you have been invaluable mentors to me and I will never forget the lessons I've learned from you. Your passion for science and the dedication and generosity you show to your students inspires me. Reflecting back on when I started in the lab as an undergraduate, I have grown so much as a critical thinker and researcher, and I have both of you to thank for that progress. Thank you for guiding me, challenging me, and believing in my abilities as a scientist.

## ABSTRACT

### COMBINING SIMULATION AND THE MSPA NANOPORE TO STUDY P53 DYNAMICS AND INTERACTIONS

SEPTEMBER 2023

SAMANTHA SCHULTZ, B.S., UNIVERSITY OF MASSACHUSETTS AMHERST

M.S., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Jianhan Chen and Professor Min Chen

p53 is a transcription factor and an important tumor suppressor protein that becomes activated due to DNA damage. Because of its role as a tumor suppressor, mutations in the gene that encodes it are found in over 50% of human cancers. The N-terminal transactivation domain (NTAD) of p53 is intrinsically disordered and modulates the function and interactions of p53 in the cell. Its disordered structure allows it to be controlled closely by post-translation modifications that regulate p53's ability to bind DNA and interact with regulatory binding partners. p53 is an attractive target for developing cancer therapeutics, but its intrinsically disordered region makes it difficult for traditional experimental techniques to resolve its heterogeneous conformational ensemble. This challenge necessitates the use of techniques that can capture the transient structural features and interactions of p53 to aid in designing effective drugs that can modulate and stabilize its activity. Hybrid-resolution (HyRes) II is a coarse-grained molecular dynamics force field that was parameterized specifically to capture the dynamics of IDPs and can give insight into secondary structure propensity and how post-translational modifications affect the structural ensemble of the protein. Nanopore experiments allow for real-time, single-molecule studies of protein dynamics and interactions with binding partners through characteristic changes in the current that passes through the nanopore. Pairing nanopore experiments with simulations can give insight into the molecular detail of IDP

ensembles and interactions, revealing a fuller picture of how p53 is controlled in stressed cell conditions and how its structure is affected due to various modifications and small molecules with therapeutic implications. Herein, we show the HyRes II force field can capture the complex, long-range dynamics of the p53 tetramer and provide molecular-level detail of the p53 autoinhibition mechanism, which is enhanced by the phosphorylation of the NTAD. Secondly, we use the MspA nanopore to capture the differences in events of the wild-type NTAD and a cancer-associated NTAD mutant. Lastly, we detect a small molecule binding to the WT NTAD using nanopore sensing. This approach of integrating MD simulations and nanopore experiments can be applied to the study of other IDPs which are prevalent in biology and integral to human health and disease.

## TABLE OF CONTENTS

<b>ACKNOWLEDGEMENTS.....</b>	<b>i</b>
<b>ABSTRACT.....</b>	<b>ii</b>
<b>TABLE OF CONTENTS.....</b>	<b>iv</b>
<b>LIST OF FIGURES.....</b>	<b>vi</b>
<b>LIST OF TABLES.....</b>	<b>vii</b>
<b>LIST OF ABBREVIATIONS.....</b>	<b>viii</b>
<b>CHAPTER 1</b>	
<b>INTRODUCTION.....</b>	<b>1</b>
<i>1.1 The role of intrinsic disorder in key signaling proteins and disease.....</i>	<i>1</i>
<i>1.2 Guardian of the genome: p53 tumor suppressor.....</i>	<i>2</i>
<i>1.3 Regulation of p53 by key binding partners and post-translational modifications.....</i>	<i>3</i>
<i>1.4 Methods used to characterize IDPs- advantages and limitations.....</i>	<i>6</i>
<i>1.5 The use of simulations to study IDPs at the molecular level.....</i>	<i>8</i>
<i>1.6 Nanopore sensing is a sensitive, single-molecule tool for studying IDPs.....</i>	<i>10</i>
<i>1.7 Statement of thesis.....</i>	<i>12</i>
<b>CHAPTER 2</b>	
<b>USING THE HYRES II FORCE FIELD TO INVESTIGATE THE EFFECTS OF PHOSPHORYLATION ON THE P53 AUTOINHIBITION MECHANISM.....</b>	<b>14</b>
Abstract.....	14
Introduction.....	15
Materials and methods.....	18
<i>Parameterization of phosphorylated serine and threonine in HyRes II.....</i>	<i>18</i>
<i>p53 System Setup and MD Simulation Details.....</i>	<i>21</i>
<i>Data analysis.....</i>	<i>23</i>
Results.....	24
<i>Using a coarse-grained force field to study complex IDP dynamics.....</i>	<i>24</i>
<i>Investigating the effects of phosphorylation on the interactions of the N- terminal domain in the p53 tetramer.....</i>	<i>27</i>
Discussion.....	29
<i>HyRes II can capture nontrivial dynamics of IDPs and the effect of PTMs on</i>	

<i>their structure</i> .....	29
<i>Phosphorylation-induced shifts in p53 conformational ensemble</i> .....	31
<i>Conclusions</i> .....	33
<b>CHAPTER 3</b>	
<b>SINGLE-MOLECULE DETECTION OF THE INTRINSICALLY DISORDERED P53</b>	
<b>NTAD USING THE MSPA NANOPORE</b> .....	<b>34</b>
Abstract.....	34
Introduction.....	35
Materials and methods.....	37
<i>Expression and Purification of p53-NTAD (1-73) and N29K/N30D mutant</i> .....	37
<i>Expression and Purification of M2-MspA-NNN</i> .....	41
<i>Single-Channel Electrophysiology Experiments</i> .....	42
<i>Data analysis</i> .....	43
<i>HyRes II simulations and analysis</i> .....	44
Results.....	45
<i>Label-free detection of disordered p53 NTAD</i> .....	45
<i>Nanopore detection of an NTAD disease-associated mutant</i> .....	49
<i>Real-time identification of a small molecule binding p53 NTAD using MspA</i>	
<i>nanopore sensor</i> .....	52
Discussion.....	55
<i>MspA nanopore sensor allows for the characterization of heterogeneous NTAD</i>	
<i>events</i> .....	55
<i>MspA can detect changes in NTAD events due to a cancer-associated mutation</i> .....	56
<i>The nanopore can be used to detect small molecule binding to an IDP</i> .....	57
<i>Conclusions</i> .....	59
<b>CHAPTER 4</b>	
<b>CONCLUSIONS AND OUTLOOK</b> .....	<b>60</b>
4.1 <i>Conclusions</i> .....	60
4.2 <i>Future directions</i> .....	61
<b>SUPPLEMENTARY FIGURES</b> .....	<b>63</b>
<b>REFERENCES</b> .....	<b>67</b>

## LIST OF FIGURES

Figure 1.1. Diagram of key p53 target genes.....	4
Figure 1.2. HyRes II force field can capture IDP dynamics.....	9
Figure 1.3. Characterizing analytes with protein nanopore sensors.....	11
Figure 2.1. Illustration of amino acid and phosphorylation representation in the HyRes II force field.....	18
Figure 2.2. Validation of bond length and angle distributions in HyRes II.....	20
Figure 2.3. Initial conformations of p53 tetramers for HyRes II simulations.....	22
Figure 2.4. Radius of gyration and helicity probability distributions of p53 NTAD.....	25
Figure 2.5. PRE effects of the NTAD induced by a spin label on S121 of p53 DBD.....	26
Figure 2.6. PCA of conformations of the WT and phosphorylated p53 NTAD.....	28
Figure 3.1. WT NTAD purification SDS-PAGE gel.....	39
Figure 3.2. N29K/N30D NTAD purification SDS-PAGE gel.....	40
Figure 3.3. Raw data from MALDI-TOF analysis of p53 NTAD.....	41
Figure 3.4. WT NTAD voltage study Ires vs dwell time plots.....	46
Figure 3.5 MspA can capture p53 NTAD events in a concentration-dependent manner.....	48
Figure 3.6. Probability distribution of the radius of gyration for WT and N29K/N30D NTAD.....	50
Figure 3.7. Detection of subtle differences in WT and mutant NTAD events.....	51
Figure 3.8. EGCG control and representative WT NTAD and EGCG “stuck” event.....	53
Figure 3.9. Detection of a small molecule binding p53 NTAD.....	54
Figure S1. WT NTAD concentration study Ires vs dwell time plots.....	64
Figure S2. Full analyzed traces of WT and N29K/N30D NTAD at 300 nM.....	65
Figure S3. Full analyzed traces of 200 nM WT NTAD alone and 200 nM WT NTAD and 40 nM EGCG.....	66

## LIST OF TABLES

Table 2.1 Equilibrium bond lengths ( $b_o$ ) and $k_b$ values for phosphorylated serine and threonine residues in HyRes II.....	19
Table 2.2 Equilibrium angle values ( $\theta_0$ ) for phosphorylated serine and threonine residues in HyRes II.....	19
Table 2.3 Nonbonded parameters for phosphorylated serine and threonine residues in HyRes II.....	21
Table S1. Fitting results for interevent duration ( $\tau_{on}$ ) at different NTAD concentrations.....	63

## LIST OF ABBREVIATIONS

N-terminal transactivation domain of p53 (NTAD); DNA binding domain of p53 (DBD); intrinsically disordered protein (IDP); intrinsically disordered region (IDR); post-translational modifications (PTMs); wild-type (WT); human double minute protein-2 (HDM2); nuclear magnetic resonance (NMR); paramagnetic resonance enhancement (PRE); Förster resonance energy transfer (FRET); small-angle X-ray scattering (SAXS); molecular dynamics (MD); phosphorylated serine (pSer); phosphorylated threonine (pThr); principal component analysis (PCA); *Mycobacterium smegmatis* porin A (MspA); residual current ( $I_{res}$ ); Epigallocatechin gallate (EGCG); matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF)

# CHAPTER 1

## INTRODUCTION

### *1.1 The role of intrinsic disorder in signaling proteins and disease*

Intrinsically disordered proteins (IDPs) are crucial components of signaling pathways within the cell and take on a variety of critical functions. IDPs do not fold into a stable tertiary structure under physiological conditions and have a heterogeneous ensemble of marginally stable conformations. They are characterized by low-complexity sequences that usually have a net charge and lack bulky, hydrophobic residues<sup>1-5</sup>. The instability of an IDP's structure, therefore, allows it to be an interaction hub for many binding partners and cellular pathways due to subtle perturbations<sup>6-8</sup>. As such, many important signaling proteins are IDPs or have extensive disordered regions, and the vast majority of eukaryotic transcription factors have significant disorder in their structure<sup>9</sup>. This also means that the dysfunction or misregulation of these important proteins is involved in many different diseases<sup>10,11</sup>.

When developing therapeutics to target intrinsically disordered proteins or regions, the bound state or most ordered conformation of the IDP is not the only structure that determines its function, but every state in its conformational ensemble is important for mediating its interactions<sup>12</sup>. Disease-related mutations and post-translational modifications can shift an IDPs conformational ensemble or stabilize one of its sub-conformations, affecting its functions and interactions in the cell. Given this important relationship between the ensemble of an IDP and its function, it is important to effectively characterize the relevant substates of an IDP to understand its mechanism of action in the cell and to design effective therapeutics that target its structure

and interactions. p53 is one such intrinsically disordered protein that is involved in critical cell functions, and studying its interactions and regulation has been an area of ongoing study since its discovery.

### *1.2 Guardian of the Genome: p53 tumor suppressor*

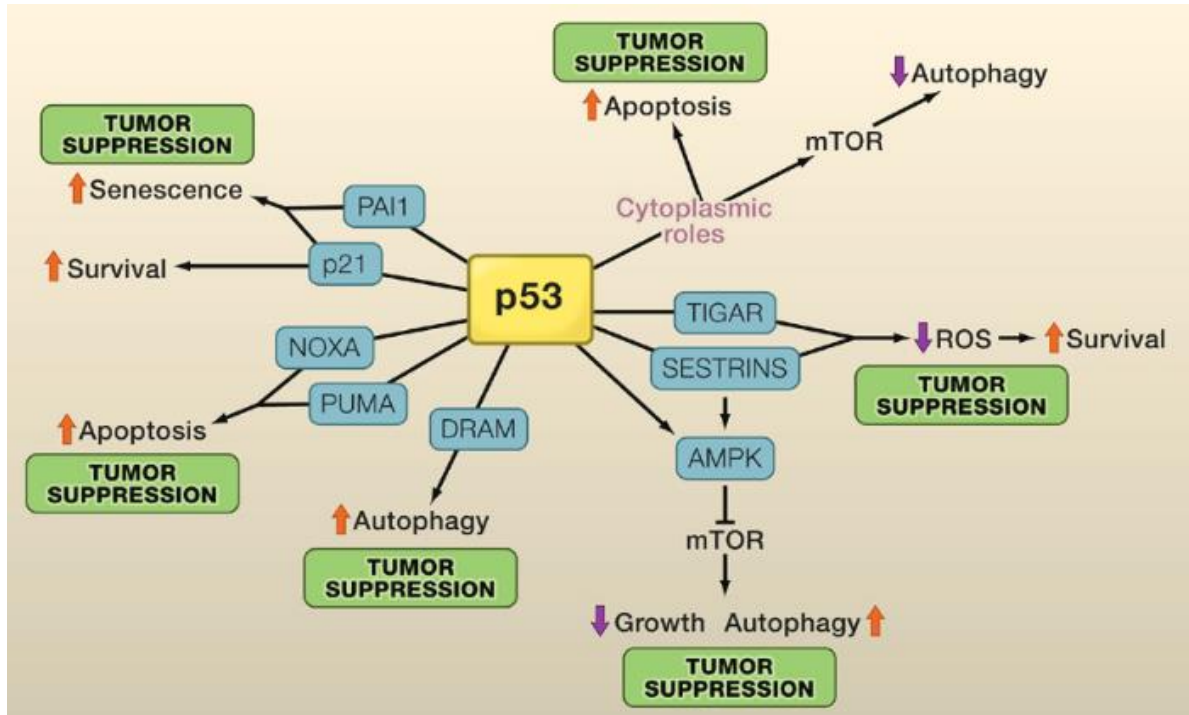
p53 is an IDP that is involved in numerous integral cellular pathways, and the gene that encodes it is the most frequently mutated gene in human cancers. It was first discovered in 1979 when a 53 kDa protein was identified via immunoprecipitation interacting with the large T-antigen of simian virus 40, a virus known to cause tumors in animals. p53 was also identified around the same time using an antiserum made against chemically induced mouse sarcoma cells, and it was found to be absent in nontransformed cells<sup>13,14</sup>. p53's presence in these tumor cells indicated a potential function related to cell growth, but its exact function and relation to cancer remained unclear.

1989 was a pivotal year in the study of p53, and it was found to be mutated in numerous types of tumors<sup>15-17</sup>. Specifically, a study by Baker and colleagues investigated mutations on the short arm of chromosome 17, which are found in over 75% of colorectal carcinomas. They identified the region of these deletions as the gene encoding p53, suggesting the mutations inactivate the tumor-suppressing activities of the wild-type protein. This study marked the beginning of p53 being classified as a tumor suppressor, and since then p53 loss-of-function mutations have been found in over 50% of human cancers, earning it the title of guardian of the genome<sup>18-21</sup>.

The intrinsic disorder in p53's structure is integral to its regulation and function in the cell. p53 is a transcription factor comprised of a disordered N-terminal transactivation domain (NTAD), a DNA binding domain (DBD), a tetramerization domain (TET), and a C-terminal domain. The NTAD is highly negatively charged and can be further divided into two interaction motifs, AD1 (residues 18-26) and AD2 (residues 40-55), and a proline rich domain (PRD, residues 64-92)<sup>22</sup>. p53 forms dimers that cooperatively bind DNA and assemble as a tetramer in a sequence-specific fashion<sup>22-26</sup>. p53 is a clear therapeutic target due to its role in cancer, and understanding its regulation in the cell, including the interactions mediated by the intrinsically disordered NTAD, is imperative for developing effective cancer treatments.

### *1.3 Regulation of p53 by key binding partners and post-translational modifications*

Due to its disordered NTAD, p53 is an interaction hub and has been implicated in a growing number of cellular pathways, namely with mechanisms involving the cell cycle and apoptosis<sup>27</sup>. p53 is known to activate the transcription of p21, a cyclin-dependent kinase inhibitor that arrests the cell cycle in the G1 phase<sup>28-30</sup>. p53 has also been shown to target genes such as *14-3-3 sigma* and *GADD45* which can induce cell cycle arrest in G2<sup>28</sup>. p53 can mediate apoptosis via the mitochondrial pathway by initiating the transcription of PUMA (p53-upregulated modulator of apoptosis) but can also target genes like that of the transcription factor SLUG that represses the expression of PUMA and promotes cell survival<sup>31,32</sup>. Some of p53's target genes are involved in pathways of autophagy and metabolism, contributing to the numerous cellular responses that are mediated by this protein (Fig. 1.1)<sup>33</sup>.



**Figure 1.1. Diagram of key p53 target genes.** p53 controls the expression of multiple genes (examples shown in blue) with roles in cell-cycle arrest and apoptosis, making it an important tumor suppressor protein. Taken from Vousden and Prives 2009<sup>33</sup>.

p53 is a vital protein for initiating responses to cellular stress, and with varying functions that can also have contradictory effects, it must be closely regulated. p53 function must be suppressed in an unstressed cell to prevent unnecessary inhibition of the cell cycle or apoptosis, but quickly activated in response to DNA damage and other stress signals. This can be accomplished by the addition of post-translational modifications to the disordered NTAD that change the structure of the protein and modulate its interactions with key binding partners<sup>34-36</sup>. p53 is targeted for proteasomal degradation in a healthy cell by E3 ubiquitin ligase human double minute protein-2 (HDM2)<sup>37,38</sup>. When DNA damage occurs, stress-activated kinases phosphorylate p53 and regulate this interaction<sup>39-42</sup>. Specifically, phosphorylation of Threonine 18 acts as a switch

mechanism to disrupt HDM2 binding by changing the secondary structure of the NTAD, preventing p53 from being targeted for degradation<sup>43,44</sup>.

Conversely, phosphorylation of key residues on the p53 N-terminus in response to cellular stress increases p53 interactions with stabilizing transcriptional coactivators and histone acetyl transferases cAMP response element binding protein-binding protein (CBP) and p300, allowing p53 to carry out its tumor suppressing function<sup>45-47</sup>. This binding is enhanced through multiple phosphorylation sites on the NTAD in a graded response<sup>48</sup>. The inhibition of HDM2-mediated degradation and the increase in p53 contact with CBP/p300 are both important steps in stabilizing and activating p53 function in the cell<sup>49</sup>.

Because of its disordered nature, the NTAD closely controls p53 function not only through interactions with binding partners but also through interactions with other domains within the p53 tetramer via its two interaction motifs, AD1 and AD2. Phosphorylation of threonine 55 on the AD2 motif can increase transient interactions of AD2 with the positively charged DNA binding domain and control p53's ability to bind to DNA and initiate gene transcription<sup>50</sup>.

Threonine 55 is constitutively phosphorylated in an unstressed cell, which is an autoinhibitory mechanism to prevent nonspecific DNA binding. Serine 46 is another residue in the AD2 motif that is thought to be phosphorylated as part of this autoinhibition process.

Although these interactions have been characterized by techniques such as NMR, there is still much to be learned about the molecular detail involved in processes such as the NTAD and DBD interaction. The tetrameric structure of p53 adds complexity to the autoinhibition mechanism, and there is evidence that increased AD2 interactions with the DBD due to phosphorylation of threonine 55 occur *in trans* across the p53 dimers to dissociate the tetramer from DNA<sup>50</sup>. This

mechanism has been hypothesized because the cooperativity of tetramer-DNA dissociation at high p53 concentrations, but the specific details of this process on the molecular level are still unknown.

Additionally, because of p53's critical role in the development of many cancers, there is much interest in studying p53 disease-associated mutants and potential therapeutics that target and stabilize p53 activity. Epigallocatechin gallate (EGCG) is a small molecule that binds to the NTAD of p53 and prevents it from interacting with HDM2, inhibiting p53 degradation and increasing its wild-type activity in the cell<sup>51</sup>. The ability to identify these therapeutics and study the effects they have on p53 function and structure is critical for advancements in cancer treatment. A higher resolution understanding of these processes could allow for a more complete picture of how p53 is regulated in the cell and how this controls its role in various pathways and diseases.

#### *1.4 Methods used to characterize IDPs: advantages and limitations*

Studying the full landscape of p53 structure presents challenges for traditional structural experimental techniques such as x-ray crystallography and cryogenic electron-microscopy (cryo-EM). X-ray crystallography has been used to solve many protein structures, but relies on a homogenous protein sample and the existence of one stable, biologically-relevant structure to crystallize, precluding the characterization of IDPs<sup>52,53</sup>. cryo-EM is another popular structural technique that has limitations in resolving highly flexible regions in proteins, again limiting its ability to characterize intrinsically disordered regions<sup>54</sup>.

