Division of Theoretical Physics

The Institute of Physics of the Polish Academy of Sciences

All-atom and coarse-grained studies of disulfide-bond role in biologically-important proteins

Pełnoatomowe i gruboziarniste badania roli wiązań disulfidowych w biologicznie istotnych białkach

Author: **Pamela Smardz**

Supervisor: dr hab. Paweł Krupa

This dissertation is submitted for the degree of

Doctor of Philosophy

2026

Abstract

Organic polymers, such as proteins, nucleic acids, and carbohydrates, form the foundation of known life. While genetic material is primarily stored as nucleotide sequences in DNA and RNA, proteins operate as the primary functional systems within living organisms. Their architecture is strictly hierarchical: building from a linear sequence of amino-acid residues, through local and global spatial arrangements, up to the complex structures of multi-subunit assemblies. Because the specific three-dimensional conformation adopted by a protein critically dictates its physicochemical properties and biological activity, studying these structural states is an essential element in understanding the biochemical processes occurring in living cells.

Among naturally occurring proteins, about 20% contain at least one disulfide bond, and 50% of all cysteine residues found in protein sequences can form such linkages. The covalent bonding of sulfur atoms within cysteine side chains is one of the most common and vital post-translational modifications in higher organisms. However, despite their prevalence, the precise role of these bonds is not yet fully understood due to the inherent difficulties in experimentally studying the transient processes of their formation and breaking. In nature, their formation and rupture require appropriate conditions, such as a specific reducing-oxidizing (redox) environment, pH, or the presence of catalysts. Protein folding, stability, and enzymatic activity are therefore heavily dependent on the presence of both native and non-native disulfide bonds that form and break over the lifetime of the molecule.

Computational biophysics makes it possible to discover and understand phenomena that remain inaccessible to experimental studies due to high costs and current technical limitations. Molecular dynamics (MD) simulations constitute an excellent extension and complement to available experimental data, offering predictive insights down to the atomic level and individual amino-acid residues.

The aim of this doctoral dissertation is to investigate the role of disulfide bonds in biologically important proteins, relating their presence, absence, or dynamic exchange to the structure, stability, and function of the systems in which they occur. To achieve this, the research employs both classical static models and novel dynamic models of disulfide bonds. In the static model, bonds are permanently defined for the duration of the simulation. In contrast, the dynamic model preselects amino-acid residues capable of forming and breaking bonds on the fly. The introduction of this capability is a profound novelty in force fields; it relies on a new pseudopotential that successfully reproduces the energy barriers associated with disulfide bond breaking under specific studied redox conditions.

This methodological framework is applied across four distinct but complementary studies. Initially, the dynamic model is utilized to probe the mechanical stability of Ribonuclease A. By employing steered molecular dynamics (SMD) under external pulling forces, the research

reveals how varying the local redox environment directly governs the probability of disulfide rupture, fundamentally altering the protein's mechanical unfolding pathways.

Building on this understanding of covalent stabilization, the focus shifts to rational drug design. Through systematic disulfide-bond engineering, a series of truncated peptide inhibitors targeting the HVEM–LIGHT immune checkpoint were developed. This study demonstrates that minimizing covalent constraints down to a single, strategically placed disulfide bond yields the optimal balance of structural integrity and interfacial adaptability required for high-affinity receptor binding.

The thesis then explores the impact of engineered disulfides on intrinsically disordered proteins (IDPs). Using the amyloid-β42 monomer as a case study, a comprehensive multiscale protocol spanning all-atom (Amber) and coarse-grained (SIRAH, Martini) force fields was established. The findings uncover a fascinating structural paradox: introducing a disulfide cross-link actually decreases the peptide's native β-strand content, yet simultaneously shifts the conformational ensemble toward fibril-like states, effectively explaining the accelerated aggregation rates observed *in vitro*.

Finally, to make sense of the complex energetic data generated by these extensive simulations, an open-source software package named MDDG (Molecular Dynamics to Disconnectivity Graphs) was developed. MDDG applies Savitzky–Golay smoothing to translate MD trajectories directly into disconnectivity graphs, mapping the underlying energy landscape without requiring exhaustive geometry optimizations. Applying this tool to cysteine-rich proteins explicitly visualized how dynamic disulfide reshuffling dictates the shape and branching of the global folding funnel.

Ultimately, this dissertation demonstrates that disulfide bonds are not merely rigid structural pillars, but highly dynamic elements that actively regulate protein thermodynamics and mechanics. By integrating multiscale simulations, dynamic bonding topologies, and novel visualization tools, this work bridges the gap between atomic-level motions and macroscopic biological functions, offering a robust computational framework for future biophysical research.

Abstrakt

Polimery organiczne, takie jak białka, kwasy nukleinowe i węglowodany, stanowią fundament znanego nam życia. Podczas gdy materiał genetyczny jest przechowywany głównie w postaci sekwencji nukleotydów w niciach DNA i RNA, białka działają jako główne układy funkcjonalne w organizmach żywych. Ich budowa jest ściśle hierarchiczna: począwszy od liniowej sekwencji reszt aminokwasowych, poprzez ich lokalne i globalne ułożenie przestrzenne, aż po złożone struktury kompleksów wielopodjednostkowych. Ponieważ specyficzna trójwymiarowa konformacja przyjmowana przez białko w decydujący sposób wpływa na jego właściwości fizykochemiczne i aktywność biologiczną, badanie tych stanów strukturalnych jest istotnym elementem w rozumieniu i wyjaśnianiu procesów biochemicznych zachodzących w komórkach.

Wśród naturalnie występujących białek około 20% zawiera co najmniej jedno wiązanie dwusiarczkowe, a 50% wszystkich reszt cysteiny występujących w sekwencjach białkowych może tworzyć takie wiązania. Kowalencyjne wiązanie atomów siarki w obrębie łańcuchów bocznych cysteiny jest jedną z najczęstszych i najważniejszych modyfikacji potranslacyjnych w organizmach wyższych. Jednakże, pomimo ich powszechności, dokładna rola tych wiązań nie jest jeszcze w pełni zrozumiała z powodu nieodłącznych trudności w eksperymentalnym badaniu przejściowych procesów ich powstawania i pękania. W naturze ich tworzenie i rozrywanie wymaga odpowiednich warunków, takich jak specyficzne środowisko redukująco-utleniające (redoks), odpowiednie pH czy obecność katalizatorów. Zwijanie białek, ich stabilność i aktywność enzymatyczna są zatem silnie zależne od obecności zarówno natywnych, jak i nienatywnych wiązań dwusiarczkowych, które tworzą się i pękają w trakcie cyklu życia cząsteczki.

Biofizyka obliczeniowa pozwala na odkrywanie i zrozumienie zjawisk, które pozostają niedostępne dla badań eksperymentalnych z powodu wysokich kosztów i obecnych ograniczeń technicznych. Symulacje dynamiki molekularnej (MD) stanowią doskonałe rozszerzenie i uzupełnienie dostępnych danych eksperymentalnych, oferując możliwości predykcyjne schodzące do poziomu atomowego i pojedynczych reszt aminokwasowych.

Celem niniejszej rozprawy doktorskiej jest zbadanie roli wiązań dwusiarczkowych w biologicznie ważnych białkach, powiązanie ich obecności, braku lub dynamicznej wymiany ze strukturą, stabilnością i funkcją układów, w których występują. Aby to osiągnąć, w badaniach wykorzystano zarówno klasyczne modele statyczne, jak i nowatorskie modele dynamiczne wiązań dwusiarczkowych. W modelu statycznym wiązania są zdefiniowane na stałe przez cały czas trwania symulacji. Z kolei model dynamiczny wstępnie typuje reszty aminokwasowe zdolne do tworzenia i zrywania wiązań w trakcie symulacji („w locie”). Wprowadzenie tej możliwości jest istotną nowością w polach siłowych; opiera się ono na

nowym pseudopotencjale, który z powodzeniem odtwarza bariery energetyczne związane z rozrywaniem wiązań dwusiarczkowych w określonych, badanych warunkach redoks.

Powyższe ramy metodologiczne zastosowano w czterech odrębnych, lecz uzupełniających się badaniach. Początkowo model dynamiczny został wykorzystany do zbadania stabilności mechanicznej rybonukleazy A. Poprzez zastosowanie ukierunkowanej dynamiki molekularnej (SMD) przy użyciu zewnętrznych sił rozciągających, badania ujawniły, w jaki sposób zróżnicowane lokalne środowisko redoks bezpośrednio determinuje prawdopodobieństwo rozerwania wiązań dwusiarczkowych, fundamentalnie zmieniając ścieżki mechanicznego rozwijania białka.

Opierając się na tym zrozumieniu kowalencyjnej stabilizacji, uwaga przenosi się na racjonalne projektowanie leków. Poprzez systematyczną inżynierię wiązań dwusiarczkowych opracowano serię skróconych inhibitorów peptydowych ukierunkowanych na punkt kontrolny układu odpornościowego HVEM–LIGHT. Badanie to wykazuje, że zminimalizowanie ograniczeń kowalencyjnych do pojedynczego, strategicznie umieszczonego wiązania dwusiarczkowego zapewnia optymalną równowagę pomiędzy integralnością strukturalną a zdolnością adaptacji na granicy faz, niezbędną do wiązania z receptorem z wysokim powinowactwem.

Następnie w rozprawie przeanalizowano wpływ zaprojektowanych wiązań dwusiarczkowych na białka wewnętrznie nieuporządkowane (IDP). Używając monomeru amyloidu $\beta 42$ jako studium przypadku, ustanowiono kompleksowy protokół wieloskalowy, obejmujący dynamikę pełnoatomową (pole siłowe Amber) oraz pola siłowe gruboziarniste (SIRAH, Martini). Odkrycia te ujawniają fascynujący paradoks strukturalny: wprowadzenie mostka dwusiarczkowego w rzeczywistości zmniejsza w peptydzie natywną zawartość struktur typu β-kartki, lecz jednocześnie przesuwa zbiór konformacyjny w kierunku stanów podobnych do fibryli, co skutecznie wyjaśnia przyspieszone tempo agregacji obserwowane *in vitro*.

Na koniec, aby ułatwić analizę złożonych danych energetycznych generowanych przez te obszerne symulacje, stworzono otwartoźródłowy pakiet oprogramowania o nazwie MDDG (ang. *Molecular Dynamics to Disconnectivity Graphs*). Program MDDG wykorzystuje wygładzanie Savitzky'ego–Golaya do bezpośredniego przekształcania trajektorii MD w grafy rozspójnień, odwzorowując leżący u ich podstaw krajobraz energetyczny bez konieczności przeprowadzania wyczerpujących optymalizacji geometrii. Zastosowanie tego narzędzia do białek bogatych w cysteinę wyraźnie zobrazowało, w jaki sposób dynamiczne tasowanie wiązań dwusiarczkowych dyktuje kształt i rozgałęzienia globalnego lejka zwijania.

Ostatecznie, niniejsza rozprawa udowadnia, że wiązania dwusiarczkowe nie są jedynie sztywnymi filarami strukturalnymi, ale wysoce dynamicznymi elementami, które aktywnie regulują termodynamikę i mechanikę białek. Poprzez integrację symulacji wieloskalowych, dynamicznych topologii wiązań oraz nowoczesnych narzędzi do wizualizacji, praca ta stanowi pomost pomiędzy ruchami na poziomie atomowym a makroskopowymi funkcjami biologicznymi, oferując solidne ramy obliczeniowe dla przyszłych badań biofizycznych.

Pamela Smardz
al. Lotników 32/46
Warszawa 02-668

Warsaw, 05.2026

Statement

by the author of a doctoral dissertation

As the author of the doctoral dissertation entitled: ”All-atom and coarse-grained studies of disulfide-bond role in biologically-important proteins”, prepared under the guidance of: dr hab. Paweł Krupa, submitted to the Scientific Council of the Institute of Physics of the Polish Academy of Sciences, I declare that:

the dissertation was written by me independently and does not contain content obtained in a way that violates the provisions of the Act of February 4, 1994, on Copyright and Related Rights (Journal of Laws of 2018 item 1191 as amended),

the dissertation has not previously been the subject of procedures related to obtaining an academic degree in other proceedings for awarding a doctoral degree,

the doctoral dissertation in the digital version placed on the attached data carrier is the same as the printed and submitted version in my procedure for awarding a doctoral degree,

I grant the Institute of Physics of the Polish Academy of Sciences a free of charge non-exclusive license to publish the full text of the above-mentioned electronic version of the doctoral dissertation on the Institute’s website in the ifpan.edu.pl domain, thus making its content available to everyone on the Internet.

I declare that the subject of the doctoral dissertation is not covered by legal confidentiality.

Pamela Smardz

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Dr. hab. Paweł Krupa, whose guidance, encouragement, and expertise have been invaluable throughout my PhD journey. His support and insightful feedback challenged me to grow both academically and personally, while his patience and understanding provided constant reassurance whenever the research tested my sanity. This work simply would not have been possible without his mentorship, and I am honoured to have had the privilege of learning under his guidance.

I would also like to extend my heartfelt thanks to Dr. hab. Adam Sieradzan from the University of Gdańsk. His unwavering support has been a source of inspiration from the very beginning, and his relentless optimism served as a brilliant reminder that every academic hurdle is just a temporary glitch. His belief in my potential has shaped my academic path, and his words of wisdom, along with his infectious positive energy, have been a constant source of motivation. I am deeply grateful for his dedication and support throughout this journey.

I am equally thankful to the professors in my division: Prof. dr hab. Mai Suan Li, Dr. hab. Bartosz Różycki, Dr. hab. Łukasz Cywiński, Prof. dr hab. Jan Mostowski, Dr. hab. Emilia Witkowska, Dr. hab. Piotr Deuar, and Dr. Piotr Szańskowski. Their expertise and insightful discussions have significantly enriched my academic experience. I am grateful for their encouragement and for the supportive environment that made this research incredibly rewarding.

Finally, for all the life-saving coffee breaks, casual chats, and shared camaraderie throughout my stay at the Institute of Physics, a massive thank you to my wonderful colleagues: Mgr. Eng. Hubert Dunikowski, Dr. Kwasi Nyandey, Mgr. Navid Kazemiseresht, Mgr. Mateusz Molenda, Mgr. Marek Teske, Mgr. Tejas Joshi, Dr. Paweł Rogowski, Dr. Quyen V. Vu, Dr. Soheil Arbabi, Dr. Luís Carnevale da Cunha, Dr. Tanausú Hernández Yanes, and Dr. King Lun Ng.

This thesis was supported by the National Science Centre, Poland, under grant Sonata 2019/35/D/ST4/03156. I gratefully acknowledge Polish high-performance computing infrastructure PLGrid (HPC Centers: ACK Cyfronet AGH) for providing computer facilities and support within computational grant no. PLG/2023/016624.

List of Abbreviations

- **IDP** – Intrinsically Disordered Protein
- **RNase A** – Ribonuclease A
- **PDB** – Protein Data Bank
- **NMR** – Nuclear Magnetic Resonance
- **D-I-TASSER** – Deep-learning-based Iterative Threading ASSEmbly Refinement
- **CD** – Circular Dichroism
- **AFM** – Atomic Force Microscopy
- **DFT** – Density Functional Theory
- **QM/MM** – Quantum Mechanics / Molecular Mechanics
- **MD** – Molecular Dynamics
- **SMD** – Steered Molecular Dynamics
- **QM** – Quantum Mechanics
- **HF** – Hartree–Fock
- **MM** – Molecular Mechanics
- **PME** – Particle Mesh Ewald
- **PBC** – Periodic Boundary Conditions
- **RF** – Reaction Field
- **GB** – Generalized Born
- **PB** – Poisson–Boltzmann
- **MM-GBSA** – Molecular Mechanics Generalized Born Surface Area
- **MM-PBSA** – Molecular Mechanics Poisson–Boltzmann Surface Area
- **SASA** – Solvent-Accessible Surface Area
- **AMBER** – Assisted Model Building with Energy Refinement
- **CHARMM** – Chemistry at HARvard Macromolecular Mechanics

- **CMAP** – Cross-Term Map
- **CG** – Coarse-Grained
- **PMF** – Potential of Mean Force
- **CASP** – Critical Assessment of Protein Structure Prediction
- **CAPRI** – Critical Assessment of PRedicted Interactions
- **MREMD** – Multiplexed Replica Exchange Molecular Dynamics
- **UNRES** – United Residue
- **SIRAH** – South American Initiative for a Rapid and Accurate Hamiltonian
- **MDDG** – Molecular Dynamics to Disconnectivity Graphs
- **aMD** – Accelerated Molecular Dynamics
- **REMD** – Replica Exchange Molecular Dynamics
- **WHAM** – Weighted Histogram Analysis Method
- **RMSD** – Root-Mean-Square Deviation
- **RMSF** – Root-Mean-Square Fluctuation
- $\mathbf{R_g}$ – Radius of Gyration
- **PCA** – Principal Component Analysis

Table of contents

1 Goal of the Thesis **1**

2 Theoretical Background and Methodology **3**

2.1 Biophysical background . . . 3

2.1.1 Proteins . . . 3

2.1.2 Disulfide Bonds . . . 7

2.2 Computational calculations . . . 12

2.2.1 Calculation Methods . . . 12

2.2.2 Molecular Mechanics . . . 15

2.2.2.1 MM-GBSA Binding Free-Energy Estimates . . . 18

2.2.2.2 Molecular Docking . . . 19

2.2.3 Force Fields . . . 21

2.2.4 All-Atom Force Fields . . . 22

2.2.4.1 AMBER Protein Force Fields . . . 24

2.2.4.1.1 AMBER ff14SBonlysc. . . . 24

2.2.4.1.2 AMBER ff19SB. . . . 25

2.2.5 Coarse-Grained Force Fields . . . 26

2.2.5.1 UNRES . . . 26

2.2.5.2 SIRAH . . . 28

2.2.5.3 Martini . . . 28

2.3 Molecular Dynamics . . . 30

2.3.1 Enhanced Sampling and Free-Energy Methods . . . 35

2.3.2 Steered Molecular Dynamics and Jarzynski Equality . . . 37

2.3.3 Umbrella Sampling . . . 39

2.3.4 Trajectory Analysis Metrics . . . 41

2.3.5 Calculation of Chemical Shifts . . . 42

2.4 Energy Landscape . . . 43

2.4.0.1 Disconnectivity Graphs . . . 44

2.5 Treatment of Disulfide Bonds in Molecular Dynamics Simulations . . . 46

2.5.1 Dynamic Disulfide Bond Treatment in UNRES 48
2.5.2 Finite-Distance Disulfide Restraints in All-Atom Simulations . . . 51

3 Results: Summary of Publications **55**
3.1 Mechanical Stability of Ribonuclease A Heavily Depends on the Redox Environment 57
3.2 Multilayered Computational Framework for Designing Peptide Inhibitors of HVEM–LIGHT Interaction 79
3.3 A Practical Guide to All-Atom and Coarse-Grained Molecular Dynamics Simulations Using Amber and Gromacs: A Case Study of Disulfide-Bond Impact on the Intrinsically Disordered Amyloid Beta 137
3.4 Visualising the Energy Landscape for a Molecular Dynamics Trajectory . . 186

4 Conclusion **205**

References **209**

Appendix Authorship statements **223**
.1 Mechanical Stability of Ribonuclease A Heavily Depends on the Redox Environment 223
.2 Multilayered Computational Framework for Designing Peptide Inhibitors of HVEM–LIGHT Interaction 227
.3 A Practical Guide to All-Atom and Coarse-Grained Molecular Dynamics Simulations Using Amber and Gromacs: A Case Study of Disulfide-Bond Impact on the Intrinsically Disordered Amyloid Beta 233
.4 Visualising the Energy Landscape for a Molecular Dynamics Trajectory . . 240

Chapter 1

Goal of the Thesis

The overarching goal of this doctoral research was to revise the traditional static interpretation of disulfide bonds[1, 2] by establishing a computational framework capable of capturing their dynamic, environment-dependent roles in protein structure, stability, and function. Conventional molecular mechanics typically treat disulfide bonds as passive stabilising elements. My motivation was to challenge this paradigm and characterise them as active modulators of protein energy landscapes that respond to environmental fluctuations, such as varying redox conditions, mechanical stress, and conformational rearrangements. Consequently, these studies sought to bridge the gap between atomistic-level details and the global thermodynamic properties of biomolecules, spanning systems that range from single proteins to multiprotein complexes and from stable folds to intrinsically disordered peptides.

The first specific objective was driven by the methodological limitations of standard force fields in modelling covalent bond rupture. My goal was to develop and apply a multiscale molecular modelling approach that accommodates dynamic disulfide bond formation and breakage[3]. By overcoming the constraints of static representations, I aimed to quantify how covalent connectivity and stability modulate the mechanical resistance of proteins on both local and global scales. Using ribonuclease A (RNase A) as a model system[4], I sought to test the hypothesis that protein mechanical stability is not an intrinsic constant but a variable governed by redox-sensitive topology, ultimately aiming to differentiate between cysteine residues preserved for structural reinforcement and those maintained for catalytic activity.

A second major objective was motivated by the practical challenges of rational drug design, specifically the synthetic complexity associated with disulfide-rich therapeutic peptides. My goal was to establish a multilayered computational protocol to identify minimal disulfide architectures capable of preserving or enhancing binding affinity[5]. I hypothesised that

for inherently unstructured ligands, the selective removal of native disulfide bonds could increase conformational adaptability. Therefore, this segment of the research aimed to determine whether reduced covalent stabilization could actually promote induced-fit binding mechanisms and improve target recognition.

This investigation into structural flexibility naturally directed my research toward the most challenging class of biomolecules: intrinsically disordered proteins (IDPs). My motivation here stemmed from a biophysical paradox observed in experimental studies. I sought to understand why the addition of an engineered disulfide bond to a disordered peptide increases its overall compactness and reduces its β-strand content, yet simultaneously accelerates its transition toward toxic fibrillar conformations[6]. My objective was to evaluate the performance of diverse force fields to quantify how disulfide bonds narrow the conformational ensemble of an IDP. Ultimately, I aimed to demonstrate that a single disulfide bond can act as a conformational switch redirecting a disordered monomer toward an amyloidogenic pathway, thereby providing theoretical insights to guide future experimental investigations into the early stages of neurodegenerative diseases.

An additional goal of the thesis, conducted in collaboration with Professor David Wales's group[7], was motivated by the need for better analytical tools to investigate the influence of covalent dynamics on protein energy landscapes. My objective was to map minimal-energy conformations to specific disulfide topologies and to define the pathways signifying the transition states of bond reshuffling. Through this mapping, I aimed to establish which specific disulfide bonds act as dominant factors in shaping the folding funnel.

Ultimately, my research aimed to provide a comprehensive computational template for the study of disulfide bonds across all levels of protein organisation. Rather than presenting an isolated collection of data, my objective was to formulate a versatile methodology that the scientific community can build upon, providing robust guidance for future investigations into disulfide-rich biomolecular systems.

Chapter 2

Theoretical Background and Methodology

2.1 Biophysical background

2.1.1 Proteins

Biophysics provides a quantitative framework for understanding biological systems by applying physical principles to molecular processes. In this context, proteins represent a central class of biomolecules whose structure and dynamics determine their biological function. The combination of experimental techniques and theoretical computational modeling enables the quantitative characterization of protein behavior across a broad range of spatial and temporal scales, thereby providing a robust foundation for molecular and structural biology studies.

Proteins are heteropolymers composed of amino-acid residues connected by peptide bonds, capable of adopting highly complex three-dimensional conformations. Their biological functions encompass virtually every process within the cell, including enzymatic catalysis, molecular recognition, intracellular transport, structural stabilization, and signal transduction. The architecture of a protein is traditionally described across multiple hierarchical levels. The primary structure is dictated by the linear sequence of amino acids. The secondary structure encompasses localized folding patterns, most notably α-helices and β-strands, which are stabilized by extensive hydrogen-bonding networks along the peptide backbone. The tertiary structure represents the global three-dimensional fold, a topology primarily driven by the burial of nonpolar residues and frequently locked into place by covalent bonds, such as disulfide bonds. Finally, the quaternary structure involves the assembly and spatial arrangement of multiple polypeptide subunits into functional multi-protein complexes.

While many proteins fold into well-defined, rigid native conformations characterized by tightly packed hydrophobic cores[8], IDPs[9] and highly flexible peptide regions lack a single stable fold. Instead, they exist as highly dynamic, fluctuating conformational ensembles. Their relatively flat and frustrated free-energy landscapes allow them to rapidly interconvert between multiple transient states. This structural plasticity enables adaptive processes such as "folding-upon-binding" during molecular recognition events, allowing short peptide sequences to mold themselves to larger target receptors[10]. However, this inherent flexibility also renders IDPs susceptible to pathological misfolding and aberrant aggregation into β-sheet-rich fibrillar structures, a process heavily implicated in neurodegenerative diseases. The relationship between amino-acid sequence, structure, dynamics, and biological function is therefore intrinsically complex due to this conformational heterogeneity. Consequently, IDPs are more accurately described as statistical ensembles of interconverting conformational states rather than static molecular entities[11].

Protein stability is governed by thermodynamic principles in which the native conformation generally corresponds to the global or local minimum on the free-energy landscape under physiological conditions. Protein folding occurs spontaneously when the Gibbs free-energy change associated with folding (ΔG) is negative, indicating that the folded state is thermodynamically more favorable than the unfolded ensemble. This relationship is expressed by the Gibbs free-energy equation:

$$\Delta G = \Delta H - T\Delta S. \tag{2.1}$$

Here, ΔH denotes the enthalpic contribution arising from noncovalent interactions, including hydrogen bonding, electrostatic interactions, and van der Waals forces. Meanwhile, ΔS represents the entropic penalty associated with restricting the flexible unfolded polypeptide chain into a defined geometry, counterbalanced by the favorable entropy gain from solvent reorganization. A principal driving force underlying protein folding is the hydrophobic effect[12], whereby nonpolar amino-acid residues become sequestered within the protein core, thereby minimizing unfavorable interactions with water and increasing the overall entropy of

the surrounding solvent molecules. Furthermore, because temperature (T) directly scales the entropic term, protein stability is highly sensitive to thermal changes[13]. High temperatures drive thermal denaturation as the $-T\Delta S$ term dominates and unravels the chain, whereas cold denaturation can also occur due to the temperature-dependent weakening of the hydrophobic effect[14].

This thermodynamic description is closely related to the Anfinsen dogma[15, 16], which postulates that the native three-dimensional structure of a protein is determined primarily by its amino-acid sequence under appropriate environmental conditions. Crucially, Anfinsen's pioneering experimental studies on ribonuclease A demonstrated that chemically denatured proteins can spontaneously regain their native conformation and full biological activity following the removal of denaturants and the reoxidation of their disulfide cross-links[17]. This seminal observation proved that the information necessary for folding, and for the correct pairing of structural constraints, is entirely encoded within the primary structure.

Protein folding can therefore be conceptualized as a self-organizing process in which a polypeptide chain explores a funnel-shaped conformational energy landscape[18, 19]. Within this multidimensional funnel, the protein navigates through various local minima (basins) separated by energy barriers (transition states) to achieve thermodynamic stability through free-energy minimization. Nevertheless, folding pathways are strongly influenced by kinetic effects, including the presence of high activation barriers and metastable intermediates. These intermediate basins may transiently trap the protein in non-native conformations prior to the attainment of the native state, dictating the timescale and trajectory of the folding process[20, 21].

While this funnel model elegantly describes folding in isolation, physiological proteins function within highly heterogeneous cellular environments characterized by continuous fluctuations in pH, ionic strength, temperature, molecular crowding, mechanical stress, and redox conditions[22]. These physicochemical parameters profoundly influence conformational stability by physically reshaping the underlying free-energy landscape. In particular, the local reducing-oxidizing (redox) environment dictates the dynamic formation and rupture of covalent disulfide bonds[23, 24]. The breaking and reforming of these bonds can drastically

tilt the free-energy landscape, altering a protein's thermodynamic stability and its resistance to external mechanical forces. Consequently, protein conformational equilibria are determined not only by amino-acid sequence but also by dynamic environmental and mechanical factors that regulate the relative populations of folded, unfolded, and intermediate states[25].

The ability to computationally investigate these intricate thermodynamic and kinetic processes fundamentally relies on the availability of accurate three-dimensional structural data. Historically, high-resolution protein structures have been determined through laborious experimental techniques such as X-ray crystallography[26], nuclear magnetic resonance (NMR) spectroscopy[27], and cryogenic electron microscopy (cryo-EM)[28], with the resulting atomic coordinates carefully curated in the Protein Data Bank (PDB)[29, 30]. More recently, artificial-intelligence-based structure prediction methods, most notably AlphaFold[31] and related implementations such as OpenFold[32], have transformed structural biology by enabling high-quality structural models to be inferred directly from amino-acid sequences. In parallel, hybrid frameworks such as D-I-TASSER (Deep-learning-based Iterative Threading ASSEmbly Refinement)[33] have demonstrated that combining deep-learning-derived spatial restraints with more traditional template-based threading, fragment assembly, and structural refinement can further strengthen predictive modeling, particularly when evolutionary information and structural analogies are informative. However, these approaches are not universally accurate, and their performance depends strongly on the nature of the studied system. In particular, flexible, intrinsically disordered, multimeric, or environmentally responsive proteins remain challenging to describe fully using static prediction alone. Consequently, although experimental structures, AI-based predictors, and hybrid modeling frameworks provide indispensable structural blueprints, they often do not capture the full conformational heterogeneity of IDPs, the transient pathways of mechanical unfolding, or the dynamic behavior of covalent constraints such as disulfide bonds over time. In order to truly understand biological function, these structures must be brought to life using computational biophysics, allowing researchers to dynamically explore the complex energy landscapes that govern protein behavior.

2.1.2 Disulfide Bonds

Disulfide bonds are covalent linkages formed between pairs of cysteine residues through the oxidation of their thiol groups, resulting in the formation of sulfur–sulfur (S–S) bonds. They can be classified as intramolecular (within a single polypeptide chain) or intermolecular (linking separate chains), and they occur in a wide variety of proteins and peptide, including extracellular enzymes, receptors, antibodies, and structural scaffolds. In many cysteine-rich families, such as defensins[34], non-specific lipid transfer proteins (nsLTPs)[35], and numerous venom peptides[36], characteristic patterns of disulfide connectivity are an integral part of the fold. In reducing environments such as the cytosol, cysteine residues are predominantly maintained in their unbound thiol form, whereas more oxidizing environments, including the endoplasmic reticulum and extracellular space, favor disulfide bond formation. Consequently, the local redox environment acts as a macroscopic environmental variable that dictates the physical topology, stability, and conformational dynamics of the protein[37, 24, 38].

At the structural level, disulfide bonds exhibit preferred physical geometries and stereochemistry. The dihedral angles around the C_α–C_β–S–S–C_β–C_α backbone define distinct disulfide conformers, which are non-equivalent in terms of structural strain and local packing [39]. These geometric preferences influence how a given linkage responds to mechanical deformation and how strictly it constrains backbone flexibility. As a result, the same pair of cysteine residues can provide very different stabilizing or regulatory physical effects depending on their spatial context within the three-dimensional fold[40, 41].

At the fundamental level, disulfide bond formation is a chemical oxidation reaction,

$$2\mathrm{R{-}SH} \rightarrow \mathrm{R{-}S{-}S{-}R} + 2\mathrm{H}^+ + 2e^-. \tag{2.2}$$

While this process involves electron transfer and is often catalyzed in vivo by dedicated redox enzymes (such as protein disulfide isomerases), from a biophysical perspective, this reaction represents a discrete shift in the physical connectivity of the polymer. The chemical event translates into a strong topological constraint, linking previously independent regions of the

polypeptide chain. The reversibility of this reaction means that proteins can dynamically switch between constrained (oxidized) and unconstrained (reduced) physical states in response to changes in the surrounding environment, altering the underlying free-energy landscape[42].

Under fixed temperature and pressure, the equilibrium between the reduced thiol state and the oxidized disulfide state is governed by the redox half-reaction

$$\mathrm{R{-}S{-}S{-}R} + 2\mathrm{H}^{+} + 2e^{-} \rightleftharpoons 2\mathrm{R{-}SH}. \tag{2.3}$$

Accordingly, the Nernst equation may be written in its explicit form as

$$E = E^{\circ} - \frac{RT}{2F} \ln \frac{[\mathrm{R{-}SH}]^2}{[\mathrm{R{-}S{-}S{-}R}][\mathrm{H}^{+}]^2}\,. \tag{2.4}$$

Here, E is the redox potential of the environment, E° is the standard potential of the disulfide/thiol couple, R is the gas constant, T is the absolute temperature, and F is the Faraday constant. The factor of 2 reflects the transfer of two electrons in the reduction of one disulfide bond, and bracketed terms denote the effective concentrations of the species.

At constant pH, the proton concentration is fixed and its contribution is absorbed into the biochemical standard potential, $E^{\circ\prime}$. In this case, the expression is rewritten as

$$E = E^{\circ\prime} - \frac{RT}{2F} \ln \frac{[\mathrm{R{-}SH}]^2}{[\mathrm{R{-}S{-}S{-}R}]}\,. \tag{2.5}$$

This form is widely utilized in biophysical discussions because it directly links the environmental redox potential to the statistical preference for reduced versus oxidized macrostates.

More reducing conditions, corresponding to lower values of E, shift the equilibrium toward the thiol state and therefore toward more flexible, topologically unconstrained protein conformations. Conversely, more oxidizing conditions favor the disulfide-bonded state, which introduces covalent constraints, narrows the accessible conformational ensemble, and may stabilize specific folded or mechanically resistant structures. In this way, disulfide bonds function as redox-responsive molecular switches, coupling changes in the chemical environment to large-scale structural and functional transitions.

In real proteins, however, the effective redox behavior of a given cysteine pair is rarely determined solely by the bulk solution potential. Local factors such as solvent exposure, the surrounding electrostatic environment, specific cysteine pK_a values[43], and geometrical strain effectively modulate the local standard redox potential ($E^{\circ\prime}$) of the site. Therefore, while the Nernst equation provides the fundamental thermodynamic framework, the actual reactivity and equilibrium shift of a specific disulfide bond remain strongly context-dependent.

Disulfide bonds perform diverse functional roles by manipulating the physical properties of the protein. In many secreted and extracellular proteins, multiple disulfide bonds act as rigid cross-links that dramatically increase physical resistance to thermal denaturation and mechanical unfolding[44]. In enzymes, disulfide bonds often stabilize specific catalytic loop geometries rather than participating in the chemistry directly[45]. In receptors and ligands, they physically shape the binding interface, maintaining necessary loop architectures required for high-affinity recognition.

In cysteine-rich families, conserved disulfide patterns define characteristic, highly robust folds, such as the cystine-knot motif[46] or the four-disulfide core seen in many allergens, including nsLTPs[47]. These motifs illustrate how evolution has exploited topological constraints to stabilize compact, mechanically resilient architectures. Conversely, the mispairing of disulfide bonds traps the protein in suboptimal physical geometries, producing frustrated, misfolded states that are often implicated in disease.

From a thermodynamic perspective, disulfide bonds introduce covalent constraints that drastically reduce the conformational entropy of the unfolded state[48]. By restricting the volume of accessible phase space for the unfolded chain, they lower the entropic penalty of folding, which serves as a primary mechanism for stabilizing the folded conformation and increasing overall resistance to mechanical deformation. However, the introduction of a covalent cross-link does not universally confer stability. If a disulfide bond, whether naturally occurring, engineered, or transiently mispaired, forces the polypeptide backbone into an unfavorable geometry, the resulting enthalpic strain can easily offset the entropic benefit[49]. In such scenarios, the constraint may destabilize the native fold, induce topological frustration, or redirect the conformational ensemble toward alternative structural basins,

including misfolded or aggregation-prone states. At the same time, because disulfide bonds are chemically reversible, they provide the protein with dynamic, regulatory capabilities, allowing the underlying free-energy landscape to be continually reshaped in response to environmental fluctuations.

Under external mechanical stress, the rupture of native disulfide bonds can physically redirect a protein's unfolding pathways. By redistributing mechanical strain across the polypeptide scaffold, these covalent links fundamentally alter the force–extension behavior of the molecule[50]. This mechanochemical coupling demonstrates that disulfide bonds do not act merely as static structural pillars, but function as highly dynamic, load-bearing mechanical elements that can yield under specific applied forces, thereby absorbing stress or regulating mechanotransduction.

Furthermore, in the context of IDPs, the presence or engineered introduction of disulfide constraints physically restricts highly flexible conformational ensembles[51]. This restriction inherently biases the disordered chain toward particular topologies, such as transient loops or hairpin-like structures, which can significantly modulate the propensity for intermolecular interactions and macroscopic aggregation. Similarly, in the rational design of small therapeutic or receptor-derived peptides, carefully placed disulfide bonds are frequently utilized to pre-organize the binding motif[52]. By constraining a peptide into its biologically active conformation prior to target engagement, the cross-link minimizes the entropic penalty of binding. In this capacity, disulfides act as precise biophysical design elements that strike a delicate balance between required structural rigidity and necessary interfacial adaptability.

Experimentally, the physical and thermodynamic effects of disulfide bonds are characterized using a combination of spectroscopic and structural techniques. NMR spectroscopy and X-ray crystallography provide direct spatial coordinates of the constrained folds, whereas circular dichroism (CD)[53] spectroscopy tracks global shifts in secondary structure upon disulfide reduction. Crucially, single-molecule techniques, such as atomic force microscopy (AFM)[54] and optical tweezers[55], enable the direct application of mechanical force to individual proteins, measuring the precise physical rupture forces and extension lengths asso-

ciated with disulfide-stabilized domains. These physical measurements directly complement the steered molecular dynamics methods utilized in this work.

While the fundamental formation and rupture of a disulfide bond are inherently quantum-chemical processes driven by electron transfer and orbital reorganization, explicitly modeling these events presents severe computational challenges. Rigorous *ab initio* wavefunction-based methods provide the highest level of physical accuracy but exhibit steep computational scaling, strictly limiting their application to small, isolated molecular fragments. Density functional theory (DFT) offers a more favorable computational scaling by operating on the electron density rather than the many-body wavefunction. However, applying full-system DFT to entire proteins, especially within explicit solvent, remains computationally unfeasible for generating meaningful dynamical trajectories.

To bridge the gap between reactive chemistry and macromolecular environments, hybrid quantum mechanics/molecular mechanics (QM/MM)[56] approaches partition the system. In this framework, the reactive site (the disulfide bond) is treated at a quantum level (e.g., using DFT), while the surrounding protein scaffold and solvent are described by classical mechanics. Although QM/MM successfully circumvents the spatial limitations of pure quantum models, it remains severely bottlenecked by the computational cost of evaluating the QM region at every integration step. Consequently, while QM/MM is excellently suited for studing localized reaction mechanisms and transition states, it cannot sample the microsecond-to-millisecond timescales required to observe macroscopic phenomena[57], such as large-scale conformational shifts, mechanical unfolding, and global free-energy-landscape reorganization.

Because simulating these extensive biological timescales is paramount, computational biophysics typically abstracts disulfide chemistry into effective physical interactions. In standard classical molecular dynamics (MD) simulations, native disulfide bonds are encoded as permanent harmonic constraints within the molecular topology. While this static representation is appropriate for stable, oxidized states, it cannot describe bond rupture, dynamic exchange, or redox-responsive structural yielding.

To overcome this limitation and recover dynamic behavior within a purely classical framework, classical force fields can be extended by introducing reactive pseudopotentials.

By utilizing distance-dependent or angle-dependent empirical terms that approximate the macroscopic energy barriers of disulfide formation and breaking, the covalent connectivity of the protein is allowed to respond dynamically to mechanical perturbation or changes in the chemical environment without explicitly calculating the underlying electronic structure.

At even lower resolutions, coarse-grained models incorporate disulfide bonds as effective interaction terms between simplified representative sites. By drastically reducing the system's degrees of freedom, these models facilitate the exploration of massive structural transitions, alternative topological connectivities, and long-timescale features of the energy landscape that remain entirely inaccessible to all-atom MD.

Across this methodological hierarchy, classical and coarse-grained molecular simulations provide a practical multiscale framework for investigating how dynamic disulfide constraints influence folding funnels, mechanical stability, and conformational heterogeneity.

2.2 Computational calculations

2.2.1 Calculation Methods

Computational modeling of molecular systems encompasses a hierarchy of methods that differ in both their level of physical approximation and computational scaling. At the most fundamental level, *ab initio* QM approaches solve the electronic Schrödinger equation explicitly without relying on empirical parameters. By operating from first principles, they provide the highest-resolution detail regarding electronic structure, orbital hybridization, chemical bonding, and charge distribution.

Different *ab initio* methods vary considerably in accuracy and computational cost. Hartree-Fock (HF) theory treats electron-electron repulsion in an averaged (mean-field) manner[58, 59], whereas more advanced post-Hartree-Fock methods explicitly account for complex electron correlation effects[60–64]. In practice, techniques such as configuration interaction (CI) and multi-configurational self-consistent field (MCSCF) extend HF by improving the treatment of electronic correlation and, in the case of MCSCF, optimizing both the configuration

coefficients and the orbitals themselves, which is especially useful for excited states. These approaches yield highly accurate predictions of molecular geometries, reaction energies, and spectroscopic properties. However, their computational cost scales steeply, often proportional to N^4 for basic HF up to N^7 for advanced coupled-cluster methods[65], where N is the number of basis functions. Consequently, the practical application of pure *ab initio* methods is strictly limited to relatively small molecular systems or highly localized reactive regions within larger complexes.

Although some authors also classify DFT[66, 67] as *ab initio*, it is often treated as a distinct quantum-chemical framework. In DFT, the electronic density rather than the full many-electron wavefunction is the central variable, and the total energy is expressed as a functional of the electron density. The method is based on the Hohenberg–Kohn theorems, which establish that the ground-state density uniquely determines the properties of the system and that the ground-state energy can be obtained variationally. Compared with wavefunction-based methods, DFT reduces the dependence from $3N$ variables to three spatial coordinates, making it much more tractable for larger systems. Common DFT families include local density approximations (LDA)[68], gradient-corrected functionals such as BLYP, and hybrid functionals such as B3LYP[69].

To overcome some of the computational limitations associated with fully quantum mechanical calculations, semi-empirical methods introduce targeted approximations and empirical parameterizations into the rigorous QM formalism. These approaches simplify the mathematical treatment by restricting calculations to valence electrons, neglecting selected computationally expensive multi-center integrals, and replacing them with parameters fitted to experimental data or high-level *ab initio* results.

As a result, semi-empirical methods substantially reduce computational overhead while retaining the general framework of quantum mechanics. They are particularly useful for studying moderately sized molecules, performing rapid conformational analyses, and exploring reaction pathways in a preliminary manner. The Polish text notes that these methods form a bridge between exact but expensive quantum chemistry methods and simplified empirical models, and that many historical schemes, such as PPP, CNDO, INDO, and NDDO, were

developed by introducing increasingly strong overlap approximations. Later parameterized methods such as MNDO[70], AM1[71], PM3[72], PM6[73], and PM7[74] are typically fitted to thermochemical data, dipole moments, ionization potentials, and molecular geometries.

A key example is AM1 (Austin Model 1), which was developed as an extension of MNDO and improved the treatment of core–core repulsion to better reproduce van der Waals interactions and hydrogen bonding. Compared with MNDO, AM1 generally gives more accurate heats of formation. PM3 further refined the approach by optimizing all parameters, while PM6 and PM7 were tuned for better reproduction of crystal structures and heats of formation across a broader chemical database[73, 74]. In general, semi-empirical methods are most accurate for chemical systems close to those used in their parameterization[74], and their performance declines for unusual or weakly represented compounds. Nevertheless, because of their computational cost, they remain valuable for peptides, small proteins, and other large molecular systems where full *ab initio* treatment is impractical.

At larger spatial and temporal scales relevant to biophysics, empirical molecular mechanics (MM) methods abandon explicit electronic structure calculations entirely[75]. Instead, atoms are represented as classical point masses interacting through parameterized potential energy functions, collectively known as force fields[76]. Within this framework, molecular behavior is described using classical mechanics, while chemical interactions are approximated through effective interaction terms derived from experimental observations and quantum mechanical calculations. These methods operate at the level of atoms rather than electrons and rely entirely on experimentally derived parameters, which means they can only be applied to systems that have been adequately parameterized in advance.

By removing explicit electronic degrees of freedom, computational efficiency is drastically improved. This simplification allows MM algorithms to scale highly favorably. Short-range bonded interactions scale linearly ($O(N)$), and long-range electrostatics, when handled by cutoff algorithms[77] such as the Particle Mesh Ewald (PME) method[78, 79], scale as $O(N \log N)$, where N is the number of atoms. Consequently, empirical methods are uniquely suited for simulating massive biomolecular systems, such as whole proteins, nucleic acids, lipid membranes, and viral assemblies, comprising hundreds of thousands to millions of

atoms, with an up-to-date example of 53M atom system run on high performin computer with 65k CPU cores[80]. They are especially useful when the main goal is to capture large-scale structural behavior rather than electronic detail.

Together, *ab initio*, semi-empirical, and empirical approaches form a flexible, hierarchical multiscale framework for computational molecular modeling. The selection of an appropriate method is strictly dictated by the balance between the physical phenomena of interest, the required level of accuracy, and available computational resources. While quantum mechanical approaches are indispensable when studying chemical reactivity or localized electron transfer, their limitations become particularly evident in biomolecular simulations. Proteins and other biological assemblies undergo macroscopic conformational transitions on microsecond-to-millisecond timescales. Even with modern high-performance computing resources, such expansive behaviors remain completely inaccessible to *ab initio* treatment.

2.2.2 Molecular Mechanics

Within molecular mechanics, a molecular system is described by a classical potential energy function that is mathematically decomposed into bonded and non-bonded contributions[81, 82]:

$$U = U_{\text{bonded}} + U_{\text{nonbonded}}. \tag{2.6}$$

Bonded interactions (U_{bonded}) preserve the local connectivity and geometry of the molecule and are typically expressed as the sum of bond stretching, angle bending, and torsional (dihedral) terms:

$$U_{\text{bonded}} = \sum_{\text{bonds}} \frac{1}{2} k_b (r - r_0)^2 + \sum_{\text{angles}} \frac{1}{2} k_\theta (\theta - \theta_0)^2 + \sum_{\text{dihedrals}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)]. \tag{2.7}$$

Here, covalent bonds and angles are modeled as harmonic oscillators with specific force constants (k_b, k_θ) and equilibrium values (r_0, θ_0). Torsional rotations are described by periodic Fourier series that dictate the rotational energy barriers along the molecular backbone. Crucially, this classical potential energy function can be easily augmented with external biasing

potentials, enabling techniques such as Steered Molecular Dynamics (SMD) to simulate single-molecule mechanical pulling experiments.

Non-bonded interactions ($U_{\text{nonbonded}}$) account for the physical forces between atoms separated by at least three bonds (or belonging to different molecules). These are divided into van der Waals (dispersion/repulsion) and electrostatic interactions:

$$U_{\text{nonbonded}} = \sum_{i<j} \left[\varepsilon_{ij} \left(\left(\frac{R_{min,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{min,ij}}{r_{ij}} \right)^{6} \right) + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right]. \tag{2.8}$$

The van der Waals forces are modeled using a Lennard-Jones 12-6 potential[83], balancing short-range Pauli exclusion repulsion ($1/r^{12}$) with long-range attractive London dispersion forces ($1/r^{6}$). The electrostatic interactions are calculated using Coulomb's law[84] based on fixed, atom-centered partial charges (q_i, q_j).

The decomposition of the potential energy into physically distinct components provides both computational efficiency and clear physical interpretability. Bonded interactions primarily dictate local structural organization, rigidity, and molecular flexibility. Non-bonded interactions govern intermolecular recognition, solvation phenomena, and long-range stabilization effects (such as the formation of hydrophobic cores). Because each energy contribution corresponds to a distinct physical interaction, force-field parameters can be systematically optimized against empirical thermodynamic data and high-level quantum mechanical calculations.

An essential distinction must be made between the computational treatment of short-range and long-range non-bonded interactions. Because van der Waals forces decay rapidly with distance (governed by the $1/r^{6}$ dispersion term), they are commonly truncated beyond a specified cutoff radius (typically 8 to 12 Å) to enhance computational efficiency with a negligible loss of accuracy[77]. In contrast, electrostatic interactions decay slowly (as $1/r$) and remain physically significant over substantially longer distances. Direct, abrupt truncation of these forces produces severe unphysical artifacts, such as artificial charge accumulation at the cutoff boundary, which fundamentally distorts structural behavior and solvent dynamics.

To address this limitation, modern macroscopic simulations employing Periodic Boundary Conditions (PBC) rely on specialized treatments of long-range electrostatics. For state-of-the-art all-atom force fields (such as AMBER (Assisted Model Building with Energy Refinement) and modern CHARMM (Chemistry at HARvard Macromolecular Mechanics)[85]), the standard approach is the Particle Mesh Ewald[78, 79] method. PME mathematically decomposes the conditionally convergent electrostatic interactions into two rapidly converging series: a short-range real-space sum evaluated directly, and a long-range reciprocal-space sum evaluated on a grid using Fast Fourier Transforms (FFT). This dual approach provides a computationally efficient and highly accurate treatment of infinite-range electrostatics.

However, PME is not universally applied across all simulation scales. Certain force fields, particularly coarse-grained models like Martini, were explicitly parameterized using specific cutoff schemes[86]. In these cases, infinite-range electrostatics are avoided, and truncation artifacts are instead mitigated using continuous shift or switch functions, which smoothly taper the electrostatic potential to zero at the cutoff radius, or through generalized Reaction Field (RF) methods, which treat the space beyond the cutoff as a continuous dielectric medium[87]. Therefore, the choice of electrostatic treatment must strictly align with the original parameterization of the utilized force field to ensure thermodynamic accuracy.

Because most biomolecular processes occur in aqueous environments, accurate representation of solvent effects is paramount. In explicit solvent models (such as TIP3P or OPC)[88–90], individual water molecules are represented directly, enabling detailed modeling of solvation structure, specific hydrogen bonding, and solvent-mediated bridging interactions. However, explicitly modeling solvent heavily inflates computational costs. Implicit solvent models provide a computationally less demanding alternative by replacing discrete water molecules with a continuous macroscopic dielectric medium that reproduces average solvation effects. Approaches such as Generalized Born (GB) and Poisson-Boltzmann (PB) models form the foundation of the Molecular Mechanics Generalized Born Surface Area (MM-GBSA) and Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) methodologies[91, 92]. In this thesis, MM-GBSA is utilized to efficiently estimate the effective binding free energies of peptide inhibitors without the massive computational overhead of explicit water sampling.

2.2.2.1 MM-GBSA Binding Free-Energy Estimates

For the efficient estimation of relative binding affinities from molecular-dynamics trajectories, end-point free-energy methods such as MM-GBSA are widely used. In this framework, the binding free energy of a ligand to a receptor is approximated as the free-energy difference between the bound complex and the two separated binding partners,

$$\Delta G_{\text{bind}} = G_{\text{complex}} - G_{\text{receptor}} - G_{\text{ligand}}. \tag{2.9}$$

In practice, the free energy of each state is decomposed into molecular-mechanics, solvation, and entropic contributions,

$$G = E_{\text{MM}} + G_{\text{solv}} - TS, \tag{2.10}$$

so that

$$\Delta G_{\text{bind}} = \Delta E_{\text{MM}} + \Delta G_{\text{solv}} - T\Delta S. \tag{2.11}$$

The molecular-mechanics term is typically written as

$$\Delta E_{\text{MM}} = \Delta E_{\text{vdW}} + \Delta E_{\text{ele}}, \tag{2.12}$$

where ΔE_{vdW} and ΔE_{ele} denote the changes in van der Waals and electrostatic interaction energies, respectively. The solvation free energy is decomposed into polar and nonpolar parts,

$$\Delta G_{\text{solv}} = \Delta G_{\text{GB}} + \Delta G_{\text{np}}, \tag{2.13}$$

where ΔG_{GB} is the polar solvation contribution estimated with the GB model, and ΔG_{np} is the nonpolar contribution, usually approximated from the solvent-accessible surface area (SASA),

$$G_{\text{np}} = \gamma \cdot \text{SASA} + \beta. \tag{2.14}$$

MM-GBSA therefore provides an approximate but computationally efficient route to rank binding poses, compare ligand variants, and identify dominant energetic contributions at the

binding interface. Because the method is based on snapshots extracted from equilibrium trajectories, it is particularly useful for comparative studies of related complexes rather than for obtaining absolute experimental free energies. Its major advantages are modest computational cost and straightforward decomposition into residue-level energetic terms, which can be used to identify hotspots of molecular recognition[93].

An important limitation of standard MM-GBSA is the treatment of entropy[94]. In many applications, the entropic term is neglected or estimated only approximately, which may bias the final binding free energy, especially for highly flexible peptides and intrinsically disordered ligands. To partially account for configurational entropy, normal-mode analysis can be applied to representative snapshots. In this approach, the vibrational modes around a local minimum are approximated harmonically, allowing estimation of translational, rotational, and vibrational entropic contributions. Although computationally demanding and sensitive to structural preprocessing, normal-mode entropy corrections provide a useful approximation when comparing ligands that differ substantially in flexibility[95].

Because MM-GBSA evaluates free energies using end-point states rather than an explicit reaction coordinate, it does not capture the full kinetic pathway of association or dissociation. Nevertheless, when combined with molecular docking, long-timescale MD refinement, and complementary methods such as steered molecular dynamics or umbrella sampling, MM-GBSA provides a robust and practical component of multiscale binding-affinity assessment.

2.2.2.2 Molecular Docking

Molecular docking is a computational method used to predict the preferred binding mode of a ligand to a receptor, generating structural hypotheses for molecular recognition. Traditionally, this process aims to identify poses that are both geometrically compatible with the receptor surface and energetically favorable[96]. As such, docking provides an efficient first-pass strategy for exploring candidate binding sites and initial complex geometries prior to more computationally expensive molecular-dynamics refinement.

Historically, and still predominantly for small organic molecules, the docking procedure relies on two conceptually distinct components: a physical search algorithm and an empirical

scoring function. The search algorithm explores the translational, rotational, and conformational degrees of freedom of the ligand. Depending on the specific software, the receptor may be treated as rigid, partially flexible,[97] or represented by an ensemble of conformations. The scoring function subsequently evaluates each candidate pose using simplified mathematical approximations of shape complementarity, hydrogen bonding, electrostatics, and steric clash.

For protein–protein and protein–peptide systems, however, this traditional search-and-score paradigm has recently been transformed by the advent of end-to-end deep learning models. Frameworks such as AlphaFold-Multimer[98] and OpenFold bypass classical rigid-body docking entirely by predicting quaternary structures directly from concatenated amino-acid sequences. By leveraging co-evolutionary information and deep-learning-derived spatial restraints, these models effectively "co-fold" the receptor and the peptide ligand. This approach implicitly accounts for large-scale induced-fit effects and substantial conformational adaptations that have historically plagued classical docking algorithms[98, 99]. Nevertheless, while revolutionary for peptidic ligands, these pure AI predictors are currently restricted to standard amino acids and cannot natively dock small-molecule drugs or non-standard chemical modifications without relying on hybrid or heavily modified extensions.

Regardless of whether a complex is generated via classical scoring functions or AI-driven co-folding, the resulting structural models possess inherent limitations. Classical scoring functions provide only a rough approximation of binding affinity, while AI-predicted poses represent static, probabilistically generated snapshots that lack true thermodynamic context. For highly flexible peptides, binding often involves transient intermediate states, dynamic structural frustration, and solvent-mediated interactions that static models cannot fully capture.

Therefore, in modern multiscale biophysical workflows, both classical docking and AI structure prediction act not as definitive endpoints, but as sophisticated generators of initial receptor–ligand geometries. Once a highly probable pose is identified, molecular dynamics simulations are required to assess its structural stability in explicit solvent. Thus, docking and AI prediction serve as the crucial starting point for detailed thermodynamic and kinetic biophysical analysis.

2.2.3 Force Fields

Despite their profound computational efficiency, classical force fields possess a strict topological limitation: covalent bonds are permanently defined prior to the simulation and are represented by harmonic potentials that assume an infinitely rising energy penalty. As a consequence, standard force fields cannot describe spontaneous bond breaking or bond formation. This topological restriction limits their applicability to systems involving chemical reactions, proton transfer, or redox processes, including the structural formation and mechanochemical rupture of disulfide bonds.

To circumvent this limitation, reactive force fields, such as ReaxFF[100], extend standard classical molecular mechanics by introducing empirical, bond-order-dependent interaction terms that enable covalent bonds to form and dissociate dynamically. An alternative and more rigorous strategy is the hybrid QM/MM approach[56]. In this multiscale framework, the chemically active region is treated explicitly using quantum mechanics, while the surrounding macromolecular scaffold and solvent are represented using computationally efficient classical force fields.

However, both methodologies possess fundamental limitations that preclude their use in studying global protein dynamics. While QM/MM provides exceptional chemical accuracy, its immense computational overhead strictly restricts its application to highly localized phenomena on extremely short timescales (typically picoseconds to nanoseconds)[101]. Conversely, while reactive force fields are computationally faster than QM/MM, their limitation is not purely a matter of timescale, but of fundamental parameterization[102]. Because they are primarily optimized to reproduce reaction energies and transition-state geometries, they frequently fail to accurately balance the delicate non-bonded interactions; such as the hydrophobic effect, hydrogen-bonding networks, and backbone torsional preferences; that dictate macromolecular folding. Consequently, even if sufficient computational resources were available to extend these simulations to microsecond timescales, reactive force fields generally lack the conformational fidelity required to correctly capture the large-scale structural changes and macroscopic thermodynamic transitions that govern protein behavior.

To bridge the gap between rigid classical topologies and the extreme computational cost of reactive potentials, highly targeted pseudopotential approaches[103] have been developed[3]. By introducing specialized, distance-dependent restraints into classical force fields, specific bonds, such as disulfide bonds and other covalent cross-links[104], can dynamically break and reform based on macroscopic energy barriers that mimic distinct redox conditions. This method entirely bypasses the calculation of explicit electron transfer, permitting the observation of dynamic covalent reshuffling and mechanochemical yielding over extensive microsecond timescales. Ultimately, this methodological framework provides the computational foundation for the dynamic simulations performed in this thesis, enabling the detailed investigation of redox-dependent mechanical unfolding in model systems such as Ribonuclease A.

Finally, the level of molecular representation strongly influences computational efficiency and accessible simulation timescales. In all-atom models, every atom is represented explicitly, providing maximal structural resolution. In contrast, reduced representations, such as coarse-grained models, decrease the system's complexity by grouping multiple atoms into larger effective particles[105]. This reduction in degrees of freedom accelerates calculations twofold: it permits the use of longer integration timesteps, and it smooths the underlying energy landscape[106], thereby accelerating the sampling of large-scale conformational transitions. Collectively, these representations form a versatile multiscale framework, allowing researchers to investigate both fine atomistic details and global thermodynamic behaviors.

2.2.4 All-Atom Force Fields

Building upon the mathematical framework of molecular mechanics, all-atom (AA) force fields provide the highest-resolution classical representation of biomolecular systems[107]. By explicitly representing every atom, including hydrogens, this approach enables the detailed modeling of subtle atomistic interactions, such as intricate hydrogen-bonding networks, specific steric clashes, and precise electrostatic salt bridges. This explicit representation facilitates direct, one-to-one comparisons with high-resolution experimental structural data (e.g.,

NMR or X-ray crystallography), capturing fine structural features that remain inaccessible to reduced-resolution models.

However, this classical formulation relies on fundamental approximations that impose strict physical limitations. Most notably, standard biomolecular force fields utilize fixed atom-centered partial charges[81], thereby completely neglecting explicit electronic polarization. Consequently, they cannot dynamically capture the redistribution of electron density in response to changing local environments, such as the binding of highly charged ions or the transition from a highly dielectric solvent to a buried hydrophobic core. Furthermore, because these force fields are parameterized empirically, historically biased toward stable, folded conformations, they can occasionally struggle to balance secondary structure propensities, which is particularly problematic when simulating highly flexible or IDPs[108].

Finally, the immense computational cost of evaluating millions of all-atom interactions inherently restricts accessible simulation timescales, risking incomplete phase-space sampling for large-scale transitions. Nevertheless, when applied with a clear understanding of these approximations, all-atom simulations remain the gold standard for high-resolution biomolecular modeling and serve as the definitive reference framework for parameterizing lower-resolution coarse-grained methodologies.

In practice, the parameterization of all-atom force fields requires the rigorous empirical or quantum-mechanical assignment of force constants and equilibrium values for the mathematical terms defined previously[109]. The harmonic approximation for bonds and angles provides a robust description of the internal architectural rigidity of the molecule near equilibrium. Dihedral terms, describing rotational motion, are meticulously parameterized using periodic functions to reproduce specific rotational energy barriers that dictate the structural folding of protein backbones (Ramachandran space preferences[110]) and side chains (rotamer libraries). Simultaneously, non-bonded parameters (Lennard-Jones ε and R_{min} values, alongside partial atomic charges derived from electrostatic potential fitting) are finely tuned to govern molecular recognition, complex solvation behavior, and the stabilization of higher-order biomolecular assemblies.

2.2.4.1 AMBER Protein Force Fields

The AMBER (Assisted Model Building with Energy Refinement) family of force fields[111] is one of the most widely utilized and extensively validated frameworks for atomistic simulations of proteins, peptides, nucleic acids, and their complexes. In the context of protein simulations, successive AMBER parameter sets have been developed iteratively to perfect the delicate energetic balance between local backbone preferences, side-chain rotamer distributions, hydrogen-bonding propensities, and interactions with the solvent environment. Consequently, the choice of a specific AMBER protein force field substantially dictates the stability of secondary-structure elements, loop flexibility, folding behavior, and the accurate representation of conformational ensembles in intrinsically disordered systems.

In practical applications, modern force fields cannot be used in isolation; their performance heavily depends on strict compatibility between the protein parameters, the ion models, and the surrounding water model[112]. Recent literature underscores that mixing incompatible parameter sets can severely distort thermodynamic behavior. This synergy is profoundly important in studies of IDPs and flexible peptides, where the subtle thermodynamic competition between protein–protein and protein–water interactions governs the sampled ensemble. While specialized force fields have been developed explicitly for IDPs (e.g., ff14IDPSFF[113] or various force-balanced variants like AMBER-FB15), general-purpose force fields remain highly desirable for systems that undergo dynamic disorder-to-order transitions upon binding. In this thesis, two specific AMBER-based protein descriptions were employed to address distinct biophysical phenomena: ff14SBonlysc and ff19SB.

2.2.4.1.1 AMBER ff14SBonlysc. The ff14SBonlysc[114] variant is conceptually linked to the broader ff14SB family but was specifically developed to improve the description of protein side-chain behavior while maintaining modular compatibility with older, established AMBER backbone parameters. Its main practical advantage lies in a highly reliable treatment of side-chain torsional preferences and rotamer populations. This refinement is critical when

the research question depends entirely on local geometric packing, the precise orientation of specific residues, or side-chain-mediated contacts.

Such atomic-level features are especially relevant in mechanically active systems, where the precise geometry of disulfide bonds dictates mechanical load transfer and where transient, non-native side-chain contacts severely influence intermediate unfolding pathways. For this reason, ff14SBonlysc (coupled with the standard TIP3P water model for which it was originally calibrated) provided the optimal atomistic starting point for the Steered Molecular Dynamics simulations focused on the mechanochemical yielding of RNase A, allowing for highly accurate modeling of local structural consequences following the rupture of covalent disulfide constraints.

2.2.4.1.2 AMBER ff19SB. The ff19SB iteration constitutes a substantial theoretical and methodological advance over earlier AMBER parameterizations by rigorously addressing sequence-dependent backbone conformational preferences. Older parameterizations often relied on generic torsional corrections applied uniformly across different residue types, which historically led to structural biases, including the severe overstabilization of α-helices[115, 114, 116]. In contrast, ff19SB employs amino-acid-specific two-dimensional CMAP (cross-term map) corrections for backbone energetics[115]. Crucially, these maps were parameterized against high-level quantum mechanical energy surfaces calculated in an aqueous solution phase (mimicked via implicit solvation), rather than in a gas-phase vacuum.