Nuclear magnetic resonance (NMR) spectroscopy is a popular and powerful technique for characterizing intrinsically disordered proteins. It can resolve IDP conformational ensembles on the secondary and tertiary levels and give information on protein-protein interactions through chemical shift, coupling constant, nuclear overhauser effect (NOE), residual dipolar coupling (RDC), paramagnetic resonance enhancement (PRE), and spin relaxation<sup>8</sup>. Additional techniques suited for characterizing dynamic proteins can be combined with NMR to give a better understanding of the disordered ensemble, such as single-molecule Förster resonance energy transfer (sm-FRET) which can resolve transient interactions between protein domains and provide distance information between two fluorescent probes. Small-angle X-ray scattering (SAXS) is another technique that is often used to study IDPs, and can give information about the overall shape and size of the sample based on its X-ray scattering pattern<sup>55,56</sup>.

Despite the advantages of these techniques, there are some limitations, such as the fact that techniques like SAXS and NMR give structural information for IDPs as a population-weighted average over their many conformations, meaning it is difficult to assign the conformational sub-states that gave rise to the data<sup>8,57-61</sup>. Although PRE and NOE can give information on transient interactions within an IDP, these techniques have a distance dependence of  $r^{-6}$ . This can misrepresent the underlying conformational distribution and give more weight to shorter interactions while masking long-range interactions<sup>8</sup>. In the case of PRE and sm-FRET, the addition of paramagnetic and fluorescent probes, respectively, can also affect the IDP conformational ensemble and misrepresent how the protein behaves under native physiological conditions. These key challenges require new experimental and computational techniques that are able to accurately probe the heterogeneous ensembles of IDPs.

### *1.5 The use of simulations to study IDPs at the molecular level*

The challenge of studying disordered proteins necessitates the use of techniques such as molecular dynamics simulations that can capture protein conformational changes and transient interactions. Trajectories from molecular dynamics simulations can be used to identify substates of intrinsically disordered proteins with more confidence than ensemble averaging techniques, which allows for the identification of multiple relevant conformations<sup>62</sup>. Therefore, the molecular-level findings from simulations can uncover how certain perturbations affect p53's conformational ensemble, revealing the role of key PTMs or disease-related mutations.

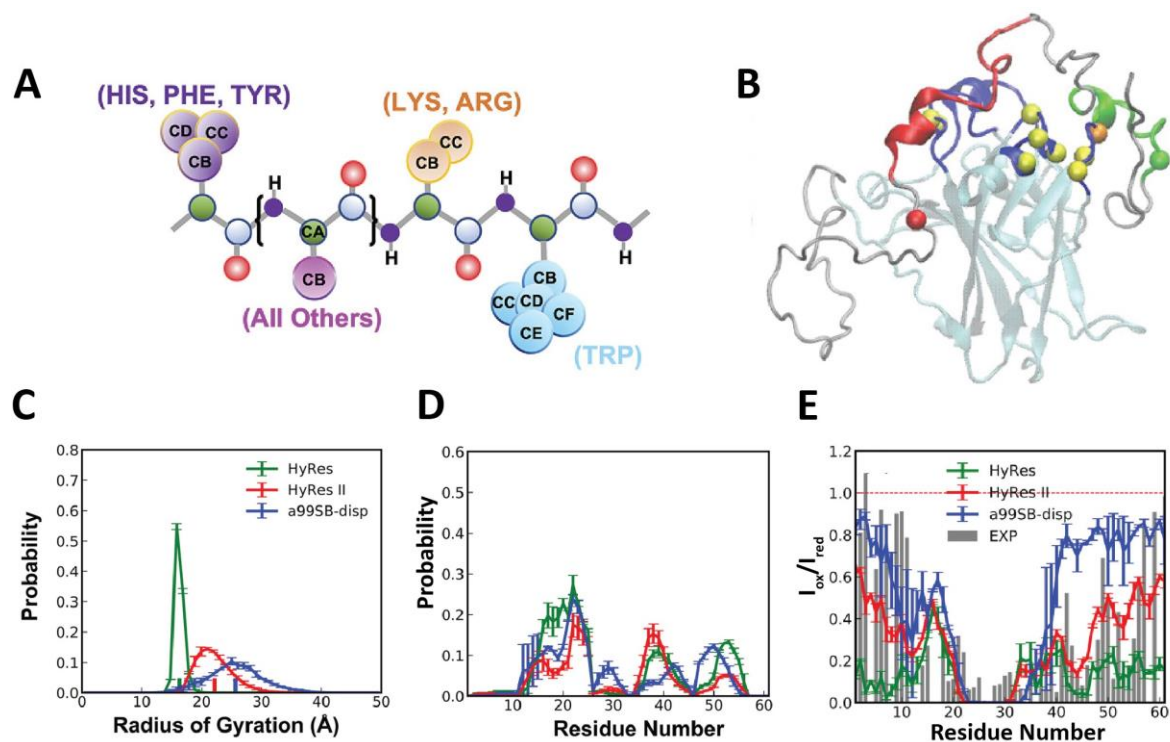
Although effective for addressing these problems, MD simulations are limited by system size and simulation timescale. Computational cost increases significantly for modeling large proteins such as the p53 tetramer and for accessing biologically relevant timescales ( $\mu\text{s}$  – seconds). Recent advancements in coarse-grained protein force fields have greatly increased the accessible size and timescale of simulations by reducing the level of protein representation compared to all-atom force fields, allowing for simulations to address more biologically relevant, complex problems<sup>63</sup>.

Hybrid-resolution (HyRes) II is a coarse-grained model that was parameterized to capture the dynamics of intrinsically disordered proteins and can give insight into secondary structure propensity and transient, long-range protein interactions. This is accomplished through the explicit representation of the protein backbone and intermediate-resolution side chains, which provide semiquantitative description of protein secondary structure with qualitative details of long-range IDP interactions and dynamics (Figure 1.2 A). HyRes II has an energy function with eight bonded and nonbonded components (eq. 1) and was improved from the original HyRes

model through reparameterization of side chain and backbone van der Waal interactions and backbone hydrogen-bonding interactions.

$$U_{\text{hyres}} = U_{\text{bond}} + U_{\text{angle}} + U_{\text{dihedral}} + U_{\text{improper}} + U_{\text{CMAP}} + U_{\text{LJ}} + U_{\text{Hbond}} + U_{\text{elec}} \quad (1)$$

These changes are implemented to better capture solvent screening effects and reduce the over-compactness of protein structures that was exhibited with the original force field<sup>64</sup>.



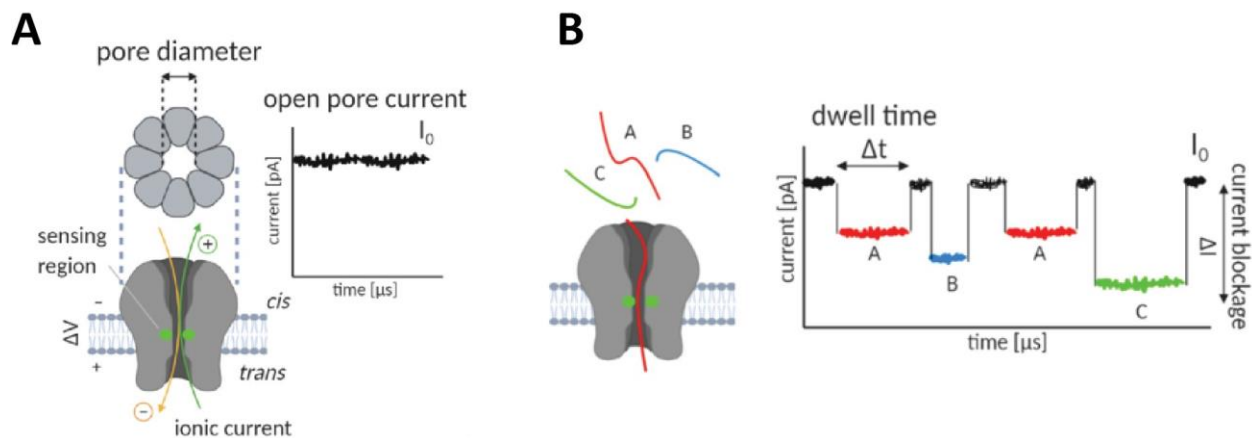
**Figure 1.2. HyRes II force field can capture IDP dynamics.** (A) HyRes representation of amino acid side chains and protein backbone. Backbone atoms are explicitly modeled, and side chains (besides glycine) are represented using one to five coarse-grained beads. (B) p53 monomer modeled with HyRes II. The structure is drawn in cartoon, with the DBD colored in cyan and TAD and PRD in gray. AD1 and AD2 are highlighted in green and red, respectively. The DNA-binding surface of DBD is highlighted in blue, with key NTAD-interacting residues T118, V112, G245, N247, A276, G279, R280, and R283 shown as yellow vdW beads. The paramagnetic spin labeling site S121 is colored orange. S15 and P58 are labeled in red and green, respectively. (C & D) Probability distribution of the radius of gyration (Å) (C) and helicity (D) of p53 TAD calculated using HyRes and HyRes II at 300 K. Results from atomistic simulations using a99SB-disp are shown in blue. (E) Calculated from simulation (blue, red and green lines) and experimental (gray bars) PRE effects induced by paramagnetic spin labeling at residue 28. Adapted from Zhang et al. 2022<sup>65</sup>.

With the improvements described above, HyRes II is able to accurately simulate the long-range interactions and dynamics of IDPs compared to the original HyRes model, and the results generated from this force field show strong agreement with experimental results. The p53 NTAD was simulated with this model, and HyRes II is able to accurately capture the radius of gyration, helicity, and PRE effects induced by paramagnetic spin labels at various residues on the peptide (Fig. 1.2 B-E). As such, simulations using the HyRes II model are suited to probe the effects of mutations or PTMs on the structural ensemble of an IDP such as p53, and it can simulate the full-length p53 tetramer much more efficiently than previous all-atom simulations to cover the conformational landscape at relevant timescales.

### *1.6 Nanopore sensing is a sensitive, single-molecule tool for studying IDPs*

Another emerging technique for capturing real-time protein conformational changes and disordered ensembles is experiments using nanopores. This technique involves two chambers of aqueous electrolyte buffer separated by Teflon film, in which a lipid bilayer can be formed. Biological pore-forming proteins can then be added to the solution and a single pore can spontaneously insert into the bilayer. Voltage is applied across the membrane, allowing the current created by ion flow through the pore to be measured (Fig. 1.3 A). An analyte of interest such as a protein can then be added to the chamber and a combination of electroosmotic force, electrophoretic force, and diffusion will cause the protein to be trapped within the nanopore, resulting in a change in current through the pore that can later be analyzed and characterized. Many analytes can be uniquely characterized by parameters such as event duration (dwell or

$\tau_{\text{off}}$ ), residual pore current ( $I_{\text{res}}$ ), and interevent time ( $\tau_{\text{on}}$ ), allowing for studies of protein dynamics and interactions with binding partners on the single-molecule level (Fig. 1.3 B)<sup>1,66–68</sup>.



**Figure 1.3. Characterizing analytes with protein nanopore sensors. (A)** Schematic of a nanopore inserted into a lipid bilayer under an applied voltage. The open pore current over time is shown as a trace to the right. **(B)** Schematic illustrating the characterization of multiple analytes (A, B, and C) using nanopore sensing. The corresponding analyte events are shown on the trace to the right and can be characterized by event duration ( $\Delta t$ , dwell) and changes in the pore current ( $\Delta I$ ). Adapted from Crnković et al. 2021<sup>69</sup>.

Nanopore experiments offer many unique advantages, such as the ability to be optimized to better represent physiological conditions without the addition of external probes or labels.

Nanopores can trap analytes for extended periods of time, allowing for the resolution of dynamics on the  $\mu\text{s}$  to minutes timescale. Nanopores have already been used for a variety of applications such as DNA sequencing<sup>70,71</sup>, protein and small molecule identification<sup>67,68,72–75</sup>, and kinetic studies of enzymes<sup>66</sup>.

Unlike NMR, nanopores allow for the trapping and observation of many single molecules over the course of a recording, which can be hours in length, eliminating the challenges that come with ensemble-averaging techniques. Nanopores also require very low concentrations of protein to detect analyte trapping and have been scaled-up and parallelized in devices like the MinION,

which uses an array of many nanopores for portable, real-time DNA sequencing<sup>76,77</sup>. Thus, nanopores offer the potential of a high-throughput technique that can be used to screen proteins and their interactions.

*Mycobacterium smegmatis* porin A (MspA) is an octameric protein pore with a funnel shape that terminates in a ~1.2 nm constriction region. The size and geometry of this pore is suited for observing smaller analytes, and it has previously been used for DNA sequencing and protein trapping<sup>72,78-81</sup>. MspA is therefore an ideal tool for trapping small peptides, such as the N-terminal domain of p53, and observing its dynamics and structural heterogeneity. These experiments can also allow for the detection of small-molecule binding to determine the effects of potential therapeutics on protein conformations in real-time.

### *1.7 Statement of thesis*

In this thesis, I pair nanopore experiments with molecular dynamics simulations to observe how p53's structure is affected due to various modifications and monitor its interactions with a small molecule. These experiments aim to better understand the molecular mechanisms of p53 regulation and show evidence for the ability of the utilized methods to study and characterize IDPs. Using the HyRes II force field to simulate the p53 tetramer, I show that HyRes II is able to capture the trends and dynamics of IDP conformational ensembles. From these simulations, I analyze the p53 autoinhibition mechanism involving the NTAD-DBD interaction and how the structure of the p53 tetramer and phosphorylation at key residues contributes to this mechanism. Secondly, I show that experiments using MspA nanopore can capture events due to the trapping of the intrinsically disordered NTAD of p53. These events can be distinguished from the events

of a cancer associated NTAD mutant, demonstrating the sensitivity of nanopore experiments. I lastly show the nanopore can be used as an effective tool for detecting small molecule binding to the NTAD, which has implications for the use of the MspA nanopore as a novel drug screening tool.

## CHAPTER 2

### USING THE HYRES II FORCE FIELD TO INVESTIGATE THE EFFECTS OF PHOSPHORYLATION ON THE P53 AUTOINHIBITION MECHANISM

#### **Abstract**

Intrinsically disordered proteins (IDPs) are critical mediators of various pathways in the cell. p53 is a tumor suppressor protein with an intrinsically disordered N-terminal region (NTAD), and one of the most frequently mutated proteins in numerous types of cancer. It is a transcription factor that controls the expression of genes related to arresting the cell cycle during conditions of cellular stress. Because of its ability to modulate cell growth and apoptosis, it is closely controlled by other proteins and post-translational modifications. p53 is regulated through multiple mechanisms by changes in the structure of the NTAD which is susceptible to subtle perturbations that occur from the addition of post-translational modifications. Specifically, phosphorylation at residue Threonine 55 increases the NTAD's interaction with the folded DNA binding domain (DBD) in an autoinhibition mechanism to prevent nonspecific DNA binding. The exact molecular mechanism of p53 autoinhibition is difficult to study using traditional structural experiments such as x-ray crystallography or ensemble-averaging techniques like NMR and FRET. Hybrid-resolution (HyRes) II is a coarse-grained molecular dynamics force field that is specifically parameterized to represent intrinsically disordered proteins. Using this force field, the full-length p53 tetramer can be simulated, and transient secondary structure and inter-domain interactions can be captured. Studying p53's conformational ensemble with different post-translational modifications can give insight into the exact mechanism of its

regulation and autoinhibition, allowing for a deeper understanding of how it is controlled in the cell and the development of potential therapeutic targets.

## **Introduction**

Intrinsically disordered proteins (IDPs) are crucial components of signaling pathways within the cell and take on a variety of functions related to cell decision making. IDPs are characterized by the absence of a stable tertiary structure under physiological conditions and instead take on a heterogeneous ensemble of marginally stable conformations. They have low-complexity sequences that usually have a net charge and lack bulky, hydrophobic residues<sup>1-5</sup>. IDPs are an interaction hub for many pathways and binding partners because of their thermodynamic instability, which allows them to undergo conformational changes due to small perturbations<sup>6-8</sup>. As such, many important signaling proteins such as transcription factors have significant disorder in their structure<sup>9</sup>. This also means that the dysfunction or misregulation of these important proteins is involved in many diseases such as cancer, cardiovascular disease, neurodegenerative diseases, and diabetes<sup>10,11</sup>.

p53 is one such transcription factor that plays an integral role in regulating the cell cycle. It has been implicated in activating genes that are involved in pathways of cell cycle arrest and apoptosis as a result of cellular stress, making it an important tumor suppressor protein<sup>28,29,31,32</sup>. As such, p53 mutations that result in a loss of function of the protein are a prerequisite in many cancers<sup>15-17</sup>. p53's critical role as a tumor suppressor and its dysregulation in disease make it an important protein to study as a potential therapeutic target, which requires a deep understanding of its interactions and modes of regulation in the cell.

p53 is a tetrameric protein comprised of an N-terminal transactivation domain (NTAD), a proline-rich domain (PRD), a DNA binding domain (DBD), a tetramerization domain (TET), and a c-terminal domain<sup>22-24,52</sup>. The NTAD is intrinsically disordered and has two interaction motifs, AD1 (18-26) and AD2 (40-55). The NTAD is an important manager of p53 and its binding interactions with regulatory proteins and DNA, which have been shown to be modulated through post-translational modifications at various sites<sup>48,50</sup>.

The NTAD has been shown to modulate p53 DNA binding through interactions with the folded DBD, which is thought to be an autoinhibition mechanism<sup>82</sup>. Studies of the molecular mechanism of the NTAD-DBD interaction reveal that the NTAD interacts with the positively charged DBD mainly through the negatively charged AD1 and AD2 motifs, which compete with DNA for DBD binding via electrostatic interactions<sup>53,83</sup>. Phosphorylation plays an important role in this mechanism, and threonine 55 was identified as a key residue that enhances the AD2/DBD interaction upon phosphorylation<sup>50</sup>.

These previous studies have used techniques like nuclear magnetic resonance (NMR), single-molecule Förster resonance energy transfer (sm-FRET), and small-angle x-ray scattering (SAXS) to probe the molecular mechanism of p53 autoinhibition and DNA binding, and although powerful, these techniques have limitations in the study of IDPs<sup>55</sup>. NMR and SAXS are ensemble averaging techniques, making it difficult to recover the heterogeneous conformational ensemble that gave rise to the data<sup>58-60</sup>. NMR observables NOE and PRE can give information on interdomain interactions but have distance limitations and are more sensitive to short-range interactions, which can misrepresent the actual conformational distribution<sup>61</sup>. PRE and sm-FRET require the addition of external paramagnetic and fluorescent probes, respectively, which can affect the native ensemble of the protein<sup>8</sup>.

These challenges necessitate the use of a technique that can capture the native disordered ensemble of the p53 NTAD on the molecular level and identify substates that are biologically relevant. Molecular dynamics (MD) simulations are a powerful tool for capturing the dynamics of IDPs and are especially effective at probing the differences that arise from mutations and PTMs like phosphorylation. MD simulations are limited by system size and time constraints due to computational cost, but these limitations can be addressed by using enhanced sampling and coarse-graining techniques that greatly expand the accessible size and timescales of simulations<sup>63</sup>.

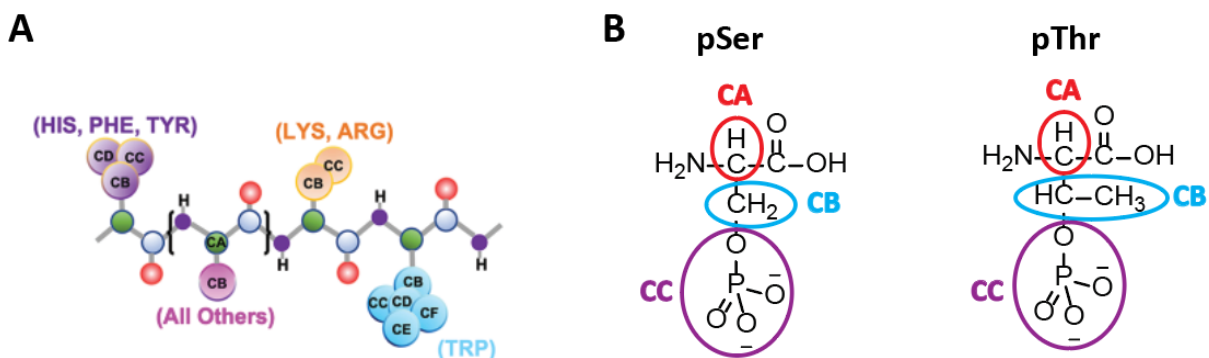
Hybrid-resolution (HyRes) II is a coarse-grained model that was parameterized specifically to capture the transient conformations and secondary structure of IDPs through the explicit representation of the protein backbone and the coarse-graining of residue side chains<sup>64,65</sup>. This allows for the semiquantitative description of secondary structure and the ability to qualitatively capture long-range IDP interactions. HyRes II has already been applied to the study of the p53 monomer and was able to capture conformational substates and long-range dynamics that recapitulate experimental results<sup>65</sup>.

Using the HyRes II force field, we simulate both the wild-type and phosphorylated p53 tetramer constructs and analyze the generated ensemble to create a more complete picture of the molecular mechanism of p53 autoinhibition. Here, we capture the effects of phosphorylation of a key residue within the NTAD and show HyRes II is an effective model for representing complex IDP dynamics.