Because these QM targets were generated in a polarized aqueous environment, ff19SB was explicitly optimized for, and must be strictly coupled with[115], the highly accurate four-point OPC (Optimal Point Charge) water model[117]. As a direct consequence of this combined parameterization, ff19SB vastly improves the energetic balance between α-helical, β-strand, and loop conformations. It faithfully differentiates the intrinsic structural propensities of individual amino acids, rendering it exceptionally well-suited for resolving dynamic disorder-to-order transitions. In the context of this thesis, the ff19SB/OPC combination was utilized for mapping the rugged free-energy landscapes of highly dynamic systems, including the

intrinsically disordered ensembles of Aβ42 peptides and the flexible loops governing peptide–protein recognition in the HVEM-LIGHT complex.

2.2.5 Coarse-Grained Force Fields

Coarse-Grained (CG) force fields[105, 106, 118] provide a reduced-resolution representation of biomolecular systems in which multiple atoms are grouped into effective interaction sites, commonly referred to as beads. This reduction in resolution decreases the number of particles and interactions that must be evaluated during the simulation, substantially improving computational efficiency. In addition, the reduction of fine atomistic detail smooths the effective energy landscape and often permits the use of longer integration timesteps, thereby accelerating the sampling of large-scale conformational transitions.

As a result, coarse-grained models are particularly useful for studying phenomena that are difficult to access with all-atom simulations alone, such as large-amplitude structural rearrangements, protein aggregation, long-timescale equilibration, and global features of conformational energy landscapes. However, this gain in efficiency is accompanied by a loss of atomistic detail, so coarse-grained methods are best viewed as complementary to all-atom approaches rather than as direct replacements. Together, all-atom and coarse-grained models form a multiscale framework that enables the investigation of biomolecular systems across different spatial and temporal resolutions.

2.2.5.1 UNRES

The United Residue (UNRES) model[119–142, 104, 143] is a sophisticated, physics-based coarse-grained force field explicitly designed to simulate protein folding thermodynamics and global conformational dynamics. Developed primarily by the groups of Harold A. Scheraga and Adam Liwo, UNRES reduces an entire amino acid to just two primary interaction centers: a peptide group center located midway between consecutive α-carbons on the backbone, and an ellipsoidal bead representing the specific side chain. Unlike empirical CG models (such as standard Martini) that rely on generic, parameterized Lennard–Jones interactions

and harmonic networks to enforce folding, the UNRES energy terms are rigorously derived from the statistical-mechanical potentials of mean force (PMF) of all-atom systems in water. By expanding the PMF via a cluster-cumulant series, UNRES mathematically integrates out the fast internal degrees of freedom. This rigorous physical derivation allows the force field to accurately model backbone hydrogen bonding and side-chain packing purely through geometry-dependent and multibody potentials, without relying on secondary structure biases or sequence-specific elastic networks.

Because of its fundamental physical grounding and a 1000- to 4000-fold computational speed-up relative to all-atom molecular dynamics, UNRES is capable of true *ab initio* protein folding. This capability has been repeatedly validated in global blind-prediction experiments. The force field has demonstrated consistent success in the Critical Assessment of protein Structure Prediction (CASP) exercises, where it has been utilized to predict novel folds and refine multidomain structures directly from sequence data without relying on homologous structural templates. Furthermore, the model has been successfully extended to simulate biomolecular interactions, correctly predicting protein–protein docking geometries in the Critical Assessment of PRedicted Interactions (CAPRI) challenges. Beyond static structure prediction, the continuous, physics-based energy landscape of UNRES has been widely applied, often coupled with multiplexed replica exchange molecular dynamics (MREMD), to elucidate explicit folding pathways, identify transient kinetic intermediates, and model complex binding trajectories.

Crucially for this work, the underlying mathematical architecture of UNRES smoothly supports the integration of dynamic, flat-bottom, and time-averaged distance restraints. Combined with its ability to generate massive, unconstrained conformational trajectories, UNRES is uniquely suited for advanced energy landscape mapping. When integrated with algorithms such as those in the Cambridge energy landscapes software (UNOPTIM)[144], UNRES facilitates the construction of disconnectivity graphs Molecular Dynamics to Disconnectivity Graphs (MDDG)[7] for highly dynamic systems, making it an ideal computational framework for investigating how complex covalent reshuffling dictates macroscopic protein thermodynamics.

2.2.5.2 SIRAH

The South American Initiative for a Rapid and Accurate Hamiltonian (SIRAH)[145, 146] provides a fundamentally different approach to coarse-graining. While many CG models operate strictly in an implicit environment, SIRAH is a "top-down" force field specifically engineered to function seamlessly within classical explicit solvent. SIRAH maps biomolecules onto a reduced set of beads while deliberately retaining the specific topological directionality of side chains and backbone hydrogen bond donors and acceptors. By preserving this electrostatic and structural anisotropy, SIRAH achieves a remarkable biophysical balance: it accurately reproduces the structural dynamics, specific solvation shells, and salt-bridge networks characteristic of all-atom simulations, but operates with a significantly reduced particle count.

A major practical advantage of the SIRAH model is its direct architectural compatibility with standard classical molecular dynamics engines. Because SIRAH relies on the same classical pairwise Hamiltonian functions used in all-atom simulations, it operates natively within major software suites such as AMBER, GROMACS, and NAMD. Furthermore, its parameterization and mapping protocols are explicitly designed to interface with the topological frameworks of both the AMBER and CHARMM force fields. This high degree of interoperability enables researchers to easily generate coarse-grained topologies from existing all-atom structures and facilitates seamless backmapping procedures for multiscale refinement.

In the context of intrinsically disordered proteins and amyloidogenic peptides, such as those studied in this thesis, SIRAH's unique explicit-solvent framework provides a distinct advantage. It enables the massive conformational sampling required to observe spontaneous aggregation and fibril formation pathways, while retaining the critical explicit water interactions that drive hydrophobic collapse and accurately mediate protein–protein recognition.

2.2.5.3 Martini

The Martini force field[147, 86] is a highly popular, parameterized coarse-grained model based on a systematic mapping strategy, which historically groups approximately four heavy

atoms into a single interaction bead. Unlike physics-based models that derive potentials from all-atom PMFs (such as UNRES), Martini is a "top-down" model parameterized primarily to reproduce macroscopic thermodynamic properties, particularly the partitioning free energies of molecules between polar and non-polar phases.

The recently released Martini 3[148] iteration provides a comprehensive re-parameterization of the model, introducing an improved interaction balance, new bead types (including smaller mapping ratios for ring structures), and the expanded ability to represent specific interactions such as hydrogen bonding and electronic polarizability. While Martini is exceptionally well-suited for modeling lipid bilayers and membrane–protein systems, it possesses a fundamental limitation regarding protein dynamics: its generic, spherically symmetric bead potentials cannot natively reproduce backbone hydrogen-bonding networks. Consequently, the secondary structure of a protein is strictly imposed by the user during the initial topology generation and is maintained using stiff bonded parameters or external elastic networks (e.g., Elnedyn or Go-Martini). Because the secondary structure is predefined rather than emergent, standard Martini models cannot be used to study true protein folding events or large-scale conformational changes in structured proteins where elements like α-helices or β-sheets transition or melt.

However, in the context of IDPs, where secondary structure does not need to be artificially maintained, this limitation is circumvented. Even so, the application of standard Martini models historically resulted in overly compact conformations that failed to match experimental radii of gyration due to unbalanced protein–protein interactions. Prior to the development of specialized IDP parameter sets, researchers successfully mitigated this over-compaction in Martini 3 by explicitly scaling the Lennard–Jones interaction strength between protein backbone virtual sites and water beads (commonly referred to as adding a water bias). This interaction scaling proved to be a highly effective, computationally simple method for recovering expanded IDP ensembles. More recently, this empirical scaling has been formalized into explicitly optimized extensions such as Martini3-IDP[149] (published in 2025), which refines bonded parameters and interaction strengths against reference atomistic simulations without requiring ad hoc water scaling.

Consequently, whether through explicit water scaling or utilizing the Martini3-IDP framework, Martini serves as an invaluable, computationally inexpensive tool for probing large-scale structural assemblies of unstructured peptides. By accurately capturing the basic physical chemistry of hydrophobic collapse, solvent partitioning, and liquid–liquid phase separation, it effectively complements all-atom frameworks, allowing researchers to explore the extensive conformational dynamics and early stages of aggregation in amyloidogenic peptides where static secondary structure bias is not required.

2.3 Molecular Dynamics

Molecular dynamics simulations extend molecular mechanics by introducing an explicit time dimension into the description of biomolecular systems. Rather than restricting analysis to static structures or local energy minima, MD generates time-dependent trajectories that capture atomic fluctuations, conformational transitions, solvent-mediated interactions, and the dynamical reorganization of intermolecular contacts. This dynamical perspective is essential for understanding processes such as protein folding, ligand recognition, conformational rearrangements, and allosteric regulation. By resolving atomic motion over time, MD establishes a direct connection between molecular structure, thermodynamic stability, and molecular function, thereby providing mechanistic insight that is often difficult to obtain from experiment alone. Owing to continuing advances in computational hardware, force fields, and simulation methodologies, MD has become a central tool in computational biophysics and structural biology.

For the systems investigated in this thesis, this time-resolved description is particularly important. Intrinsically disordered peptides such as Aβ42 do not occupy a single well-defined structure but rather an ensemble of interconverting conformations, whereas mechanically perturbed proteins such as RNase A can access rare pathways that are not evident from static structures alone. Similarly, peptide–protein recognition cannot be understood solely from a single docked pose, because binding stability depends on fluctuations, solvent response, and the accessibility of competing conformational states.

The theoretical foundation of molecular dynamics is classical mechanics, specifically the numerical solution of Newton's equations of motion for a system of interacting particles,

$$m_i \frac{d^2\mathbf{r}_i}{dt^2} = \mathbf{F}_i, \tag{2.15}$$

where m_i denotes the mass of atom i, $\mathbf{r}_i$ its position vector, and $\mathbf{F}_i$ the force acting on that atom. The forces are obtained from the gradient of the potential energy function,

$$\mathbf{F}_i = -\nabla_i U(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N), \tag{2.16}$$

where U is the force-field potential energy introduced in the previous section. Integration of these coupled equations produces trajectories describing the time evolution of atomic coordinates and velocities.

Although the underlying equations of motion are deterministic, the physical interpretation of MD is rooted in statistical mechanics. In practice, thermodynamic observables are obtained as time or ensemble averages over trajectories, invoking the ergodic hypothesis, which assumes that a sufficiently long trajectory will sample phase space according to the corresponding statistical ensemble. For biomolecules, this connection between microscopic dynamics and macroscopic observables is fundamental, because experimentally measurable properties, such as average structures, fluctuations, free energies, and response to perturbation, all emerge from the underlying configurational ensemble.

In practical simulations, the continuous equations of motion must be discretized for numerical integration. This is achieved by dividing time into small intervals of duration Δt and approximating positions and velocities using Taylor expansions around a reference time point,

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \frac{1}{2}\mathbf{a}(t)\Delta t^2 + \mathscr{O}(\Delta t^3), \tag{2.17}$$

where $\mathbf{v}(t)$ and $\mathbf{a}(t)$ denote velocity and acceleration, respectively. Truncation of this expansion yields practical propagation schemes for advancing the system in time. The numerical accuracy of an MD trajectory depends on both the integration algorithm and the selected

time step, so every simulation involves a careful balance between computational efficiency, numerical stability, and physical fidelity[150].

The requirement for stable and efficient time propagation has led to the development of specialized integration algorithms. Among these, the Verlet family of methods is particularly important because of its simplicity, numerical stability, and favorable conservation properties. Verlet-type algorithms[151] are time-reversible and symplectic, meaning that they approximately preserve the Hamiltonian structure of the equations of motion and therefore exhibit good long-term energy conservation[152].

The standard Verlet algorithm updates particle positions according to

$$\mathbf{r}(t+\Delta t) = 2\mathbf{r}(t) - \mathbf{r}(t-\Delta t) + \mathbf{a}(t)\Delta t^2. \tag{2.18}$$

A widely used alternative is the velocity Verlet algorithm[153], which explicitly propagates both positions and velocities,

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \frac{1}{2}\mathbf{a}(t)\Delta t^2, \tag{2.19}$$

$$\mathbf{v}(t+\Delta t) = \mathbf{v}(t) + \frac{1}{2}\left[\mathbf{a}(t) + \mathbf{a}(t+\Delta t)\right]\Delta t. \tag{2.20}$$

The explicit treatment of velocities enables direct calculation of kinetic energy and instantaneous temperature, facilitating coupling to thermodynamic control algorithms.

An important numerical parameter in molecular dynamics simulations is the integration time step, which strictly determines the temporal resolution of the generated trajectory. To ensure that the numerical integration of the equations of motion remains mathematically stable, the time step must be significantly smaller than the period of the fastest motions present in the system. In all-atom biomolecular simulations, this intrinsic limit is governed by the high-frequency vibrations of bonds involving light hydrogen atoms. Consequently, the all-atom time step is typically restricted to approximately 1–2 fs[154]. To slightly alleviate this limitation, constraint algorithms such as SHAKE or LINCS are routinely employed to fix selected covalent bond lengths involving hydrogens. By effectively freezing these

highest-frequency internal motions, such constraints permit the use of the 2 fs upper limit while preserving total energy conservation and overall dynamic stability.

In contrast, coarse-grained methodologies bypass this fundamental limit by physically integrating out the fastest internal atomic degrees of freedom. Because coarse-grained beads represent the center of mass of multiple heavy atoms, their effective masses are significantly larger, and the empirical potentials connecting them are correspondingly softer. This fundamental shift in system mechanics permits an increase in the integration time step without inducing numerical instability. For instance, the physics-based UNRES model mathematically averages out fast vibrations[155], safely allowing stable integration time steps of nearly 5 fs. Highly empirical coarse-grained models such as Martini utilize even softer potentials, routinely enabling integration time steps of 20–30 fs.

However, the true power of coarse-graining extends beyond merely increasing the size of the numerical integration step. By eliminating the fine-grained degrees of freedom and internal atomic friction, the underlying free-energy landscape is fundamentally smoothed. Consequently, conformational transitions occur much faster in the simulation than they do in physical reality. This phenomenon is known as dynamical or temporal scaling. For example, rigorous benchmarking by the developers of UNRES demonstrated that 1 unit of UNRES simulated time corresponds to approximately 1000 units of real laboratory time. When this massive dynamical scaling is combined with the reduced computational cost of evaluating fewer interaction centers, UNRES offers roughly a 1000- to 4000-fold total speed-up relative to explicit-solvent all-atom molecular dynamics. Ultimately, it is this combination of larger time steps, reduced computational overhead, and accelerated effective dynamics that allows coarse-grained models to access the extensive microsecond-to-millisecond timescales required to observe global conformational transitions.

In the microcanonical ensemble (NVE), Newtonian dynamics conserves the total energy of the system. However, biomolecular processes occur under constant temperature and pressure. Sampling the canonical (NVT) or isothermal–isobaric (NPT) ensembles is achieved through thermostats and barostats, which regulate the exchange of heat and volume with an external

reservoir. The instantaneous temperature is related to the kinetic energy through

$$T = \frac{2\langle E_{\text{kin}} \rangle}{f k_{\text{B}}}, \tag{2.21}$$

where f is the number of active degrees of freedom and k_{B} is the Boltzmann constant.

While weak-coupling algorithms like the Berendsen thermostat[156] are efficient for initial equilibration, since they can create temperature artefacts[157], rigorous canonical sampling requires more advanced formulations, such as Langevin dynamics. In Langevin dynamics[158], the system interacts with an implicit thermal environment through dissipative and stochastic forces:

$$m_i \frac{d^2 \mathbf{r}_i}{dt^2} = \mathbf{F}_i - \gamma_i m_i \frac{d\mathbf{r}_i}{dt} + \mathbf{R}_i(t), \tag{2.22}$$

where γ_i is a friction coefficient and $\mathbf{R}_i(t)$ is a random force satisfying the fluctuation–dissipation theorem[159]. This approach ensures a correct Boltzmann distribution of conformations and is particularly vital when using implicit solvent models (like MM-GBSA) to mimic the hydrodynamic drag of missing explicit water molecules.

For NPT simulations, barostats allow the simulation volume to fluctuate dynamically in response to pressure differences. Methods such as the Berendsen[156] and Parrinello–Rahman barostats[160] adjust the dimensions of the simulation box to maintain a target pressure. Pressure is related to the virial expression

$$P = \frac{N k_{\text{B}} T}{V} + \frac{1}{3V} \left\langle \sum_i \mathbf{r}_i \cdot \mathbf{F}_i \right\rangle, \tag{2.23}$$

where V denotes the system volume. Accurate pressure regulation is essential for reproducing experimentally relevant densities and structural organization in explicitly solvated proteins.

Despite its detailed atomistic resolution, conventional molecular dynamics is often limited by the disparity between computationally accessible timescales and macroscopic biological reality. However, the severity of this limitation depends strictly on the topology of the underlying free-energy landscape.

For structured, folded proteins, the landscape is notoriously rugged, characterized by deep metastable basins separated by high kinetic activation barriers ($\Delta G^{\ddagger} \gg k_B T$). Because the probability of spontaneously crossing these barriers is exponentially suppressed, conventional simulations of these systems frequently suffer from severe kinetic trapping. A trajectory may oscillate indefinitely within a single local minimum, failing to capture the true thermodynamic equilibrium ensemble.

Conversely, for highly flexible or IDPs, the free-energy landscape is considerably shallower and highly degenerate[161]. Because the kinetic barriers separating distinct microstates in IDPs are often comparable to thermal energy ($\Delta G^{\ddagger} \approx k_B T$), conventional microsecond-scale MD is frequently sufficient to effectively sample their rapidly fluctuating, unstructured ensembles without artificial biasing.

Consequently, while classical MD is highly effective for characterizing IDP ensembles or local spontaneous fluctuations, it remains fundamentally inadequate for processes with long physical waiting times. To investigate substantial thermodynamic hurdles, such as complete protein unfolding, large-scale domain rearrangements, or slow ligand dissociation, advanced enhanced sampling methodologies must be employed to force the system over high energy barriers.

2.3.1 Enhanced Sampling and Free-Energy Methods

To overcome the timescale limitations of conventional MD, enhanced sampling techniques are designed to accelerate the exploration of configurational space. These methods facilitate barrier crossing either by modifying the effective potential energy surface, introducing external biasing potentials, or altering the statistical sampling procedure.

In full configurational space, the equilibrium probability of observing a configuration $\mathbf{x}$ is governed by the Boltzmann distribution, $P(\mathbf{x}) \propto e^{-\beta U(\mathbf{x})}$, where $\beta = 1/(k_{\mathrm{B}} T)$. When the system is projected onto a reduced, low-dimensional coordinate, often termed a Collective Variable (CV) or reaction coordinate, denoted as s, the corresponding free-energy profile (or

PMF) is given by:

$$F(s) = -k_{\mathrm{B}} T \ln P(s) + C, \tag{2.24}$$

where C is an arbitrary constant. The core objective of most enhanced sampling methods is to artificially flatten $F(s)$ so the system can freely diffuse across barriers, followed by rigorous statistical reweighting to recover the true unbiased thermodynamic landscape.

Representative enhanced sampling approaches include Replica Exchange Molecular Dynamics (REMD)[162, 163], Accelerated Molecular Dynamics (aMD)[164], and Metadynamics [165]. REMD overcomes kinetic trapping by simulating multiple, independent replicas of the system simultaneously across a ladder of increasing temperatures. At predefined intervals, exchanges of configurations between adjacent temperature replicas are attempted. To ensure rigorous thermodynamic sampling and preserve detailed balance, the probability of a successful exchange is governed by the Metropolis criterion:

$$P(i \leftrightarrow j) = \min\left(1, \exp\left[(\beta_i - \beta_j)(U_i - U_j)\right]\right) \tag{2.25}$$

where $\beta = 1/k_B T$ and U represents the instantaneous potential energy of the respective replica. Through successful exchanges, configurations trapped in deep local minima at low temperatures transition to higher temperatures, acquire sufficient thermal energy to cross macroscopic activation barriers, and eventually cool back down to explore newly discovered conformational basins. To further maximize sampling efficiency, extensions such as MREMD[166, 167] deploy multiple independent walkers at each temperature node. By significantly increasing the frequency of successful exchanges, MREMD prevents temperature bottlenecks and is heavily utilized in conjunction with the UNRES coarse-grained force field to exhaustively sample complex folding pathways and compute highly converged free-energy landscapes.

Conversely, Metadynamics operates by directly altering the topological features of the free-energy landscape rather than manipulating the system's thermal energy. It progressively fills previously visited free-energy minima by depositing history-dependent Gaussian bias potentials along a small set of predefined collective variables. This accumulating artificial

penalty physically forces the simulation out of deep thermodynamic wells and drives it to continuously explore new, unvisited regions of phase space, ultimately allowing the underlying unbiased free-energy surface to be mathematically reconstructed.

When evaluating the absolute binding affinity of a ligand or the effect of a specific amino acid mutation, "alchemical" free-energy methods such as Free-Energy Perturbation (FEP)[168] or Thermodynamic Integration (TI)[169] are frequently employed. Rather than sampling geometric transitions, these methods calculate the free-energy difference by non-physically mutating one chemical state into another (e.g., transforming a wild-type residue into a mutant) through a coupling parameter λ. Based on the Zwanzig equation[168], the free-energy difference is evaluated over discrete λ-windows, providing highly accurate estimations of relative binding affinities that guide rational peptide drug design.

2.3.2 Steered Molecular Dynamics and Jarzynski Equality

While alchemical methods transform chemical topologies, geometric transitions, such as mechanical protein unfolding, ligand unbinding, or domain separation, require specific directional biasing. Steered Molecular Dynamics[170, 171] is a non-equilibrium technique in which external forces are applied to a molecular system to induce transitions along a predefined spatial reaction coordinate. Typically, a selected atom, center of mass, or molecular domain is coupled to a virtual harmonic spring whose equilibrium position is translated at a constant velocity v, mimicking AFM pulling experiments or optical tweezers.

The external biasing potential is commonly represented as:

$$U_{\mathrm{SMD}}(t) = \frac{1}{2}k\left[x(t) - x_0 - vt\right]^2, \tag{2.26}$$

where k is the spring constant, $x(t)$ is the instantaneous coordinate, and x_0 is the initial reference position. The corresponding external force exerted on the system is:

$$F_{\mathrm{SMD}}(t) = -\frac{\partial U_{\mathrm{SMD}}}{\partial x} = -k\left[x(t) - x_0 - vt\right]. \tag{2.27}$$

By monitoring the force–extension profile, SMD provides direct mechanistic insight into mechanical stability, intermediate states, and the rupture forces of structural elements. In the context of this thesis, SMD is critically employed for evaluating the mechanochemical yielding of RNase A, where the externally applied force dynamically competes with the topological constraints of native disulfide bonds to reveal distinct unfolding pathways.

Because SMD simulations operate inherently out of equilibrium, the system experiences irreversible frictional dissipation. Consequently, the measured non-equilibrium work (W) performed on the system systematically exceeds the true reversible free-energy difference (ΔG, or the Potential of Mean Force, PMF). However, Christopher Jarzynski proved that these non-equilibrium trajectories can be rigorously related to equilibrium thermodynamic quantities through the Jarzynski equality[172]:

$$\exp\left(-\frac{\Delta G}{k_{\mathrm{B}}T}\right) = \left\langle \exp\left(-\frac{W}{k_{\mathrm{B}}T}\right)\right\rangle. \tag{2.28}$$

The angular brackets denote an ensemble average over multiple independent pulling trajectories, where the total work performed along a given trajectory of duration τ is $W = \int_0^\tau \frac{\partial U_{\mathrm{SMD}}(t)}{\partial t}\, dt$.

While this mathematical relationship is exact in theory, achieving statistical convergence of the exponential average in practice is notoriously difficult. The average is heavily dominated by rare trajectories that experience unusually low dissipation. Furthermore, the magnitude of the dissipated work is inextricably linked to the pulling velocity v. High pulling velocities drive the system far from equilibrium, drastically increasing the irreversible work and broadening the work distribution, which renders the free-energy reconstruction statistically unreliable.

Conversely, utilizing slower pulling speeds allows the internal degrees of freedom of the protein to relax, minimizing frictional dissipation and pushing the simulation toward a quasi-reversible, equilibrium-like regime. However, employing the extremely low pulling velocities or the massive ensemble sampling required to adequately capture the low-dissipation tail of the work distribution is computationally prohibitive for large biomolecular systems. Consequently, while SMD is an unparalleled tool for generating non-equilibrium unfolding

trajectories and identifying mechanical intermediates, relying on Jarzynski's equality to compute robust free-energy profiles is often practically inefficient. To obtain highly converged and reliable PMFs at a manageable computational cost, researchers instead rely on equilibrium stratification techniques, most notably Umbrella Sampling.

2.3.3 Umbrella Sampling

To obtain accurate free-energy profiles without the dissipation issues and convergence failures inherent to fast SMD, Umbrella Sampling provides a highly robust, equilibrium-based alternative.[173] Originally introduced by Torrie and Valleau, umbrella sampling employs a static biasing potential to restrain the system at specific, energetically unfavorable points along a predefined reaction coordinate (x). This artificial restraint forces the simulation to thoroughly sample high-energy transition states and barriers that would otherwise be thermodynamically inaccessible during conventional MD due to low spontaneous crossing probabilities.

Most commonly, this bias is implemented as a simple harmonic restraint:

$$U_{\text{bias}}(x) = \frac{1}{2}k(x - x_0)^2, \tag{2.29}$$

where k is the force constant and x_0 defines the target center of a specific sampling window. The total potential energy governing the system dynamics is therefore modified to:

$$U_{\text{tot}}(x) = U(x) + U_{\text{bias}}(x), \tag{2.30}$$

where $U(x)$ denotes the true, unbiased potential energy of the system. To map an extensive conformational transition, the overall reaction coordinate is partitioned into a series of multiple overlapping windows, each simulated independently at a different x_0. Ensuring sufficient geometric overlap between the resulting coordinate distributions of adjacent windows is strictly required to guarantee continuous phase-space coverage.

Because the simulation is artificially restrained, the observed probability distribution within each window, $P_{\text{bias}}(x)$, is inherently biased. However, it is mathematically related to

the true, unbiased thermodynamic probability $P(x)$ through the Boltzmann relationship:

$$P_{\text{bias}}(x) \propto P(x) \exp\left[-\beta U_{\text{bias}}(x)\right], \tag{2.31}$$

where $\beta = 1/k_B T$. Rearranging this expression allows for the exact mathematical unbiasing of the local distribution to extract the local free energy, or PMF, denoted as $F(x)$:

$$F(x) = -\frac{1}{\beta} \ln P_{\text{bias}}(x) - U_{\text{bias}}(x) + C_i, \tag{2.32}$$

where C_i is a window-specific, mathematically unknown constant.

Because each sampling window yields its own independent free-energy segment with an arbitrary vertical offset C_i, statistical reweighting procedures must be employed to carefully align the isolated simulation windows. The universally utilized approach is the Weighted Histogram Analysis Method (WHAM).[174] By solving a set of self-consistent equations, WHAM iteratively combines the histograms generated from all overlapping sampling windows. It optimally weights the contribution of each window to minimize global statistical error, ultimately generating a single, continuous estimate of the unbiased probability distribution along the entire reaction coordinate.

The resulting PMF provides an absolute thermodynamic map of the conformational transition, detailing precise activation barrier heights and the relative free energies of intermediate states. Compared to the massive non-equilibrium dissipation inherent to fast SMD pulling, equilibrium-based umbrella sampling substantially improves the quantitative accuracy and reproducibility of free-energy calculations.

However, the physical validity of the reconstructed PMF is strictly dependent on the equilibration of the simulation within each window.[175] Because the harmonic bias only restrains the system along the predefined one-dimensional collective variable, the simulation must run long enough for all orthogonal, un-biased degrees of freedom, such as side-chain repacking, solvent reorganization, and secondary structural shifts, to thoroughly relax and sample their local equilibrium states. Provided this complete orthogonal sampling is achieved,

umbrella sampling serves as a robust and highly reliable standard for biomolecular free-energy determination.

2.3.4 Trajectory Analysis Metrics

The interpretation of molecular-dynamics trajectories relies on a set of structural observables that quantify conformational stability, flexibility, compactness, and secondary-structure content. Among the most widely used descriptors are the root-mean-square deviation (RMSD)[176], root-mean-square fluctuation (RMSF)[107], and radius of gyration (R_g)[177], which together provide a compact but informative characterization of the conformational ensemble. The RMSD measures the average structural deviation of a conformation from a reference structure after optimal rigid-body superposition. For a structure containing N atoms, it is defined as

$$\mathrm{RMSD}(t) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left|\mathbf{r}_i(t) - \mathbf{r}_i^{\mathrm{ref}}\right|^2}. \tag{2.33}$$

RMSD is commonly used to assess overall structural stability during a simulation, monitor equilibration, and compare the relative persistence of different folds or binding poses. Low RMSD values indicate close similarity to the reference state, whereas large values suggest structural rearrangement or transitions between metastable basins.

Whereas RMSD characterizes global deviation, RMSF quantifies the local flexibility of each atom or residue around its time-averaged position. It is defined as

$$\mathrm{RMSF}_i = \sqrt{\left\langle \left|\mathbf{r}_i(t) - \langle \mathbf{r}_i \rangle\right|^2 \right\rangle}, \tag{2.34}$$

where $\langle \mathbf{r}_i \rangle$ is the time-averaged position of atom or residue i. RMSF therefore provides a residue-resolved measure of mobility and is especially useful for identifying flexible loops, terminal regions, or binding-interface segments that undergo significant conformational adaptation. The radius of gyration describes the spatial compactness of the molecular structure and is given by

$$R_g = \sqrt{\frac{\sum_i m_i |\mathbf{r}_i - \mathbf{r}_{\mathrm{cm}}|^2}{\sum_i m_i}}, \tag{2.35}$$

where m_i is the mass of atom i and $\mathbf{r}_{\mathrm{cm}}$ is the center of mass of the molecule. Smaller values of R_g correspond to more compact conformations, whereas larger values indicate more expanded states. This observable is particularly informative in the study of intrinsically disordered proteins, peptide collapse, and aggregation-prone systems.

To characterize secondary structure, the DSSP algorithm[178] is frequently employed. DSSP assigns local structural motifs such as α-helices, β-strands, turns, bends, and coils on the basis of backbone hydrogen-bond patterns and local geometry. In this way, time-dependent DSSP analysis allows one to quantify the persistence, gain, or loss of secondary-structure elements throughout the trajectory and to compare secondary-structure propensities across force fields or sequence variants.

Because molecular-dynamics trajectories often sample heterogeneous ensembles rather than a single dominant structure, clustering methods are commonly used to identify representative conformational states. In hierarchical clustering, pairwise structural dissimilarities, typically measured by RMSD, are used to group similar conformations into clusters. This procedure provides a reduced description of the trajectory in terms of major structural families and their populations, allowing one to identify dominant substates, representative structures, and transitions between metastable conformational basins.

Together, RMSD, RMSF, R_g, DSSP analysis, and clustering form a coherent set of tools for transforming raw MD trajectories into physically interpretable structural observables. These descriptors are particularly valuable in studies of flexible peptides, disordered proteins, and binding interfaces, where function emerges from an ensemble of interconverting conformations rather than from a single static structure.

2.3.5 Calculation of Chemical Shifts

To compare simulated conformational ensembles with NMR data, experimentally observable quantities can be back-calculated from structural snapshots. Among the most commonly used observables are backbone chemical shifts[179], which are highly sensitive to local geometry, hydrogen bonding, secondary structure, and solvent exposure. In particular, $^{13}C_\alpha$ chemical

shifts provide a convenient probe of local backbone conformation and are therefore frequently used to assess the realism of simulated peptide and protein ensembles.

In practice, chemical shifts are typically predicted from atomic coordinates using empirical or machine-learning-based programs trained on high-resolution experimental structures and NMR datasets. Such methods estimate the chemical shift of each nucleus as a function of local structural descriptors, including backbone dihedral angles, neighboring residue types, hydrogen-bond geometry, and aromatic ring-current effects. Ensemble-averaged predicted chemical shifts can then be compared directly with experiment.

Because chemical shifts are averaged over the underlying conformational ensemble, they are particularly valuable for validating simulations of flexible peptides and intrinsically disordered proteins. Good agreement with experimental chemical shifts indicates that the simulated ensemble reproduces realistic local structural preferences, even when no single dominant structure exists. For this reason, chemical-shift back-calculation provides an important link between atomistic simulation and experimentally measurable observables.

2.4 Energy Landscape

The concept of the free-energy landscape is a central paradigm in modern biophysics, providing a rigorous statistical-mechanical framework for understanding protein folding, conformational dynamics, and allosteric regulation. A biomolecule explores a highly complex, multidimensional phase space governed by its potential energy function. When projected onto relevant collective variables, such as the radius of gyration, fraction of native contacts, or end-to-end distance, this phase space is represented as a free-energy landscape.

For highly stable, globular proteins, the landscape is typically described as a multidimensional "folding funnel." The funnel guides the conformational ensemble toward a distinct, deep global free-energy minimum corresponding to the native state. However, the systems investigated in this thesis exhibit fundamentally different landscape topologies that require specialized theoretical and computational treatment.

Within the potential energy landscape framework, the conformational space can be coarse-grained in terms of stationary points of the potential energy surface, where the gradient of the potential energy vanishes. The most important stationary points are local minima, characterized by Hessian matrices containing only non-negative eigenvalues, and transition states connecting these minima, which are first-order saddle points possessing exactly one negative Hessian eigenvalue. In discrete path sampling approaches[180], these stationary points are iteratively generated and assembled into a database describing the underlying kinetic transition network. Programs such as PATHSAMPLE[181] and OPTIM[182] employ geometry optimization techniques to identify minima, transition states, and the pathways connecting them. Combined with unimolecular rate theory[183], these methods provide a detailed description of the kinetic organization of the landscape. However, because the number of minima and transition states is expected to increase exponentially with system size, such approaches become computationally demanding for large biomolecular systems.

Intrinsically disordered proteins and amyloidogenic peptides, such as Aβ42, possess highly rugged, frustrated landscapes lacking a single dominant minimum. Instead, they navigate a broad, relatively flat basin containing multiple interconverting metastable states. Topological constraints, such as disulfide bonds, heavily restrict the accessible volume of the landscape by entirely cutting off certain regions of phase space. A single disulfide cross-link acts as an entropic topological boundary, dictating folding pathways and providing immense mechanical resilience. Consequently, visualizing and quantifying these complex multidimensional spaces, often through techniques such as MDDG or Principal Component Analysis (PCA), is essential for translating raw trajectory data into a clear mechanistic understanding of biomolecular stability and mechanical yielding.

2.4.0.1 Disconnectivity Graphs

Because biomolecular free-energy landscapes are intrinsically high-dimensional, direct visualization of all minima and transition pathways is generally impossible. Disconnectivity graphs provide a compact topological representation of such landscapes by organizing local minima according to the energy barriers that separate them. In this representation, each leaf of the

graph corresponds to a local minimum, while branching points indicate the lowest energy level at which distinct basins become connected through accessible transition pathways.

Conceptually, a disconnectivity graph answers the following question: if all regions of the landscape above a given energy threshold are removed, which minima still belong to the same connected superbasin? By progressively lowering the threshold, one can resolve the hierarchical partitioning of the landscape into subfunnels and metastable basins. The vertical axis of the graph represents energy, whereas the horizontal arrangement is chosen mainly for visual clarity rather than for direct physical interpretation. In this way, disconnectivity graphs reveal whether the system is organized into a single dominant funnel, multiple competing funnels, or a highly frustrated landscape containing many comparably deep minima separated by substantial barriers.

For protein folding and conformational transitions, this representation is especially useful. A simple funnel-shaped landscape appears as a graph in which many branches merge rapidly into a single low-energy basin, consistent with efficient convergence toward one dominant native structure. By contrast, frustrated or heterogeneous systems produce more highly branched graphs with several competing low-energy basins, reflecting conformational diversity and kinetic trapping. Intrinsically disordered proteins, aggregation-prone peptides, and systems with dynamic covalent rearrangements often display precisely this more complex organization.

Traditionally, disconnectivity graphs are constructed from databases of local minima and transition states obtained through dedicated energy-landscape exploration and geometry optimization. However, approximate graph-based descriptions can also be derived directly from molecular-dynamics trajectories by identifying proxy minima and transition-state-like events in suitably processed energy time series. In such cases, the resulting graph no longer represents the complete stationary-point network of the full potential-energy surface, but rather an interpretable summary of the part of the landscape actually sampled during the trajectory.

For systems containing dynamic disulfide bonds, disconnectivity graphs are particularly informative because they allow structural basins to be classified not only by geometry but

also by covalent topology. Distinct disulfide-connectivity patterns may generate separate subfunnels, and the barriers between them reflect the energetic cost of reshuffling between alternative covalent states. Thus, disconnectivity graphs provide a natural link between microscopic trajectory data and the global organization of the conformational landscape.

2.5 Treatment of Disulfide Bonds in Molecular Dynamics Simulations

Standard molecular mechanics force fields are intrinsically limited in their ability to describe chemical reactions involving bond breaking and bond formation. In conventional biomolecular simulations, covalent bonds are represented by fixed harmonic or Morse potentials defined rigidly within the molecular topology. Because these classical potentials cannot represent electron redistribution, conventional MD is fundamentally blind to spontaneous changes in covalent connectivity, meaning processes such as disulfide bond formation, rupture, or reshuffling lie entirely outside the scope of a strictly standard classical description.

This limitation is particularly pronounced for disulfide bonds, which play a critical structural and regulatory role in many proteins, including RNase A. In biological reality, the reshuffling of these bonds involve complex, coupled proton- and electron-transfer processes. The thermodynamic stability and chemical reactivity of a given disulfide bond depend strongly on its local environment, including specific geometric factors such as the $C\beta - S\gamma - S\gamma - C\beta$ dihedral angle and the sulfur–sulfur bond length, as well as backbone strain. Furthermore, solvent accessibility and the local electrostatic context heavily dictate the protonation states of the resulting thiolate species depending on the environmental pH.

Under externally applied force, the physical elongation of the sulfur–sulfur linkage mechanically strains the bond, which can significantly alter the activation barrier for rupture. Explicitly modeling this mechanochemical rupture requires quantum-chemical or reactive approaches (such as QM/MM or ReaxFF) to properly redistribute electronic density. However, these methods are computationally prohibitive, effectively restricting simulations to localized

reaction mechanisms on the picosecond timescale. This entirely precludes the microsecond-to-millisecond sampling necessary to observe macroscopic phenomena, such as full mechanical unfolding or large-scale conformational reshuffling.

Fortunately, for macroscopic biophysical investigations, explicitly resolving the quantum-chemical transition state is neither computationally feasible nor physically necessary. Molecular dynamics inherently relies on the abstraction of complex electronic behaviors into effective empirical potentials. Because the timescale of actual electron transfer (femtoseconds) is fundamentally decoupled from the timescale of global protein domain separation (microseconds to milliseconds), the microscopic chemical event can be safely abstracted into a macroscopic topological transition. Provided the applied pseudopotential correctly captures the macroscopic free-energy boundaries between the oxidized (bonded) and reduced (unbonded) macrostates, the precise quantum trajectory of the electrons can be integrated out.

Implementing this dynamic topological switch at the all-atom level does introduce specific representational challenges. In physical reality, disulfide reduction requires the acquisition of protons to form thiol groups ($-$SH). Dynamically creating or annihilating explicit atomic coordinates during a classical simulation introduces discontinuities in the integration engines. To resolve this, dynamic all-atom models frequently approximate the reduced state by absorbing the newly acquired hydrogen into a unified sulfur center. While this momentarily diverges from strict all-atom explicitness, the use of unified atoms, where non-polar hydrogens are absorbed into heavy atoms, is a rigorously validated approximation in biomolecular simulation (e.g., the GROMOS force field). Because the mass, charge differential, and steric bulk of a single proton are thermodynamically negligible compared to the massive mechanical forces driving global domain separation, this effective representation captures the physical mechanics of the ruptured state without compromising the structural integrity of the simulation.

Conversely, representing dynamic bond reshuffling at the coarse-grained level aligns inherently with the fundamental design of these models. Because CG methodologies already abstract groups of atoms into single, effective interaction centers, the absence of explicit protons is a native feature of the architecture. Transitioning between bonded and unbonded macrostates simply involves modulating the phenomenological interaction potentials between

existing beads. Consequently, modeling dynamic covalent topologies represents a natural extension of the coarse-grained abstraction.

Ultimately, the theoretical treatment of dynamic disulfide bonds in classical molecular dynamics requires a deliberate shift in focus from quantum chemistry to structural biophysics. The primary scientific objective in such methodologies is not to elucidate the elementary reaction kinetics of bond formation and rupture. Rather, the goal is to capture the profound macroscopic influence of these covalent constraints on the accessible conformational ensemble, structural stability, and mechanochemical yielding of proteins. By treating disulfide transitions through theoretically grounded pseudopotential approaches at both the all-atom and coarse-grained levels, classical simulations can rationally bridge the gap between dynamic covalent topologies and the extensive timescale sampling required to map complex biological free-energy landscapes.

2.5.1 Dynamic Disulfide Bond Treatment in UNRES

Dynamic disulfide bonding in the coarse-grained UNRES force field is implemented by replacing the ordinary side-chain interaction between two cysteines with a bimodal potential by Chinchio et al. in 2007[3]. This potential allows the system to switch continuously between a nonbonded state and a disulfide-bonded state during the simulation. Crucially, bond formation and breaking are not imposed externally by the user or defined by fixed distance cutoffs; instead, they occur naturally via thermal crossing of an artificial macroscopic barrier placed between the two minima of the interaction potential.

For a pair of cysteine residues, the disulfide interaction is described by a continuous piecewise function:

$$U_{\mathrm{dynSS}}(r_{ij}) = \begin{cases} U_{\mathrm{SS}}(r_{ij}), & r_{ij} < r_{ij}^{(\mathrm{SS})}, \\ U_{\mathrm{SS}}^{*}(r_{ij}), & r_{ij}^{(\mathrm{SS})} \le r_{ij} < R_{ij}, \\ U_{\mathrm{GB}}^{*}(r_{ij}), & R_{ij} \le r_{ij} < r_{ij}^{(\mathrm{GB})}, \\ w_{\mathrm{SC}}\, U_{\mathrm{GB}}(r_{ij}), & r_{ij} \ge r_{ij}^{(\mathrm{GB})}. \end{cases} \tag{2.36}$$

Here, r_{ij} is the distance between the side-chain centers of the two cysteines, $r_{ij}^{(\mathrm{SS})}$ is the distance at the bonded minimum, $r_{ij}^{(\mathrm{GB})}$ is the distance at the nonbonded minimum, and R_{ij} is the position of the transition barrier, located exactly between both minima.

The bonded-state potential incorporates the highly anisotropic and directional nature of the disulfide interaction:

$$\begin{aligned} U_{\mathrm{SS}}(r_{ij}) = \varepsilon_0^{(\mathrm{SS})} + \sum_{n=1}^{3} V_n \xi_{ij}^{\,n} + k_0 \left[(\eta_{ij}^{(1)})^2 + (\eta_{ij}^{(2)})^2 \right] \\ + k_1 (\eta_{ij}^{(1)} + \eta_{ij}^{(2)})(r_{ij} - d_0) + k_2 (r_{ij} - d_0)^2. \end{aligned} \tag{2.37}$$

with the orientational variables defined as:

$$\eta_{ij}^{(1)} = 1 - \hat{u}_i \cdot \hat{r}_{ij}, \qquad \eta_{ij}^{(2)} = 1 + \hat{u}_j \cdot \hat{r}_{ij}, \qquad \xi_{ij} = \hat{u}_i \cdot \hat{u}_j - (\hat{u}_i \cdot \hat{r}_{ij})(\hat{u}_j \cdot \hat{r}_{ij}). \tag{2.38}$$

This mathematical form stabilizes the geometry of the S–S bond not only through the inter-centroid distance but also heavily through side-chain orientation (η, ξ). This orientational dependence is absolutely critical to accurately represent the strict dihedral angular preferences characteristic of native disulfide bonds, ensuring that non-physiological bond geometries are heavily penalized.

The nonbonded state is described by the standard UNRES side-chain interaction potential:

$$U_{\mathrm{GB}}(r_{ij}) = 4 \left(\left| \varepsilon_{ij}^{(\mathrm{GB})} \right| x_{ij}^{12} - \varepsilon_{ij}^{(\mathrm{GB})} x_{ij}^{6} \right), \qquad x_{ij} = \frac{\sigma_{0,ij}}{r_{ij} - (\sigma_{ij} - \sigma_{0,ij})}. \tag{2.39}$$

The dynamic disulfide model connects this nonbonded state with the bonded state using a smooth barrier, ensuring the potential energy surface remains continuously differentiable. The switching region is governed by a smoothing function:

$$g(x;a,b) = \left(\frac{x-a}{b-a}\right)^2 \left[3 - 2\left(\frac{x-a}{b-a}\right)\right], \tag{2.40}$$

which satisfies:

$$g(a;a,b) = 0, \qquad g(b;a,b) = 1, \qquad \left.\frac{\partial g}{\partial x}\right|_{x=a} = 0, \qquad \left.\frac{\partial g}{\partial x}\right|_{x=b} = 0. \tag{2.41}$$

Because the function and its first derivative strictly vanish at the boundaries, the transition from bonded to nonbonded configurations does not introduce unphysical numerical discontinuities into the MD numerical integration, thus preserving energy conservation.

The thermodynamics and kinetics of the transition are controlled by two adjustable parameters:

$$\varepsilon_0^{(\mathrm{SS})} = w_{\mathrm{SC}}\varepsilon_{0,ij}^{(\mathrm{GB})} - \Delta\varepsilon, \tag{2.42}$$

$$H_t = w_{\mathrm{SC}}\varepsilon_{0,ij}^{(\mathrm{GB})} + \Delta H. \tag{2.43}$$

The parameter $\Delta\varepsilon$ determines the thermodynamic free-energy preference for the bonded state, whereas ΔH controls the kinetic activation barrier. In practice, a larger ΔH slows both bond formation and rupture, allowing the user to implicitly mimic specific redox potentials or environmental oxidation states, while $\Delta\varepsilon$ alters the equilibrium population of bonded versus nonbonded states.

In this dynamic formulation, the ordinary pairwise side-chain term is replaced by this switching potential for every possible cysteine pair ($w_{\mathrm{SC}}U_{\mathrm{SC}_i\mathrm{SC}_j} \longrightarrow U_{\mathrm{dynSS}}(r_{ij})$). Consequently, any cysteine pair can transiently cross the barrier, allowing disulfide pairing patterns to become an emergent property of the trajectory. A practical consequence is that a disulfide bond is operationally identified by a distance criterion on the Cys–Cys side-chain centroids,

while its structural stability is assessed from the fraction of simulation time spent below the transition threshold.

To prevent highly unphysical transition states–such as three cysteines simultaneously occupying bonding proximity during reshuffling events–an additional repulsive penalty term is introduced for cysteine triplets[184]:

$$U_{\mathrm{sss}} = \sum_{i,j,k\in C} d\,\frac{1}{1+a(r_{ij}-r_{\mathrm{ts}})^2+b(r_{ij}-r_{\mathrm{ts}})^6} \times \frac{1}{1+a(r_{ik}-r_{\mathrm{ts}})^2+b(r_{ik}-r_{\mathrm{ts}})^6}\,\frac{1}{1+a(r_{jk}-r_{\mathrm{ts}})^2+b(r_{jk}-r_{\mathrm{ts}})^6}, \tag{2.44}$$

where C is the set of cysteine residues and r_{ts} is a contact-distance scale. This successfully suppresses artificial trisulfide formation, ensuring physically realistic reshuffling pathways.

Ultimately, this dynamic model allows oxidative rearrangement to occur naturally over the course of an MD trajectory. It is particularly advantageous for mechanically unfolding proteins containing multiple cysteines, where several alternative pairing patterns compete dynamically, and the correct structural connectivity must emerge directly from the perturbed free-energy landscape without ad-hoc topological updates.

2.5.2 Finite-Distance Disulfide Restraints in All-Atom Simulations

While coarse-grained approaches can efficiently govern spontaneous bond formation and breaking during global folding or large-scale reshuffling, a different approach is required when studying the mechanically induced rupture of specific native bonds at atomistic resolution[4]. In all-atom SMD simulations, native disulfide bonds must be treated using a targeted finite-distance restraint scheme designed to permit controlled mechanochemical bond dissociation without the catastrophic force singularities inherent to standard harmonic springs.

The restraint energy in such dynamic models is often defined as a piecewise function featuring multiple switching regions, enabling a smooth, physically realistic transition from a bonded harmonic well to a constant-force regime at large distances. The restraint potential

between two interacting sulfur atoms, i and j, can be formulated as:

$$U_{ij} = \begin{cases} k(r_{ij} - r_1)^2, & r_{ij} < r_1, \\ 0, & r_1 \leq r_{ij} < r_2, \\ k(r_{ij} - r_2)^2, & r_2 \leq r_{ij} < r_3, \\ 2k(r_3 - r_2)(r_{ij} - r_3) + k(r_3 - r_2)^2, & r_{ij} \geq r_3, \end{cases} \tag{2.45}$$

where r_{ij} is the absolute distance between the sulfur atoms, k is the force constant, and r_1, r_2, and r_3 define the spatial switching points of the restraint. The first and third regions correspond to harmonic stabilization around the native disulfide geometry, whereas the intermediate regions gradually relax the interaction as the protein domain is mechanically extended.

Crucially, for separations greater than r_3, the potential becomes strictly linear with respect to r_{ij}, which yields a constant restoring force mathematically derived from its spatial derivative:

$$F = 2k(r_3 - r_2). \tag{2.46}$$

This linear regime completely avoids the unphysical, infinite increase in force that results from purely harmonic molecular mechanics restraints at large distances. During non-equilibrium pulling simulations, when the applied mechanical force exceeds this constant restoring force threshold, the simulated bond effectively "yields." This provides a robust, classical approximation of mechanochemical dissociation, allowing the surrounding protein domain to unfold realistically after the topological constraint is broken, thereby exposing mechanochemically relevant intermediate states.

To ensure the physical validity of this pseudo-reactive behavior, the specific parameters of the piecewise function (k and the spatial thresholds) cannot be assigned arbitrarily. In rigorous computational protocols, these values are typically derived by systematically curve-fitting the classical restraint against highly accurate reference potentials, such as the PMF profiles of disulfide dissociation generated from high-level quantum mechanical calculations

or rigorously calibrated coarse-grained models. By utilizing a least-squares fitting approach to match the activation energy barriers (E_a) and maximum rupture forces of the reference data, the resulting classical restraint provides an accurate atomistic representation of covalent yielding, seamlessly complementing long-timescale thermodynamic simulations.

Chapter 3

Results: Summary of Publications

The research presented in this thesis addresses a fundamental challenge in computational biophysics: understanding how dynamic structural constraints, particularly disulfide bonds, dictate the conformational landscapes, mechanical stability, and functional interactions of proteins. While experimental techniques provide crucial macroscopic observables, exploring the transient, high-resolution dynamics of these systems requires advanced computational frameworks. Therefore, the overarching objective of the studies conducted during my doctoral research was to leverage multiscale MD, spanning all-atom to coarse-grained resolutions, coupled with novel analytical tools to dissect the role of covalent constraints across a diverse spectrum of protein systems.

The core results of this research are presented across four complementary, peer-reviewed publications. Together, these works traverse different scales of protein ordering and complexity. The research progresses from elucidating the redox-dependent mechanical unfolding of highly structured enzymes (Ribonuclease A), to applying these insights toward the rational design of disulfide-stabilized peptide inhibitors (HVEM–LIGHT). It then extends to defining robust multiscale simulation protocols for intrinsically disordered proteins (amyloid-β42), and culminates in the development of a novel computational framework and software package (MDDG) to directly visualize the complex energy landscapes governed by dynamic disulfide reshuffling from MD simulations.

The primary aim of this chapter is to provide the scientific rationale, methodology, and major conclusions for each of these studies, contextualizing them within the broader narrative of the thesis.

Each section corresponds to one publication and begins with a short introductory description to situate the findings. Because the main manuscripts focus on the core discoveries,

relevant supplementary information published alongside them has also been integrated to provide a comprehensive view of the research process. Furthermore, to fulfill the formal requirements of this doctoral thesis, each section is accompanied by statements from co-authors detailing their specific contributions to the work.

All four papers included in this thesis are peer-reviewed publications indexed in the Web of Science. They are reproduced here in accordance with the Creative Commons Attribution 4.0 International License (CC BY 4.0).

3.1 Mechanical Stability of Ribonuclease A Heavily Depends on the Redox Environment

Despite extensive characterisation of disulfide bonds, their precise role in modulating protein stability and dynamics remains context-dependent and subject to ongoing debate. In this study, we present a computational investigation into the mechanical stability of Ribonuclease A, a model enzyme containing four distinct disulfide bonds. Our focus was to determine how different oxidative and reductive environments influence the protein's response to externally applied mechanical stress.

To achieve this, we employed a multiscale computational framework to simulate the mechanical stretching of RNase A under a range of redox conditions. At the coarse-grained level, I carried out steered molecular dynamics simulations using the UNRES model. I selected this model because it enables the dynamic formation and rupture of disulfide bonds during the simulation trajectory, a capability essential for the realistic modelling of redox-dependent behaviour. By systematically tuning the depth of the disulfide bond formation potential between 5.5 and 11.0 kcal/mol, I represented environments ranging from weakly to strongly oxidising. In parallel, we performed all-atom simulations using the Amber ff14SBonlysc force field coupled with a Generalized Born implicit solvent model. Because classical all-atom force fields do not natively support covalent bond rupture, I implemented a custom protocol based on finite-distance restraints between sulfur atoms to mimic disulfide bond breakage energy barriers. While the coarse-grained simulations employed a constant velocity pulling mode, these atomistic simulations were conducted under a constant force setup, serving as a high-resolution benchmark to compare with the coarse-grained results.

My analysis of the force–extension profiles revealed that the mechanical stability of RNase A is highly sensitive to the surrounding redox environment. I found that disulfide bonds exert a significant stabilising effect only under strongly oxidising conditions. In these simulations, I observed pronounced force peaks that directly correlated with abrupt losses in disulfide bond populations. In contrast, under reductive or weakly oxidising conditions, cysteine–cysteine

contacts were rapidly lost and contributed negligibly to the overall mechanical resistance of the protein.

Detailed tracking of individual disulfide bonds showed specific roles for each linkage. We identified that the Cys26–Cys84 and Cys58–Cys110 disulfide bonds function as the primary structural bridges resisting mechanical deformation, whereas the Cys65–Cys72 bond exhibited the highest intrinsic stability, securing the protein core. Interestingly, our results suggest that the Cys40–Cys95 bond is more closely associated with functional dynamics and enzymatic activity rather than long-term structural integrity, particularly in cytoplasmic environments. Additionally, my UNRES simulations revealed the formation of non-native disulfide contacts during the pulling process, most notably between Cys84 and Cys95, which provided transient stabilisation during the late stages of unfolding.

A comparative analysis of the coarse-grained and all-atom simulations demonstrated strong qualitative consistency across both levels of resolution, validating my restraint-based approach in Amber. Specifically, both models predicted the identical breaking order of native disulfide bonds in a highly oxidising environment: Cys26–Cys84, followed by Cys40–Cys95, Cys58–Cys110, and finally Cys65–Cys72. Furthermore, while the general sequence of secondary structure unfolding was consistent in both models (initiating rapidly from the terminal regions with the $\beta 6$–$\beta 7$ sheet and N-terminal helices), I observed minor, resolution-dependent deviations. For instance, the subsequent unfolding of the $\beta 1$–$\beta 4$ and $\beta 3$–$\beta 6$ sheets occurred in a slightly different sequence between the Amber and UNRES trajectories. Moreover, the all-atom simulations inherently lacked the ability to sample the non-native Cys84–Cys95 contact observed in UNRES, highlighting the complementary value of our multiscale approach. Despite these minor structural nuances, the excellent agreement between Amber and UNRES results confirms that force peak positions serve as reliable signatures of specific disulfide rupture events.

In conclusion, although disulfide bonds are intrinsically stabilising, we demonstrated that their contribution to the mechanical stability of RNase A is strictly governed by the cellular redox potential. These findings provide important insights into how physiological conditions

regulate the mechanical properties of proteins and contribute to a deeper understanding of protein folding and stability under varying environmental stresses.

Supporting Information for "Mechanical Stability of Ribonuclease A Heavily Depends on the Redox Environment"

Pamela Smardz,[a] Adam K. Sieradzan,[b] Paweł Krupa[a]

[a]Institute of Physics Polish Academy of Sciences, Al. Lotników 32/46, 02-668 Warsaw, Poland
[b]Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, 80-308 Gdańsk, Poland

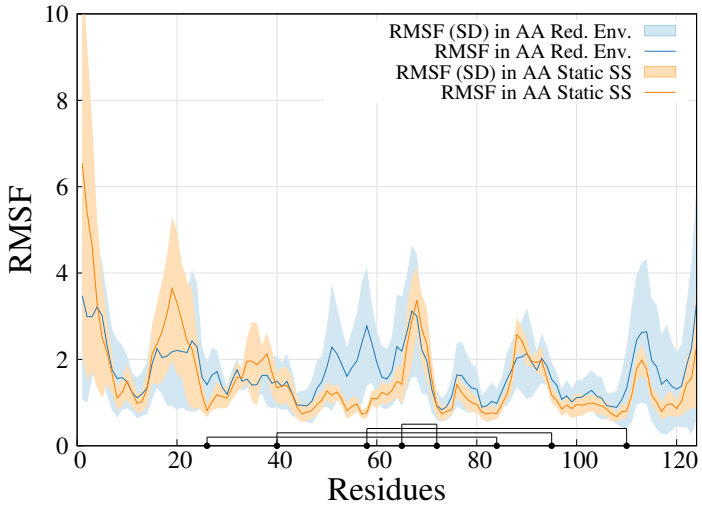


Fig. S1: Root-mean-square fluctuations (RMSF) of each C α atom, for canonical all-atom simulations without (solid blue line) and disulfide bonds (solid orange line). Cysteine residues (black dots) involved in disulfide bridges are marked with black lines.

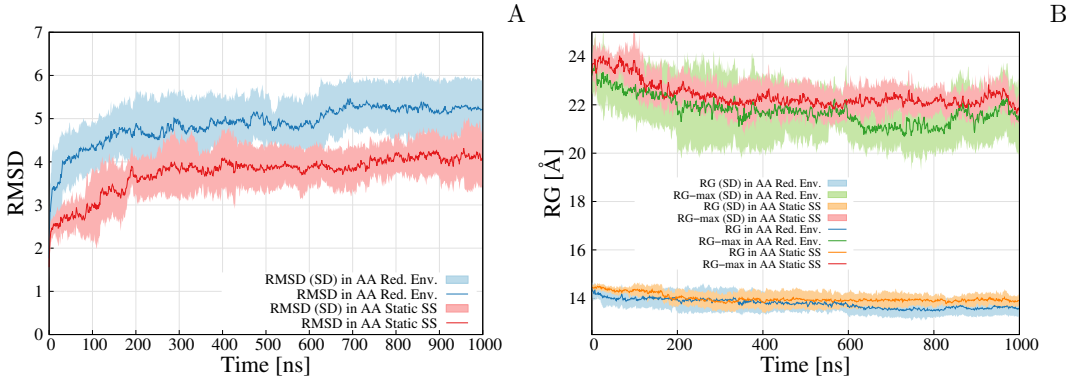


Fig. S2: RMSD (A) and plot of RG and RG max (B) of each C α atom, for canonical all-atom simulations without (solid blue and green line) and disulfide bonds (solid red and orange line).

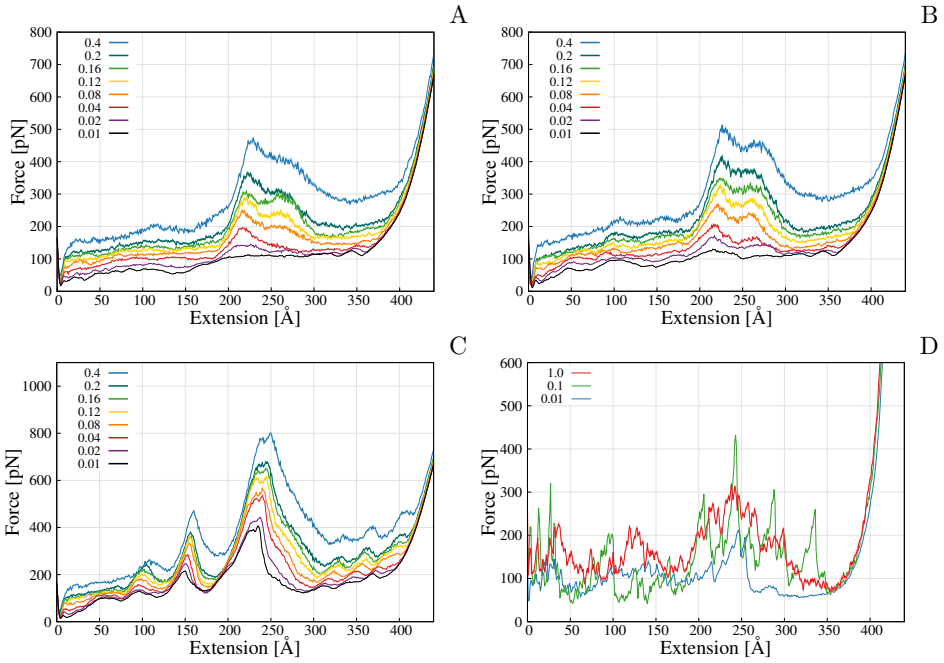


Fig. S3: Comparison of simulations in coarse-grained reductive (A), weak (B) and highly (C) oxidativ enviroment of different velocities from 0.04 to 0.4 m/s at 300K, and in reductive enviroment for different velocities in all-atom simulations (D) from 0.01 to 1 m/s at 300K.

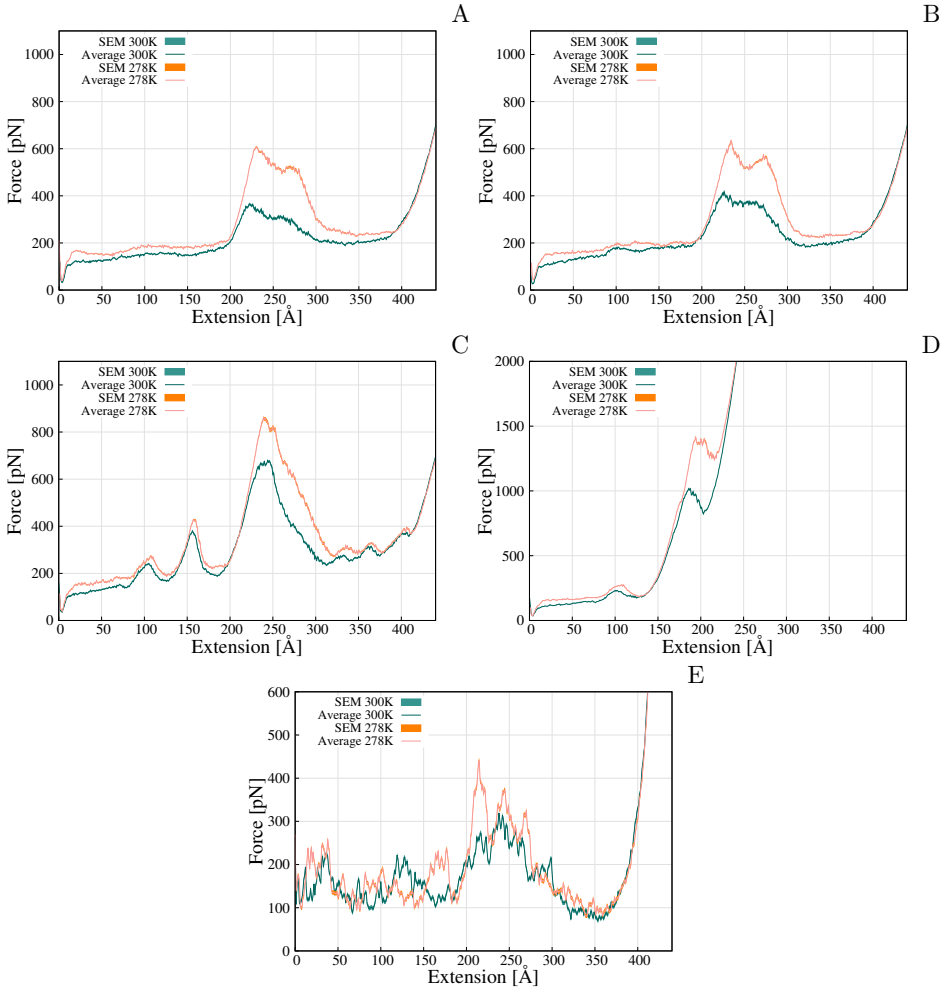


Fig. S4: Comparison of simulations in coarse-grained reductive (A), weak (B), and highly oxidative (C) enviroment and simulations with static disulfide bond (D) for velocity 0.2 m/s, and in all-atom representation for velocity 1 m/s (E) at temperatures 278K and 300K.