## Materials and Methods

### *Parameterization of phosphorylated serine and threonine in HyRes II*

The HyRes II force field includes a united-atom representation of the protein backbone and represents the amino acid side chains as varying numbers of beads depending on the properties of the amino acid (Fig. 2.1 A). For the parameterization of phosphorylated serine (pSer) and threonine (pThr) in HyRes II, the atoms were grouped according to the scheme in Figure 2.1 B, with the sidechain carbon atoms grouped as one bead (CB) and the phosphate group assigned as a second bead (CC).



**Figure 2.1. Illustration of amino acid and phosphorylation representation in the HyRes II force field.** (A) (taken from Zhang et al. 2022<sup>65</sup>) Amino acids in HyRes II are represented with explicit atoms for the backbone and five (tryptophan), three (histidine, phenylalanine, and tyrosine), two (lysine and arginine), or one (all others except glycine) beads to represent the amino acid side chain. (B) For parameterization of phosphorylated serine (left) and threonine (right) in HyRes II, the grouping of the coarse-grained beads is illustrated as circles around the all-atom representation. The CB bead encompasses the carbon atoms and their associated hydrogens in the side chain (cyan circle) and the CC bead encompasses the entire phosphate group ( $\text{PO}_4^{2-}$ , purple circle).

To determine the equilibrium bond length between the CB and CC beads and angle values between the beads and the backbone atoms, phosphorylated serine and threonine dipeptides were

simulated using the CHARMM22 all-atom force field and GBSW implicit solvent<sup>84</sup>. The equilibrium bond lengths and angles were calculated as the average of the values for the pseudo-beads in the all-atom simulations and are listed in Table 2.1 and 2.2 below. The parameter  $k_b$  (Table 2.1), which controls the flexibility of the bond between two beads, was calculated using the standard deviation of the average values from the all-atom dipeptide simulations according to the following equation:

$$k_b = \frac{kT}{2\sigma^2} \quad (2)$$

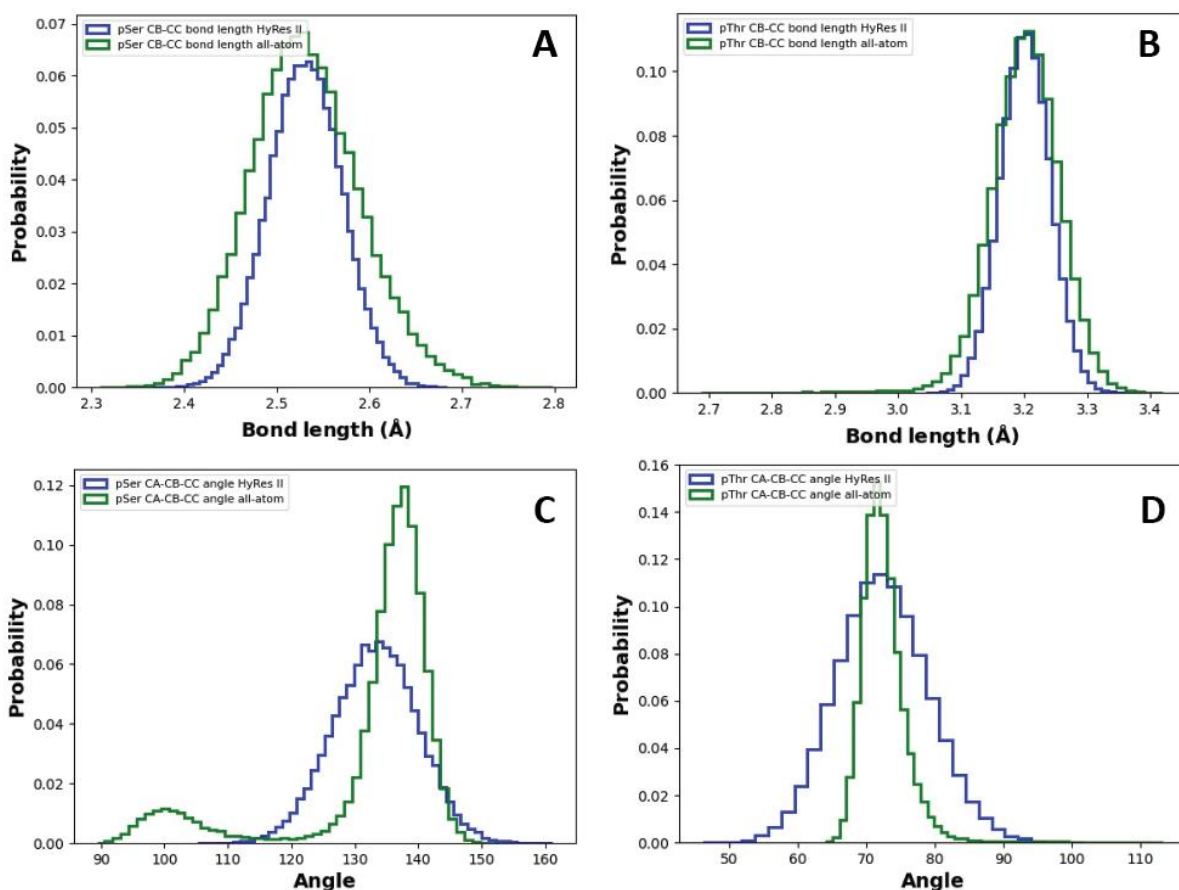
**Table 2.1. Equilibrium bond lengths ( $b_0$ ) and  $k_b$  values for phosphorylated serine and threonine residues in HyRes II.** Obtained from atomistic simulations of phosphorylated serine (pSer) or threonine (pThr) dipeptides using GBSW implicit solvent.

Bond type	$k_b$ (kcal/mol/Å <sup>2</sup> )	$b_0$ (Å)
pSer CA-CB	262.9	1.61
pSer CB-CC	200.0	2.53
pThr CA-CB	200.0	2.03
pThr CB-CC	200.0	3.20

**Table 2.2. Equilibrium angle values ( $\theta_0$ ) for phosphorylated serine and threonine residues in HyRes II.** Obtained from atomistic simulations of phosphorylated serine (pSer) or threonine (pThr) dipeptides using GBSW implicit solvent.

Angle type	$\theta_0$ (°)
pSer C-CA-CB	112.3
pSer N-CA-CB	110.3
pSer CA-CB-CC	132.8
pThr C-CA-CB	101.2
pThr N-CA-CB	98.4
pThr CA-CB-CC	72.6

The distributions of the bond lengths and angles of phosphorylated serine and threonine dipeptides simulated for 10-20 ns using HyRes II were calculated, and the results for the CB-CC bond length and CA-CB-CC angle for both dipeptides are shown in Figure 2.2. These distributions were compared to the results of atomistic dipeptide simulations using the CHARMM22 force field and GBSW implicit solvent to ensure the correct residue geometry was reproduced in the coarse-grained simulations.



**Figure 2.2. Validation of bond length and angle distributions in HyRes II.** The bond length (A and B) and angle (C and D) distributions from 10-20 ns HyRes II simulations are plotted in blue with the results from CHARMM22 all-atom simulations in green. The results for the phosphorylated serine dipeptide are shown in A and C and the phosphorylated threonine dipeptide is shown in B and D.

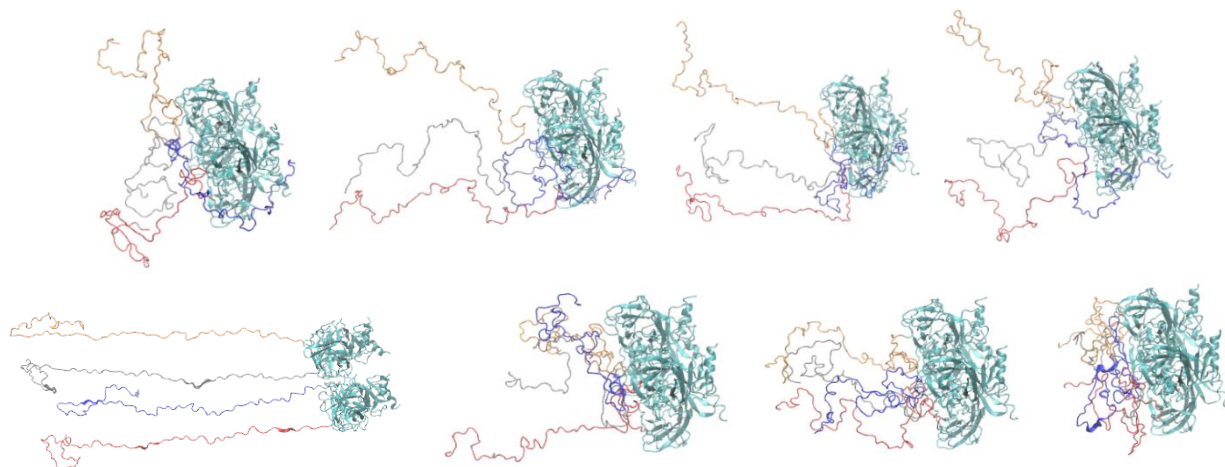
The nonbonded parameters (Table 2.3) in HyRes II for the CB and CC beads of phosphorylated serine and threonine include the van der Waals interaction strength ( $\epsilon_i$ ) and van der Waals radius. The  $\epsilon_i$  of the phosphate bead (CC) for both serine and threonine was approximated based on the existing glutamate  $\epsilon_i$  parameterized in HyRes II<sup>65</sup>. The van der Waals radius for the phosphate bead for both residues was calculated by adding the phosphorus-oxygen bond length taken from the atomistic dipeptide simulations and the atomic radius of oxygen. The  $\epsilon_i$  and radius for the phosphorylated serine CB bead were approximated to be the same as the existing values for alanine, and the  $\epsilon_i$  and radius values for the threonine CB bead were calculated as an average between the values for alanine and valine.

**Table 2.3. Nonbonded parameters for phosphorylated serine and threonine residues in HyRes II.**

<b>Bead type</b>	<b><math>\epsilon_i</math> (kcal/mol)</b>	<b>vdW radius (Å)</b>
pSer CB	-0.15	2.12
pSer CC	-0.07	2.25
pThr CB	-0.22	2.44
pThr CC	-0.07	2.25

### *p53 System Setup and MD Simulation Details*

The p53 tetramer (PDB ID: 1TUP) was converted from an atomistic PDB to the HyRes II coarse-grained model using an in-house Python script. Each monomer of the tetramer includes residues 1-291 of p53. To generate distinct initial conformations for simulation replicates in order to assess convergence, a combination of steer MD and high temperature (500 K and 800 K) equilibration simulations were used, and random frames were chosen as initial conformations (Fig. 2.3). A CHARMM script was used to mutate residues in the wild-type tetramer simulations to the desired phosphorylated residues in order to directly compare the results of the simulations.



**Figure 2.3. Initial conformations of p53 tetramers for HyRes II simulations.** A combination of steer MD (force applied to the NTAD in the opposite direction of the DBD) and high temperature (500 K and 800 K) equilibration simulations were used to generate eight distinct conformations for the WT and pThr55 tetramer simulations. The structures are drawn in cartoon with the DBDs colored in cyan and the NTAD domains in four different colors (orange, silver, blue, and red).

Simulations were run using the HyRes II GPU-implemented force field in CHARMM. The production simulations were run at a constant temperature (300 K unless otherwise specified) using a 2 fs timestep with a Langevin thermostat. All bonds involving hydrogen atoms were constrained by SHAKE algorithm. Nonbonded interactions were smoothly switched off from 1.6 to 1.8 nm. For each condition (wild-type and phosphorylated tetramer), eight replicates of 400 ns were run starting from distinct initial conformations. The C $\alpha$  atoms of the p53 DBD (residues 95-291) were harmonically restrained with a force constant of 1.0 kcal/(mol  $\text{\AA}^2$ ). Simulations of the truncated NTAD (residues 1-61) were run using the HyRes II force field and the above conditions for 1  $\mu$ s. Two replicates were run from dissimilar initial conformations for assessment of convergence and calculation of radius of gyration (Rg) and helicity.

## Data Analysis

Analysis of the generated trajectories was performed using CHARMM and in-house Python scripts that utilize the MDTraj library<sup>85</sup>. The  $\alpha$ -helical propensity was identified by the standard Dictionary of Secondary Structure of Proteins (DSSP) protocol<sup>86</sup>. Rg values from HyRes II simulations were increased by 4 Å due to the larger coarse-grained side-chain beads compared to atomistic sampling<sup>64,65</sup>.

The interactions between the NTAD and DBD of p53 and the effect of phosphorylation were evaluated experimentally in previous studies using paramagnetic relaxation enhancement (PRE).

This data was calculated from the simulated ensemble using the following equations:

$$\frac{I_{ox}}{I_{red}} = \frac{R_2 \exp(-R_2^{sp} t)}{R_2 + R_2^{sp}} \quad (3)$$

$$R_2^{sp} = K \langle r^{-6} \rangle \left( 4\tau_c + \frac{3\tau_c}{1 + \omega_H^2 \tau_c^2} \right) \quad (4)$$

( $K = 1.23 \times 10^{-32} \text{ cm}^6/\text{s}^2$ ,  $\omega_H = 600 \text{ MHz}$ ,  $\tau_c = 3.3 \text{ ns}$ ,  $R_2 = 16 \text{ s}^{-1}$ ,  $t = 9.8 \text{ ms}$ )

Where  $r$  is the  $C\alpha - C\alpha$  distance from the other residues of the protein to the chosen paramagnetic spin label and the other parameters were chosen, as listed, to replicate experimental conditions.

The error bars shown were estimated from differences between results calculated from eight independent simulations with distinct initial conformations. Trajectory and structure visualizations were performed using Visual Molecular Dynamics (VMD)<sup>87</sup>.

Principal component analysis (PCA) was performed to assess simulation convergence and identify conformational substates using the python SciKit-learn package<sup>88</sup>. For this analysis, the coordinates of the  $C\alpha$  atoms from the protein structure were obtained for 10 ps increments from the 400 ns trajectories to collect all conformations sampled. For p53 tetramer simulations of

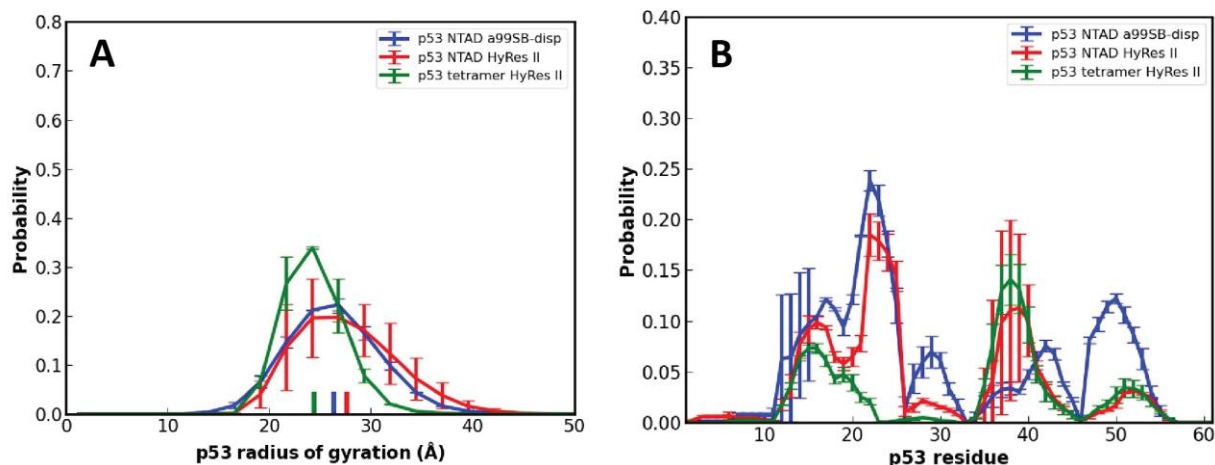
wild-type and phosphorylated constructs, all sixteen trajectories of the WT and phosphorylated replicates were aligned using the backbone atoms of the folded DBD region (residues 95–291), and the PCA was calculated based on the coordinates of C $\alpha$  atoms of the disordered N-terminal domain (residues 1–61). PCA analysis allowed for energetically favorable conformations to be identified in different tetramer constructs to provide molecular detail on the effects of phosphorylation. The free energy surfaces were derived directly from the two-dimensional (2D) probability distributions along the first two principal components (PCs).

## Results

### *Using a coarse-grained force field to study complex IDP dynamics*

The original HyRes II model does not include phosphorylated residues, so we first needed to parameterize phosphorylated serine and threonine in the coarse-grained force field (see Materials and Methods for details). We compared the distribution of the bond lengths and angles between the HyRes II phosphorylated residue beads to their atomistic equivalent in phosphorylated serine or threonine dipeptide simulations to validate that the model captures the correct values (Fig. 2.2).

Before simulating the effects of phosphorylation, we assessed the ability of the HyRes II force field to recapitulate biologically relevant p53 NTAD dynamics and structure. One of the ways we assessed this is by comparing the distributions of the radius of gyration and helicity of the NTAD from HyRes II simulations of the truncated NTAD (residues 1-61), HyRes II tetramer simulations, and a99SB-disp all-atom simulations of the isolated p53 NTAD (Fig. 2.4). Data for the a99SB-disp atomistic simulations was obtained from the original HyRes II publication<sup>65</sup>.

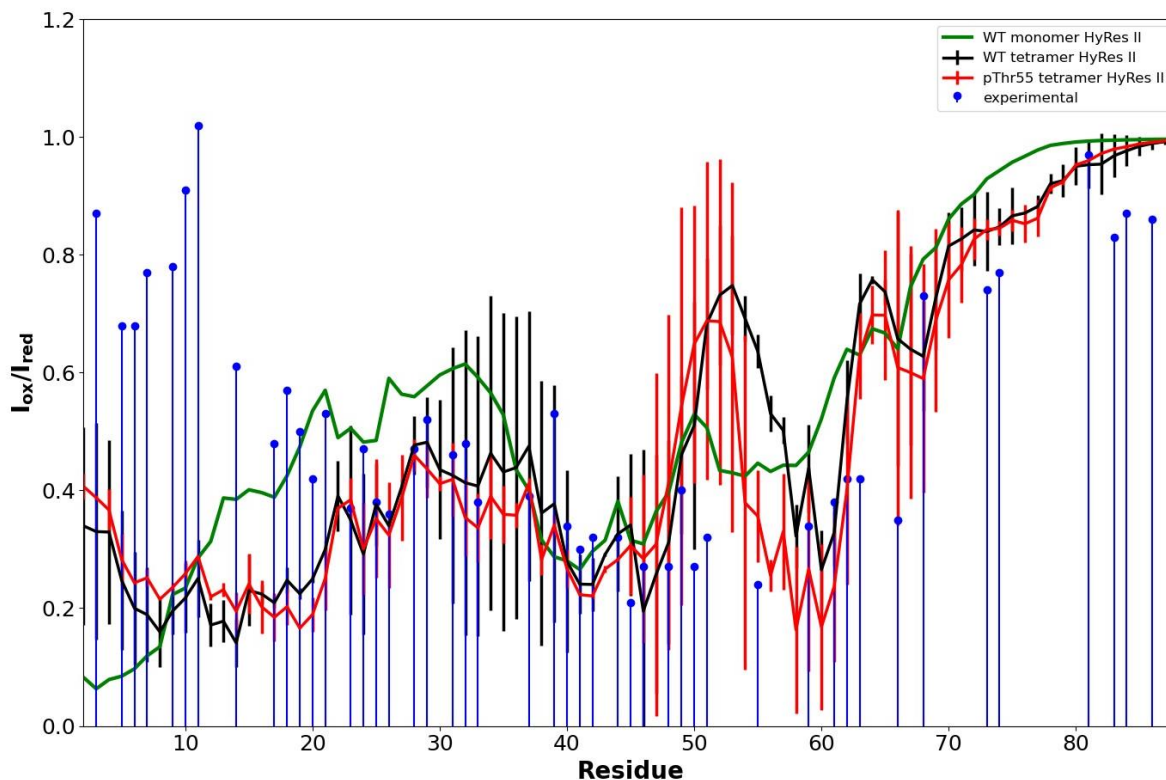


**Figure 2.4. Radius of gyration and helicity probability distributions of p53 NTAD.** (A) Radius of gyration ( $\text{\AA}$ ) of the p53 NTAD (residues 1-61) calculated from a99SB-disp atomistic simulations of the isolated NTAD (blue), HyRes II simulations of the p53 tetramer (green), and HyRes II simulations of the isolated NTAD (red). (B) Helicity profile of the p53 NTAD (residues 1-61) calculated from a99SB-disp atomistic simulations of isolated NTAD (blue), HyRes II simulations of the p53 tetramer (green), and HyRes II simulations of the isolated NTAD (red). For the HyRes II simulations, eight replicates of 400 ns each were performed at 300K, and the plotted distributions are the average of the eight replicates and the error bars represent the standard deviation between the replicates.

We next sought to determine if HyRes II can capture long-range IDP interactions by using the generated conformational ensemble from our tetramer simulations to calculate the PRE effects of a paramagnetic spin label on residue S121 (eq. 3 and 4). Residue S121 is located on the folded DBD of p53 and the PRE of residues 1-90 reports on the disordered NTADs transient interactions with the DBD, which has already been determined experimentally and can be used to validate the HyRes II simulation results.