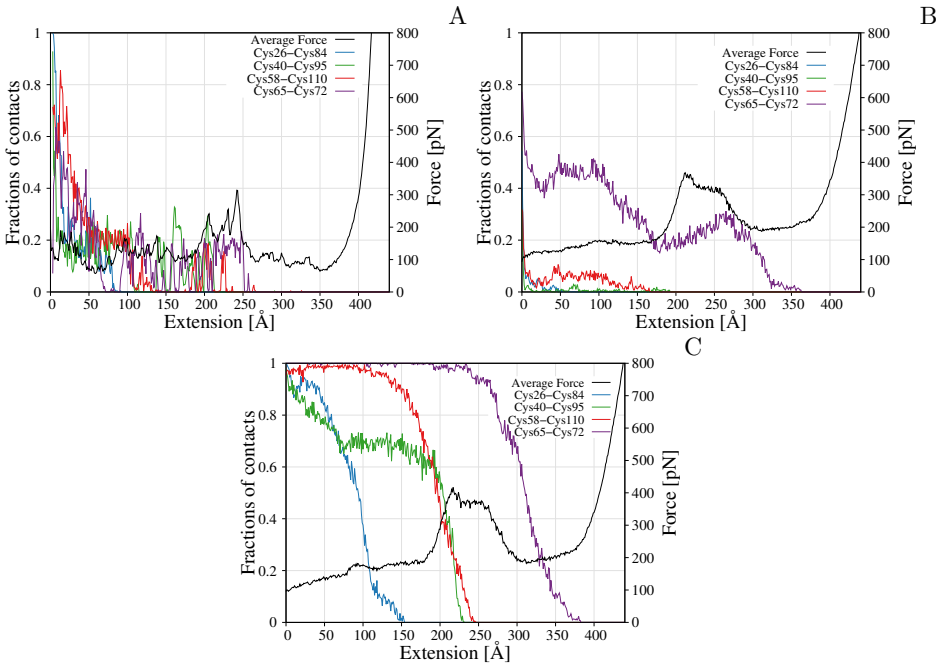


Fig. S5: Fraction of contacts between Cys residues forming native disulfide bonds in all-atom simulation without predefined disulfide bonds at 300K (A), and fraction of native disulfide bonds in coarse-grained simulation in reductive (B) and weak (C) oxidation enviroment at 300K.

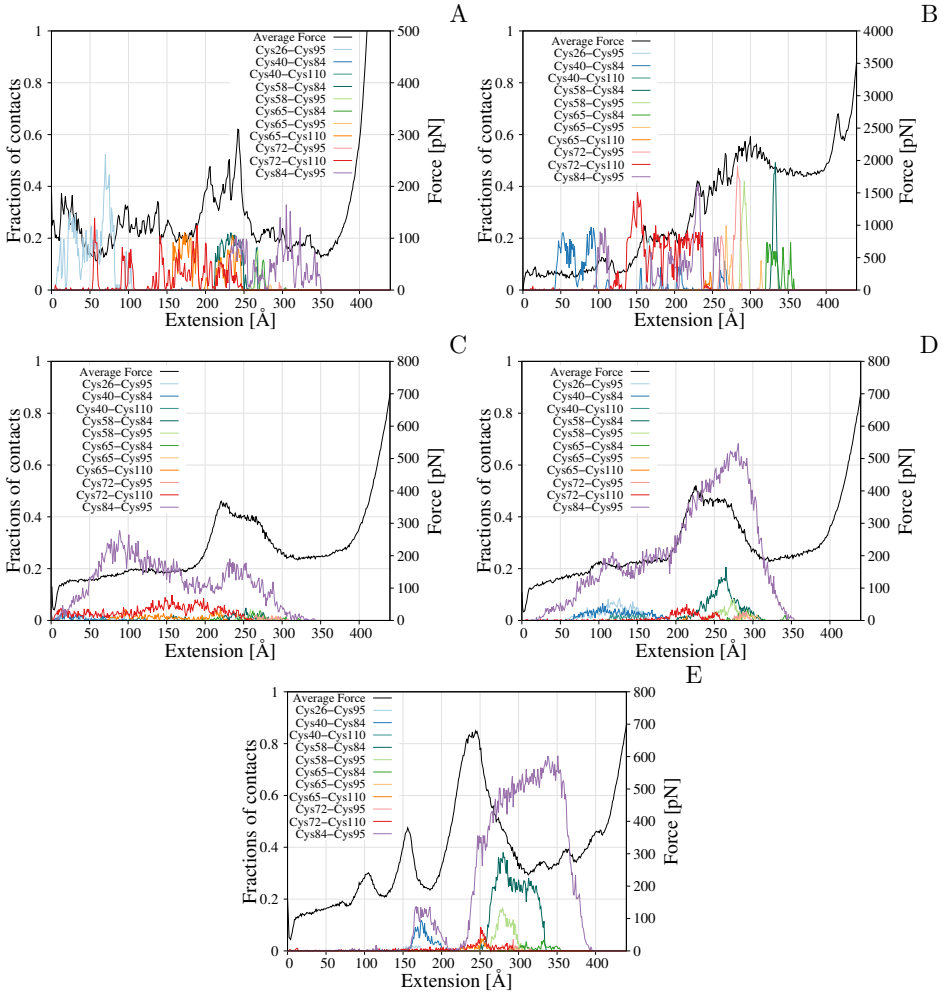


Fig. S6: Fraction of nonnative disulfide bonds in all-atom simulation without (A) and with disulfide bonds restraints (B), and in coarse-grained simulation in reductive (C), weak (D) and highly (E) oxidative enviroment at 300K.

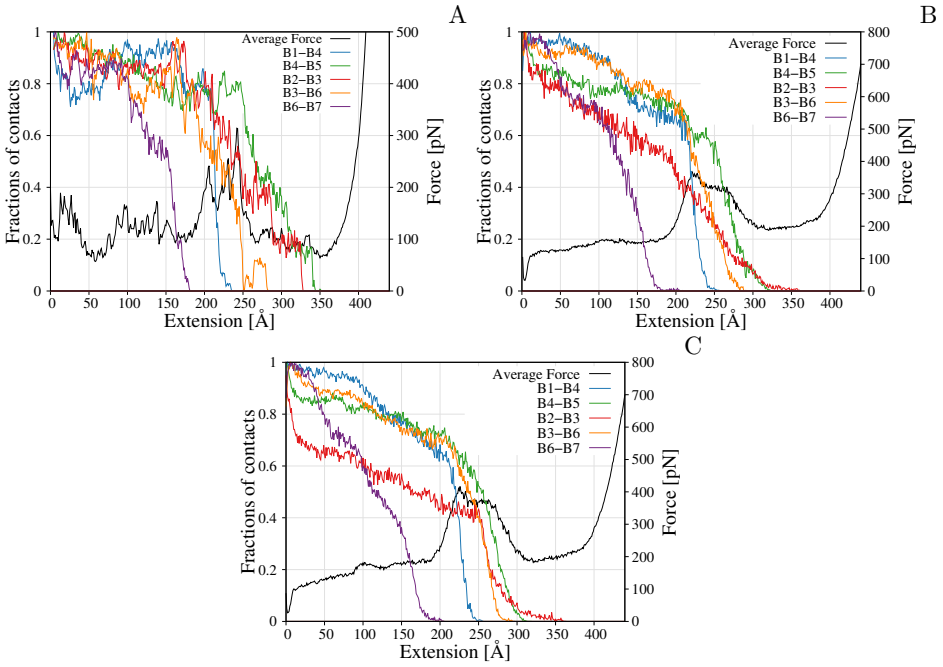


Fig. S7: Fraction of beta sheets in all-atom simulation without disulfide bonds (A) and fraction of beta sheets in coarse-grained simulation in reductive (B) and weak oxidation (C) enviroment at 300K.

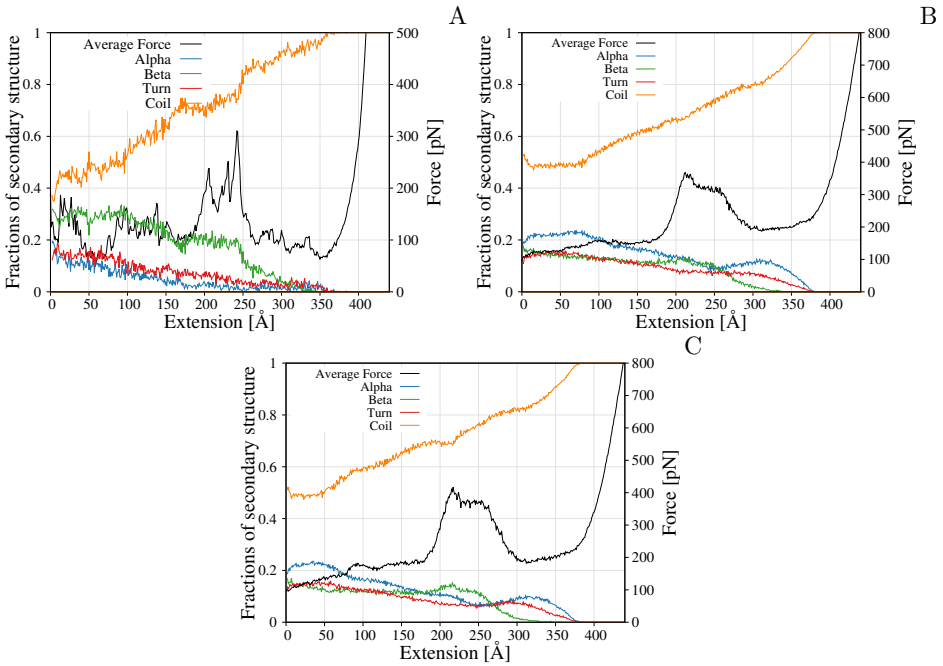


Fig. S8: Fraction of secondary structure in all-atom simulation without disulfide bonds (A), in coarse-grained simulation without disulfide bonds (B), and in weak (C) oxidizing environment at 300K.

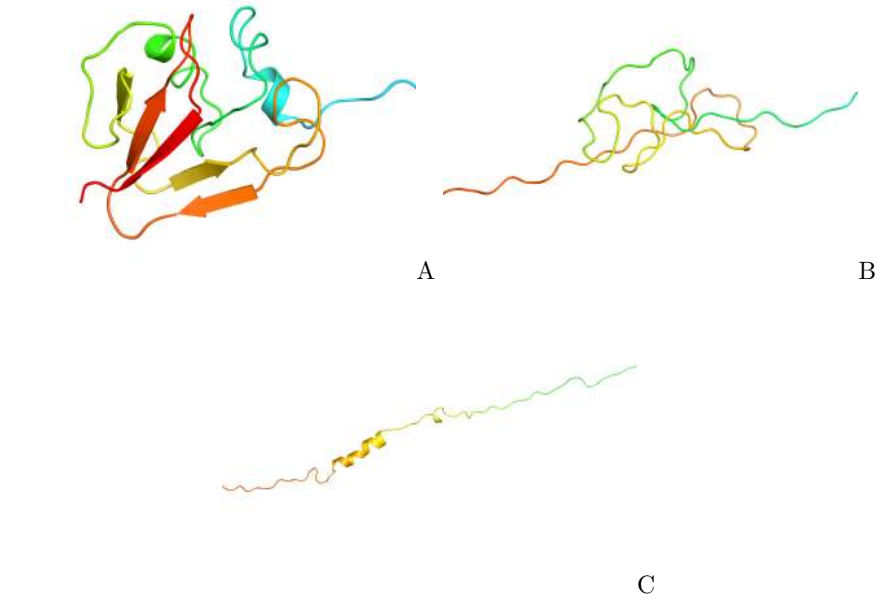


Fig. S9: Cartoon representation of the stable structures (for extensions: 68Å (A), 189Å (B), 272Å (C)) of RNase A otained during stretching in all-atom simulations without disulfide bonds.

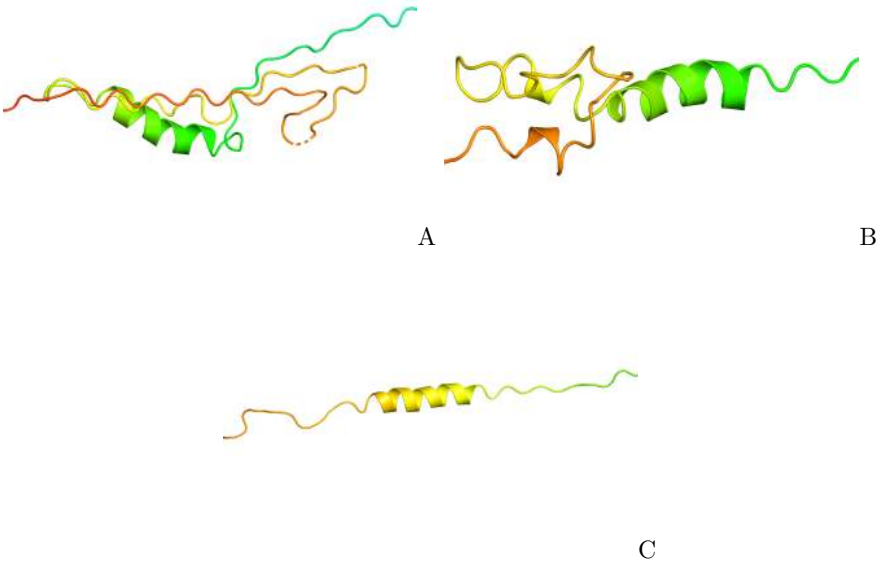
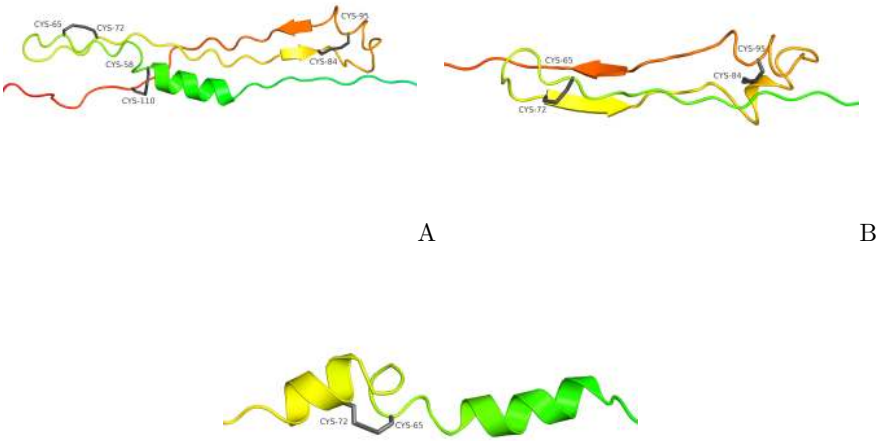


Fig. S10: Cartoon representation of the stable structures (for extensions: 174Å (A), 247Å (B), 317Å (C)) of RNase A otained during stretching in coarse-grained simulations without disulfide bonds.



A B

C

Fig. S11: Cartoon representation of the stable structures (for extensions: 180Å (A), 247Å (B), 327Å (C)) of RNase A otained during stretching in coarse-grained simulations in weak oxidizing environemnt.

3.2 Multilayered Computational Framework for Designing Peptide Inhibitors of HVEM–LIGHT Interaction

We developed a comprehensive multiscale computational framework integrating all-atom MD, SMD, and umbrella sampling to design HVEM-derived peptide inhibitors targeting the HVEM–LIGHT interaction, a central costimulatory pathway in T-cell activation and inflammatory signaling. Systematic disulfide-bond engineering, point mutations, and truncation strategies were applied across 65 HVEM-based variants. This workflow identified HVEM's second cysteine-rich domain (CRD2) as the principal LIGHT-binding module and highlighted two lead candidates: CRD2e_K54E, which contains a single Cys58–Cys73 disulfide bond together with a K54E mutation, and CRD2(39–73)e, a truncated construct that retains the same conserved disulfide linkage. Together, these results showed that a minimal disulfide architecture can support potent and selective inhibition without requiring the full native cysteine network.

We performed microsecond-scale MD simulations on HVEM–LIGHT complexes and on unbound peptide variants using the Amber ff19SB force field with the OPC water model. Each system was simulated for 600 ns across three independent replicas, with selected systems extended to 3 μs. We assessed structural stability and conformational behavior using RMSD, RMSF, (R_g), SASA, and DSSP-based secondary-structure analysis. We estimated binding free energies using MM-GBSA with per-residue decomposition, augmented by conformational work derived from unbound simulations and by normal-mode entropy contributions. Mechanical unbinding was characterized using SMD with 25 independent trajectories at a pulling velocity of 0.05 Å/ps and 10 independent trajectories at 0.01 Å/ps, yielding force and work profiles. Umbrella sampling along the center-of-mass separation coordinate employed at least 44 windows with 100 ns per window to construct PMFs.

Our disulfide-bond analysis across the cysteine-rich domains revealed pronounced differences in stabilizing effects. Within CRD2, the Cys58–Cys73 disulfide bond was the only single linkage that substantially stabilized the isolated domain, reducing RMSD from 10.52

±0.28 Å to 4.93 ±0.65 Å. Although the fully bonded CRD2 reference remained more rigid at 2.60 ±0.18 Å, this result demonstrated that one strategically retained disulfide bond can preserve much of the fold while avoiding unnecessary covalent constraints. These observations, together with interdomain stabilization effects observed in the full receptor context, guided our truncation strategy and focused our design on preserving the Cys58–Cys73 linkage as the minimal stabilizing element.

CRD2-derived variants consistently dominated the affinity landscape. Notably, the disulfide-free CRD2no construct outperformed the fully bonded CRD2all variant, indicating that increased flexibility can facilitate binding-induced adaptation at the LIGHT interface. At the same time, complete removal of covalent constraints was not the most practical design solution. Once the Cys58–Cys73 disulfide bond was retained, the K54E substitution transformed CRD2e into the strongest full-domain binder by MM-GBSA, with an effective binding free energy of -58.55 ± 9.45 kcal/mol. Further truncation to CRD2(39–73)e improved the effective binding free energy to -61.43 ± kcal/mol and enabled deeper insertion into the LIGHT binding groove, indicating that the single disulfide bond was sufficient to support a highly optimized binding geometry.

SMD simulations revealed clear mechanical differences between the lead candidates. Under fast pulling conditions, CRD2(39–73)e exhibited the highest rupture force (4.14 ±1.22 kcal/mol/Å) and unbinding work (66.18 ±11.07 kcal/mol), consistent with especially tight interlocking at the protein–protein interface. In parallel, the K54E substitution reshaped the interfacial interaction network and improved the stability of the single-disulfide CRD2e scaffold. PMF calculations ranked CRD2e_K54E as the thermodynamically strongest of the closely examined variants, indicating that CRD2(39–73)e and CRD2e_K54E are favored by partially distinct criteria: the former by mechanical resistance and compact groove engagement, and the latter by equilibrium free-energy analysis. Docking studies further supported preferential targeting of the LIGHT groove, with the truncated single-disulfide peptide showing the clearest selectivity over the competing BTLA interface, while circular dichroism spectroscopy confirmed that these peptides remain predominantly disordered in solution.

Overall, our multilayered workflow identified CRD2e_K54E and CRD2(39–73)e as complementary CRD2-based inhibitors of the HVEM–LIGHT interaction. Most importantly, our results showed that disulfide bonds should be optimized rather than maximized: the Cys58–Cys73 linkage was sufficient to stabilize the active CRD2 core, whereas additional disulfide bonds could reduce the conformational adaptability that benefits LIGHT recognition. More broadly, these findings establish a generalizable computational template for disulfide-aware peptide design aimed at modulating immune signaling pathways.

Supporting Information for "Multi-Layered Computational Framework for Designing Peptide Inhibitors of HVEM-LIGHT Interaction"

Piotr Ciura[1], Pamela Smardz[2], Marta Spodzieja[1], Adam K. Sieradzan*[1], and Pawel Krupa[2]

[1]Faculty of Chemistry, University of Gdańsk, Fahrenheit Union of Universities in Gdańsk, Bażyńskiego 8, 80-309, Gdańsk, Poland
[2]Institute of Physics, Polish Academy of Sciences, Al. Lotników 32/46, 02-668, Warsaw, Poland

1 Influence of disulfide bonds on CRD stability

The HVEM domains demonstrate a very different role of particular disulfide bonds on their stability. In CRD1, out of all simulations with only one bond, Cys16-Cys29 (marked as b) lowered RMSD to 7.47Å from 11.20Å in simulations without bonds. Yet, only when coupled with Cys19-Cys29 (c), the RMSD lowers to 3.27Å, which is close to the result of simulations with all bonds present (2.89Å). Both the combinations "ab" and "ac" have results comparable to the one bond "b". It can be concluded that "a" has the smallest importance in the overall stability of CRD1.

In CRD3, which has up to two disulfide bonds, the presence of Cys83-Cys100 (g) has a greater effect on stability (3.89 Å) than the presence of Cys89-Cys97 (h). The measured end-to-end distance for CRD3h is comparable to that without bonds, and for CRD3g, it is comparable to all bonds present. The RMSF of residues in both the C- and N-terminal in simulations without (g) are much higher compared to those with (h) present (Fig. S5).

Further increase in the stability of the domain was observed when the rest of the HVEM was present, indicating that these are rather protein fragments than structurally independent moieties.

The radius of gyration (RG) and RGmax analyses provide additional confirmation that, for CRD1, the most crucial bond is (b). This is evident from the similarity in values between simulations where it is present and those where all bonds are present, indicating a consistent level of compactness. In contrast, the compactness of CRD2 was not significantly affected by the presence or absence of disulfide bonds, except for CRD2no and CRD2d (Fig. S2), which exhibited a more relaxed or loose structure, in comparison to the structures with disulfide bonds in position (e) and (f). Regarding CRD3, disulfide bond (g) appears to contribute more significantly to structural stabilization than (h). It is worth noting that for molecules without disulfide bonds the rigidify occurs upon binding. The higher the bonding affinity, the more structure rigidity is observed (Fig S2, S3). The SASA of bound and unbound HVEM variants (Fig. S4) are similar (the variants bound to LIGHT trimer have slightly larger SASA), this indicate that hydrophobic interaction are not the key players in HVEM-LIGHT interactions.

*adam.sieradzan@ug.edu.pl

List of Figures

S1 Distance between S atoms in disulfide bonds averaged for 3 trajectories for the second halves of the simulations (500-1000ns) for CRD1, CRD2 and CRD3 without the LIGHT trimer. S10
S2 (A) RMSD, (C) RG, (E) RGmax, (G) End-to-end distance calculated from simulations of free domains and (B) RMSD, (D) RG, (F) RGmax and (H) End-to-end distance values calculated from simulations with LIGHT, averaged over the second halves of the 3 trajectories. S11
S2 continued: (I) SASA values calculated from simulation of free domains and (J) theoretical SASA values calculated from simulation with the LIGHT trimer (the LIGHT trimer was excluded in this analysis, therefore, the whole HVEM variant surface was computed) averaged over the second halves of the 3 trajectories. . . . S12
S3 Experimental (A) and simulated (B) CD spectra of the selected designed peptides.S12
S4 Heatmap of per-residue $\Delta SASA = SASA_{\text{free}} - SASA_{\text{complex}}$ (right-hand bar [Å^2]) calculated for all combinations of disulfide bonds present and absent in the simulations for the full HVEM molecule, its domains (CRD1-3) and it variant, and CRD2e_K54E, and fragment (CRD(39-73)e), averaged over the second halves of the 3 trajectories. S13
S5 Heatmap of per-residue $\Delta RMSF = RMSF_{\text{free}} - RMSF_{\text{complex}}$ (right-hand bar [Å]) calculated for HVEM with and without disulfide bonds, and its selected variants forming stable complexes with LIGHT, averaged over the second halves of the 3 trajectories. S14
S6 Heatmap of the change in the content of secondary structures expressed in percentage points calculated for HVEM with and without disulfide bonds, and its selected variants forming stable complexes with LIGHT, averaged over the second halves of the 3 trajectories S15
S7 Number of native contacts between HVEM variant and LIGHT trimer as a function of time for last 20ns for (A): CRD1no (red), CRD1a (blue), and CRD1b (grey); (B): CRD1c (red), CRD1ab (blue), and CRD1ac (grey); (C): CRD1bc (red), CRD1abc (blue), and CRD1(16-38)no (grey); (D): CRD1(16-38)bc (red), CRD1-2(16-73)bce (blue), and CRD1-2(16-73)bce_K54E (grey) (E): CRD2no (red) CRD2no_K54E (blue), CRD2no_K54D (grey); (F): CRD2no_K54I (red); (F): CRD2no_K54L (blue), and CRD2no_K54S (grey); (G): CRD2no_K54V (red) CRD2no_K54Y (blue), and CRD2d (grey); (H): CRD2e (red), CRD2e_K54E (blue), and CRD2e_K54D (grey). Different trajectories are depicted as various color tone. S16
S7 continued: (I): CRD2e_K54I (red), CRD2e_K54L (blue), and CRD2e_K54S (grey); (J) CRD2e_K54V (red), CRD2e_K54Y (blue), and CRD2e_K54E_D62A (grey); (K): CRD2e_K54E_D62K (red), CRD2e_K54E_D62L (blue), and CRD2e_K54E_D62S (grey); (L): CRD2f (red); CRD2de (blue) and CRD2de_K54E (grey); (M): CRD2de_K54D (red), CRD2de_K54I (blue), and CRD2de_K54L (grey); (N): CRD2de_K54S (red), CRD2df (blue), and CRD2ef (grey); (O): CRD2def (red), CRD2(39-73)d (blue), and CRD2(39-73)de (grey); (P): CRD2(39-73)e (red), CRD2(39-73)e_K54E(blue), and CRD2(39-73)e_P39A (grey). Different trajectories are depicted as various color tone. S17

S7 continued (Q): CRD2(39-73)e_P39V(red), CRD2(39-73)e_P39W (blue), and CRD2(39-73)e_P39L (grey); (R): CRD2(39-73)e_P39I (red), CRD2(39-73)e_P39x(blue), and CRD2(39-73)e_N72A (grey); (S): CRD2(39-73)e_N72G(red), CRD2(39-73)e_C55A (blue), and CRD2(39-73)e_C55V (grey); (T): CRD2(39-73)e_C61S (red), CRD2(39-73)no (blue) and CRD3no (grey); (U): CRD3g (red), CRD3h (blue) and CRD3gh(grey); (V): CRD3(89-97)no (red), and CRD3(89-97)h (blue) . Different trajectories are depicted as various color tone. S18
S8 Bar plot of the effective free energies of the various stable variants of the HVEM molecule interacting with the LIGHT trimer. S19
S9 MM-GBSA effective binding energy decomposition results for (A) CRD1no, (B) CRD1(16-38)no, (C) CRD1bc, (D) CRD1-2(16-73)bce, (E) CRD1-2(16-73)bce_K54E, (F) CRD2no variants S20
S9 continued: MM-GBSA effective binding energy decomposition results for (G) CRD2no_K54D, (H) CRD2no_K54I, (I) CRD2no_K54L, (J) CRD2no_K54S, (K) CRD2no_K54V, (L) CRD2no_K54Y variants S21
S9 continued: MM-GBSA effective binding energy decomposition results for: (M) CRD2d, (N) CRD2e, (O) CRD2e_K54D, (P) CRD2e_K54E, (Q) CRD2e_K54I, (R) CRD2e_K54L variants S22
S9 continued: MM-GBSA effective binding energy decomposition results for: (S) CRD2e_K54S, (T) CRD2e_K54V, (U) CRD2e_K54Y, (V) CRD2e_K54E_D62A, (W) CRD2e_K54E_D62K, (X) CRD2e_K54E_D62L variants S23
S9 continued: MM-GBSA effective binding energy decomposition results for: (Y) CRD2f, (Z) CRD2de, (AA) CRD2de_K54D, (AB) CRD2de_K54E, (AC) CRD2de_K54I, (AD) CRD2de_K54S variants S24
S9 continued: MM-GBSA effective binding energy decomposition results for: (AE) CRD2df, (AF) CRD2ef, (AG) CRD2def, (AH) CRD2(39-73)no, (AI) CRD2(39-73)d, (AJ) CRD2(39-73)e variants S25
S9 continued: MM-GBSA effective binding energy decomposition results for: (AK) CRD2(39-73)e_P39A, (AL) CRD2(39-73)e_P39I, (AM) CRD2(39-73)e_P39L, (AN) CRD2(39-73)e_P39V, (AO) CRD2(39-73)e_P39W, (AP) CRD2(39-73)e_P39x variants S26
S9 continued: MM-GBSA effective binding energy decomposition results for: (AQ) CRD2(39-73)e_K54E, (AR) CRD2(39-73)e_C55A, (AS) CRD2(39-73)e_C55V, (AT) CRD2(39-73)e_C61S, (AU) CRD2(39-73)e_N72A, (AV) CRD2(39-73)e_N72G variants S27
S9 continued: MM-GBSA effective binding energy decomposition results for: (AW) CRD2(39-73)de, (AX) CRD3no, (AY) CRD3(39-73)no S28
S9 continued: MM-GBSA effective binding energy decomposition results for: (AZ) HVEMno variant, (AAA) HVEMall S29
S10 Contact heatmap averaged over 25 SMD simulations for HVEM variant CRD2e with the LIGHT trimer. The values are normalized in the range of increasing extension (relative difference of the distance between centers of mass of HVEM variant and LIGHT trimer) averaged for 4 Å. S30
S10 continued: Contact heatmap averaged over 25 SMD simulations for HVEM variant CRD2def with the LIGHT trimer. The values are normalized in the range of increasing extension (relative difference of the distance between centers of mass of HVEM variant and LIGHT trimer) averaged for 4 Å. S31
S10 continued: Contact heatmap averaged over 25 SMD simulations for HVEM variant CRD2(39-73)e with the LIGHT trimer. The values are normalized in the range of increasing extension (relative difference of the distance between centers of mass of HVEM variant and LIGHT trimer) averaged for 4 Å. S32

S10 continued: Contact heatmap averaged over 25 SMD simulations for HVEM variant CRD2e_K54E with the LIGHT trimer. The values are normalized in the range of increasing extension (relative difference of the distance between centers of mass of HVEM variant and LIGHT trimer) averaged for 4 Å. S33
S10 Representative structures upon extending (relative difference of the distance between centers of mass of HVEM variant and LIGHT trimer) averaged for 4 Å. . . S34
S11 Change of: Binding free energy (ΔG, **blue**), internal work (ΔW, **orange**), summary of Gibbs free energy and internal work (ΔG+ΔW, **green**), enthalpy (ΔH, **red**), entropy (ΔS, **violet**) in 200 ns and 1000 ns long simulations ("1μs" postscript), total work (Wtotal, pulling speed 0.05 m/s **brown**, 0.01 m/s **pink**), force max (Fmax, pulling speed 0.05 m/s **gray**, 0.01 m/s **yellowish green**) and potential mean force (PMF, **cyan**) calculated on SMD basis of eleven selected systems with highest affinity to LIGHT. CRD2e_1μs, due to dissociating, is not shown. S35
S12 Bar plots of the stable complexes of the structural properties: (A) Complex CαRMSD, (B) Receptor CαRMSD, and (C) Ligand CαRMSD values averaged for 3 trajectories for the last 20 ns of simulation. For the RMSD calculation, an initial conformation adapted from the PDB file was used as a reference. S36
S12 continued: Bar plots of the stable complexes of the structural properties: (D) RG and (E) RGmax values averaged for 3 trajectories for the last 20 ns of simulation. For the RMSD calculation, an initial conformation adapted from the PDB file was used as a reference. S37
S12 continued: Bar plots of the stable complexes of the structural properties: (F) End-to-end distance and (G) lRMSD values averaged for 3 trajectories for the last 20 ns of simulation. For the RMSD calculation, an initial conformation adapted from the PDB file was used as a reference. S38
S13 Number of native contacts between HVEM variant and LIGHT trimer as a function of time for 1μs simulations (A): CRD2e (red), CRD2e_K54E (blue), and CRD2(39-73)e (grey); (B): CRD2def (red). Different trajectories are depicted as various color tone. S38
S14 Comparison between the all-atom dominant structure (left, cartoon representation) with the UNRES-dock cluster with lowest RMSD to respective all-atom structure (right, ribbon representation). Each LIGHT chain is colored green, cyan, magenta and peptide is colored in red for: A) CRD2e (RMSD 4.80Å), B) CRD2def (RMSD 4.39Å), C) CRD2_K54E (RMSD 6.30Å) and D) CRD2(39-73)e (RMSD 4.95Å). S39

Table S1: Names and amino acid sequences of the designed peptides

name	disulfide bonds	residues
CRD1no	without	$L^{1}PSCKEDEYPV^{11}GSECCPKCSP^{21}GYRVKEACG^{31}ELTGTVCE$
CRD1a	C^{4}-C^{15}	$L^{1}PSCKEDEYPV^{11}GSECCPKCSP^{21}GYRVKEACG^{31}ELTGTVCE$
CRD1b	C^{16}-C^{29}	$L^{1}PSCKEDEYPV^{11}GSECCKCSP^{21}GYRVKEACG^{31}ELTGTVCE$
CRD1c	C^{19}-C^{37}	$L^{1}PSCKEDEYPV^{11}GSECCPKCSP^{21}GYRVKEACG^{31}ELTGTVCE$
CRD1ab	C^{4}-C^{15};C^{16}-C^{29}	$L^{1}PSCKEDEYPV^{11}GSECCPKCSP^{21}GYRVKEACG^{31}ELTGTVCE$
CRD1ac	C^{4}-C^{15};C^{19}-C^{37}	$L^{1}PSCKEDEYPV^{11}GSECCPKCSP^{21}GYRVKEACG^{31}ELTGTVCE$
CRD1bc	C^{16}-C^{29};C^{19}-C^{37}	$L^{1}PSCKEDEYPV^{11}GSECCPKCSP^{21}GYRVKEACG^{31}ELTGTVCE$
CRD1abc	C^{4}-C^{15};C^{16}-C^{29};C^{19}-C^{37}	$L^{1}PSCKEDEYPV^{11}GSECCPKCSP^{21}GYRVKEACG^{31}ELTGTVCE$
CRD1(16-38)no	without	$CPKCSP^{21}GYRVKEACG^{31}ELTGTVCE$
CRD1(16-38)bc	C^{16}-C^{29};C^{19}-C^{37}	$CPKCSP^{21}GYRVKEACG^{31}ELTGTVCE$
CRD1-2(16-73)bce	C^{16}-C^{29};C^{19}-C^{37};C^{58}-C^{73}	$CPKCSP^{21}GYRVKEACG^{31}ELTGTVCEPCP^{41}PGTYIHLNG^{51}LSKCLQCQMC^{61}DPAMGLRASR^{71}NC^{73}$
CRD1-2(16-73)bce_K54E	C^{16}-C^{29};C^{19}-C^{37};C^{58}-C^{73}	$CPKCSP^{21}GYRVKEACG^{31}ELTGTVCEPCP^{41}PGTYIHLNG^{51}LSECLQCQMC^{61}DPAMGLRASR^{71}NC^{73}$
CRD2no	without	$PCP^{41}PGTYIHLNG^{51}LSKCLQCQMC^{61}DPAMGLRASR^{71}NCSRTENAVC^{81}$
CRD2no_K54D	without	$PCP^{41}PGTYIHLNG^{51}LSDCLQCQMC^{61}DPAMGLRASR^{71}NCSRTENAVC^{81}$
CRD2no_K54E	without	$PCP^{41}PGTYIHLNG^{51}LSECLQCQMC^{61}DPAMGLRASR^{71}NCSRTENAVC^{81}$
CRD2no_K54I	without	$PCP^{41}PGTYIHLNG^{51}LSICLQCQMC^{61}DPAMGLRASR^{71}NCSRTENAVC^{81}$
CRD2no_K54L	without	$PCP^{41}PGTYIHLNG^{51}LSLCLQCQMC^{61}DPAMGLRASR^{71}NCSRTENAVC^{81}$
CRD2no_K54S	without	$PCP^{41}PGTYIHLNG^{51}LSSCLQCQMC^{61}DPAMGLRASR^{71}NCSRTENAVC^{81}$
CRD2no_K54V	without	$PCP^{41}PGTYIHLNG^{51}LSVCLQCQMC^{61}DPAMGLRASR^{71}NCSRTENAVC^{81}$
CRD2no_K54Y	without	$PCP^{41}PGTYIHLNG^{51}LSYCLQCQMC^{61}DPAMGLRASR^{71}NCSRTENAVC^{81}$
CRD2d	C^{40}-C^{55}	$PCP^{41}PGTYIHLNG^{51}LSKCLQCQMC^{61}DPAMGLRASR^{71}NCSRTENAVC^{81}$
CRD2e	C^{58}-C^{73}	$PCP^{41}PGTYIHLNG^{51}LSKCLQCQMC^{61}DPAMGLRASR^{71}NCSRTENAVC^{81}$
CRD2e_K54D	C^{58}-C^{73}	$PCP^{41}PGTYIHLNG^{51}LSDCLQCQMC^{61}DPAMGLRASR^{71}NCSRTENAVC^{81}$
CRD2e_K54E	C^{58}-C^{73}	$PCP^{41}PGTYIHLNG^{51}LSECLQCQMC^{61}DPAMGLRASR^{71}NCSRTENAVC^{81}$
CRD2e_K54I	C^{58}-C^{73}	$PCP^{41}PGTYIHLNG^{51}LSICLQCQMC^{61}DPAMGLRASR^{71}NCSRTENAVC^{81}$
CRD2e_K54L	C^{58}-C^{73}	$PCP^{41}PGTYIHLNG^{51}LSLCLQCQMC^{61}DPAMGLRASR^{71}NCSRTENAVC^{81}$
CRD2e_K54S	C^{58}-C^{73}	$PCP^{41}PGTYIHLNG^{51}LSSCLQCQMC^{61}DPAMGLRASR^{71}NCSRTENAVC^{81}$
CRD2e_K54V	C^{58}-C^{73}	$PCP^{41}PGTYIHLNG^{51}LSVCLQCQMC^{61}DPAMGLRASR^{71}NCSRTENAVC^{81}$
CRD2e_K54Y	C^{58}-C^{73}	$PCP^{41}PGTYIHLNG^{51}LSYCLQCQMC^{61}DPAMGLRASR^{71}NCSRTENAVC^{81}$
CRD2e_K54E_D62A	C^{58}-C^{73}	$PCP^{41}PGTYIHLNG^{51}LSECLQCQMC^{61}APAMGLRASR^{71}NCSRTENAVC^{81}$
CRD2e_K54E_D62K	C^{58}-C^{73}	$PCP^{41}PGTYIHLNG^{51}LSECLQCQMC^{61}KPAMGLRASR^{71}NCSRTENAVC^{81}$
CRD2e_K54E_D62L	C^{58}-C^{73}	$PCP^{41}PGTYIHLNG^{51}LSECLQCQMC^{61}LPAMGLRASR^{71}NCSRTENAVC^{81}$
CRD2e_K54E_D62S	C^{58}-C^{73}	$PCP^{41}PGTYIHLNG^{51}LSECLQCQMC^{61}SPAMGLRASR^{71}NCSRTENAVC^{81}$
CRD2f	C^{61}-C^{81}	$PCP^{41}PGTYIHLNG^{51}LSKCLQCQMC^{61}DPAMGLRASR^{71}NCSRTENAVC^{81}$

Table S1: Continued: Names and amino acid sequences of the designed peptides (x denotes point deletion)

CRD2de	C^{40}-C^{55};C^{58}-C^{73}	PC**P**^{41}PGTYIHLNG^{51}LSK**C**LQ**C**QMC^{61}DPAMGLRASR^{71}N**C**SRTENAVC^{81}
CRD2de_K54D	C^{40}-C^{55};C^{58}-C^{73}	PC**P**^{41}PGTYIHLNG^{51}LSD**C**LQ**C**QMC^{61}DPAMGLRASR^{71}N**C**SRTENAVC^{81}
CRD2de_K54E	C^{40}-C^{55};C^{58}-C^{73}	PC**P**^{41}PGTYIHLNG^{51}LSE**C**LQ**C**QMC^{61}DPAMGLRASR^{71}N**C**SRTENAVC^{81}
CRD2de_K54I	C^{40}-C^{55};C^{58}-C^{73}	PC**P**^{41}PGTYIHLNG^{51}LSI**C**LQ**C**QMC^{61}DPAMGLRASR^{71}N**C**SRTENAVC^{81}
CRD2de_K54L	C^{40}-C^{55};C^{58}-C^{73}	PC**P**^{41}PGTYIHLNG^{51}LSL**C**LQ**C**QMC^{61}DPAMGLRASR^{71}N**C**SRTENAVC^{81}
CRD2de_K54S	C^{40}-C^{55};C^{58}-C^{73}	PC**P**^{41}PGTYIHLNG^{51}LSS**C**LQ**C**QMC^{61}DPAMGLRASR^{71}N**C**SRTENAVC^{81}
CRD2df	C^{40}-C^{55};C^{61}-C^{81}	PC**P**^{41}PGTYIHLNG^{51}LSK**C**LQ**C**QMC^{61}DPAMGLRASR^{71}N**C**SRTENAVC^{81}
CRD2ef	C^{58}-C^{73};C^{61}-C^{81}	PC**P**^{41}PGTYIHLNG^{51}LSK**C**LQ**C**QMC^{61}DPAMGLRASR^{71}N**C**SRTENAVC^{81}
CRD2def	C^{40}-C^{55};C^{58}-C^{73};C^{61}-C^{81}	PC**P**^{41}PGTYIHLNG^{51}LSK**C**LQ**C**QMC^{61}DPAMGLRASR^{71}N**C**SRTENAVC^{81}
CRD2(39-73)d	C^{40}-C^{55}	PC**P**^{41}PGTYIHLNG^{51}LSK**C**LQ**C**QMC^{61}DPAMGLRASR^{71}NC^{73}
CRD2(39-73)e	C^{58}-C^{73}	PC**P**^{41}PGTYIHLNG^{51}LSK**C**LQ**C**QMC^{61}DPAMGLRASR^{71}N**C**^{73}
CRD2(39-73)e_P39A	C^{58}-C^{73}	A**C**P^{41}PGTYIHLNG^{51}LSK**C**LQ**C**QMC^{61}DPAMGLRASR^{71}N**C**^{73}
CRD2(39-73)e_P39L	C^{58}-C^{73}	L**C**P^{41}PGTYIHLNG^{51}LSK**C**LQ**C**QMC^{61}DPAMGLRASR^{71}N**C**^{73}
CRD2(39-73)e_P39I	C^{58}-C^{73}	I**C**P^{41}PGTYIHLNG^{51}LSK**C**LQ**C**QMC^{61}DPAMGLRASR^{71}N**C**^{73}
CRD2(39-73)e_P39V	C^{58}-C^{73}	V**C**P^{41}PGTYIHLNG^{51}LSK**C**LQ**C**QMC^{61}DPAMGLRASR^{71}N**C**^{73}
CRD2(39-73)e_P39W	C^{58}-C^{73}	W**C**P^{41}PGTYIHLNG^{51}LSK**C**LQ**C**QMC^{61}DPAMGLRASR^{71}N**C**^{73}
CRD2(39-73)e_P39x	C^{57}-C^{72}	**C**P^{40}PGTYIHLNG^{50}LSK**C**LQ**C**QMC^{60}DPAMGLRASR^{70}N**C**^{72}
CRD2(39-73)e_K54E	C^{58}-C^{73}	PC**P**^{41}PGTYIHLNG^{51}LSE**C**LQ**C**QMC^{61}DPAMGLRASR^{71}N**C**^{73}
CRD2(39-73)e_C55A	C^{58}-C^{73}	PC**P**^{41}PGTYIHLNG^{51}LSK**A**LQ**C**QMC^{61}DPAMGLRASR^{71}N**C**^{73}
CRD2(39-73)e_C55V	C^{58}-C^{73}	PC**P**^{41}PGTYIHLNG^{51}LSK**V**LQ**C**QMC^{61}DPAMGLRASR^{71}N**C**^{73}
CRD2(39-73)e_C61S	C^{58}-C^{73}	PC**P**^{41}PGTYIHLNG^{51}LSK**C**LQ**C**QMS^{61}DPAMGLRASR^{71}N**C**^{73}
CRD2(39-73)e_N72A	C^{58}-C^{73}	PC**P**^{41}PGTYIHLNG^{51}LSK**C**LQ**C**QMC^{61}DPAMGLRASR^{71}**A**C^{73}
CRD2(39-73)e_N72G	C^{58}-C^{73}	PC**P**^{41}PGTYIHLNG^{51}LSK**C**LQ**C**QMC^{61}DPAMGLRASR^{71}G**C**^{73}
CRD2(39-73)de	C^{40}-C^{55};C^{58}-C^{73}	PC**P**^{41}PGTYIHLNG^{51}LSK**C**LQ**C**QMC^{61}DPAMGLRASR^{71}N**C**^{73}
CRD2(39-73)no	without	PC**P**^{41}PGTYIHLNG^{51}LSK**C**LQ**C**QMS^{61}DPAMGLRASR^{71}NC^{73}
CRD3no	without	GCSPQHFCIV^{91}QDGDHCAAC**R**^{101}AYA
CRD3g	C^{83}-C^{100}	G**C**SPQHFCIV^{91}QDGDH**C**AAC**R**^{101}AYA
CRD3h	C^{89}-C^{97}	GCSPQH**F**CIV^{91}QDGD**H**CAAC**R**^{101}AYA
CRD3gh	C^{83}-C^{100};C^{89}-C^{97}	G**C**SPQH**F**CIV^{91}QDGD**H**CAAC**R**^{101}AYA
CRD3(89-97)h	C^{89}-C^{97}	**C**IV^{91}QDG**D**H**C**
CRD3(89-97)no	without	CIV^{91}QDGDHC

Table S2: Details of the simulations performed in the project. cMD stands for conventional Molecular Dynamics simulations, SMD for Steered Molecular Dynamics simulations with 0.05 and 0.01 m/s for fast and slow pulling, respectively, and US for Umbrella Sampling. All designed variants of HVEM-based molecules are treated as possible inhibitors, while complete HVEM simulations were run for comparison only.

Description	**No. of Molecules**	**Number and Length of Trajectories**	**Total Time [µs]**
cMD of inhibitors with LIGHT	65	3 × 200ns	37.8
cMD of HVEM-LIGHT complex	2	3 × 200ns	1.2
cMD of top inhibitors with LIGHT	4	3 × 800ns (continuation from 200ns)	9.6
cMD of unbound inhibitors	29	3 × 1000ns	87
cMD of complete unbound HVEM	2	3 × 1000ns	6
SMD_{fast} of inhibitors with LIGHT	4	25 × 80ns	8
SMD_{slow} of inhibitors with LIGHT	4	10 × 250ns	10
US of inhibitors with LIGHT	4	44 × 100ns	17.6
			177.2

Table S3: Estimated secondary structure content

	Estimated secondary structure content (%)															
	α-Helix				**β-sheet**				**Turn**				**Coil**			
Peptide	DSSP	STRIDE	KAKSI	CD	DSSP	STRIDE	KAKSI	CD	DSSP	STRIDE	KAKSI	CD	DSSP	STRIDE	KAKSI	CD
HVEM-CRD1no	0.60	0.93			35.33	30.20			27.48	36.06			36.58	32.62		
CRD1no	4.49	8.89	5.50		10.26	12.47	9.20		34.32	38.20			50.94	45.55	85.29	
CRD1a	4.58	10.82	7.77		21.78	21.94	13.41		30.28	35.22			43.36	37.60	78.82	
CRD1b	1.42	6.24	0.31		20.55	16.22	8.23		32.63	40.93			45.40	41.16	91.46	
CRD1c	2.45	5.64	0.35		12.84	7.17	2.61		33.80	43.39			50.91	49.60	97.04	
CRD1ab	3.02	6.18	0.73		23.15	27.39	17.83		30.98	34.84			42.85	36.36	81.44	
CRD1ac	1.90	6.49	1.87		26.08	24.30	13.81		31.42	42.30			40.59	30.75	84.32	
CRD1bc	0.54	3.39	0.57		26.70	22.26	13.17		27.87	41.68			44.89	33.75	86.26	
CRD1abc	0.02	3.07	0.00		26.65	24.04	15.18		27.66	42.82			45.66	32.84	84.82	
HVEM-CRD1abc	0.00	0.00			39.97	37.22			23.23	35.38			36.80	27.14		
HVEM-CRD2no	1.87	3.07			12.39	10.89			29.86	31.55			55.87	52.28		
CRD2no	5.22	8.13	6.46	3.9	10.44	4.48	0.34	28	32.44	39.50		17.6	51.90	47.57	93.20	50.5
CRD2d	4.14	4.98	2.87	10.9	12.86	10.21	3.33	26.4	33.71	37.72		15.4	49.30	46.38	93.81	47.3
CRD2e	2.58	0.89	0.09	13.3	20.28	17.48	16.21	18.8	28.65	37.92		15	48.48	40.24	83.70	52.9
CRD2e_K54E	3.23	2.80	0.53	14.1	19.90	17.66	14.07	21.6	27.61	30.97		15.1	49.26	46.70	85.40	49.2
CRD2f	2.84	3.44	3.76	7.9	13.28	13.71	11.27	23.3	31.97	40.68		17.3	51.92	39.71	84.97	51.5
CRD2de	1.56	0.32	0.02	3.2	19.48	18.07	8.88	32.1	30.96	37.65		17.4	48.00	42.97	91.10	47.3
CRD2df	1.38	2.95	0.00	1.2	20.84	20.34	7.44	39.5	34.66	42.10		15.1	43.12	33.37	92.56	44.2
CRD2ef	1.32	1.08	0.10	2.5	19.63	19.77	16.11	33.4	28.15	32.18		17.4	50.90	44.84	83.79	46.7
CRD2def	1.72	0.87	0.00		19.67	18.46	9.88		31.86	39.83			46.75	38.18	90.12	
CRD2(39-73)e	6.61	9.42	11.10		9.09	2.13	5.67		32.29	28.69			52.10	88.44	54.37	
HVEM-CRD2def	5.09	6.01			17.45	13.72			24.77	25.31			52.69	50.56		
HVEM-CRD3no	2.31	4.29			20.03	8.55			31.66	37.64			46.00	49.32		
CRD3no	7.56	21.78	13.58		9.44	3.55	0.05		33.10	35.75			49.90	54.49	86.37	
CRD3g	6.31	14.76	0.00		29.96	27.90	16.19		35.73	43.12			28.00	40.41	83.81	
CRD3h	1.45	14.38	1.13		24.19	21.23	0.24		30.29	32.36			44.06	54.11	98.63	
CRD3gh	0.03	14.60	0.00		31.88	38.86	18.22		28.17	35.28			39.92	37.32	81.78	
HVEM-CRD3gh	0.00	0.00			46.69	50.89			24.62	24.24			28.68	24.66		

Table S4: Solvent accessible surfaces area (SASA) [Å^2] obtained for last 20 ns averaged over three trajectories and their standard deviation between trajectories. Δ Sasa hydrophobic indicates the difference in SASA when only hydrophobic residues (GLY,ALA,VAL,LEU,ILE,PRO,PHE,MET,TRP) are taken into consideration

Name	SASA of complex	Sum of SASA of components	Ratio	ΔSASA hydrophobic
CRD1no	21793 ± 123	22566 ± 259	0.966 ± 0.006	121.89 ± 65.17
CRD1bc	25262 ± 2090	25763 ± 1755	0.980 ± 0.015	85.53 ± 67.89
CRD1(16-38)no	23833 ± 1703	24237 ± 2002	0.984 ± 0.012	16.26 ± 26.36
CRD1-2(16-73)bce	25231 ± 1528	25737 ± 1404	0.980 ± 0.016	108.92 ± 104.12
CRD1-2(16-73)bce_K54E	24095 ± 578	25235 ± 578	0.955 ± 0.024	328.86 ± 192.27
CRD2no	25123 ± 2189	25582 ± 1986	0.982 ± 0.023	56.50 ± 63.87
CRD2no_K54D	24743 ± 1492	25361 ± 1314	0.975 ± 0.018	155.94 ± 200.26
CRD2no_K54I	24971 ± 2016	25456 ± 1592	0.980 ± 0.028	122.83 ± 206.07
CRD2no_K54L	21883 ± 403	23306 ± 445	0.939 ± 0.002	416.64 ± 35.44
CRD2no_K54S	24718 ± 2741	25081 ± 2421	0.984 ± 0.015	80.03 ± 95.60
CRD2no_K54V	24767 ± 2205	25474 ± 1805	0.971 ± 0.021	260.70 ± 223.09
CRD2no_K54Y	26345 ± 2527	26944 ± 2218	0.977 ± 0.016	133.45 ± 89.89
CRD2def	24550 ± 1650	25751 ± 1548	0.953 ± 0.007	317.82 ± 94.98
CRD2d	23337 ± 1531	24674 ± 1702	0.946 ± 0.004	319.52 ± 134.31
CRD2e	24237 ± 2004	24922 ± 1860	0.972 ± 0.024	128.37 ± 6.76
CRD2e_K54D	23486 ± 324	23842 ± 439	0.985 ± 0.007	83.56 ± 46.29
CRD2e_K54E	24103 ± 1751	24662 ± 1499	0.977 ± 0.025	307.50 ± 216.24
CRD2e_K54I	25385 ± 1477	26133 ± 1708	0.972 ± 0.007	320.30 ± 75.40
CRD2e_K54L	22874 ± 1091	24023 ± 554	0.952 ± 0.037	262.49 ± 229.68
CRD2e_K54S	24145 ± 1782	24429 ± 1542	0.988 ± 0.017	85.41 ± 147.93
CRD2e_K54V	23837 ± 1628	24581 ± 1306	0.969 ± 0.016	201.54 ± 121.90
CRD2e_K54Y	23456 ± 3045	24879 ± 2996	0.942 ± 0.011	368.41 ± 90.04
CRD2e_K54E_D62A	22354 ± 595	23329 ± 557	0.959 ± 0.030	229.90 ± 199.63
CRD2e_K54E_D62K	25340 ± 2252	25662 ± 2122	0.987 ± 0.012	66.02 ± 83.04
CRD2e_K54E_D62L	23559 ± 608	24065 ± 750	0.979 ± 0.019	94.17 ± 101.94
CRD2f	24639 ± 1542	25361 ± 1181	0.971 ± 0.021	207.58 ± 179.94
CRD2de	23300 ± 1938	24183 ± 1870	0.963 ± 0.019	210.06 ± 183.06
CRD2de_K54D	22298 ± 949	22956 ± 719	0.971 ± 0.012	148.10 ± 94.97
CRD2de_K54S	21582 ± 744	22400 ± 501	0.963 ± 0.018	164.07 ± 53.50
CRD2de_K54E	22901 ± 464	24198 ± 594	0.947 ± 0.013	341.12 ± 74.60
CRD2de_K54I	22798 ± 376	23902 ± 88	0.954 ± 0.018	313.70 ± 131.19
CRD2df	21272 ± 165	22639 ± 254	0.940 ± 0.004	232.40 ± 154.92
CRD2ef	24146 ± 1967	25097 ± 2028	0.962 ± 0.017	235.08 ± 230.81
CRD2(39-73)no	22343 ± 209	23489 ± 492	0.951 ± 0.012	215.44 ± 39.36
CRD2(39-73)d	26290 ± 1109	26524 ± 981	0.991 ± 0.011	42.95 ± 62.62
CRD2(39-73)e	23765 ± 2148	24621 ± 1688	0.964 ± 0.034	447.85 ± 87.44
CRD2(39-73)e_P39A	23400 ± 2209	23953 ± 2368	0.977 ± 0.017	103.85 ± 92.87
CRD2(39-73)e_P39I	24892 ± 1753	25883 ± 1624	0.961 ± 0.011	183.16 ± 92.87
CRD2(39-73)e_P39L	22699 ± 133	23398 ± 57	0.970 ± 0.004	137.40 ± 30.19
CRD2(39-73)e_P39V	25384 ± 1813	25852 ± 2016	0.983 ± 0.015	74.50 ± 64.61
CRD2(39-73)e_P39W	23523 ± 1889	24169 ± 2042	0.974 ± 0.019	95.89 ± 83.10
CRD2(39-73)e_P39x	24127 ± 1826	24842 ± 1282	0.970 ± 0.024	216.31 ± 261.48
CRD2(39-73)e_C55A	24474 ± 1943	25498 ± 1605	0.959 ± 0.022	268.11 ± 239.23
CRD2(39-73)e_C55V	23251 ± 2839	24533 ± 2756	0.947 ± 0.009	359.40 ± 45.79
CRD2(39-73)e_C61S	23123 ± 365	24232 ± 557	0.954 ± 0.015	326.65 ± 145.24
CRD2(39-73)e_N72A	23896 ± 2472	24779 ± 2119	0.963 ± 0.033	234.76 ± 227.51
CRD2(39-73)e_N72G	22300 ± 454	23785 ± 470	0.938 ± 0.006	374.15 ± 48.42
CRD2(39-73)de	21193 ± 347	22760 ± 453	0.931 ± 0.012	425.29 ± 104.52
CRD3no	24141 ± 849	24272 ± 857	0.995 ± 0.005	34.82 ± 41.84
CRD3(89-97)no	22657 ± 1922	22871 ± 1645	0.990 ± 0.017	0.45 ± 0.54
HVEMno	26750 ± 223	28726 ± 229	0.931 ± 0.011	505.543 ± 166.06
HVEMall	25892 ± 2028	27787 ± 1332	0.931 ± 0.030	496.38 ± 349.50

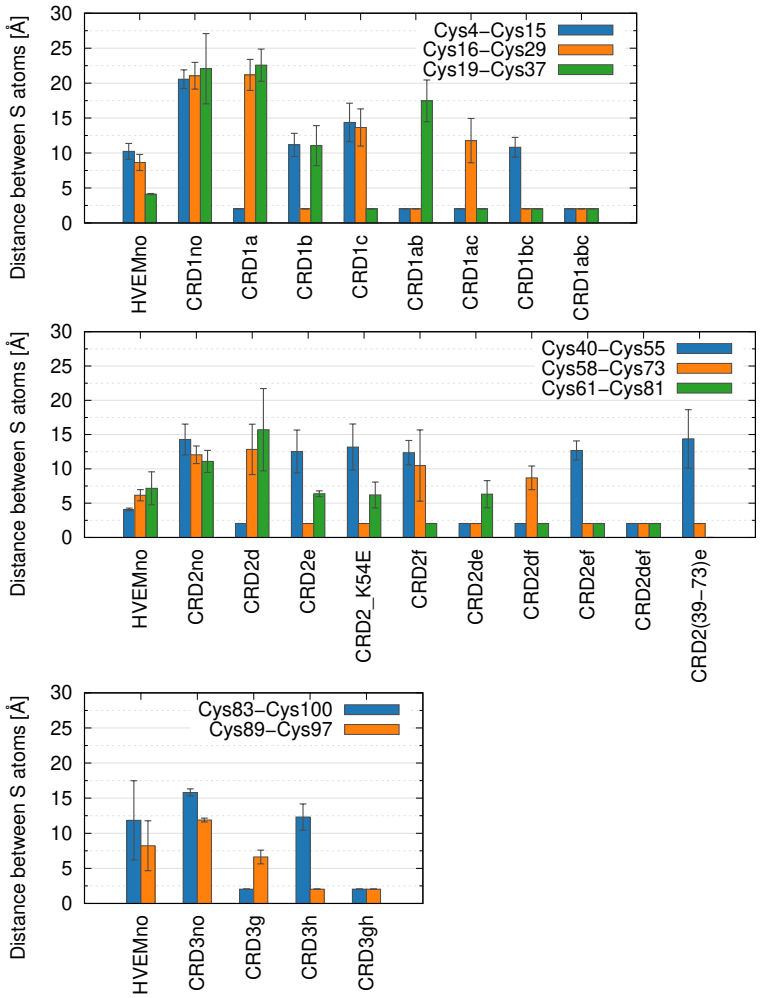


Figure S1: Distance between S atoms in disulfide bonds averaged for 3 trajectories for the second halves of the simulations (500-1000ns) for CRD1, CRD2 and CRD3 without the LIGHT trimer.

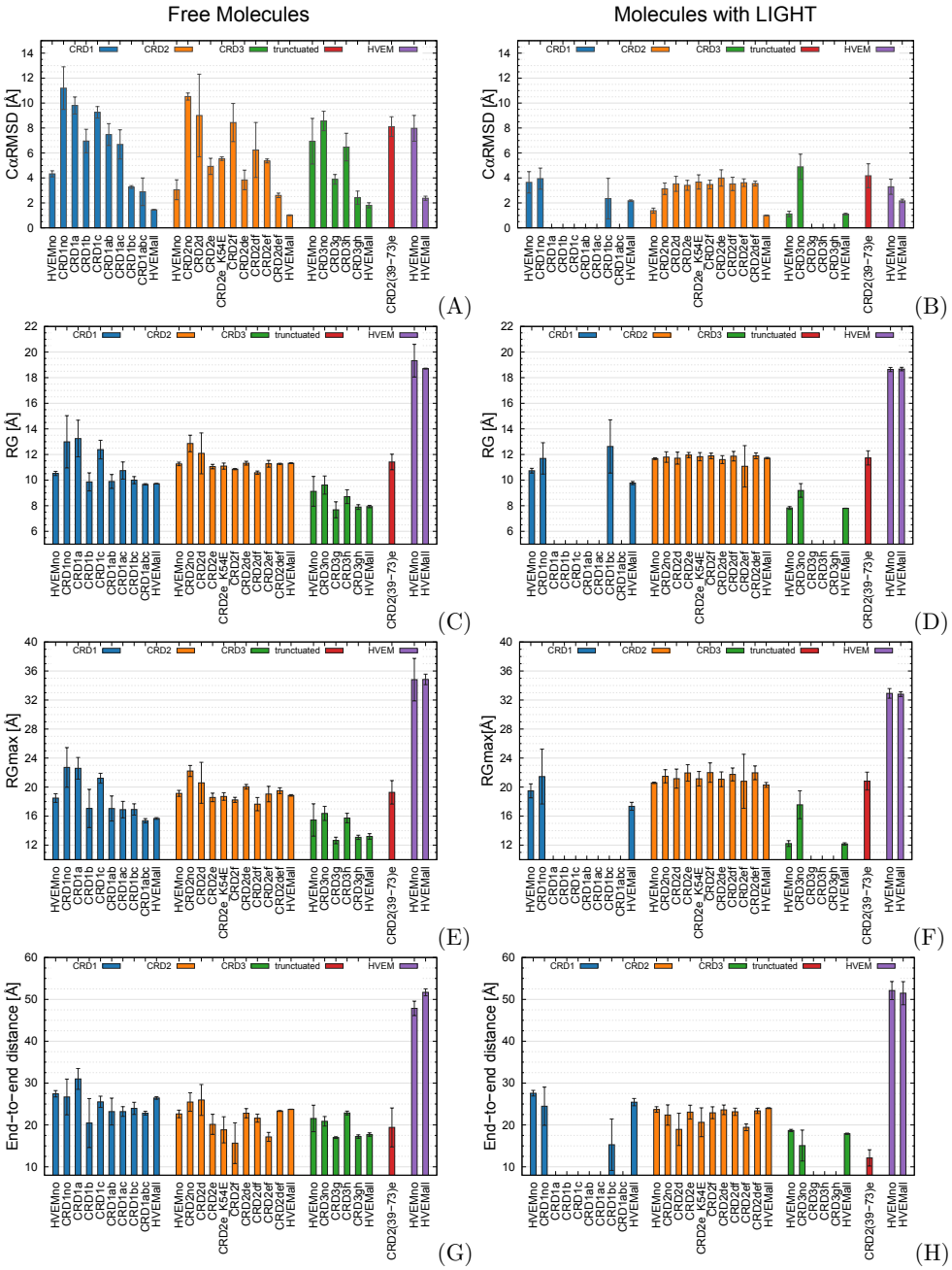


Figure S2: (A) RMSD, (C) RG, (E) RGmax, (G) End-to-end distance calculated from simulations of free domains and (B) RMSD, (D) RG, (F) RGmax and (H) End-to-end distance values calculated from simulations with LIGHT, averaged over the second halves of the 3 trajectories.