HyRes II has previously been used to simulate the full length p53 monomer, but these simulations exhibited an overestimation of the interaction between the N-terminal region of the NTAD and the DBD. This was hypothesized to be due to the four negative residues located on the NTAD in this region that can compete electrostatically with AD1 and AD2 (Fig. 2.5, green line). We hypothesize that simulating the p53 tetramer will reduce this artifact by eliminating

some of the accessible DBD surface area and adding additional NTADs that are negatively charged and can add a repulsive force when competing for DBD binding. The PRE calculated for the WT tetramer from our HyRes II simulations is shown as a black line in Figure 2.5 and is plotted with experimental values for comparison.



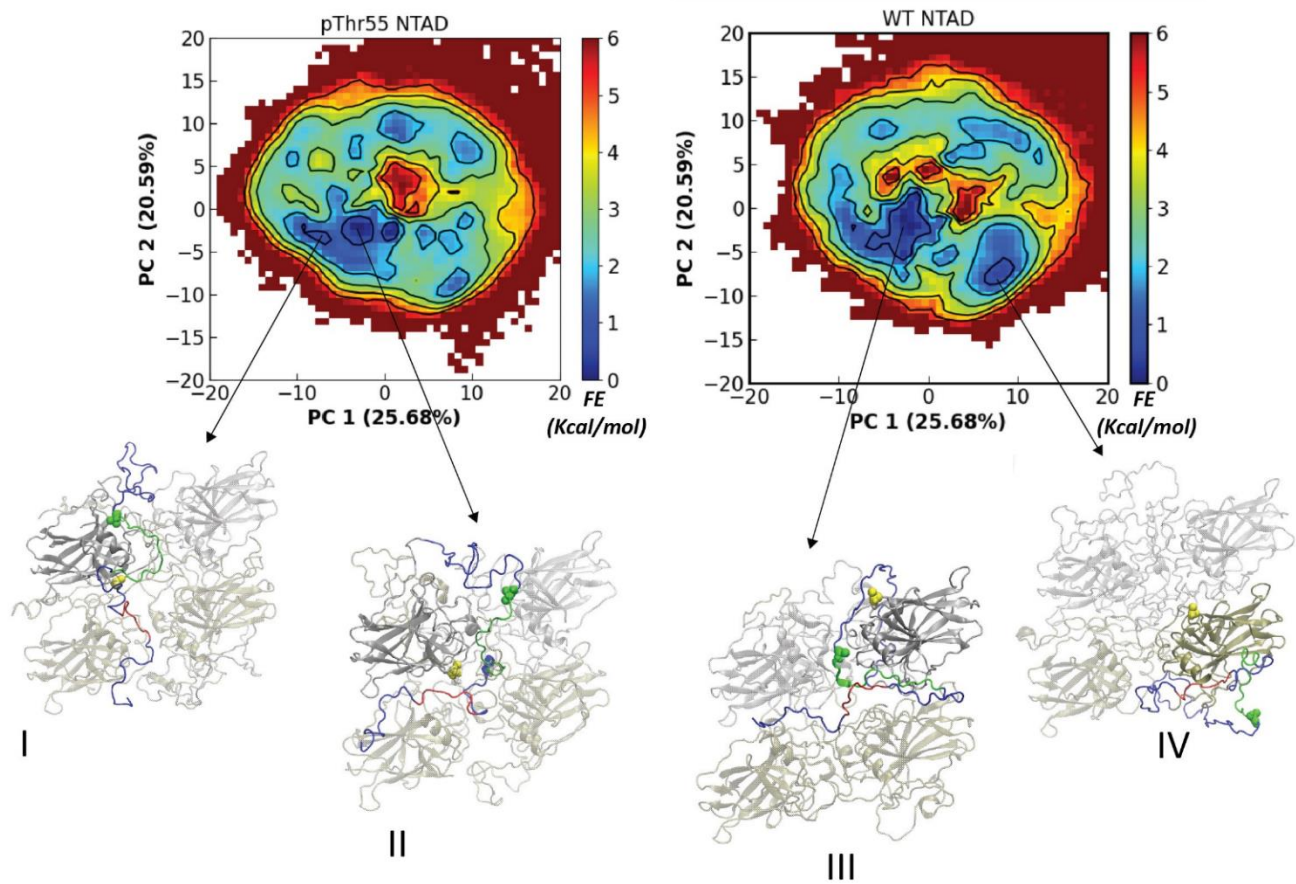
**Figure 2.5. PRE effects of the NTAD induced by a spin label on S121 of p53 DBD.** PRE of p53 residues 1-90 from experimental results (blue bars) and calculated from HyRes II monomer (green line) and tetramer (black and red lines) simulations induced by paramagnetic spin labeling on residue S121 of the p53 DBD. The black and green lines and blue bars show the results for WT p53 and the red line shows the PRE for the phosphorylated p53 tetramer. The PRE calculated from tetramer simulations is averaged over eight replicates and accounts for all NTAD contacts between segments with the spin label. The error bars reflect the difference in PRE calculated with two sets of four replicates.

We next simulated the p53 tetramer with residue threonine 55 phosphorylated, which has been shown experimentally to increase the NTAD's interactions with the DBD<sup>50</sup>. We used the same

eight initial conformations used for the WT tetramer simulations so the results could be directly compared to observe the effect of phosphorylation on the calculated PRE, which is shown in Figure 2.5 in red.

*Investigating the effects of phosphorylation on the interactions of the N-terminal domain in the p53 tetramer*

Once the ensembles generated from the HyRes II simulations of the WT and phosphorylated p53 tetramer were validated using experimental results, we were able to further analyze the generated trajectories to gain insight into the effects of phosphorylation on the NTAD structure and interactions with the DBD. Figure 2.6 shows the principal component analysis (PCA) calculated using the C $\alpha$  coordinates of the NTAD (residues 1-60), where the free energy basins in the PC1-PC2 space correspond to frequently visited conformations of the NTAD that can be further investigated for trends related to the p53 autoinhibition mechanism.



**Figure 2.6. PCA of conformations of the WT and phosphorylated p53 NTAD.** Calculated using the  $C\alpha$  coordinates of residues 1-60. Snapshots from eight replicates of 400 ns HyRes II simulations of the phosphorylated threonine 55 tetramer and from eight replicates of 400 ns HyRes II simulations of the WT tetramer were combined to derive the first two PCs and calculate the projections (pThr55 shown on the left, WT on the right). The conformations corresponding to two of the basins for the pThr55 tetramer (I and II) and the WT tetramer (III and IV) are shown with the DBDs (residues 90-291) colored with one dimer in silver and the other in tan, the highlighted NTAD in blue (residues 1-90), AD1 in red (residues 18-26), AD2 in green (residues 40-55), residue Thr55 in green and vdW representation, and residue Ser121 in yellow vdW representation.

Comparing the PCA projection of the WT and phosphorylated threonine 55 tetramer simulations reveals shifts in the energetic basins that can be analyzed to determine the corresponding conformation. Comparing these structures allows for molecular detail of the NTAD interaction with the DBD, which is known to be strengthened through phosphorylation, and can provide

more information about the modes of AD1 and AD2 interaction with the DNA binding surface, as well as other DBDs within the tetramer structure. There is a shift in the conformations sampled between the WT tetramer and the phosphorylated construct at residue 55, and the conformations corresponding to two major basins from both PCA projections are shown in Figure 2.6 and labeled I-IV.

## Discussion

### *HyRes II can capture nontrivial dynamics of IDPs and the effect of PTMs on their structure*

Here, we assess the ability of the HyRes II force field to capture complex, biologically relevant dynamics of the p53 tetramer. p53 has previously been studied using experimental techniques such as NMR, and the p53 monomer has been simulated using all-atom force fields as well as HyRes II, providing data to validate our findings<sup>50,53,65</sup>.

The probability distributions of the NTAD radius of gyration and helicity generated from our HyRes II simulations of the truncated NTAD and tetramer agree reasonably well with the distributions calculated from all-atom simulations (Fig. 2.4). The average radius of gyration of the NTAD predicted from our tetramer simulations, 24.4 Å, is also consistent with the experimentally determined Stokes radius of the NTAD peptide (residues 1-73), 23.8 Å<sup>89</sup>.

The helicity distribution and radius of gyration generated from the HyRes II tetramer simulations seem to underestimate both the helicity in some regions (residues 10-35) and the average radius of gyration compared to the atomistic results (Fig. 2.4B, green line). This could be due to the binding interaction of the NTAD to the DBD that may affect secondary structure and overall size of the NTAD, and this NTAD-DBD interaction is not present in the truncated atomistic

simulations (residues 1-61). The helicity and radius of gyration calculated from HyRes II simulations of the truncated NTAD are able to better recapitulate secondary structure and average radius of gyration compared to atomistic results, indicating that the discrepancies in the data calculated from the tetramer NTAD is likely due to the DBD interaction rather than a reflection of the force-field or sampling quality (Fig. 2.4B, red line). These results show that our coarse-grained simulations can capture overall NTAD size propensity and relevant IDP secondary structure in the correct regions of the NTAD.

Figure 2.5 shows the PRE effects on the NTAD (residues 1-90) induced by a paramagnetic spin label on residue S121, and compares the results calculated from experimental data to results back calculated from HyRes II simulations of the p53 monomer, WT tetramer, and tetramer with phosphorylation on threonine 55. All of the results from HyRes II simulations show significant agreement with experimental results, providing evidence that HyRes II is able to capture the long-range effects of interactions between the p53 NTAD and DBD.

Furthermore, the results from HyRes II tetramer simulations are able to more accurately capture the PRE effects on the N-terminal of the NTAD (residues 1-15) than the previous monomer simulations. As discussed previously, the HyRes II p53 monomer simulations overestimate the interaction of the NTAD N-terminal region with the DBD, which was hypothesized to be a result of the strong electrostatic interaction between the negative residues of the N-terminal region and the positive DBD. We hypothesized that simulating the full-length p53 tetramer would decrease the favorability of this interaction by adding competing negative charges and repulsive forces in the form of additional NTADs and by occluding some regions of the DBD that become inaccessible in the tetramer conformation. We indeed observe these effects, and the PRE data calculated from the WT tetramer simulations has a root-mean-square deviation (RMSD) of 0.28

from the experimental results, which is an improvement from the RMSD of 0.30 for the monomer PRE data compared to the same experimental dataset. Specifically, the far N-terminal region of the NTAD appears to have decreased interactions with the DBD in the tetramer structure that more closely recapitulates the experimental results than the previous monomer simulations (Fig. 2.5, black line).

In addition to the results generated from the WT tetramer simulations, HyRes II is able to accurately model the effects of phosphorylation on the NTAD, which is also reflected in the PRE data shown in Figure 2.5. Phosphorylation at residue Thr55 is known to increase the NTAD interaction with the DBD through the additional negative charge, and this mechanism is thought to prevent nonspecific p53-DNA binding in the absence of cellular stress<sup>50</sup>. We also observe this effect in our tetramer simulations, and the pThr55 tetramer (Fig. 2.5, red line) shows a significant increase in the paramagnetic broadening and DBD interaction within the AD2 motif (residues 40-55) compared to the WT tetramer (black line). These results provide evidence that the HyRes II force field is able to generate biologically relevant IDP ensembles that agree with experimental data and can capture the subtle effects of PTMs such as phosphorylation on the conformational ensemble.

#### *Phosphorylation-induced shifts in p53 conformational ensemble*

Experimental studies using phosphorylated p53 have shown that there is a cooperative DNA disassociation effect at high p53 concentrations that seems to be induced by the formation of the tetramer. This effect was found to be more prominent than the reduction of dimer-DNA  $K_D$  due to phosphorylation, suggesting that the NTAD interacts significantly with the neighboring dimer

in the tetramer structure, and this effect is more important for the reduction of nonspecific DNA binding<sup>50</sup>.

To investigate the effects of phosphorylation on the p53 tetramer ensemble, we simulated the full-length tetramer using eight replicas initiated from different starting conformations to sample the full conformational landscape (Fig. 2.3). We used the same initial conformations generated for the WT replicas and mutated Thr55 to the phosphorylated residue to probe the specific effects of phosphorylation. Once the WT and pThr55 ensembles were generated, we analyzed the conformational space sampled throughout their eight replicas projected on the same first two PCs derived from all of the combined conformations sampled (Fig. 2.6). This allowed us to visualize any shifts in the energetic space that would indicate differences in the favored conformations between the WT and phosphorylated construct.

There is a shift in the conformational energetic landscape induced by phosphorylation, and figure 2.6 highlights four prominent structures sampled by the pThr55 tetramer (I and II) and the WT tetramer (III and IV). The conformations shown seem to be sampled by both tetramer constructs, but it appears that phosphorylation increases or decreases the probability of some key states.

State III and IV are sampled with a higher probability (indicated by a lower free energy) in the WT ensemble than in the phosphorylated ensemble, and these states show the NTAD interacting *in cis* with the DBD. States III and IV show the AD1 and AD2 motifs flanking the DNA binding surface within their respective dimers.

This is distinct from the states that are more prominently sampled in the phosphorylated tetramer and states I and II show an increased interaction of the AD2 motif directly with the DNA binding surface. The AD2 motif is still interacting *in cis* with the tetramer structure, but in both conformations, this allows the tail of the NTAD to project across the dimers to interact with the

DBDs *in trans*. This suggests an interesting mechanism where the AD1-DBD interaction occurs *in trans*, effectively competing with DNA for DBD binding and preventing nonspecific DNA interactions. These conformations would agree with the mechanism proposed by Sun and colleagues in which the NTAD interacts preferentially *in trans* to dissociate p53 from DNA in the tetramer conformation.

### *Conclusions*

Here, we demonstrate that HyRes II coarse-grained simulations of the p53 tetramer are able to capture biologically relevant IDP dynamics that can recapitulate experimental PRE data. The simulations of the full p53 tetramer improve the PRE results compared to previous monomer simulations by adding additional repulsive forces through multiple negatively charged NTADs, which better captures the transient interactions between the NTAD and DBD. We also show that the HyRes II force field is able to capture the subtle effect of phosphorylation of threonine 55 on the interaction between the NTAD and the DBD.

By further analyzing the conformations sampled in these trajectories, we show the molecular-level detail of the autoinhibition mechanism induced by Thr55 phosphorylation and see an increased interaction of the NTAD *in trans* in the phosphorylated tetramer simulations compared to the WT tetramer. This mechanism of NTAD-DBD interaction *in trans* was proposed experimentally, illustrating the ability of our model to simulate these biologically relevant mechanisms of IDP regulation. These promising results indicate the use of the HyRes II force field for continued studies of the effects of PTMs on p53 regulation, as well as studies of other IDPs that are prevalent in human health and disease.

## CHAPTER 3

### SINGLE-MOLECULE DETECTION OF THE INTRINSICALLY DISORDERED P53 NTAD USING THE MSPA NANOPORE

#### Abstract

Intrinsically disordered proteins (IDPs) lack stable tertiary structure under physiological conditions and can therefore act as signaling hubs in the cell via interactions with various binding partners and pathways. p53 is a protein with an extensive disordered region that enables it to modulate crucial cell functions. It is a transcription factor that is mutated in over 50% of cancers due to its vital role in arresting the cell cycle after DNA damage and other forms of cellular stress. Its prominent role in the development of cancer means characterizing the structure and interactions of its disease-associated mutants is an ongoing area of study. Because of its activity as a tumor suppressor protein, p53 is an attractive target for the development of cancer therapeutics that aim to stabilize and increase its wild-type function. One of the ways that this could be achieved is through targeting p53's interactions with the protein binding partners that modulate its activation. In an unstressed cell, p53 is continually targeted for degradation by E3 ubiquitin ligase HDM2, and this interaction is disrupted under stress to activate p53's tumor suppressing activity. A small molecule commonly found in green tea, epigallocatechin gallate (EGCG), has been found to bind the disordered N-terminal domain (NTAD) of p53 and disrupt HDM2-p53 binding, effectively stabilizing and activating p53. Nanopore technology is an emerging technique for capturing the heterogeneous events of intrinsically disordered proteins and for identifying protein-protein interactions. In this study, we show that trapping experiments using the *Mycobacterium smegmatis* porin A (MspA) nanopore are able to capture signals from

the NTAD and to distinguish the wild-type protein events from a cancer-associated mutant. The nanopore can also capture the NTAD's binding interaction small molecule EGCG, showing promising evidence that the MspA nanopore can be used to characterize and screen IDP disease-related mutants and potential therapeutics.

## **Introduction**

Intrinsically disordered proteins (IDPs) are crucial components of signaling pathways within the cell due to their ensemble of marginally stable conformations under physiological conditions that enables them to act as an interaction hub for multiple binding partners<sup>1-5</sup>. This structural instability also allows IDPs to undergo conformational changes due to small perturbations such as post-translational modifications (PTMs) and mutations<sup>6-8</sup>. As such, many important signaling proteins contain significant disorder in their structure, meaning the dysfunction or misregulation of these proteins is involved in many different diseases<sup>9-11</sup>.

p53 is a transcription factor with an intrinsically disordered N-terminal transactivation domain (NTAD) that activates the expression of numerous genes with roles in cellular stress responses and apoptosis. p53's role as a tumor suppressor has earned it the nickname "guardian of the genome," and as such, its dysfunction can be linked to many types of human cancers<sup>15-17</sup>.

With such a prominent role in disease, p53 has long been a therapeutic target, and candidate drugs seek to reactivate mutant p53 or stabilize and increase wild-type p53 function in tumor cells<sup>33,90,91</sup>. One approach for stabilizing and activating wild-type p53 is targeting its interaction with HDM2. HDM2 is a ubiquitin ligase that targets p53 for degradation in the absence of cellular stress<sup>37,38,45</sup>. This interaction is mediated by binding to p53's NTAD, which is disrupted

during the cellular stress response, activating p53 function<sup>43</sup>. Inhibitors of the HDM2/p53 NTAD interaction can thus have therapeutic implications by preventing the degradation of p53 and increasing its tumor-suppressing function.

Epigallocatechin gallate (EGCG) is a small molecule found in green tea that has been shown to have antitumor properties via apoptosis and inhibiting cell growth, but the exact molecular mechanism of this function was unclear until recently<sup>92</sup>. Techniques such as nuclear magnetic resonance (NMR), small angle x-ray scattering (SAXS), and molecular dynamics simulations were used to determine that EGCG binds to the NTAD of p53 with micromolar affinity and inhibits HDM2 binding, preventing the ubiquitination and degradation of p53<sup>51</sup>.

The identification of this small molecule inhibitor is an important step in the development of therapeutics that target p53's interactions, but the techniques used to characterize this molecule such as NMR and SAXS are ensemble averaging techniques which are not well suited for determining biologically-relevant substates of IDPs<sup>8,58-60</sup>. Techniques such as paramagnetic relaxation enhancement (PRE) and single-molecule Förster resonance energy transfer (sm-FRET) also require the addition of external probes that could interfere with the native physiological ensemble of an IDP<sup>61</sup>.

Nanopores are an emerging label-free, single-molecule technique that are uniquely suited for capturing the heterogeneity of IDP events, and have been previously used to characterize protein analytes and their binding partner interactions<sup>1,66,67,72,93</sup>. Nanopores are highly sensitive and can allow for the characterization of multiple event types for a protein, which can reveal differences in the ensembles of an IDP and its disease-related mutants. They can also detect real-time binding partner interactions and the changes of these interactions due to the addition of small molecules<sup>94</sup>. The MinION sequencer uses nanopore technology for cost-effective, portable, and

real-time DNA sequencing, which suggests an exciting application of nanopores as high-throughput protein sensing or drug screening tools using a similar array platform<sup>76,95</sup>.

Here, we use the biological nanopore *Mycobacterium smegmatis* porin A (MspA) to trap and observe p53's NTAD. In previous studies, the cancer associated NTAD mutant N29K/N30D has been found to have more compact substates than the WT NTAD, and we use this mutant to show the sensitivity of the nanopore to subtle changes in the NTAD's conformational ensemble<sup>19</sup>. We further show evidence for nanopore detection of the wild-type NTAD binding to small molecule EGCG. These experiments show that nanopores are a sensitive and label-free method of studying IDPs and their interactions and provide evidence that the MspA nanopore may be a viable tool for disease-related mutant characterization and small molecule screening.

## **Materials and Methods**

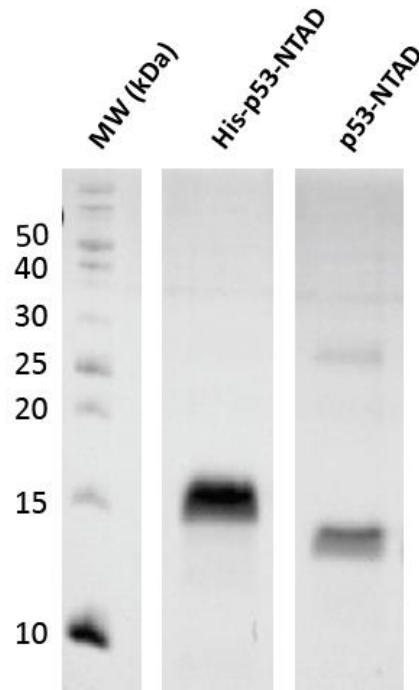
### *Expression and Purification of p53-NTAD (1-73) and N29K/N30D mutant*

A pET28a vector with sub-cloned cDNA of *Hsp53*-(1-73) (72R) and the N-terminal His-tag and thrombin cleavage site (Addgene plasmid #62082) transformed into BL21(DE3) low-background strain (LOBSTR, Kerafast, Boston, Massachusetts) chemically competent cells were a gift from Dr. Michal Zolkiewski. The transformed cells were grown in 4L total volume cultures of LB media and Kanamycin, shaking at 200 rpm at 30°C to OD<sub>600</sub> = 0.6. The cells were then induced with 1 mM Isopropyl-b-D-thiogalactoside (IPTG, Sigma) and allowed to express overnight at 16°C.