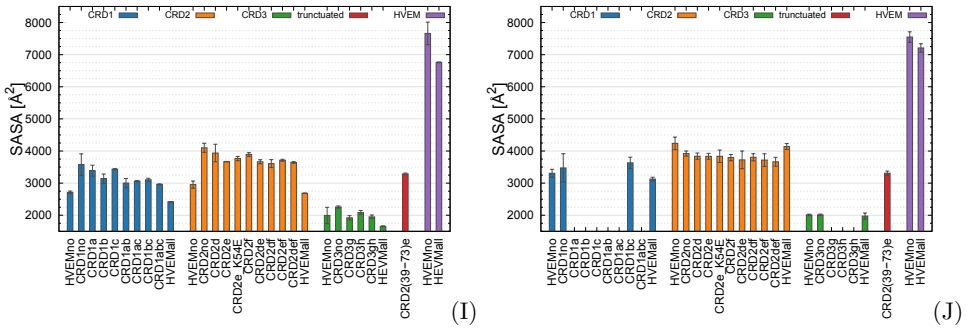


Figure S2: continued: (I) SASA values calculated from simulation of free domains and (J) theoretical SASA values calculated from simulation with the LIGHT trimer (the LIGHT trimer was excluded in this analysis, therefore, the whole HVEM variant surface was computed) averaged over the second halves of the 3 trajectories.

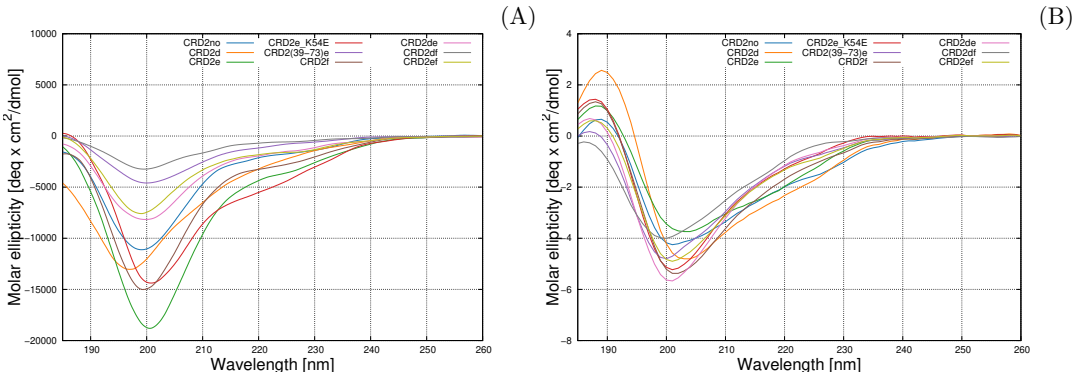


Figure S3: Experimental (A) and simulated (B) CD spectra of the selected designed peptides.

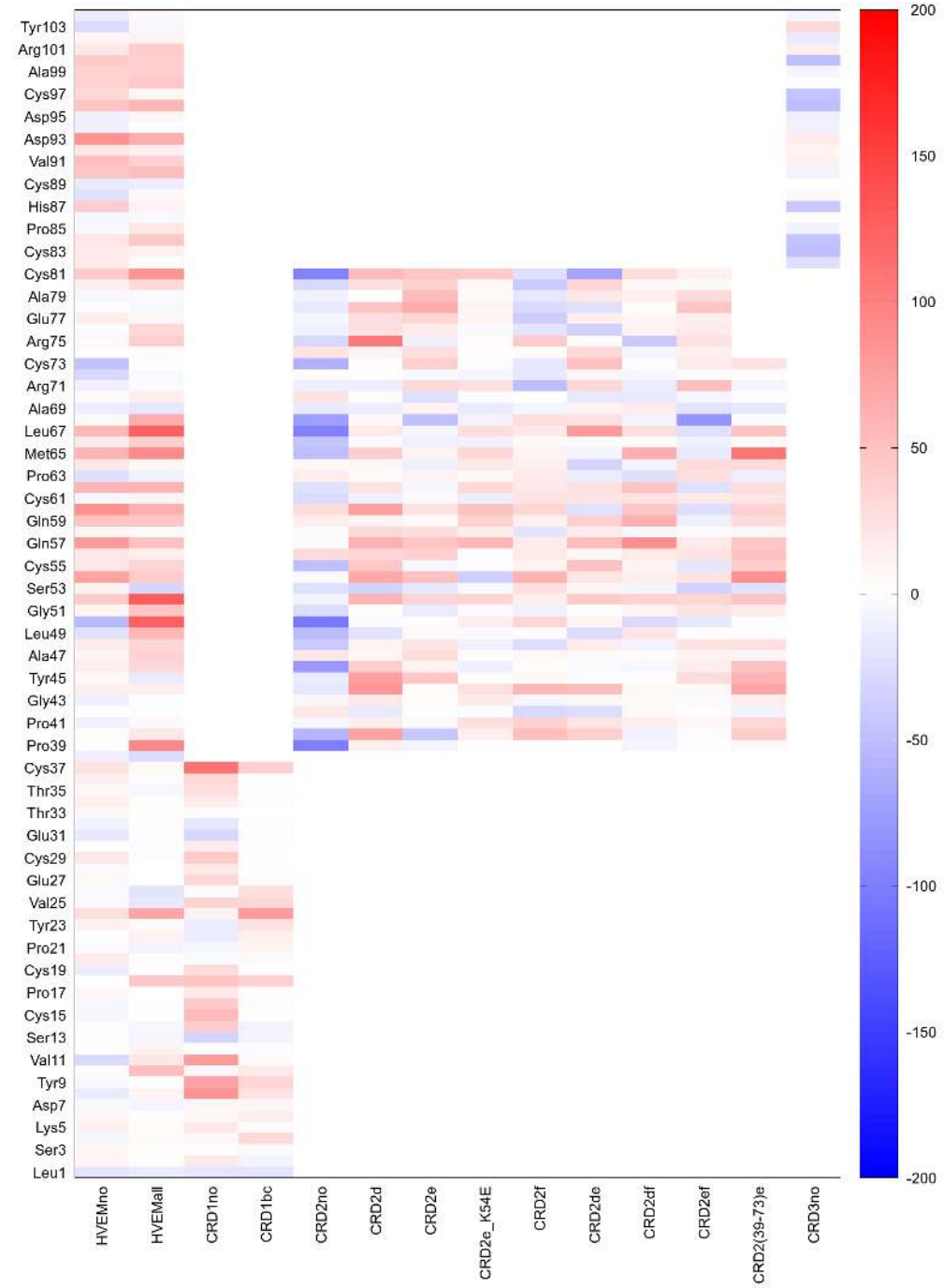


Figure S4: Heatmap of per-residue $\Delta SASA = SASA_{\text{free}} - SASA_{\text{complex}}$ (right-hand bar [Å^2]) calculated for all combinations of disulfide bonds present and absent in the simulations for the full HVEM molecule, its domains (CRD1-3) and it variant, and CRD2e_K54E, and fragment (CRD(39-73)e), averaged over the second halves of the 3 trajectories.

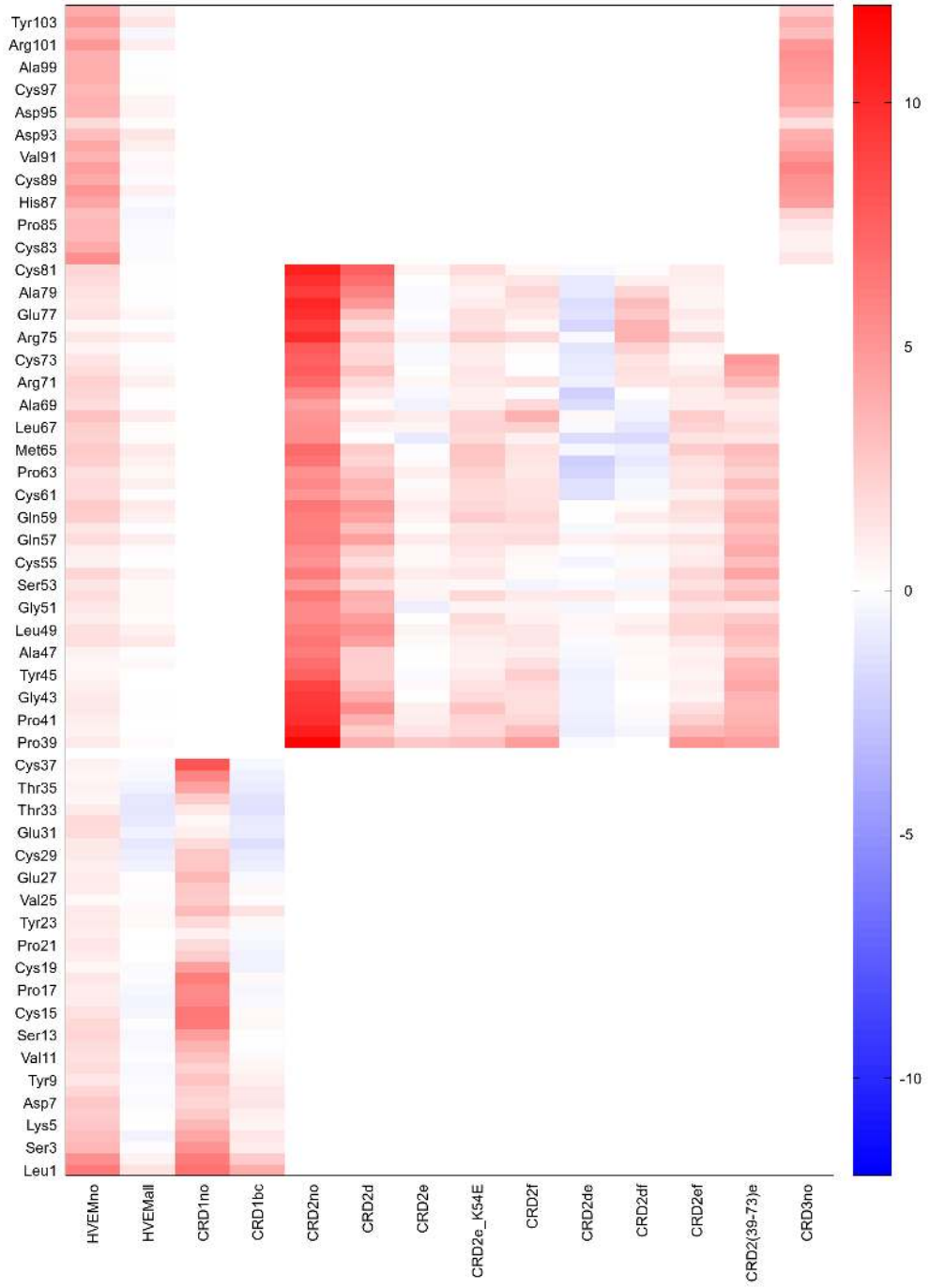


Figure S5: Heatmap of per-residue $\Delta RMSF = RMSF_{\text{free}} - RMSF_{\text{complex}}$ (right-hand bar [Å]) calculated for HVEM with and without disulfide bonds, and its selected variants forming stable complexes with LIGHT, averaged over the second halves of the 3 trajectories.

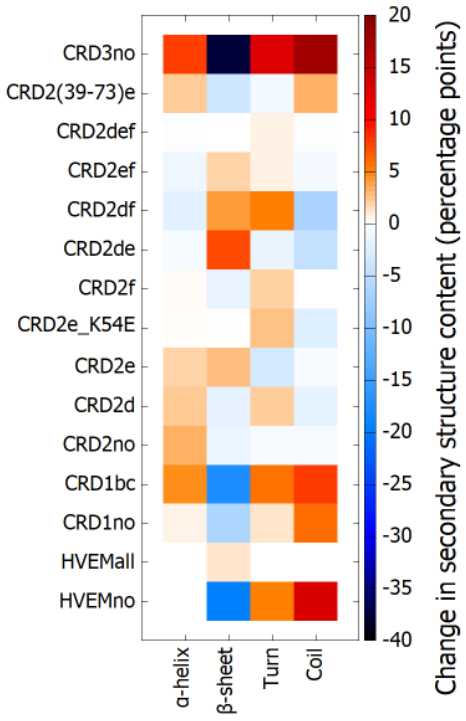


Figure S6: Heatmap of the change in the content of secondary structures expressed in percentage points calculated for HVEM with and without disulfide bonds, and its selected variants forming stable complexes with LIGHT, averaged over the second halves of the 3 trajectories

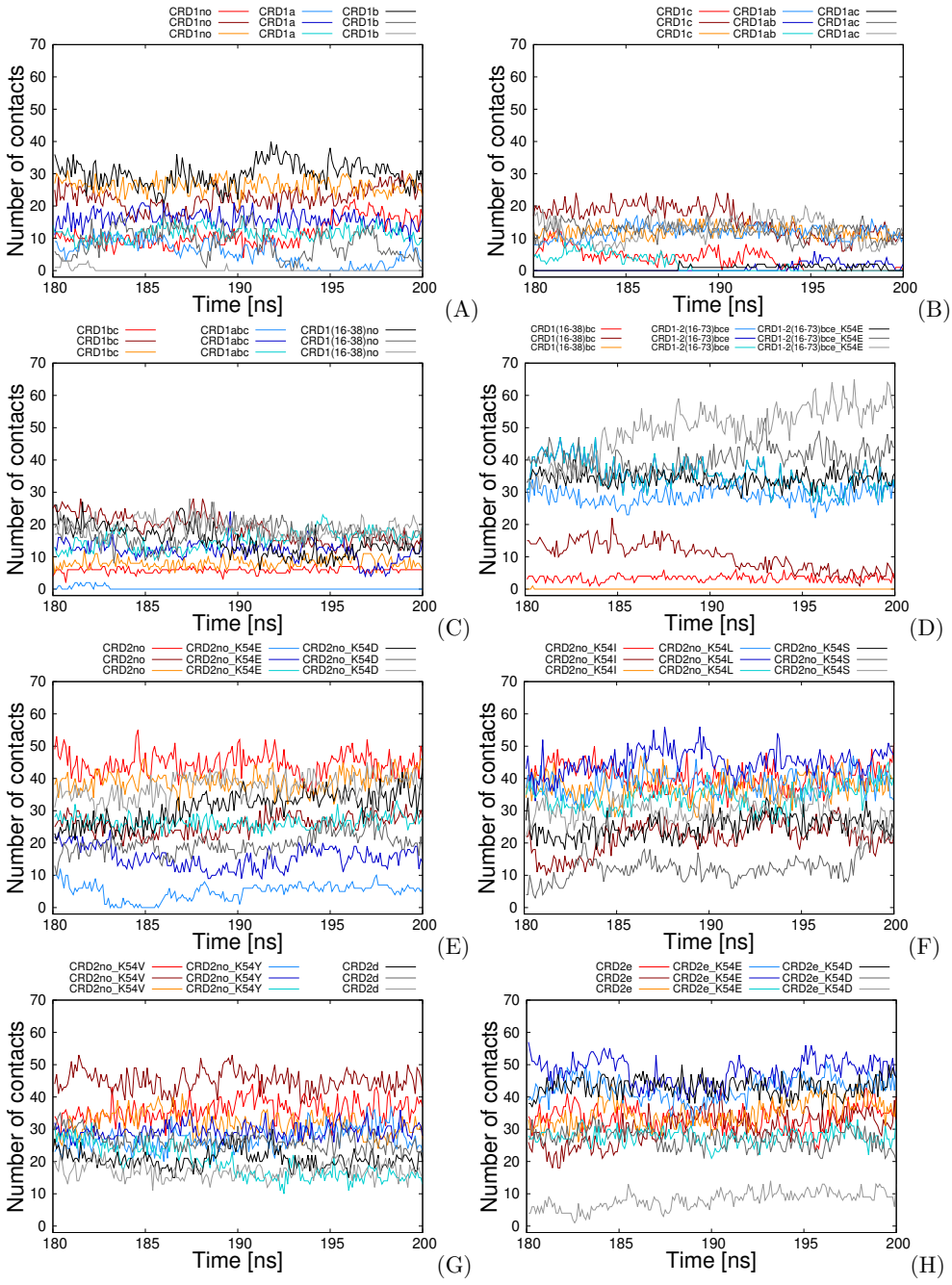


Figure S7: Number of native contacts between HVEM variant and LIGHT trimer as a function of time for last 20ns for (A): CRD1no (red), CRD1a (blue), and CRD1b (grey); (B): CRD1c (red), CRD1ab (blue), and CRD1ac (grey); (C): CRD1bc (red), CRD1abc (blue), and CRD1(16-38)no (grey); (D): CRD1(16-38)bc (red), CRD1-2(16-73)bce (blue), and CRD1-2(16-73)bce_K54E (grey) (E): CRD2no (red) CRD2no_K54E (blue), CRD2no_K54D (grey); (F): CRD2no_K54I (red); (F): CRD2no_K54L (blue), and CRD2no_K54S (grey); (G): CRD2no_K54V (red) CRD2no_K54Y (blue), and CRD2d (grey); (H): CRD2e (red), CRD2e_K54E (blue), and CRD2e_K54D (grey). Different trajectories are depicted as various color tone.

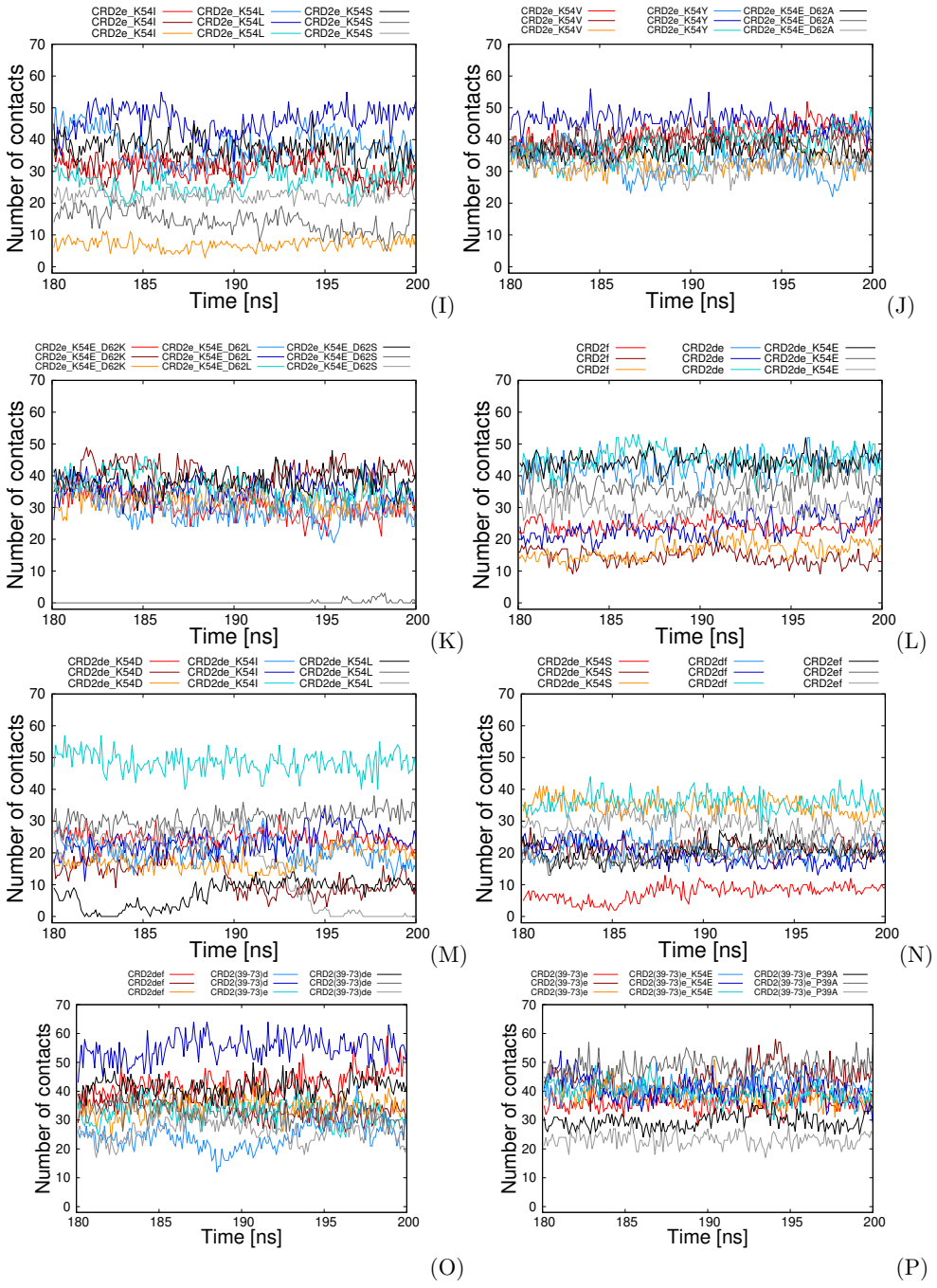


Figure S7: continued: (I): CRD2e_K54I (red), CRD2e_K54L (blue), and CRD2e_K54S (grey); (J) CRD2e_K54V (red), CRD2e_K54Y (blue), and CRD2e_K54E_D62A (grey); (K): CRD2e_K54E_D62K (red), CRD2e_K54E_D62L (blue), and CRD2e_K54E_D62S (grey); (L): CRD2f (red); CRD2de (blue) and CRD2de_K54E (grey); (M): CRD2de_K54D (red), CRD2de_K54I (blue), and CRD2de_K54L (grey); (N): CRD2de_K54S (red), CRD2df (blue), and CRD2ef (grey); (O): CRD2def (red), CRD2(39-73)d (blue), and CRD2(39-73)de (grey); (P): CRD2(39-73)e (red), CRD2(39-73)e_K54E(blue), and CRD2(39-73)e_P39A (grey). Different trajectories are depicted as various color tone.

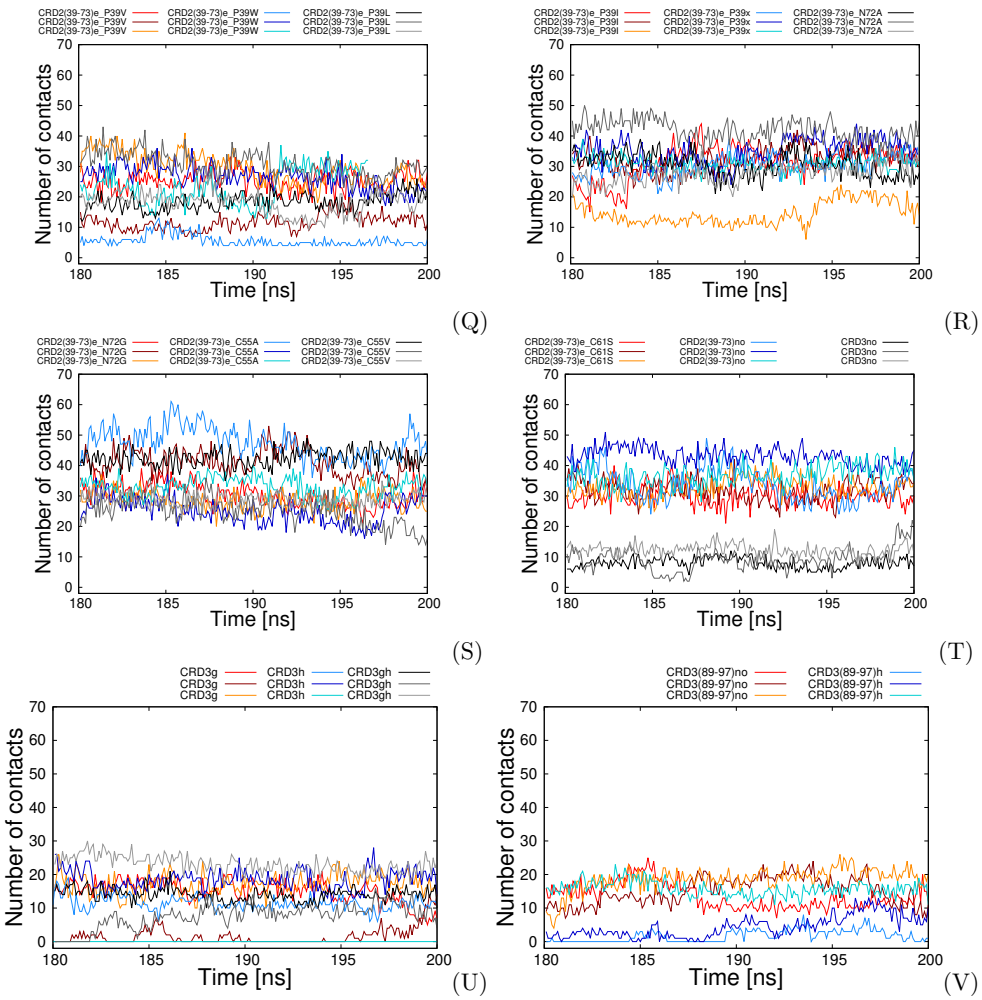


Figure S7: continued (Q): CRD2(39-73)e_P39V(red), CRD2(39-73)e_P39W (blue), and CRD2(39-73)e_P39L (grey); (R): CRD2(39-73)e_P39I (red), CRD2(39-73)e_P39x(blue), and CRD2(39-73)e_N72A (grey); (S): CRD2(39-73)e_N72G(red), CRD2(39-73)e_C55A (blue), and CRD2(39-73)e_C55V (grey); (T): CRD2(39-73)e_C61S (red), CRD2(39-73)no (blue) and CRD3no (grey); (U): CRD3g (red), CRD3h (blue) and CRD3gh(grey); (V): CRD3(89-97)no (red), and CRD3(89-97)h (blue) . Different trajectories are depicted as various color tone.

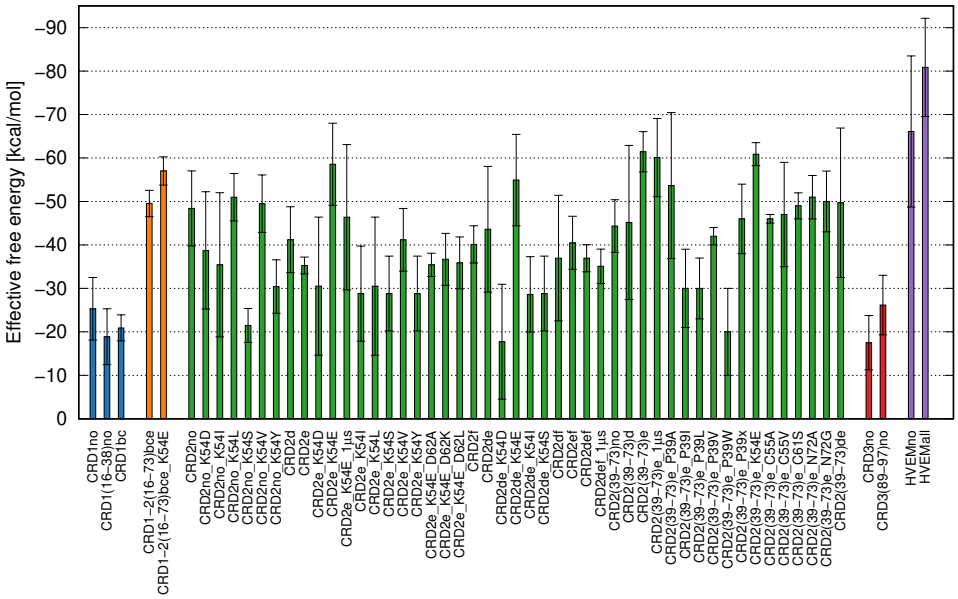


Figure S8: Bar plot of the effective free energies of the various stable variants of the HVEM molecule interacting with the LIGHT trimer.