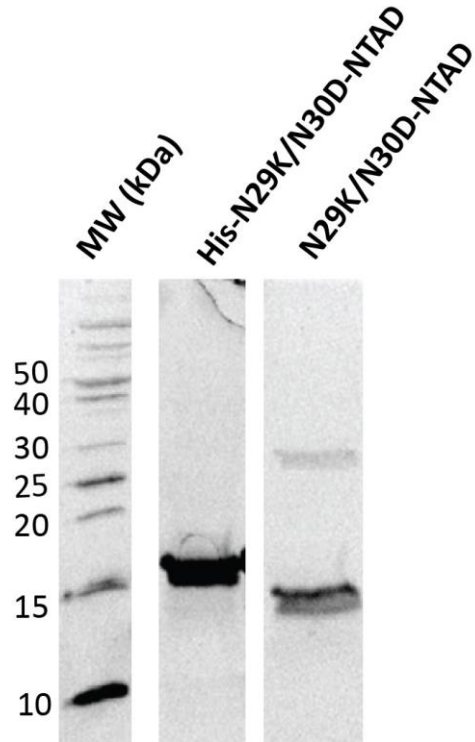
The cells were pelleted via centrifugation at 4,000 rpm for 30 minutes at 4°C. The pellet was resuspended in lysis buffer (50 mM Tris HCl pH 8.0, 60 mM EDTA, 0.5 mM PMSF) at a ratio of

30 mL buffer per liter of culture. The cells were lysed via sonication and the lysate was clarified by centrifuging at 13,000 rpm for 30 minutes at 4°C. p53-NTAD was purified from the clarified lysate using Nickel Nitrilotriacetic Acid (Ni-NTA, GE healthcare Ni Sepharose 6 Fast Flow resin) column resin in a gravity column. The column resin and lysate were incubated together on ice for 20 minutes before allowing the lysate to flow through the column. The column was then washed with ten column volumes of wash 1 buffer (150 mM NaCl, 50 mM Tris-HCl (pH 8), and 5% glycerol), and ten column volumes of wash 2 buffer (10mM imidazole, 150 mM NaCl, 50 mM Tris-HCl (pH 8), and 5% glycerol). The protein was eluted with 10 mL of elution buffer (250 mM imidazole, 150 mM NaCl, 50 mM Tris-HCl (pH 8), 5% glycerol).

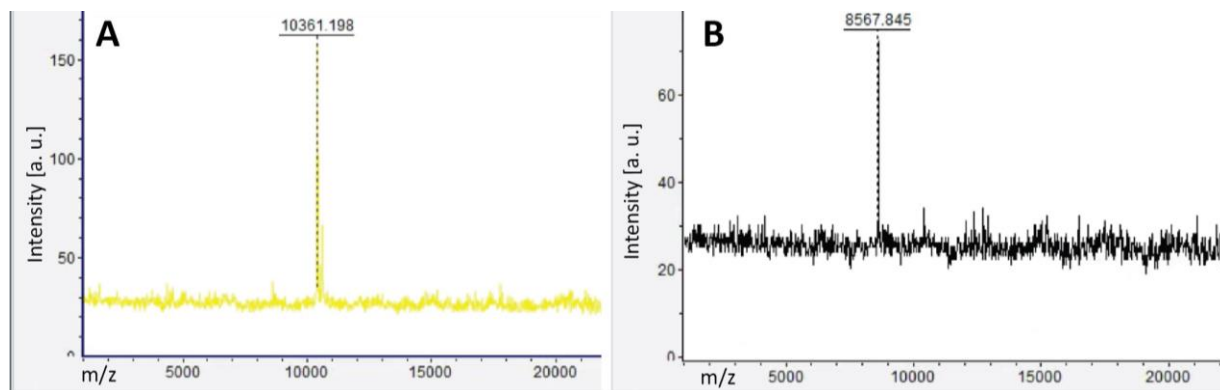
The p53-NTAD-containing fractions were then dialyzed overnight at 4°C into 300 mM NaCl, 50 mM Tris HCl, 5% glycerol, pH 8. A portion of the fractions were incubated with 100 units of thrombin protease during dialysis overnight at 4°C for His-tag cleavage. The dialyzed fractions were then run through a Superdex 75 Increase size-exclusion chromatography column (Cytiva Life Sciences) using 300 mM NaCl, 50 mM Tris HCl, 5% glycerol, pH 8 buffer to isolate p53-NTAD and mutants from aggregates and background contaminants in the sample (Fig. 3.1 and 3.2). The molecular weights of both WT constructs (His-tagged and untagged) were assessed with MALDI-TOF (Figure 3.3). Purified p53-NTAD fractions were then flash-frozen with liquid nitrogen prior to storage at -80°C until their use in electrophysiology experiments. Unless otherwise noted, the NTAD construct used for the following electrophysiology experiments in the NTAD (residues 1-73) with N-terminal 6xHis-tag.



**Figure 3.1. WT NTAD purification SDS-PAGE gel.** The His-tagged WT NTAD was first isolated using a Nickel Nitrilotriacetic Acid (Ni-NTA) column. A portion of the concentrated fractions were then incubated with thrombin protease at 4°C overnight for His-tag cleavage. Both His-cleaved and uncleaved fractions were then run through a Superdex 75 Increase column for separation from the protease and other protein contaminants. The gel shows fractions of His-tagged and untagged NTAD constructs after Superdex purification that were then flash frozen and stored at -80°C.



**Figure 3.2. N29K/N30D NTAD purification SDS-PAGE gel.** The His-tagged N29K/N30D NTAD was first isolated using a Nickel Nitrilotriacetic Acid (Ni-NTA) column. A portion of the concentrated fractions were then incubated with thrombin protease at 4°C overnight for His-tag cleavage. Both His-cleaved (p53 residues 1-73) and uncleaved (p53 residues 1-73, thrombin cleavage site, 6xHis tag) fractions were then run through a Superdex 75 Increase column for separation from the protease and other protein contaminants. The gel shows fractions of His-tagged and untagged NTAD constructs after Superdex purification that were then flash frozen and stored at -80°C.



**Figure 3.3. Raw data from MALDI-TOF analysis of p53 NTAD.** (A) NTAD with N-terminal His tag (p53 residues 1-73, thrombin protease cleavage site, 6xHis-tag, 94 residues) (B) untagged p53 NTAD (p53 residues 1-73). The molecular weight (Da) of the main peak is noted on the spectra.

Epigallocatechin gallate (EGCG) used for electrophysiology experiments was purchased from ThermoFisher.

#### *Expression and Purification of M2-MspA-NNN*

The M2-MspA-NNN recombinant plasmid with 6xHis tag (adapted from PDB: 1UUN with D90N, D91N, D93N, D118R, D134R, E139K mutations) was transformed into BL2 (DE3) pLysS competent *E. coli* cells for expression<sup>96</sup>. Cells were grown in 1L LB media, shaking at 200 rpm at 30°C to OD<sub>600</sub>=0.7. The culture was then induced with 1 mM isopropyl β-D-thiogalactoside (IPTG) and allowed to express overnight at 16 °C. Cells were pelleted by centrifugation at 4,000 rpm for 15 minutes at 4 °C and resuspended in 30 mL lysis buffer (100 mM Na<sub>2</sub>HPO<sub>4</sub>/NaH<sub>2</sub>PO<sub>4</sub>, 0.1 mM EDTA, 150 mM NaCl, pH 6.5). The cellular suspension was heated to 60 °C for 10 minutes and then centrifuged at 13,000 rpm for 20 minutes at 4 °C. The cell pellet was then resuspended in solubilize buffer (100 mM Na<sub>2</sub>HPO<sub>4</sub>/NaH<sub>2</sub>PO<sub>4</sub>, 0.1 mM EDTA, 150 mM NaCl, 0.5% Genapol X-80, pH 6.5) at a ratio of 30 mL per 500 mL of cell

culture. The solubilized cells were then centrifuged at 13,000 rpm for 20 minutes at 4 °C and the supernatant was additionally clarified by passing it through a 0.22 µm syringe filter.

The filtered supernatant was then loaded onto Nickel Nitrilotriacetic Acid (Ni-NTA) resin in a gravity column and 6xHis-tagged MspA was incubated with the resin for 20 minutes on ice to fully bind the column. The column was then washed with 10 column-volumes of Wash 1 buffer (0.5 M NaCl, 20 mM HEPES, 0.5% Genapol X-80, pH=8.0) and subsequently Wash 2 buffer (50 mM imidazole, 0.5 M NaCl, 20 mM HEPES, 0.5% Genapol X-80, pH=8.0). MspA was then eluted from the column with 10 mL of Elution buffer (200 mM imidazole, 0.5 M NaCl, 20 mM HEPES, 0.5% Genapol X-80, pH=8.0).

The MspA elution fractions were subsequently analyzed via SDS-PAGE on 7.5% acrylamide gels. The MspA oligomer gel band was cut out and transferred to a 2 mL tube containing 500 µL of Gel-extraction buffer (150 mM NaCl, 50 mM Tris-HCl, 0.5% Genapol X-80, pH 7.5). The MspA oligomer gel band was homogenized into the extraction buffer using a glass rod. The tube with the gel-homogenate was then shaken at 50 rpm for 2 hours at 23 °C. The mixture was centrifuged at 14,000 rpm for 30 minutes at 4°C. The supernatant containing the pure MspA oligomer was collected and stored at -80 °C.

#### *Single-Channel Electrophysiology Experiments*

Single-channel electrophysiology experiments were performed in an apparatus containing two, 1 mL chambers separated by a 25 µm thick Teflon film. An aperture of approximately 60 µm diameter was created near the center of the film by first puncturing with a needle and then smoothing the edges with an electric spark. n-hexadecane in pentane (2% v/v) solution was applied directly to the aperture. After allowing the pentane to evaporate, each chamber was filled

with 900  $\mu\text{L}$  of buffer (150mM NaCl, 20 mM HEPES, pH 7.4). An Ag/AgCl electrode was immersed in each chamber with the cis chamber grounded. 1,2-Diphytanoyl-sn-glycerol-3-phosphocholine (Avanti Polar Lipids, USA) dissolved in pentane to a concentration of 10 mg/ml was deposited on the surface of the buffer in both chambers. The lipid bilayer was formed by pipetting the liquid level up and down across the aperture and capacitance was monitored to observe the formation of a functional membrane. MspA was added from the cis chamber and inserted into the bilayer under an applied voltage of +300 mV. Cis-inserted MspA pores exhibited an open pore conductance of  $\sim 52\text{-}55$  pA at the primary recording voltage of +200 mV, and  $\sim 60$  pA at -200 mV.

p53 NTAD and EGCG were also added to the cis chamber for nanopore analysis. Following each analyte addition, a 100  $\mu\text{L}$  volume was pipetted 15x in to mix the solution. For analysis of event concentration dependence, analytes were sequentially added and recorded in the pore until the completion of the titration. Voltage potentials were applied across the bilayer and ionic current through nanopore was monitored in voltage-clamp mode by an integrated patch clamp amplifier (Axopatch 200B, Molecular Devices). The signal was acquired by an analog-to-digital converter (Digidata 1440A, Molecular Devices) at a sampling rate of 50 kHz after processing with a 4-pole lowpass Bessel filter at 2 kHz. Data was recorded by Clampex 11.2 software (Molecular Devices). Experiments were conducted at 23  $^{\circ}\text{C}$ .

### *Data Analysis*

Recordings were performed using a standard gap-free protocol. The recordings were analyzed for events using single-channel search in the Clampfit 11.2 software, which extracts parameters such as dwell (ms), amplitude of the event relative to the defined baseline (pA) the standard deviation of the event (pA), and the interevent time (ms). Residual current ( $I_{res}$ ) was defined as a fraction

of the open pore current using the blocked current amplitude and the open pore amplitude as:

$$I_{res} = I_{Blocked}/I_{open}.$$

To obtain  $\tau_{on}$ , the interevent times of the events identified with single-channel search were binned and fitted to a single-exponential function to extract  $\tau$  at each NTAD concentration. For the plots of  $I_{res}$  vs log dwell, the same amount of trace time was analyzed for each analyte condition that is shown in direct comparison (Figure 3.4, Figure 3.7 A and B and Figure 3.9 A and B, respectively).

#### *HyRes II simulations and analysis*

Simulations were performed using the construct of p53 NTAD that was studied experimentally, p53 residues 1-73 with thrombin cleavage site and N-terminal 6xHis-tag (Addgene plasmid #62082). The sequence was read into the hybrid-resolution (HyRes) II force field and two initial conformations were generated using a 2 ns equilibration simulation at 600 K. The two initial conformations generated for the WT simulations were mutated to the N29K/N30D sequence and used for the mutant production simulations for consistency.

Simulations were run using the HyRes II GPU-implemented force field in CHARMM<sup>65</sup>. The production simulations were run at a constant temperature (300 K unless otherwise specified) using a 2 fs timestep with a Langevin thermostat. All bonds involving hydrogen atoms were constrained by SHAKE algorithm. Nonbonded interactions were smoothly switched off from 1.6 to 1.8 nm. For each condition (WT and N29K/N30D NTAD), two replicates of 400 ns were run starting from distinct initial conformations. Analysis of the generated trajectories was performed using CHARMM and in-house Python scripts that utilize the MDTraj library<sup>85</sup>. Radius of

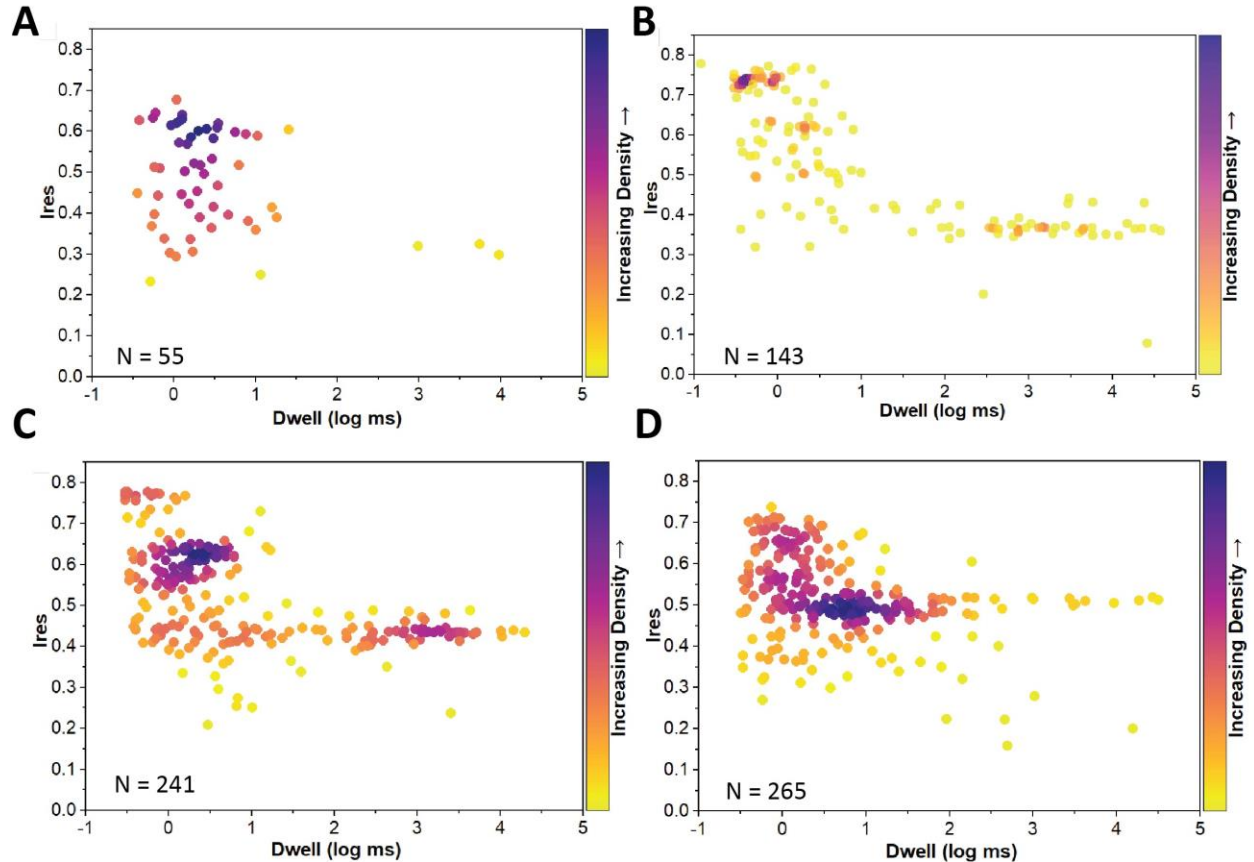
gyration ( $R_g$ ) values from HyRes II simulations were increased by 4 Å due to the larger coarse-grained side-chain beads compared to atomistic sampling<sup>64,65</sup>.

## Results

### *Label-free detection of disordered p53 NTAD*

To detect the intrinsically disordered p53 NTAD with a nanopore sensor, we first needed to select a nanopore that would be suited to trap and characterize this protein. We selected MspA because of its conical geometry that has a ~5 nm opening and a 1.2 nm constriction region. MspA is a relatively small nanopore compared to other common nanopore sensors such as Cytolysin A (ClyA), meaning MspA has a more conducive geometry for trapping small analytes such as the NTAD (~10 kDa), whereas the peptide may translocate through a larger pore like ClyA. Longer trapping events are ideal because increased dwell time can provide more information to characterize and study the trapped protein compared to translocation events.

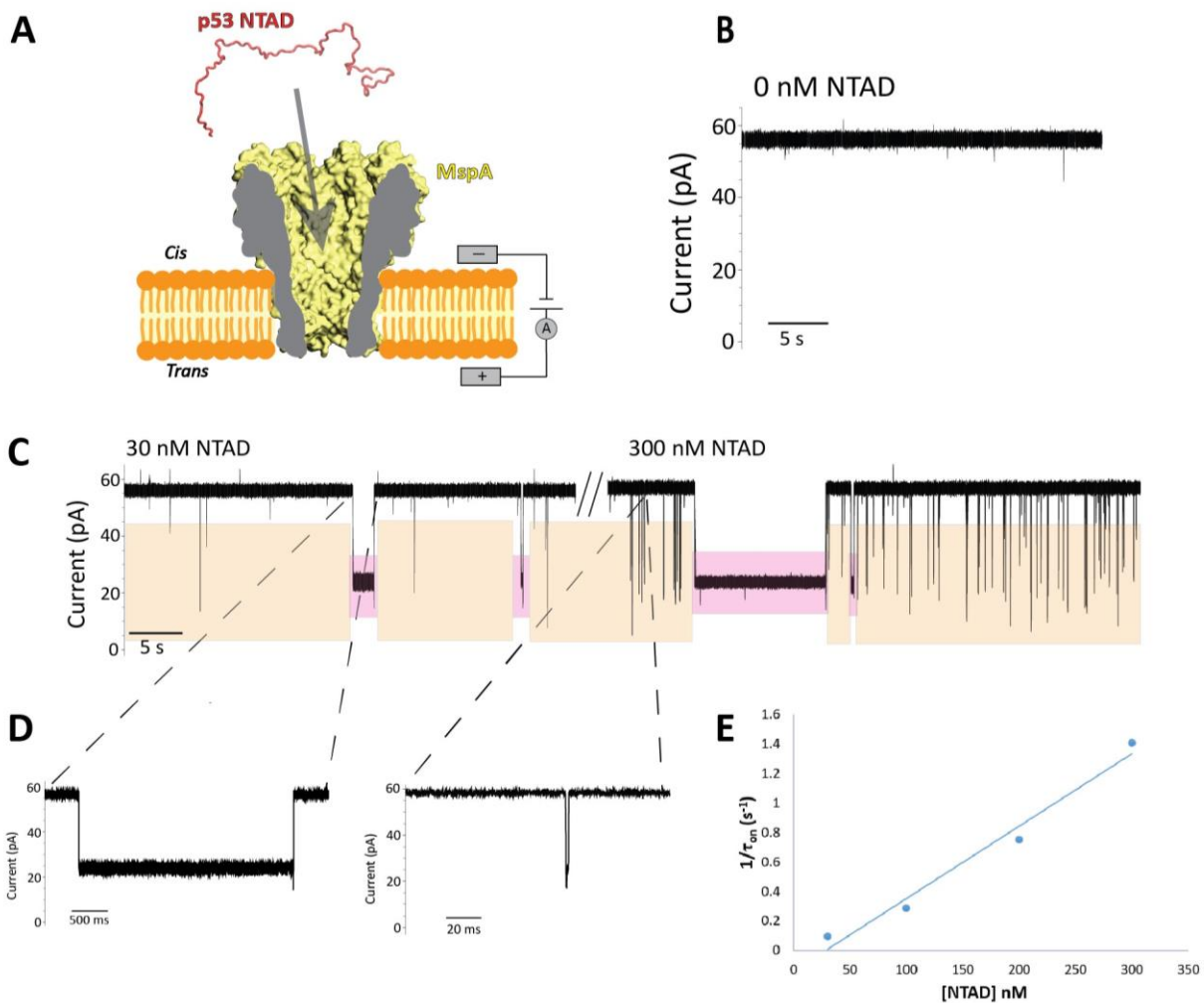
For nanopore detection of the p53 NTAD, we first characterized the event signal of the peptide in single-channel recordings using the MspA nanopore under an applied positive voltage (Fig. 3.5 A). Because of the NTAD construct's negative charge (-16 at pH 7.4), a positive *trans* applied voltage will induce an electrophoretic force on the peptide and pull it toward the opening of cis-inserted MspA. To determine the exact primary recording voltage for the following assay, a voltage study was performed on the NTAD from 100 mV – 250 mV, and +200 mV chosen for its increased event frequency while maintaining bilayer stability (Fig. 3.4).



**Figure 3.4. WT NTAD voltage study Ires vs dwell time plots.** Voltage study of 200 nM WT NTAD at (A) 100 mV, (B) 150 mV, (C) 200 mV, and (D) 250 mV. The events at each voltage were detected using Clampfit 11.2 single-channel search from 10-minute portions of the traces and the number of events at each voltage is noted on the plots. The events are characterized by Ires and the log of the dwell time (log ms) and colored by increasing event density. The experiments were done in buffer conditions of 150 mM NaCl, 20 mM HEPES at pH 7.4.

To distinguish the NTAD events from pore noise and self-gating, the empty MspA nanopore was recorded at our primary recording voltage (+200 mV) and buffer conditions. Under the buffer conditions used for these experiments (150 mM NaCl at pH 7.4 to mimic physiological conditions) and positive *trans* voltage, the MspA nanopore exhibits minimal self-gating and an open-pore current of ~52-55 pA (Fig. 3.5 B). NTAD peptide was then titrated into the *cis* chamber starting at a concentration of 30 nM and ending at 300 nM to assess the concentration dependence of events resulting from the addition of the NTAD (Fig. 3.5 C).

The events were then analyzed and characterized by properties such as residual pore current ( $I_{res}$ ) and dwell time (Fig. S1). Two main event types were observed for the NTAD at +200 mV, transient events (0.3 ms – 20 ms) that range in  $I_{res}$  from 0.7-0.2, and long events (20 ms -10s) at 0.4-0.5  $I_{res}$ . The two main event types are shown highlighted in pink and orange in Figure 3.5 C and as close-ups in Figure 3.5 D. At +200 mV, some of the long events remain in the nanopore for minutes and must be terminated by reversing the applied voltage to a negative potential which presumably allows the NTAD to leave the pore and returns the pore to baseline. Those permanent blockage events are excluded from the dwell time analysis in the following assays.

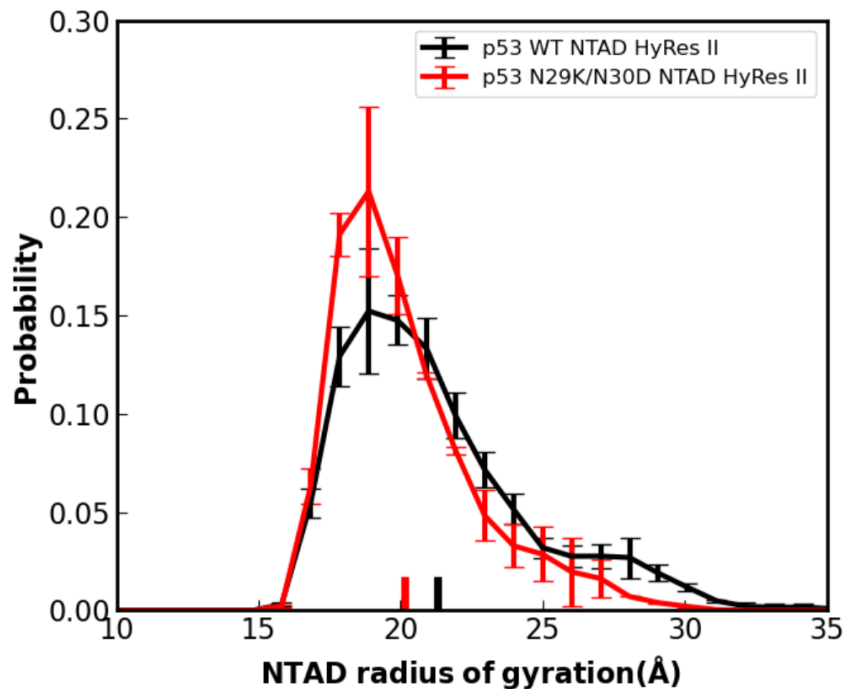


**Figure 3.5. MspA can capture p53 NTAD events in a concentration-dependent manner.** (A) Schematic of MspA (yellow) inserted in the lipid bilayer (orange) under positive applied *trans* voltage. The p53 NTAD (red) is shown outside of the pore. (B) Representative MspA current traces with 0 nM analyte (WT NTAD) added to the recording chamber. (C) Representative MspA current traces when 30 nM and 300 nM NTAD are added to the recording chamber. Events are color-coded by type according to dwell time, with short events (0.3 - 20 ms) in orange and long duration events (20 ms - 10 s) in pink. For Ires vs dwell event plots for each concentration (30 nM, 100 nM, 200 nM, and 300 nM) see Figure S1. (D) Close-ups of the two event types of the WT NTAD. (E) Plot of NTAD event  $1/\tau_{\text{on}}$  ( $\text{s}^{-1}$ ) for increasing NTAD concentrations (30 nM – 300 nM). The recordings were performed in a buffer containing 150 mM NaCl, 20 mM HEPES at pH 7.4 and with an applied voltage of +200 mV. The events were detected using Clampfit single-channel search.

The frequency of NTAD events was assessed by exponentially fitting the interevent time between all event types to obtain the  $\tau_{\text{on}}$  at each concentration, and the results of this fitting are shown in Table S1. The event frequency ( $1/\tau_{\text{on}}$ ) shows a linear relationship with increasing NTAD concentration (Fig. 3.5 E).

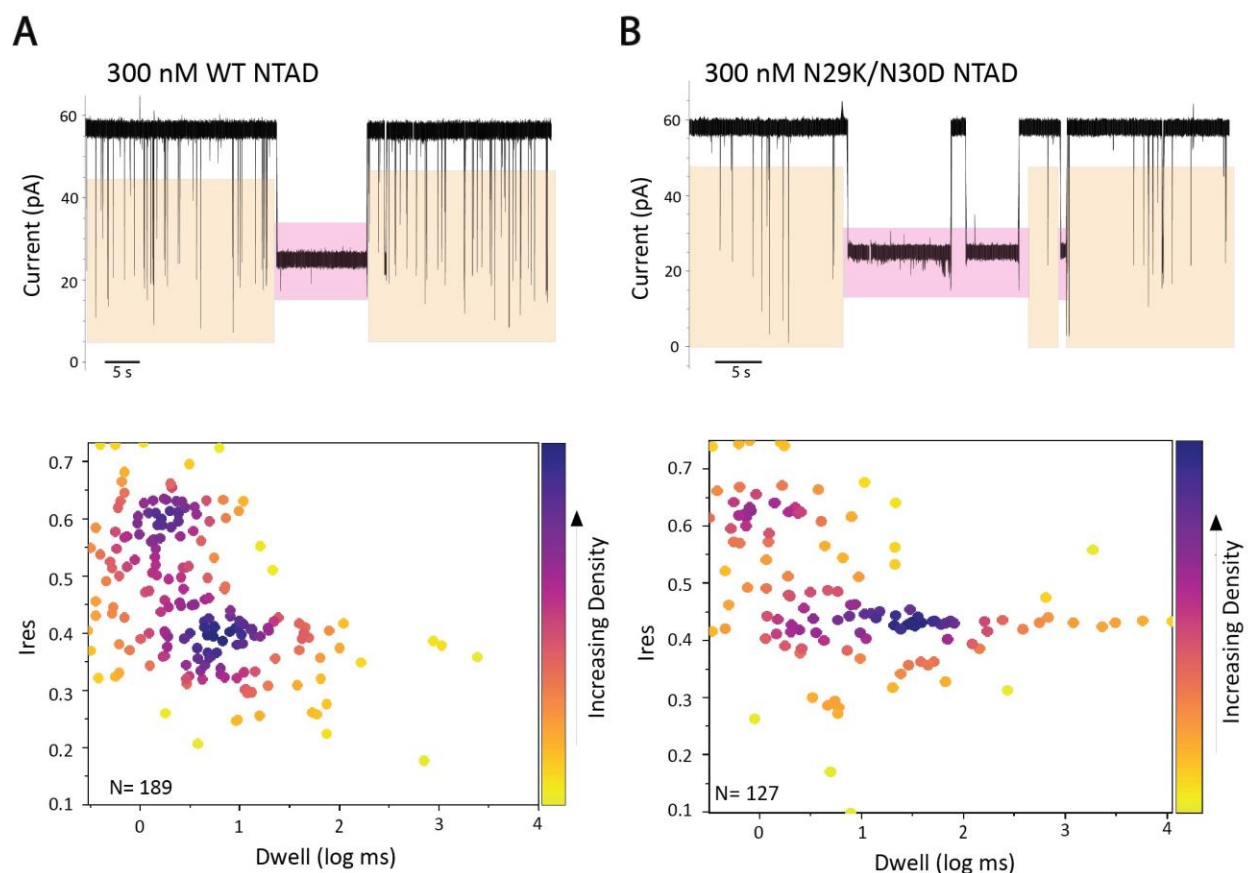
#### *Nanopore detection of an NTAD disease-associated mutant*

Prior studies have reported that the N29K/N30D cancer-associated NTAD mutant has an increased subpopulation of compact conformations<sup>19</sup>. To confirm this, we performed molecular dynamics (MD) simulations with the WT NTAD and N29K/N30D NTAD (p53 residues 1-73) using the hybrid-resolution (HyRes) II force field (Fig. 3.6)<sup>65</sup>. The probability distributions of the radius of gyration (Rg) for the two peptides reveal a slight decrease in the average in the Rg for the N29K/N30D NTAD, supporting the prior conclusions that the mutant has a slight increase in the propensity toward collapsed conformations.



**Figure 3.6. Probability distribution of the radius of gyration for WT and N29K/N30D NTAD.** Radius of gyration ( $\text{\AA}$ ) calculated from HyRes II simulations of WT (black) and N29K/N30D (red) NTAD. The NTAD construct used for the simulations is the same as what is used for the nanopore experiments, p53 residues 1-73 with N-terminal 6xHis tag. The distributions are calculated from the average of two independent 400 ns simulations and the error bars represent the standard deviation between these replicates. The lines at the bottom of the plot represent the average rg for the two distributions.

Herein, we sought to determine if we could use the MspA nanopore to distinguish between the events of WT NTAD and N29K/N30D. WT NTAD and N29K/N30D were added to separate pores at 300 nM and recorded to capture a sufficient number of events to analyze the equilibrium population  $I_{\text{res}}$  and dwell time. Representative portions of the traces are shown in Figure 3.7 A and B with the two NTAD event types, as described previously, highlighted in pink and orange.



**Figure 3.7. Detection of subtle differences in WT and mutant NTAD events.** Event comparison between WT (A) and N29K/N30D (B) NTAD at 300 nM concentration. (Top) characteristic traces of both analytes. Short events (0.3 ms – 20 ms) are highlighted in orange and long events (20 ms – 10 s) are highlighted in pink. Shorter portions of the analyzed traces are pictured to show more event detail, to see the full analyzed traces (4.2 min) see Fig. S2. (Bottom) Plots of event dwell time (log ms) and residual current ( $I_{res}$ ). The events are colored according to increasing event type density. The recordings were performed in a buffer containing 150 mM NaCl, 20 mM HEPES at pH 7.4 and with an applied voltage of +200 mV. The events were detected using Clampfit single-channel search.

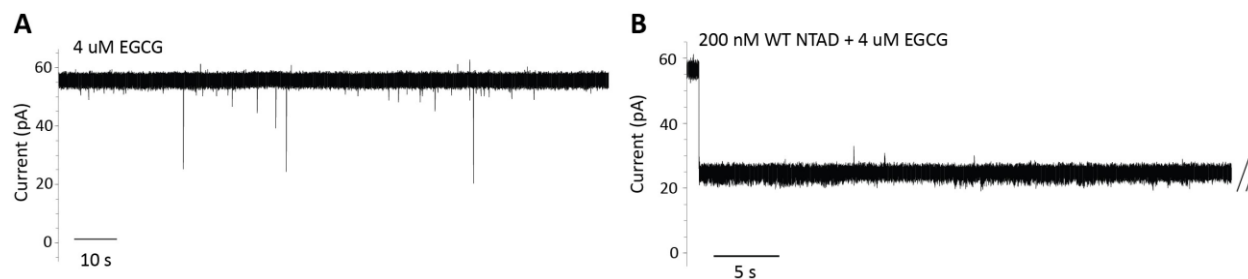
The events were then analyzed to determine their  $I_{res}$  and dwell time using Clampfit single-channel-search, and the scatter plots showing the statistics for all analyzed events are shown at the bottom of Figure 3.7 A and B, with the points for each event colored according to data point density. The full traces used to generate these plots are shown in Figure S2. The WT NTAD events show the highest density around 0.6  $I_{res}$  and 0.5 – 1 ms dwell and 0.3-0.4  $I_{res}$  and 1 – 10

ms dwell. The events observed for N29K/N30D are depleted in the 0.6  $I_{res}$  0.5 – 1 ms population and enriched in the 0.4  $I_{res}$  population, and the dwell time for these events is increased to 10 – 100 ms.

*Real time identification of a small molecule binding p53 NTAD using MspA nanopore sensor*

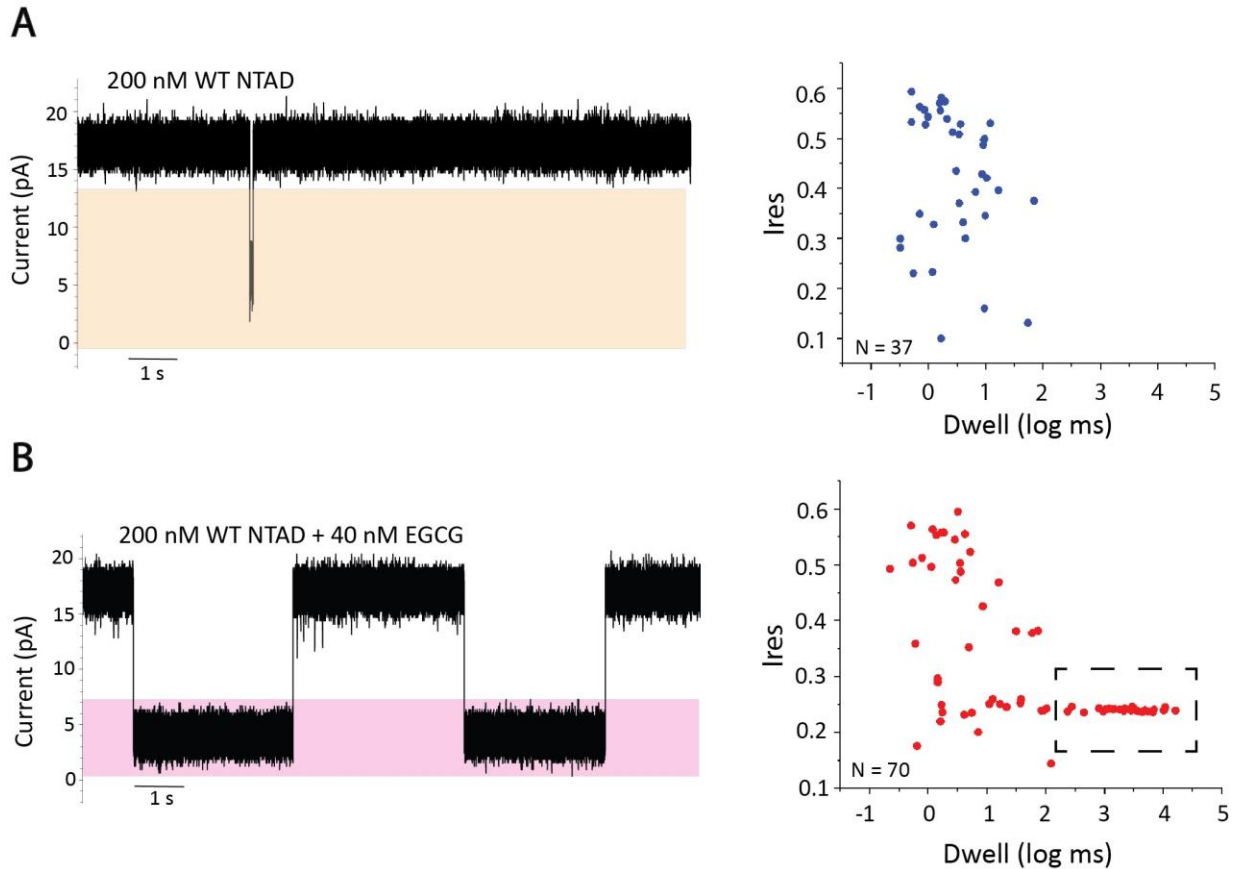
Previous MD simulation and SAXS results have shown that EGCG binding to the disordered NTAD results in a more collapsed NTAD conformation which inhibits the binding interaction between p53 and HDM2<sup>51</sup>. To see if we could resolve this difference using the nanopore, we carried out recordings of MspA in the presence of NTAD and EGCG. EGCG was first added to the nanopore alone to the concentration of the 4  $\mu$ M which is the published  $K_D$  of the NTAD EGCG interaction. The control was carried out for several minutes at +200 mV and no significant events or pore gating was observed, indicating EGCG does not affect the MspA nanopore independently (Fig. 3.8 A).

Next, MspA was recorded in the presence of 200 nM NTAD and 4  $\mu$ M EGCG together, and a permanent “stuck” event type at 0.4  $I_{res}$  was observed at +200 mV (Figure 3.8 B). This event type also exhibited an enhanced recapture rate, and when the pore returned to baseline, this event would capture in the pore within milliseconds. This signal persisted even in experiments using 200 nM NTAD and an EGCG concentration two orders of magnitude below the  $K_D$  (40 nM). These events last in the pore for minutes and can only be terminated by reversing the applied voltage to a negative potential manually, which prevents the proper sampling of the equilibrium population in the chamber and the quantification of the event dwell time.



**Figure 3.8. EGCG control and representative WT NTAD and EGCG “stuck” event. (A)** Portion of a trace from a recording with 4  $\mu\text{M}$  EGCG alone in MspA. This experiment is performed at +200 mV in buffer containing 150 mM NaCl and 20 mM HEPES at pH 7.4. No significant events or gating was observed over the course of the recording, indicating EGCG alone does not induce events like those observed in the presence of NTAD and EGCG together. **(B)** Representative event from a trace of 200 nM WT NTAD and 4  $\mu\text{M}$  EGCG in MspA. This event type has to be manually ejected through reversal of the applied voltage to recover the baseline open pore current. This experiment is performed at +200 mV in buffer containing 150 mM NaCl and 20 mM HEPES at pH 7.4.

Thus, we sought to find a voltage where reversible NTAD EGCG events could be sampled by decreasing the electrophoretic force on the protein. At +75 mV, both transient event types and the long dwell events could be sampled without manually interfering with the applied voltage, allowing for the characterization of the  $I_{\text{res}}$  and dwell time of the captured events.



**Figure 3.9. Detection of a small molecule binding p53 NTAD.** (A) Left: Portion of a representative trace of 200 nM WT NTAD. Right:  $I_{res}$  vs log dwell plot in blue shows the statistics for events detected from 9.6 min of the trace ( $N=37$ ). (B) Left: Portion of a representative trace of 200 nM WT NTAD with 40 nM EGCG. Right:  $I_{res}$  vs log dwell plot in red shows the statistics for events detected from 9.6 min of the trace ( $N=70$ ). The box on the scatter plot highlights the population of events that are unique to the addition of EGCG to WT NTAD. On the traces, short events (0.3 ms – 20 ms) are highlighted in orange and long events (20 ms – 10 s) are highlighted in pink. The recordings were performed in buffer containing 150 mM NaCl, 20 mM HEPES, at pH 7.4 with an applied voltage of +75 mV. The events were detected using Clampfit single-channel search.

Figure 3.9 A shows a representative portion of a trace with 200 nM NTAD alone at +75 mV, and a scatter plot of the captured events dwell time and  $I_{res}$  is shown to the right. At a lower voltage, the NTAD exhibits only short event types ( $I_{res}$  0.6 - 0.1 and dwell 0.3 - 20 ms) which are highlighted in orange on the trace. Figure 3.9 B shows the results of 200 nM NTAD with the

addition of 40 nM EGCG, and a significant event population appears between 0.2 and 0.3  $I_{res}$  with a dwell time between 100 ms and 10 s, which is highlighted using a box on the scatter plot. The full analyzed traces for NTAD alone and NTAD with EGCG at +75 mV are shown in figure S3.

## Discussion

### *MspA nanopore sensor allows for the characterization of heterogeneous NTAD events*

The MspA nanopore has previously been used to trap protein analytes and sequence DNA, and here we sought to determine if it could be used as a sensor to detect events and interactions of an IDP, the NTAD of p53<sup>1,66,67,72,93</sup>. Cis-inserted MspA in 150 mM NaCl buffer and a positive applied *trans* voltage is minimally noisy and exhibits only occasional self-gating, making it a sensitive probe for the trapping of protein analytes (Fig 3.5 B). The addition of the NTAD to the pore shows two distinct types of events, transient interactions with a dwell of 0.3 – 20 ms and long dwell events that remain in the pore for 20 ms – 10 s (Fig. 3.5 C and D). Both long and short event types are concentration-dependent, and the frequency of NTAD events, or the inverse of the  $\tau_{on}$ , increases linearly with increasing concentrations of NTAD (Fig. 3.5 E).

We hypothesize that these distinct event types are due to the heterogeneous conformational ensemble of the disordered NTAD, and some conformations are more favorable to be trapped within the pore lumen, while others may only transiently interact with the opening of the pore. Short event types at high voltages may be evidence of translocation of the extended peptide, but translocation events would have a nearly 0  $I_{res}$  due to the polypeptide chain passing through the constriction region, and the short events characterized at 250 mV have a minimum  $I_{res}$  of ~0.25.

Additionally, if the short events were translocation, we hypothesized that decreasing the applied voltage would increase the dwell time of the event population because the electrophoretic force on the peptide is decreased, allowing it to remain trapped in the pore lumen. The voltage study performed on the WT NTAD did not show a significant shift in dwell time of the short event population characterized at +200 mV at lower voltages, and the population remained at 0.2-0.8  $I_{res}$  and 0.3 – 20 ms at 150 mV and 100 mV (Fig. 3.4 A and B). Thus, the short dwell events are likely not a result of translocation, but rather the result of transient pore interactions.

*MspA can detect changes in NTAD events due to a cancer associated mutation*

Previous results of molecular dynamics simulations of WT NTAD and cancer associated NTAD mutant N29K/N30D have shown the appearance of a subpopulation of compact N29K/N30D conformations. From these simulations, the collapsed N29K/N30D population was found to have an average radius of gyration of 17.8 Å, compared to the average radius of gyration of the major WT NTAD population, which was calculated to be 33.2 Å<sup>19</sup>. Our own simulation results using the HyRes II force field also show a decrease in the average radius of gyration of the N29K/N30D mutant, supporting the conclusion that there is an increased population of more collapsed conformations (Fig. 3.6). This subpopulation, therefore, has a significant difference in conformation compared to WT NTAD, and we sought to determine if the events captured using the nanopore could distinguish between the WT and mutated peptide.

The traces and the density plots shown in Figure 3.7 show a shift between the event populations present with the WT NTAD and with N29K/N30D NTAD. The WT peptide has two main populations that center around 0.6  $I_{res}$  and 0.5 – 1 ms dwell and 0.3-0.4  $I_{res}$  and 1 – 10 ms dwell.

N29K/N30D events have one highest-density population around  $0.4 I_{res}$  with a dwell that ranges from 10 – 100 ms. The event population at  $0.6 I_{res}$  and 0.5 – 1 ms dwell is still present but has decreased density compared to the WT population. These findings suggest an enrichment of events with a longer dwell time on average for N29K/N30D compared to the WT NTAD.

We hypothesize that the short pore interactions represent a more extended NTAD conformation and the long duration events are a partially folded or more collapsed NTAD conformation.

N29K/N30D is known to have a decreased radius of gyration on average, and the events that arise from this peptide show a shift toward a longer dwell time in the pore. We suspect this could be due to an entropic effect. A more collapsed protein conformation has a lower entropic cost for entering the nanopore while an extended conformation would experience a higher, unfavorable loss of entropy, deterring it from entering or remaining in the pore lumen and resulting in the transient events that appear as “spikes” of blockage on the trace.

The results of these experiments indicate that the MspA nanopore is sensitive to even subtle changes in an IDPs conformational ensemble, and these differences can be characterized by analyzing the shifts in event populations.

#### *The nanopore can be used to detect small molecule binding to an IDP*

Similar to the effects of the N29K/N30D mutation on the NTAD ensemble, SAXS experiments and MD simulations have shown that EGCG binding to the disordered NTAD causes the peptide to adopt a more collapsed conformation, which disrupts NTAD binding to HDM2<sup>51</sup>. These studies revealed that EGCG binding to NTAD caused a dramatic shift in the pairwise distance distribution of the NTAD, and a significant subpopulation of a more compact state was

uncovered using principal component analysis (PCA) of NTAD conformations in all-atom simulations with the presence of EGCG.

Here, we show this difference induced by EGCG binding can be observed using the MspA nanopore. The addition of EGCG and the WT NTAD in the nanopore result in a significant increase in event duration and capture rate, with trapped events lasting for up to minutes at +200 mV (Figure 3.8). We hypothesize that the dramatic increase in event duration and  $\tau_{\text{on}}$  of the NTAD in the presence of EGCG is the result of two effects. Firstly, as described in the case of the N29K/N30D mutant, the more collapsed structure of the EGCG-bound NTAD will reduce the entropic cost of entering the confinement of the nanopore, which can enhance the ability of the peptide to enter the pore lumen. Secondly, the published pKa of EGCG is 7.68, meaning that EGCG carries a slight negative charge at physiological pH due to deprotonation of its C3'-OH or C4''-OH moieties<sup>97</sup>. This additional negative charge can increase the electrophoretic force experienced by the NTAD at positive *trans* applied voltage, pulling it toward cis-inserted MspA with more force than the unbound NTAD, and increasing its event duration within the pore.

Because of these long-lived events, the NTAD in the presence of EGCG needed to be manually ejected from the pore by the reversal of the applied voltage throughout the recording, which artificially truncates events and prevents the characterization of the recording chamber population. In order to sample the equilibrium population of the NTAD in the presence of EGCG and to quantify the dwell time of the events, we performed a voltage study to identify a voltage at which the electrophoretic force on the NTAD is decreased enough to observe both unbound NTAD events and reversible events of NTAD and EGCG together.

We identified +75 mV as the ideal voltage for these experiments (Fig. 3.9). At this voltage, the force on the unbound WT NTAD is significantly decreased, and only short duration (0.3 – 20 ms)

events are captured. When EGCG is added to a concentration of 40 nM, a significant event population is observed between 0.2 and 0.3  $I_{res}$  with a dwell time between 100 ms and 10 s that is unique to the NTAD and EGCG together and does not appear with the NTAD alone or with EGCG and MspA alone (Fig. 3.8 A). This indicates that this population is a result of NTAD bound to EGCG, and the increased dwell time of these events supports the theory that long duration events represent more compact conformations of the NTAD.

### *Conclusions*

In this work, we have shown that the MspA nanopore is a sensitive tool for the detection of an IDP, the N-terminal domain of p53. WT NTAD events show heterogeneity in their duration, and the nanopore can resolve a shift in these event types toward a longer dwell population due to a cancer-associated NTAD mutation, N29K/N30D, and a small-molecule binding interaction with EGCG. These results lead us to hypothesize that long dwell events correspond to a more collapsed conformation of the NTAD, which is known to be induced by the N29K/N30D mutation and EGCG binding.

Together, the data presented here indicate that the MspA nanopore could be used as a sensitive screening tool for IDPs and their mutations or small molecule interactions and shifts in the event populations may even indicate the structural consequences of these modifications. This tool could be further developed to report on the functional consequences of mutations or small molecule binding by characterizing NTAD protein binding partner interactions in MspA. Thus, the nanopore could be a powerful tool for studying IDPs and their role in disease, as well as developing drugs to target these dynamic proteins.

## CHAPTER 4

### CONCLUSIONS AND OUTLOOK

#### *4.1 Conclusions*

The computational results shown in this work provide promising evidence that the HyRes II force field is able to effectively simulate IDPs and capture their transient, long-range interactions, as well as shifts in their conformational ensembles induced by the addition of PTMs. The ability to capture transient interactions between the p53 NTAD and DBD that agree with experimental PRE data demonstrates the accuracy of this force field in simulating complex IDP dynamics. Simulating an IDP such as the p53 tetramer allows for the molecular-level study of its modes of regulation in the cell, including the effects of threonine 55 phosphorylation, which is not easily determined experimentally or with all-atom simulations.

Through experiments using nanopore sensing to study the p53 NTAD, we have shown results that the MspA nanopore sensor can detect the events of an IDP and may be able to sense differences in the conformational ensemble by two distinct modes of interaction with the nanopore. Our findings testing two conditions known to induce conformational changes in the NTAD, a disease-associated mutant (N29K/N30D) and interaction with small molecule EGCG, support our theory that the differences in event dwell are due to the degree of structural compaction of the NTAD. These results are sensitive, single-molecule studies of an IDP and have implications for the use of nanopores as a low-cost, high-throughput method to screen mutants and potential small molecule therapeutics.

#### *4.2 Future directions*

The computational data presented here shows good agreement with experimental results but is still dependent on the initial structures chosen for the simulation replicas to sample sufficient conformational space. Therefore, further expansion of this project could include using an enhanced sampling technique such as temperature-replica exchange to increase the convergence of the simulations, which will reduce simulation artifacts and allow for more solid conclusions to be drawn from the generated ensembles. This can also allow for the simulation of p53 tetramer constructs with PTMs whose effects have not been experimentally determined. These novel findings could provide valuable insights into p53 signaling, structure, and regulation, which are important for understanding its role in disease.

Additional experiments and controls could also be performed to strengthen the nanopore results, including characterizing the signal of WT NTAD and N29K/N30D NTAD in the presence of a binding partner such as HDM2. Nanopores have previously been used to detect protein-protein interactions, so MspA may be able to detect a unique signal for the HDM2-NTAD complex.

Experiments could then be performed with disease-associated mutants such as N29K/N30D that are known to affect binding partner interactions, and this difference could be detected through the lack of the complex signal in the nanopore. This would provide proof of concept for the use of nanopore sensors as tools to characterize and determine the functional consequences of disease related IDP mutants and their binding interactions, which would give further insight into the mechanisms that lead to the development of diseases such as cancer.

Similarly, the NTAD could be tested for HDM2 binding in the presence of EGCG using nanopore sensing, and if an ablation of the NTAD-HDM2 binding signal is observed, that would suggest that nanopores can be a powerful tool for screening the effects of small molecules on

protein-protein interactions. Further controls should also be performed, including testing the NTAD with a small molecule that should not have a binding interaction as a negative control, as well as testing other known small molecule p53 NTAD binders besides EGCG to characterize additional small molecule signals in MspA.

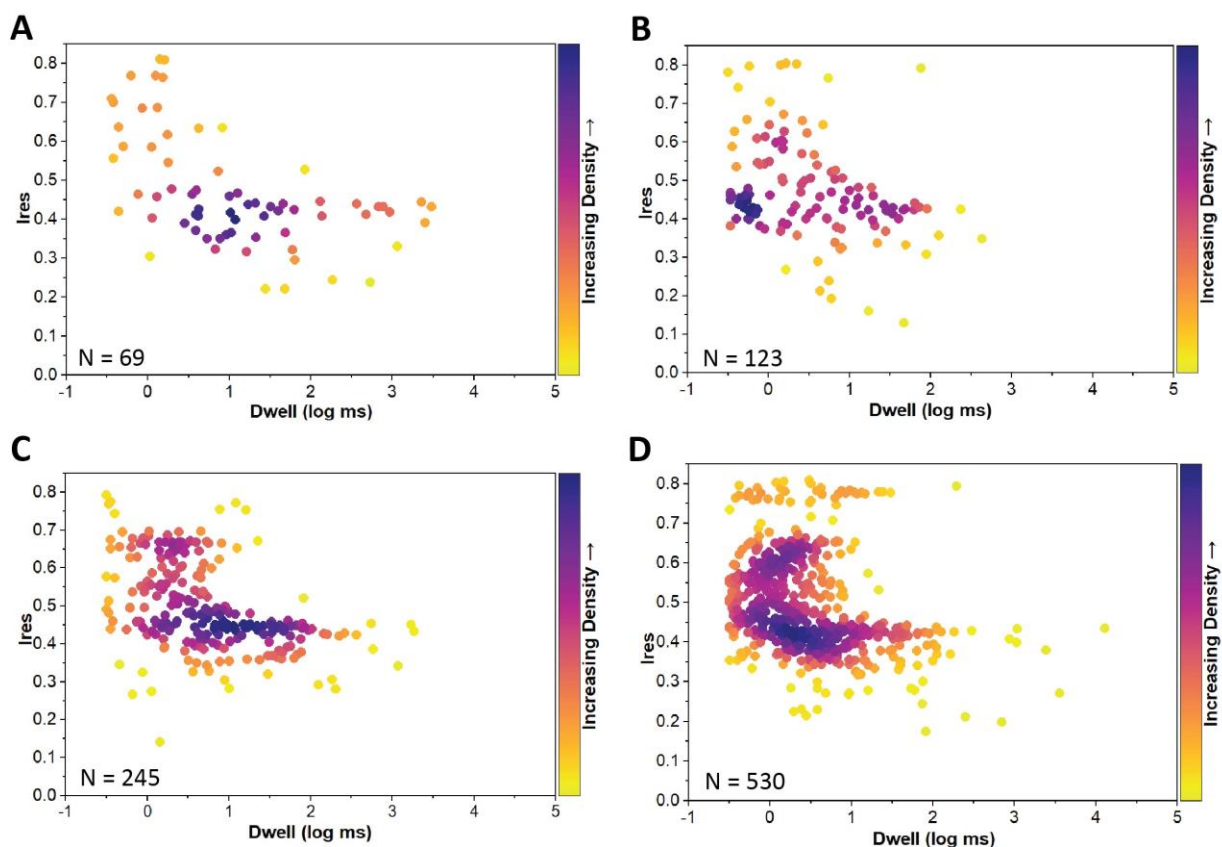
There is also the opportunity for computation and simulation to be combined in these studies, and the HyRes II force field could be used to simulate the NTAD interaction with the MspA nanopore. These simulations could reveal molecular details of what is happening within the pore, allowing us to draw more conclusions from our experimental data.

These results and the continuation of these projects can provide a foundation for exciting advancements in drug discovery for targeting IDPs and protein-protein interactions and could accelerate testing for effective cancer treatments. The strategy of combining MD simulations and nanopore sensing can be applied to other biologically relevant IDPs and allow for more detailed studies of their role in human health and disease.

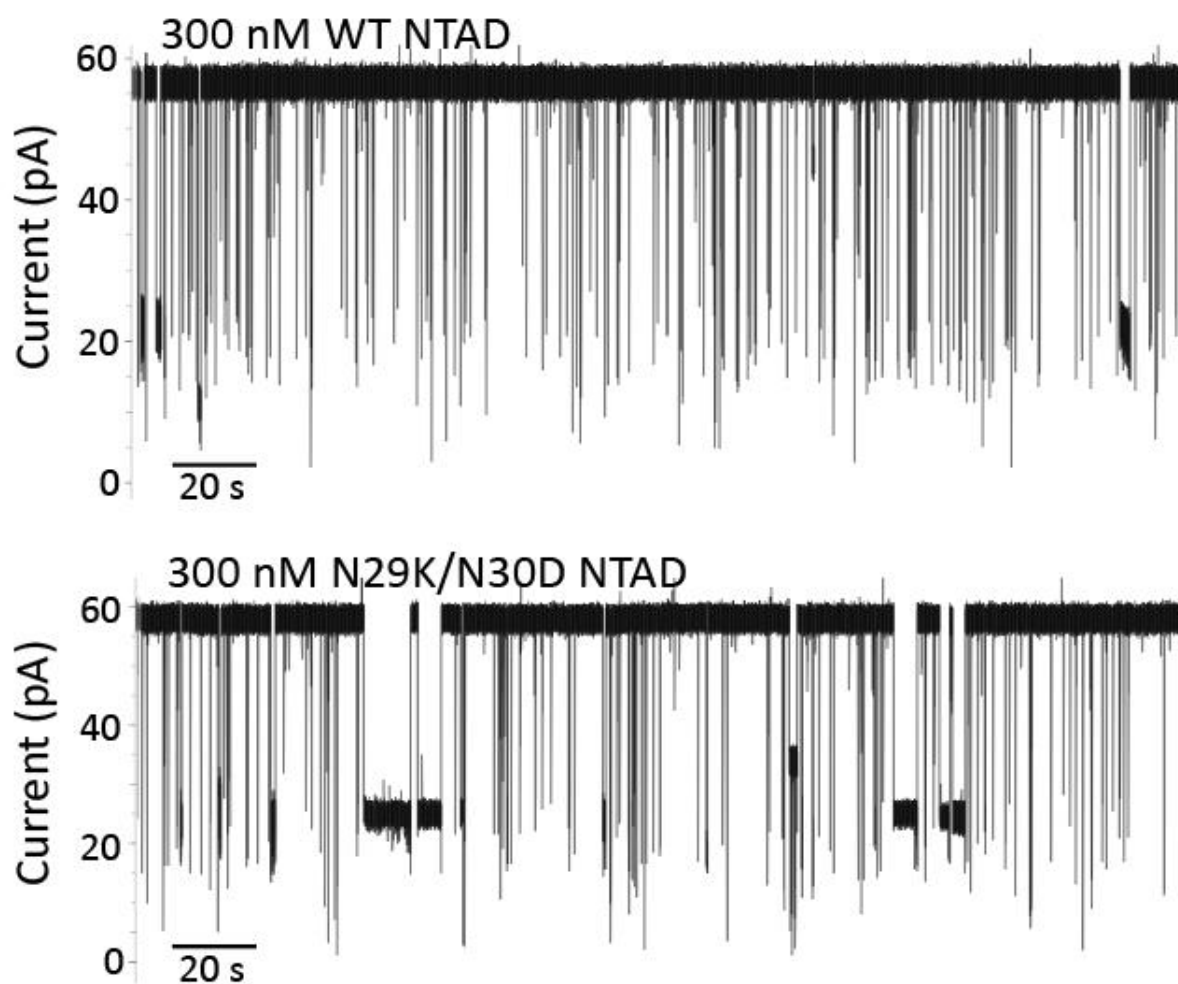
## SUPPLEMENTARY FIGURES

**Table S1. Fitting results for interevent duration ( $\tau_{on}$ ) at different NTAD concentrations.** The interevent time was fitted exponentially to obtain the  $\tau_{on}$  (ms) at increasing NTAD concentrations. The experiments were performed in buffer conditions of 150 mM NaCl, 20 mM HEPES at pH 7.4 and an applied voltage of +200 mV. The results shown are from one nanopore recording.

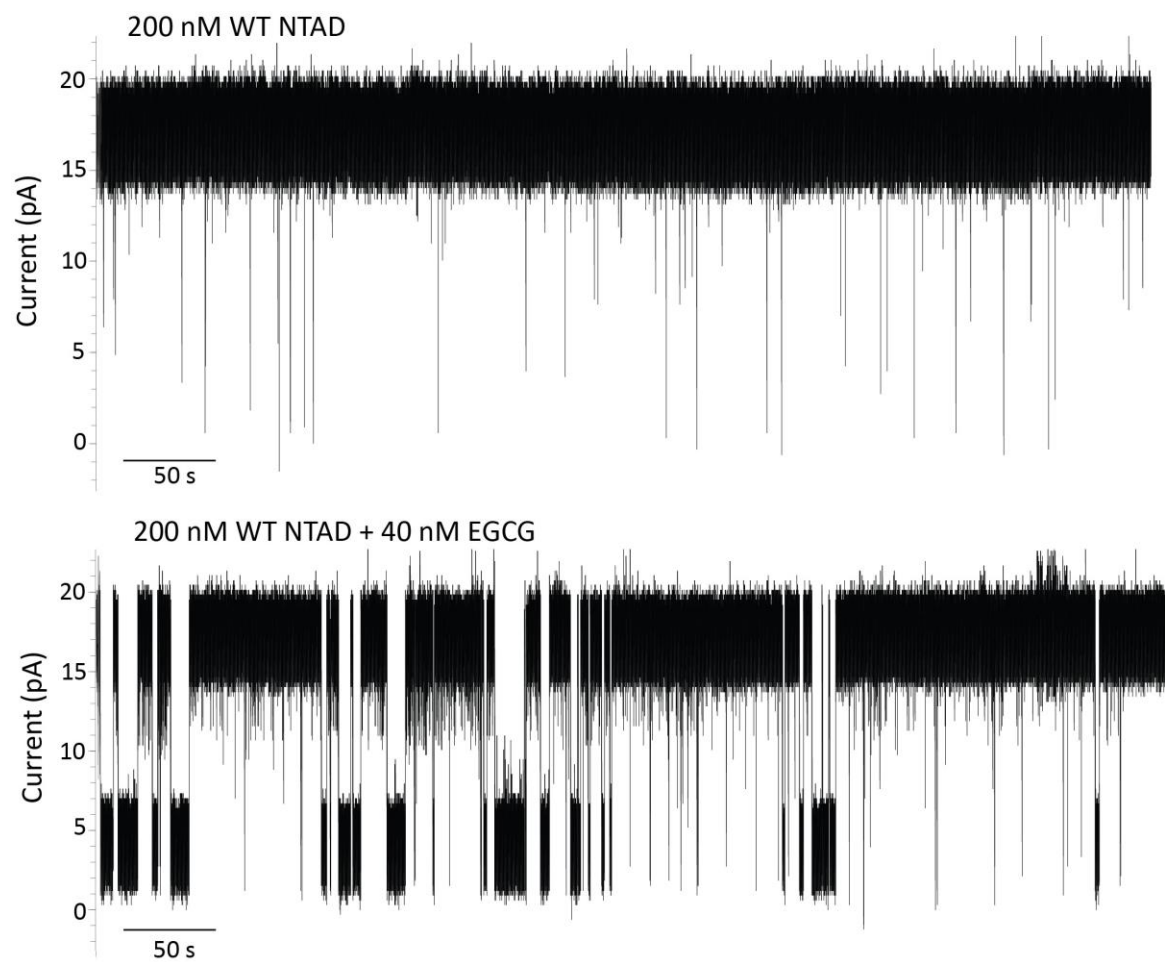
NTAD concentration (nM)	$\tau_{on} \pm \text{std}$ (ms)
30	10,552 $\pm$ 933
100	3,505 $\pm$ 156
200	1,339 $\pm$ 18
300	713 $\pm$ 17



**Figure S1. WT NTAD concentration study Ires vs dwell time plots.** Concentration study of WT NTAD at (A) 30 nM (B) 100 nM (C) 200 nM and (D) 300 nM. The events at each concentration were detected using Clampfit 11.2 single-channel search from 10-minute portions of the traces, and the number of events at each concentration is noted on the plots. The events are characterized by Ires and the log of the dwell time (log ms) and colored by increasing event density. The experiments were performed in buffer conditions of 150 mM NaCl, 20 mM HEPES at pH 7.4 and at an applied voltage of +200 mV.



**Figure S2. Full analyzed traces of WT and N29K/N30D NTAD at 300 nM.** The full analyzed trace (4.2 minutes) of 300 nM WT NTAD (top) and N29K/N30D (bottom) in MspA. This experiment is performed at +200 mV in buffer containing 150 mM NaCl and 20 mM HEPES at pH 7.4.



**Figure S3. Full analyzed traces of 200 nM WT NTAD alone and 200 nM WT NTAD and 40 nM EGCG.** The full analyzed traces (9.6 mins) of 200 nM WT NTAD (top) and 200 nM WT NTAD and 40 nM EGCG (bottom) in MspA. This experiment is performed at +75 mV in buffer containing 150 mM NaCl and 20 mM HEPES at pH 7.4.

## References

1. Brucale, M., Schuler, B. & Samorì, B. Single-Molecule Studies of Intrinsically Disordered Proteins. *Chem. Rev.* **114**, 3281–3317 (2014).
2. Wright, P. E. & Dyson, H. J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol* **16**, 18–29 (2015).
3. Romero, P. *et al.* Sequence complexity of disordered protein. *Proteins* **42**, 38–48 (2001).
4. Vucetic, S., Brown, C. J., Dunker, A. K. & Obradovic, Z. Flavors of protein disorder. *Proteins* **52**, 573–584 (2003).
5. Wright, P. E. & Dyson, H. J. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *Journal of Molecular Biology* **293**, 321–331 (1999).
6. Ferreon, A. C. M., Ferreon, J. C., Wright, P. E. & Deniz, A. A. Modulation of allostery by protein intrinsic disorder. *Nature* **498**, 390–394 (2013).
7. Berlow, R. B., Dyson, H. J. & Wright, P. E. Expanding the Paradigm: Intrinsically Disordered Proteins and Allosteric Regulation. *Journal of Molecular Biology* **430**, 2309–2320 (2018).
8. Chen, J. Towards the physical basis of how intrinsic disorder mediates protein function. *Archives of Biochemistry and Biophysics* **524**, 123–131 (2012).
9. Liu, J. *et al.* Intrinsic Disorder in Transcription Factors. *Biochemistry* **45**, 6873–6888 (2006).
10. Babu, M. M., Van Der Lee, R., De Groot, N. S. & Gsponer, J. Intrinsically disordered proteins: regulation and disease. *Current Opinion in Structural Biology* **21**, 432–440 (2011).
11. Uversky, V. N. *et al.* Pathological Unfoldomics of Uncontrolled Chaos: Intrinsically Disordered Proteins and Human Diseases. *Chem. Rev.* **114**, 6844–6879 (2014).

12. Das, R. K., Ruff, K. M. & Pappu, R. V. Relating sequence encoded information to form and function of intrinsically disordered proteins. *Current Opinion in Structural Biology* **32**, 102–112 (2015).
13. DeLeo, A. B. *et al.* Detection of a transformation-related antigen in chemically induced sarcomas and other transformed cells of the mouse. *Proc. Natl. Acad. Sci. U.S.A.* **76**, 2420–2424 (1979).
14. Lane, D. P. & Crawford, L. V. T antigen is bound to a host protein in SY40-transformed cells. *Nature* **278**, 261–263 (1979).
15. Finlay, C. A., Hinds, P. W. & Levine, A. J. The p53 proto-oncogene can act as a suppressor of transformation. *Cell* **57**, 1083–1093 (1989).
16. Baker, S. J. *et al.* Chromosome 17 Deletions and p53 Gene Mutations in Colorectal Carcinomas. *Science* **244**, 217–221 (1989).
17. Takahashi, T. *et al.* p53: A Frequent Target for Genetic Abnormalities in Lung Cancer. *Science* **246**, 491–494 (1989).
18. Joerger, A. C. & Fersht, A. R. Structure–function–rescue: the diverse nature of common p53 cancer mutants. *Oncogene* **26**, 2226–2242 (2007).
19. Schrag, L. G. *et al.* Cancer-Associated Mutations Perturb the Disordered Ensemble and Interactions of the Intrinsically Disordered p53 Transactivation Domain. *Journal of Molecular Biology* **433**, 167048 (2021).
20. Olivier, M., Hollstein, M. & Hainaut, P. TP53 Mutations in Human Cancers: Origins, Consequences, and Clinical Use. *Cold Spring Harbor Perspectives in Biology* **2**, a001008–a001008 (2010).

21. Petitjean, A., Achatz, M. I. W., Borresen-Dale, A. L., Hainaut, P. & Olivier, M. TP53 mutations in human cancers: functional selection and impact on cancer prognosis and outcomes. *Oncogene* **26**, 2157–2165 (2007).
22. Joerger, A. C. & Fersht, A. R. Structural Biology of the Tumor Suppressor p53. *Annu. Rev. Biochem.* **77**, 557–582 (2008).
23. Weinberg, R. L., Veprintsev, D. B. & Fersht, A. R. Cooperative Binding of Tetrameric p53 to DNA. *Journal of Molecular Biology* **341**, 1145–1159 (2004).
24. Dawson, R. *et al.* The N-terminal Domain of p53 is Natively Unfolded. *Journal of Molecular Biology* **332**, 1131–1141 (2003).
25. Kitayner, M. *et al.* Structural Basis of DNA Recognition by p53 Tetramers. *Molecular Cell* **22**, 741–753 (2006).
26. Petty, T. J. *et al.* An induced fit mechanism regulates p53 DNA binding kinetics to confer sequence specificity: p53 conformational switch. *The EMBO Journal* **30**, 2167–2176 (2011).
27. Espinosa, J. M., Verdun, R. E. & Emerson, B. M. p53 Functions through Stress- and Promoter-Specific Recruitment of Transcription Initiation Components before and after DNA Damage. *Molecular Cell* **12**, 1015–1027 (2003).
28. El-Deiry, W. S. Regulation of p53 downstream genes. *Seminars in Cancer Biology* **8**, 345–357 (1998).
29. Waldman, T., Kinzler, K. W. & Vogelstein, B. p21 is necessary for the p53-mediated G1 arrest in human cancer cells. *Cancer Res* **55**, 5187–5190 (1995).
30. Barboza, J. A., Liu, G., Ju, Z., El-Naggar, A. K. & Lozano, G. p21 delays tumor onset by preservation of chromosomal stability. *Proc Natl Acad Sci U S A* **103**, 19842–19847 (2006).

31. Yu, J., Wang, Z., Kinzler, K. W., Vogelstein, B. & Zhang, L. *PUMA* mediates the apoptotic response to p53 in colorectal cancer cells. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 1931–1936 (2003).
32. Jänicke, R. U., Sohn, D. & Schulze-Osthoff, K. The dark side of a tumor suppressor: anti-apoptotic p53. *Cell Death Differ* **15**, 959–976 (2008).
33. Vousden, K. H. & Prives, C. Blinded by the Light: The Growing Complexity of p53. *Cell* **137**, 413–431 (2009).
34. Appella, E. & Anderson, C. W. Post-translational modifications and activation of p53 by genotoxic stresses: p53 post-translational modifications. *European Journal of Biochemistry* **268**, 2764–2772 (2001).
35. Bode, A. M. & Dong, Z. Post-translational modification of p53 in tumorigenesis. *Nat Rev Cancer* **4**, 793–805 (2004).
36. Gu, W. & Roeder, R. G. Activation of p53 Sequence-Specific DNA Binding by Acetylation of the p53 C-Terminal Domain. *Cell* **90**, 595–606 (1997).
37. Haupt, Y., Maya, R., Kazaz, A. & Oren, M. Mdm2 promotes the rapid degradation of p53. *Nature* **387**, 296–299 (1997).
38. Kubbutat, M. H. G., Jones, S. N. & Vousden, K. H. Regulation of p53 stability by Mdm2. *Nature* **387**, 299–303 (1997).
39. Wang, Y. & Eckhart, W. Phosphorylation sites in the amino-terminal region of mouse p53. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 4231–4235 (1992).
40. Shieh, S.-Y., Ikeda, M., Taya, Y. & Prives, C. DNA Damage-Induced Phosphorylation of p53 Alleviates Inhibition by MDM2. *Cell* **91**, 325–334 (1997).

41. Banin, S. *et al.* Enhanced Phosphorylation of p53 by ATM in Response to DNA Damage. *Science* **281**, 1674–1677 (1998).
42. Shieh, S. Y., Ahn, J., Tamai, K., Taya, Y. & Prives, C. The human homologs of checkpoint kinases Chk1 and Cds1 (Chk2) phosphorylate p53 at multiple DNA damage-inducible sites. *Genes Dev* **14**, 289–300 (2000).
43. Jabbur, J. R. *et al.* Mdm-2 binding and TAFII31 recruitment is regulated by hydrogen bond disruption between the p53 residues Thr18 and Asp21. *Oncogene* **21**, 7100–7113 (2002).
44. Sakaguchi, K. *et al.* Damage-mediated phosphorylation of human p53 threonine 18 through a cascade mediated by a casein 1-like kinase. Effect on Mdm2 binding. *J Biol Chem* **275**, 9278–9283 (2000).
45. Lambert, P. F., Kashanchi, F., Radonovich, M. F., Shiekhattar, R. & Brady, J. N. Phosphorylation of p53 Serine 15 Increases Interaction with CBP. *Journal of Biological Chemistry* **273**, 33048–33053 (1998).
46. Ferreon, J. C. *et al.* Cooperative regulation of p53 by modulation of ternary complex formation with CBP/p300 and HDM2. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 6591–6596 (2009).
47. Krois, A. S., Ferreon, J. C., Martinez-Yamout, M. A., Dyson, H. J. & Wright, P. E. Recognition of the disordered p53 transactivation domain by the transcriptional adapter zinc finger domains of CREB-binding protein. *Proc. Natl. Acad. Sci. U.S.A.* **113**, (2016).
48. Lee, C. W., Ferreon, J. C., Ferreon, A. C. M., Arai, M. & Wright, P. E. Graded enhancement of p53 binding to CREB-binding protein (CBP) by multisite phosphorylation. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 19290–19295 (2010).
49. Sakaguchi, K. *et al.* DNA damage activates p53 through a phosphorylation-acetylation cascade. *Genes Dev* **12**, 2831–2841 (1998).

50. Sun, X., Dyson, H. J. & Wright, P. E. A phosphorylation-dependent switch in the disordered p53 transactivation domain regulates DNA binding. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2021456118 (2021).
51. Zhao, J. *et al.* EGCG binds intrinsically disordered N-terminal domain of p53 and disrupts p53-MDM2 interaction. *Nat Commun* **12**, 986 (2021).
52. Joerger, A. C., Allen, M. D. & Fersht, A. R. Crystal Structure of a Superstable Mutant of Human p53 Core Domain. *Journal of Biological Chemistry* **279**, 1291–1296 (2004).
53. Krois, A. S., Dyson, H. J. & Wright, P. E. Long-range regulation of p53 DNA binding by its intrinsically disordered N-terminal transactivation domain. *Proc. Natl. Acad. Sci. U.S.A.* **115**, (2018).
54. Nwanochie, E. & Uversky, V. N. Structure Determination by Single-Particle Cryo-Electron Microscopy: Only the Sky (and Intrinsic Disorder) is the Limit. *IJMS* **20**, 4186 (2019).
55. Huang, F. *et al.* Multiple conformations of full-length p53 detected with single-molecule fluorescence resonance energy transfer. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 20758–20763 (2009).
56. Kikhney, A. G. & Svergun, D. I. A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Letters* **589**, 2570–2577 (2015).
57. Camacho-Zarco, A. R. *et al.* NMR Provides Unique Insight into the Functional Dynamics and Interactions of Intrinsically Disordered Proteins. *Chem. Rev.* **122**, 9331–9356 (2022).
58. Ravera, E., Sgheri, L., Parigi, G. & Luchinat, C. A critical assessment of methods to recover information from averaged data. *Phys. Chem. Chem. Phys.* **18**, 5686–5701 (2016).
59. Bonomi, M., Heller, G. T., Camilloni, C. & Vendruscolo, M. Principles of protein structural ensemble determination. *Current Opinion in Structural Biology* **42**, 106–116 (2017).

60. Fisher, C. K. & Stultz, C. M. Constructing ensembles for intrinsically disordered proteins. *Current Opinion in Structural Biology* **21**, 426–431 (2011).
61. Ganguly, D. & Chen, J. Structural Interpretation of Paramagnetic Relaxation Enhancement-Derived Distances for Disordered Protein States. *Journal of Molecular Biology* **390**, 467–477 (2009).
62. Shrestha, U. R., Smith, J. C. & Petridis, L. Full structural ensembles of intrinsically disordered proteins from unbiased molecular dynamics simulations. *Commun Biol* **4**, 243 (2021).
63. Kmiecik, S. *et al.* Coarse-Grained Protein Models and Their Applications. *Chem. Rev.* **116**, 7898–7936 (2016).
64. Liu, X. & Chen, J. HyRes: a coarse-grained model for multi-scale enhanced sampling of disordered protein conformations. *Phys. Chem. Chem. Phys.* **19**, 32421–32432 (2017).
65. Zhang, Y., Liu, X. & Chen, J. Toward Accurate Coarse-Grained Simulations of Disordered Proteins and Their Dynamic Interactions. *J. Chem. Inf. Model.* **62**, 4523–4536 (2022).
66. Li, F., Fahie, M. A., Gilliam, K. M., Pham, R. & Chen, M. Mapping the conformational energy landscape of Abl kinase using ClyA nanopore tweezers. *Nat Commun* **13**, 3541 (2022).
67. Pham, B. *et al.* A Nanopore Approach for Analysis of Caspase-7 Activity in Cell Lysates. *Biophysical Journal* **117**, 844–855 (2019).
68. Ramsay, W. J. & Bayley, H. Single-Molecule Determination of the Isomers of D -Glucose and D -Fructose that Bind to Boronic Acids. *Angew. Chem. Int. Ed.* **57**, 2841–2845 (2018).
69. Crnković, A., Srnko, M. & Anderluh, G. Biological Nanopores: Engineering on Demand. *Life* **11**, 27 (2021).

70. Manrao, E. A. *et al.* Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase. *Nat Biotechnol* **30**, 349–353 (2012).
71. Wang, Y., Zhao, Y., Bollas, A., Wang, Y. & Au, K. F. Nanopore sequencing technology, bioinformatics and applications. *Nat Biotechnol* **39**, 1348–1365 (2021).
72. Shorkey, S. A., Du, J., Pham, R., Strieter, E. R. & Chen, M. Real-Time and Label-Free Measurement of Deubiquitinase Activity with a MspA Nanopore. *ChemBioChem* **22**, 2688–2692 (2021).
73. Fahie, M., Chisholm, C. & Chen, M. Resolved Single-Molecule Detection of Individual Species within a Mixture of *anti* -Biotin Antibodies Using an Engineered Monomeric Nanopore. *ACS Nano* **9**, 1089–1098 (2015).
74. Li, X., Lee, K. H., Shorkey, S., Chen, J. & Chen, M. Different Anomeric Sugar Bound States of Maltose Binding Protein Resolved by a Cytolysin A Nanopore Tweezer. *ACS Nano* **14**, 1727–1737 (2020).
75. Schmid, S., Stömmer, P., Dietz, H. & Dekker, C. Nanopore electro-osmotic trap for the label-free study of single proteins and their conformations. *Nat. Nanotechnol.* **16**, 1244–1250 (2021).
76. Lu, H., Giordano, F. & Ning, Z. Oxford Nanopore MinION Sequencing and Genome Assembly. *Genomics, Proteomics & Bioinformatics* **14**, 265–279 (2016).
77. Garalde, D. R. *et al.* Highly parallel direct RNA sequencing on an array of nanopores. *Nat Methods* **15**, 201–206 (2018).
78. Liu, Y. *et al.* Allosteric Switching of Calmodulin in a *Mycobacterium smegmatis* porin A (MspA) Nanopore-Trap. *Angew. Chem. Int. Ed.* **60**, 23863–23870 (2021).

79. Derrington, I. M. *et al.* Nanopore DNA sequencing with MspA. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 16060–16065 (2010).
80. Laszlo, A. H., Derrington, I. M. & Gundlach, J. H. MspA nanopore as a single-molecule tool: From sequencing to SPRNT. *Methods* **105**, 75–89 (2016).
81. Faller, M., Niederweis, M. & Schulz, G. E. The Structure of a Mycobacterial Outer-Membrane Channel. *Science* **303**, 1189–1192 (2004).
82. Cain, C., Miller, S., Ahn, J. & Prives, C. The N Terminus of p53 Regulates Its Dissociation from DNA. *Journal of Biological Chemistry* **275**, 39944–39953 (2000).
83. He, F. *et al.* Interaction between p53 N terminus and core domain regulates specific and nonspecific DNA binding. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 8859–8868 (2019).
84. Chen, J., Im, W. & Brooks, C. L. Balancing Solvation and Intramolecular Interactions: Toward a Consistent Generalized Born Force Field. *J. Am. Chem. Soc.* **128**, 3728–3736 (2006).
85. McGibbon, R. T. *et al.* MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophysical Journal* **109**, 1528–1532 (2015).
86. Kabsch, W. & Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637 (1983).
87. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **14**, 33–38 (1996).
88. Pedregosa, Fabian *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830 (2011).

89. Lowry, D. F., Stancik, A., Shrestha, R. M. & Daughdrill, G. W. Modeling the accessible conformations of the intrinsically unstructured transactivation domain of p53. *Proteins* **71**, 587–598 (2008).
90. Selivanova, G. & Wiman, K. G. Reactivation of mutant p53: molecular mechanisms and therapeutic potential. *Oncogene* **26**, 2243–2254 (2007).
91. Brown, C. J., Lain, S., Verma, C. S., Fersht, A. R. & Lane, D. P. Awakening guardian angels: drugging the p53 pathway. *Nat Rev Cancer* **9**, 862–873 (2009).
92. Chen, B.-H., Hsieh, C.-H., Tsai, S.-Y., Wang, C.-Y. & Wang, C.-C. Anticancer effects of epigallocatechin-3-gallate nanoemulsion on lung cancer cells through the activation of AMP-activated protein kinase signaling pathway. *Sci Rep* **10**, 5163 (2020).
93. Liu, Y. *et al.* Machine Learning Assisted Simultaneous Structural Profiling of Differently Charged Proteins in a *Mycobacterium smegmatis* Porin A (MspA) Electroosmotic Trap. *J. Am. Chem. Soc.* **144**, 757–768 (2022).
94. Oh, S., Lee, M.-K. & Chi, S.-W. Single-molecule analysis of interaction between p53TAD and MDM2 using aerolysin nanopores. *Chem. Sci.* **12**, 5883–5891 (2021).
95. Tyler, A. D. *et al.* Evaluation of Oxford Nanopore’s MinION Sequencing Device for Microbial Whole Genome Sequencing Applications. *Sci Rep* **8**, 10931 (2018).
96. Butler, T. Z., Pavlenok, M., Derrington, I. M., Niederweis, M. & Gundlach, J. H. Single-molecule DNA detection with an engineered MspA protein nanopore. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 20647–20652 (2008).
97. Muzolf, M., Szymusiak, H., Gliszczyńska-Świgło, A., Rietjens, I. M. C. M. & Tyrakowska, B. pH-Dependent Radical Scavenging Capacity of Green Tea Catechins. *J. Agric. Food Chem.* **56**, 816–823 (2008).