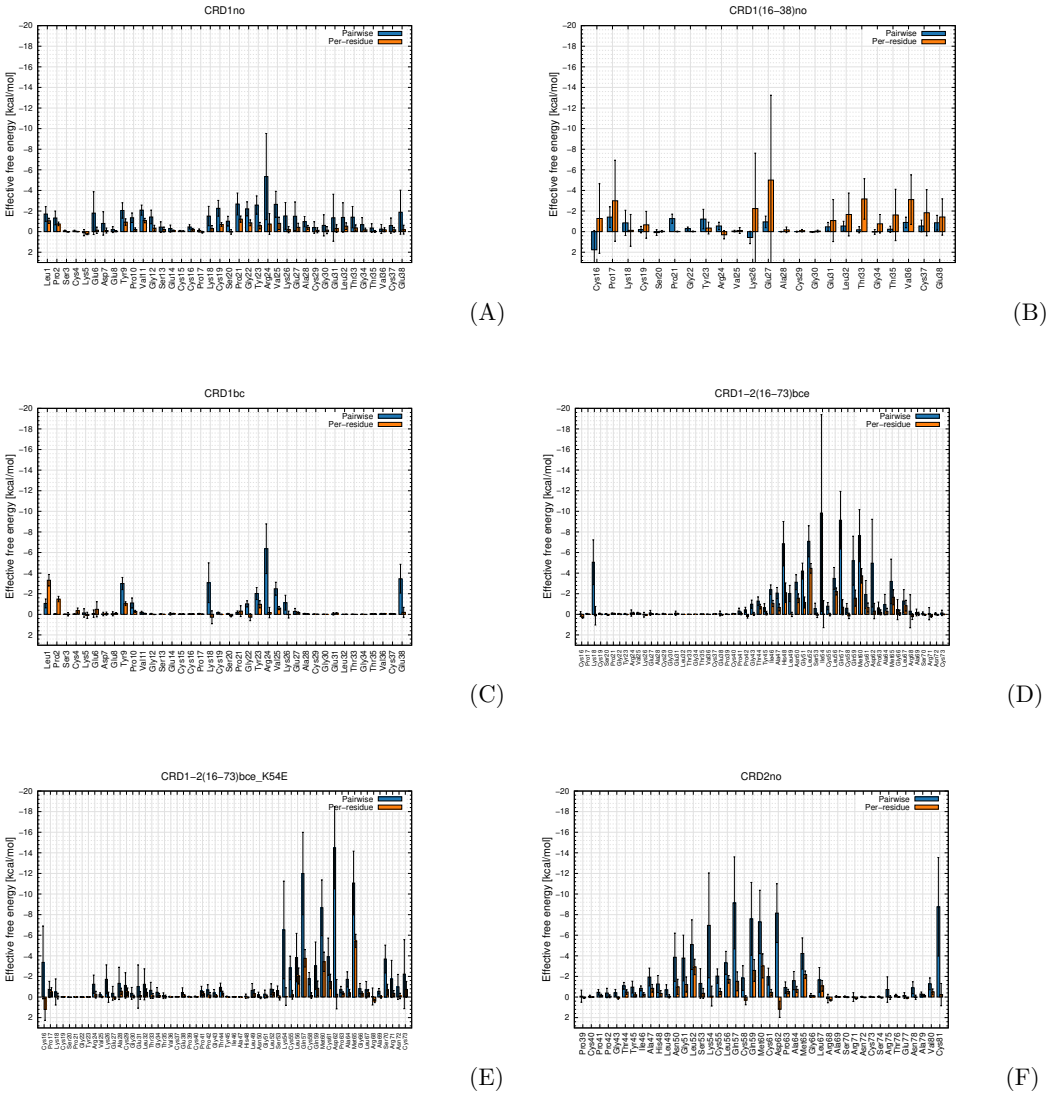


Figure S9: MM-GBSA effective binding energy decomposition results for (A) CRD1no, (B) CRD1(16-38)no, (C) CRD1bc, (D) CRD1-2(16-73)bce, (E) CRD1-2(16-73)bce_K54E, (F) CRD2no variants

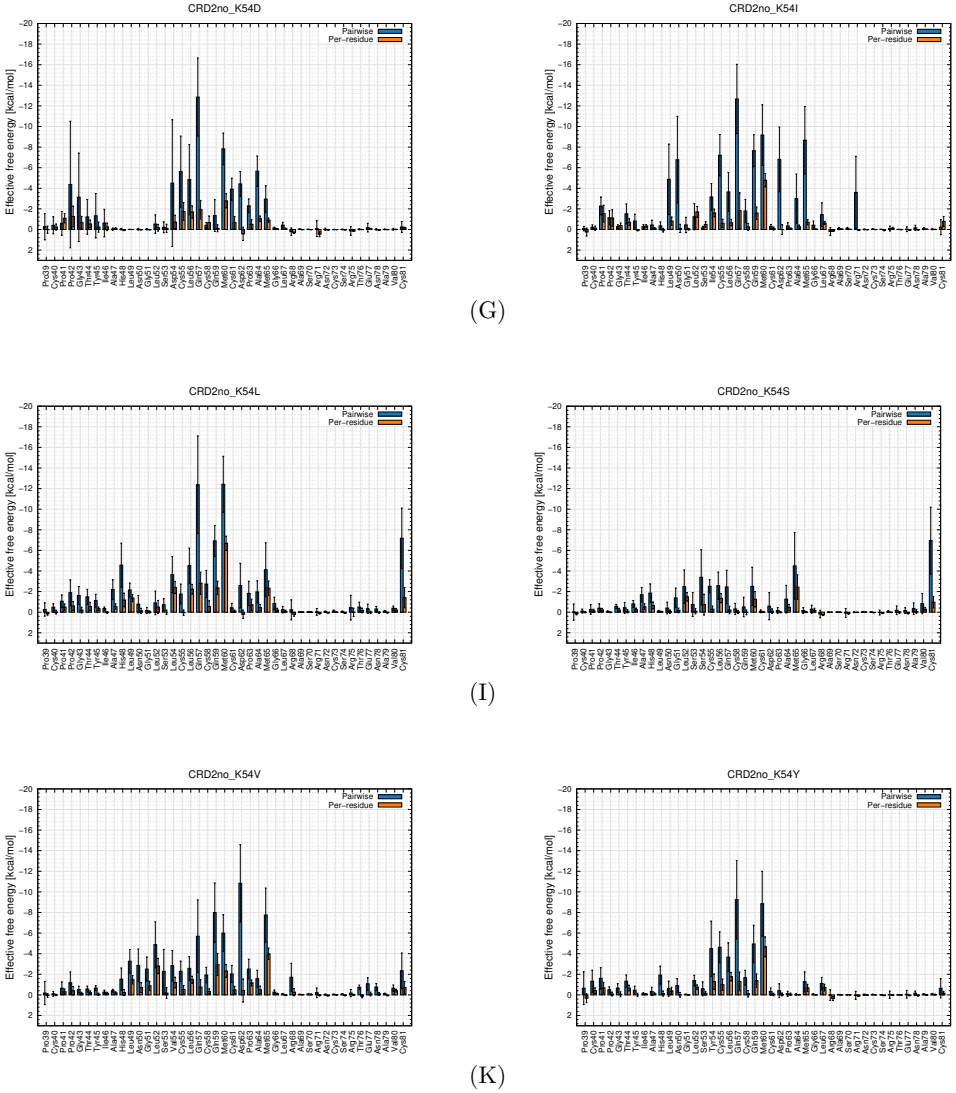


Figure S9: continued: MM-GBSA effective binding energy decomposition results for (G) CRD2no_K54D, (H) CRD2no_K54I, (I) CRD2no_K54L, (J) CRD2no_K54S, (K) CRD2no_K54V, (L) CRD2no_K54Y variants

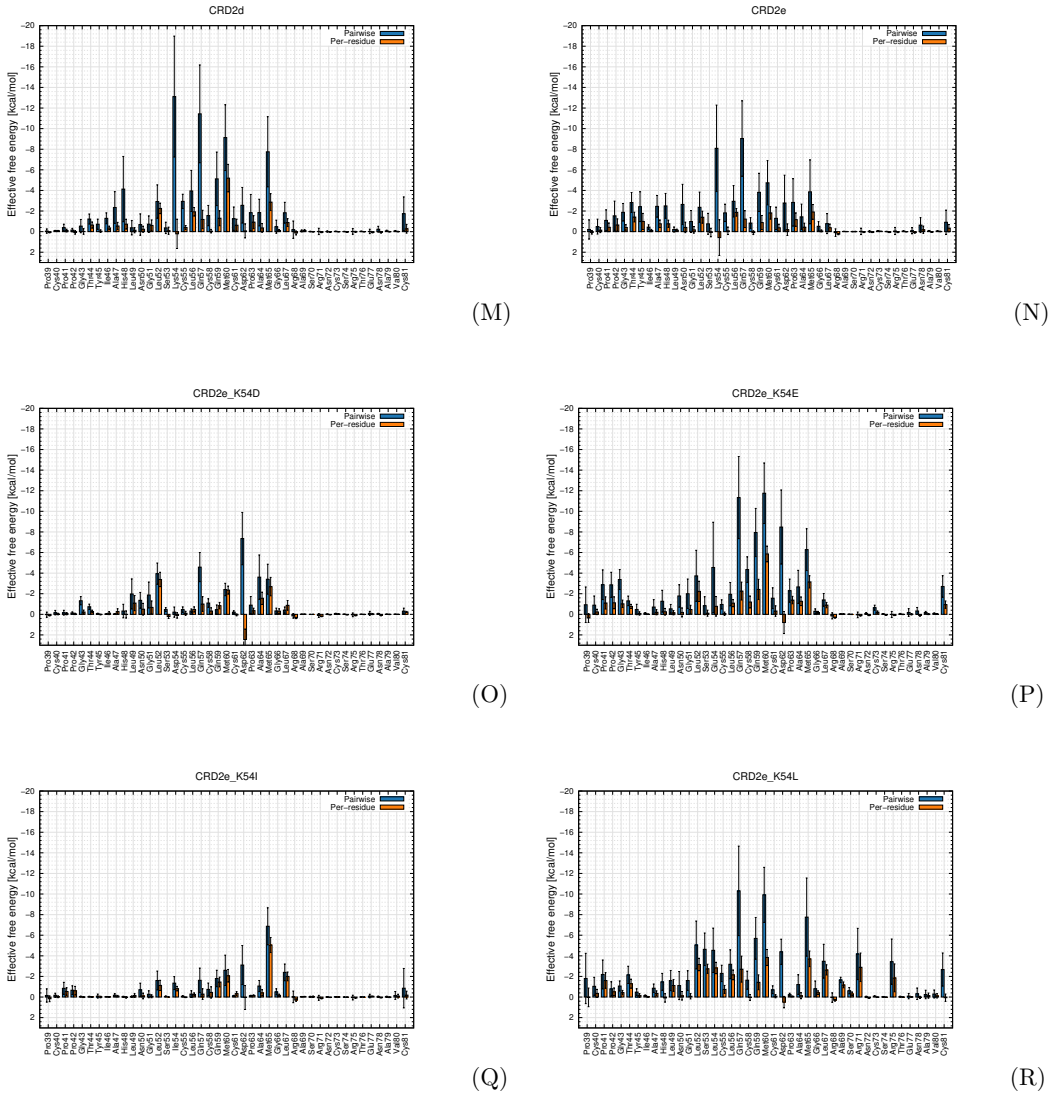


Figure S9: continued: MM-GBSA effective binding energy decomposition results for: (M) CRD2d, (N) CRD2e, (O) CRD2e_K54D, (P) CRD2e_K54E, (Q) CRD2e_K54I, (R) CRD2e_K54L variants

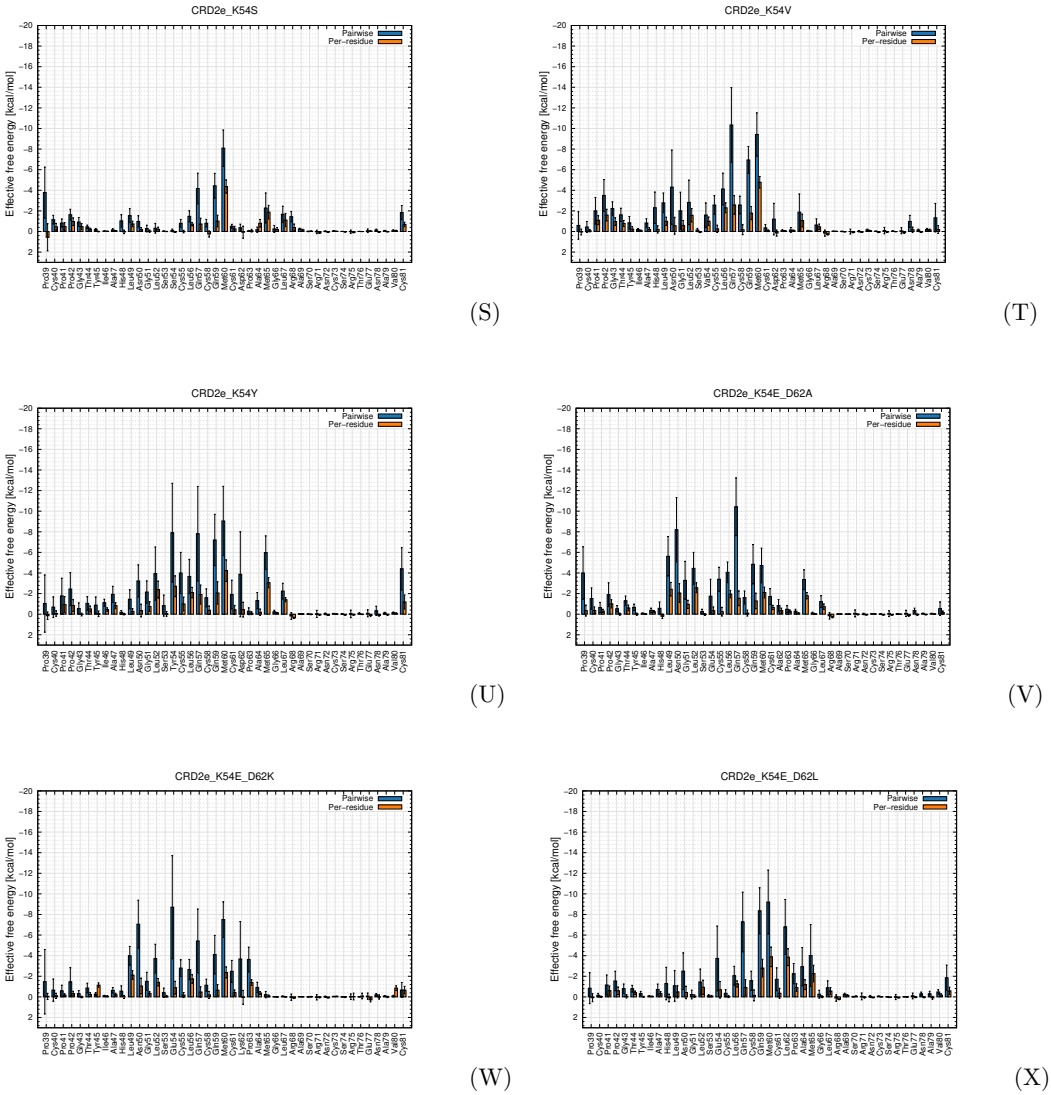


Figure S9: continued: MM-GBSA effective binding energy decomposition results for: (S) CRD2e_K54S, (T) CRD2e_K54V, (U) CRD2e_K54Y, (V) CRD2e_K54E_D62A, (W) CRD2e_K54E_D62K, (X) CRD2e_K54E_D62L variants

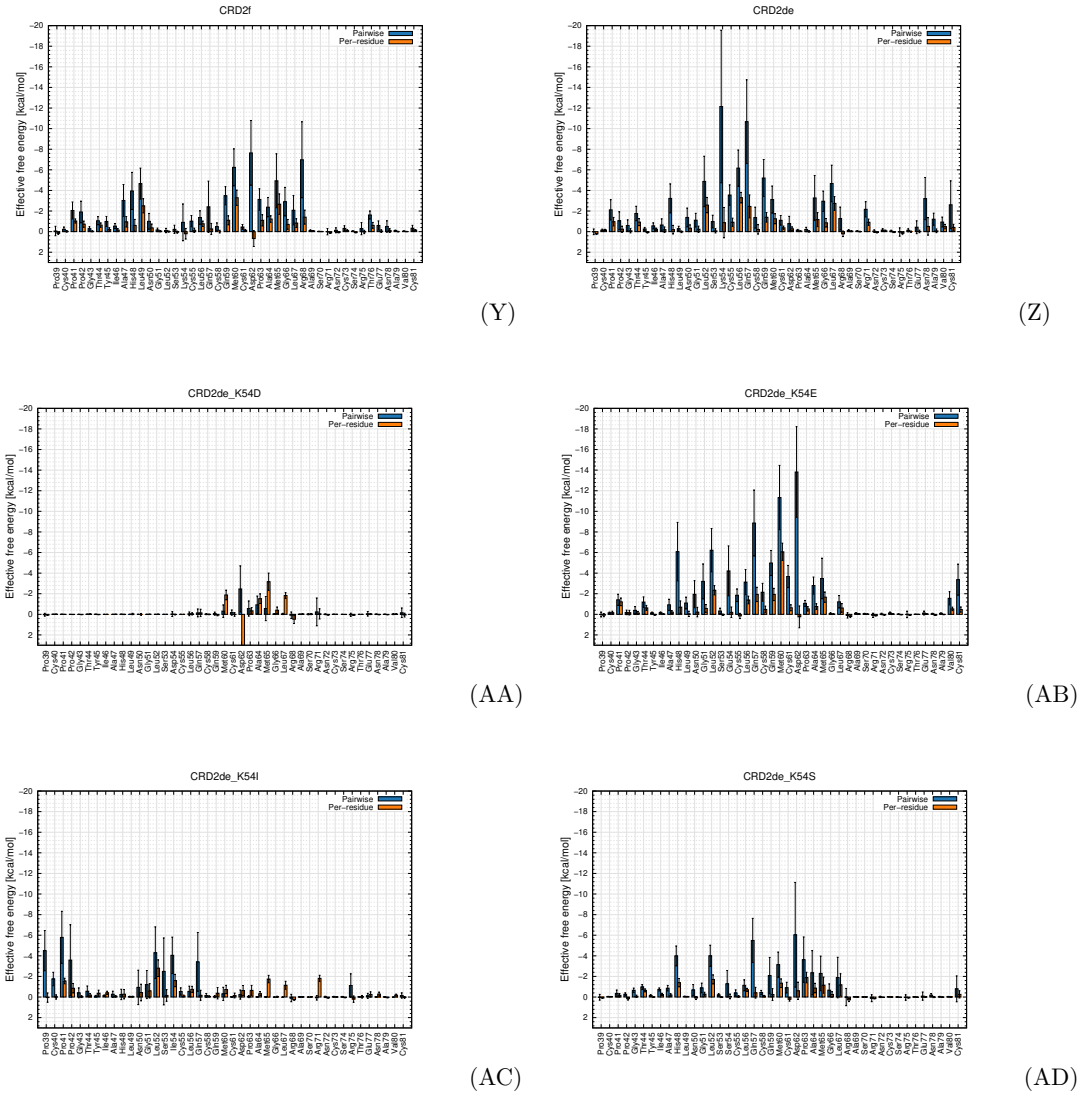
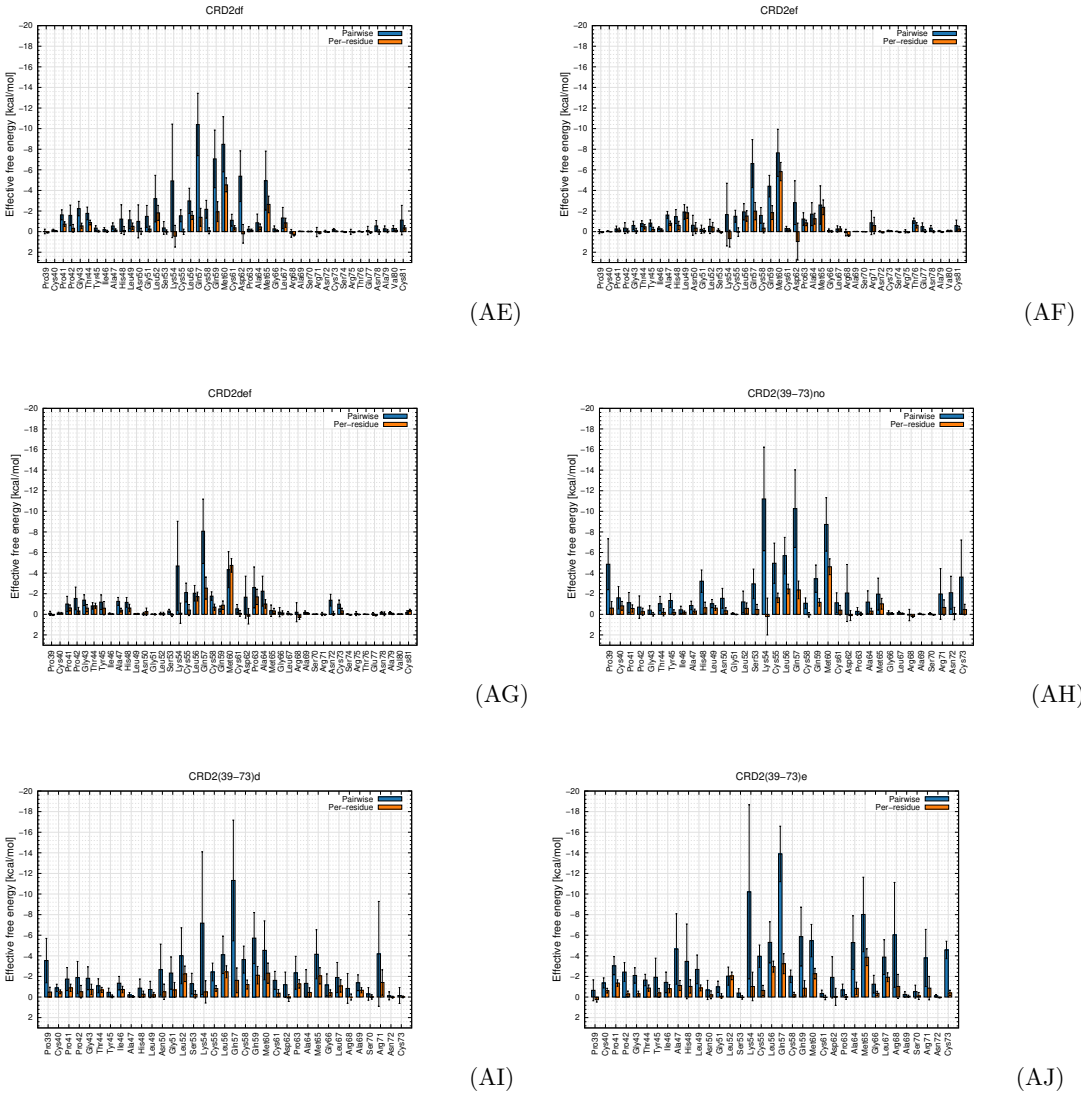


Figure S9: continued: MM-GBSA effective binding energy decomposition results for: (Y) CRD2f, (Z) CRD2de, (AA) CRD2de_K54D, (AB) CRD2de_K54E, (AC) CRD2de_K54I, (AD) CRD2de_K54S variants

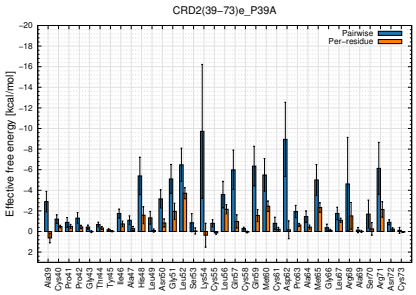


(AE) (AF)

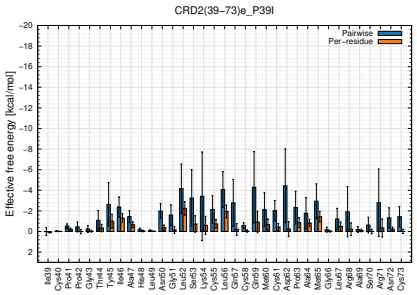
(AG) (AH)

(AI) (AJ)

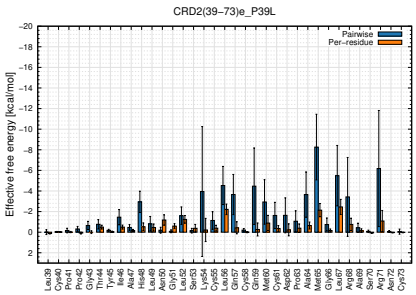
Figure S9: continued: MM-GBSA effective binding energy decomposition results for: (AE) CRD2df, (AF) CRD2ef, (AG) CRD2def, (AH) CRD2(39-73)no, (AI) CRD2(39-73)d, (AJ) CRD2(39-73)e variants



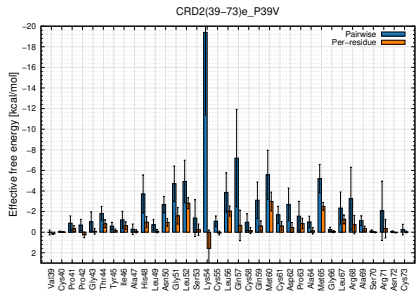
(AK)



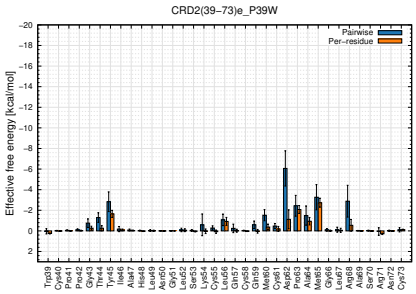
(AL)



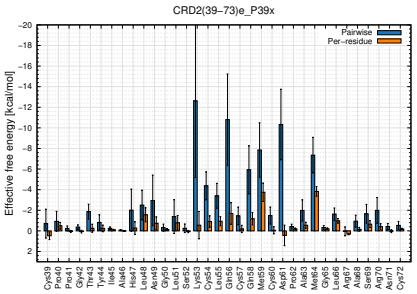
(AM)



(AN)

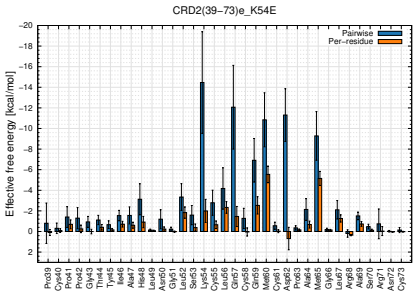


(AO)

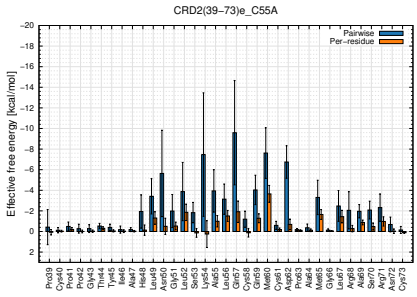


(AP)

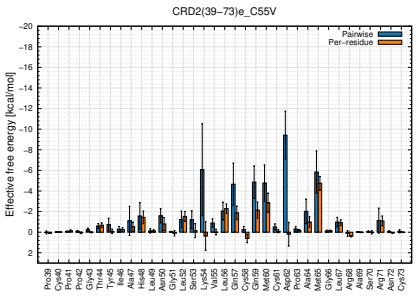
Figure S9: continued: MM-GBSA effective binding energy decomposition results for: (AK) CRD2(39-73)e_P39A, (AL) CRD2(39-73)e_P39I, (AM) CRD2(39-73)e_P39L, (AN) CRD2(39-73)e_P39V, (AO) CRD2(39-73)e_P39W, (AP) CRD2(39-73)e_P39x variants



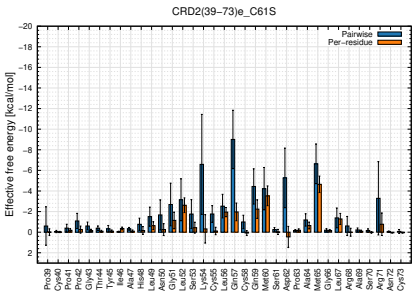
(AQ)



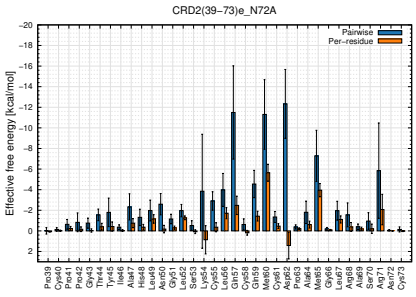
(AR)



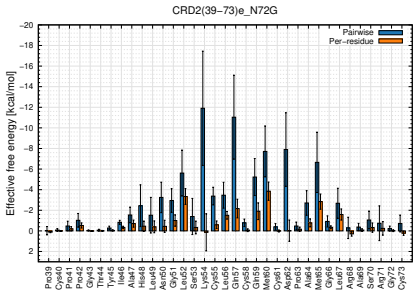
(AS)



(AT)

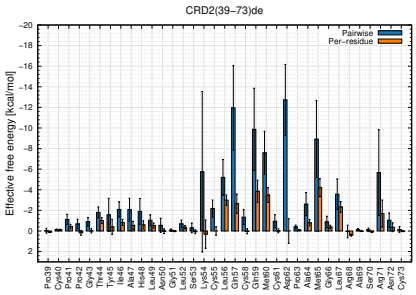


(AU)

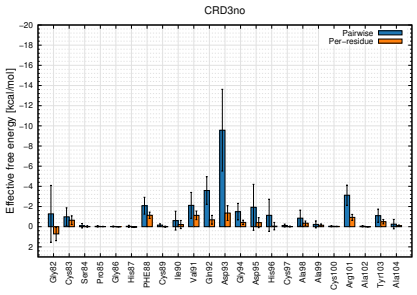


(AV)

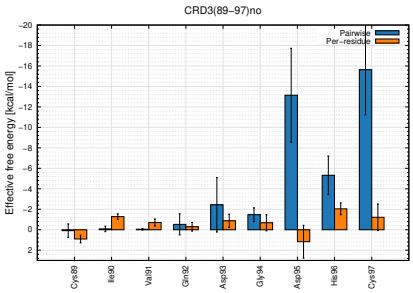
Figure S9: continued: MM-GBSA effective binding energy decomposition results for: (AQ) CRD2(39-73)e_K54E, (AR) CRD2(39-73)e_C55A, (AS) CRD2(39-73)e_C55V, (AT) CRD2(39-73)e_C61S, (AU) CRD2(39-73)e_N72A, (AV) CRD2(39-73)e_N72G variants



(AW)

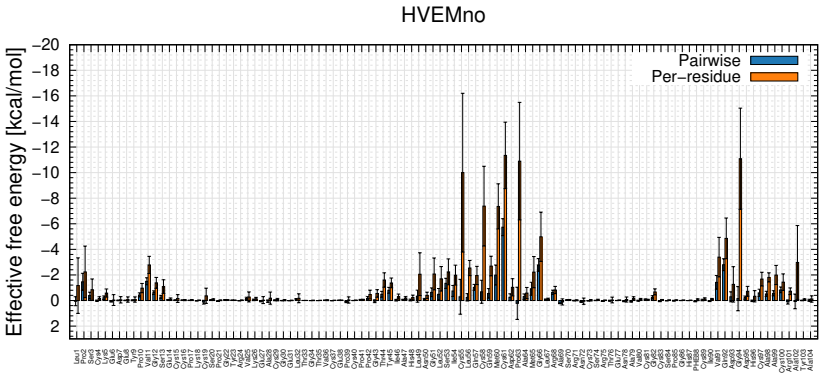


(AX)

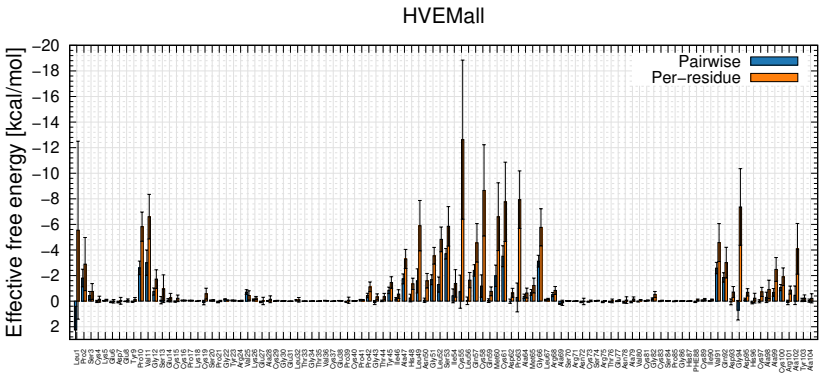


(AY)

Figure S9: continued: MM-GBSA effective binding energy decomposition results for: (AW) CRD2(39-73)de, (AX) CRD3no, (AY) CRD3(39-73)no



(AZ)



(AAA)

Figure S9: continued: MM-GBSA effective binding energy decomposition results for: (AZ) HVEMno variant, (AAA) HVEMall

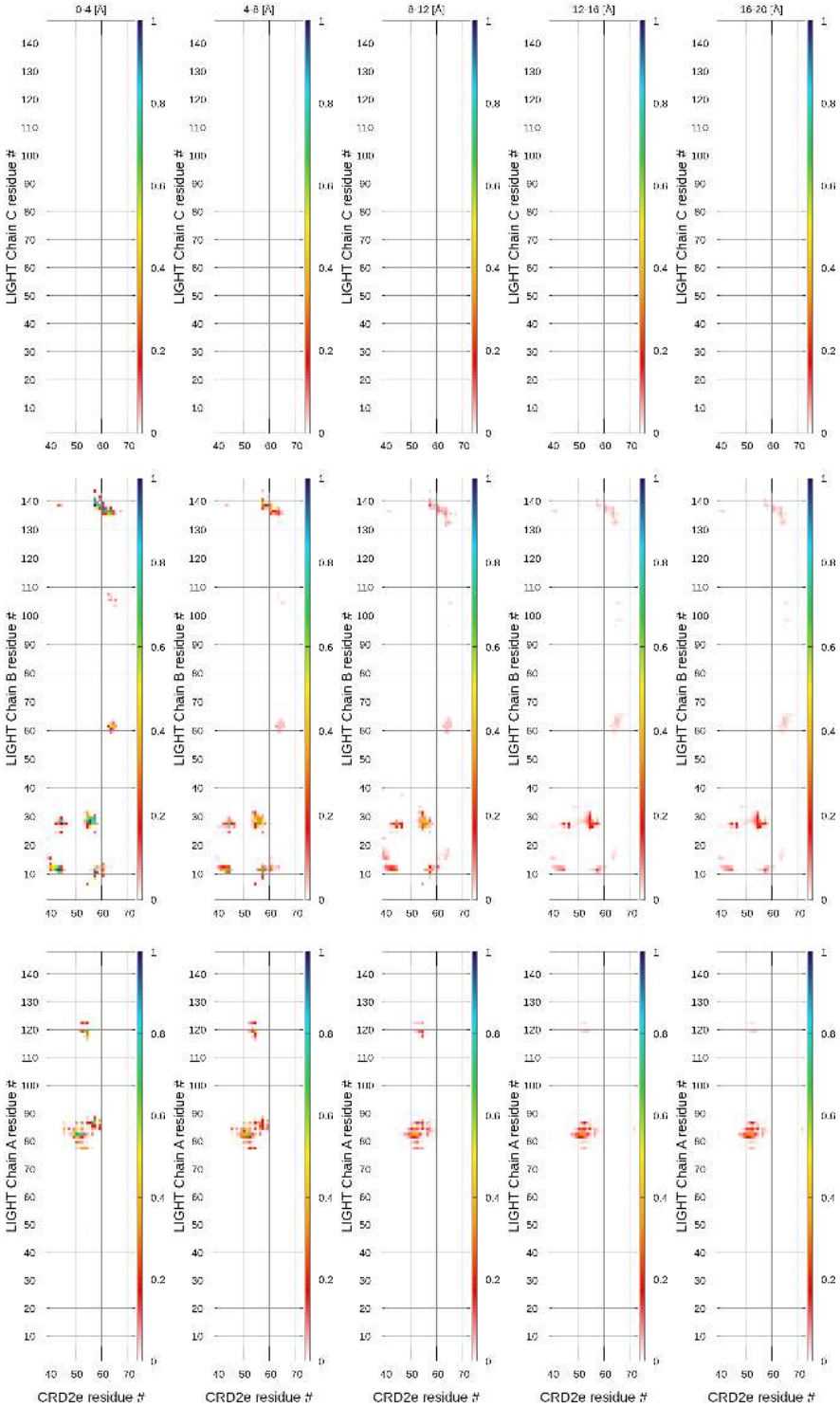


Figure S10: Contact heatmap averaged over 25 SMD simulations for HVEM variant CRD2e with the LIGHT trimer. The values are normalized in the range of increasing extension (relative difference of the distance between centers of mass of HVEM variant and LIGHT trimer) averaged for 4 Å.

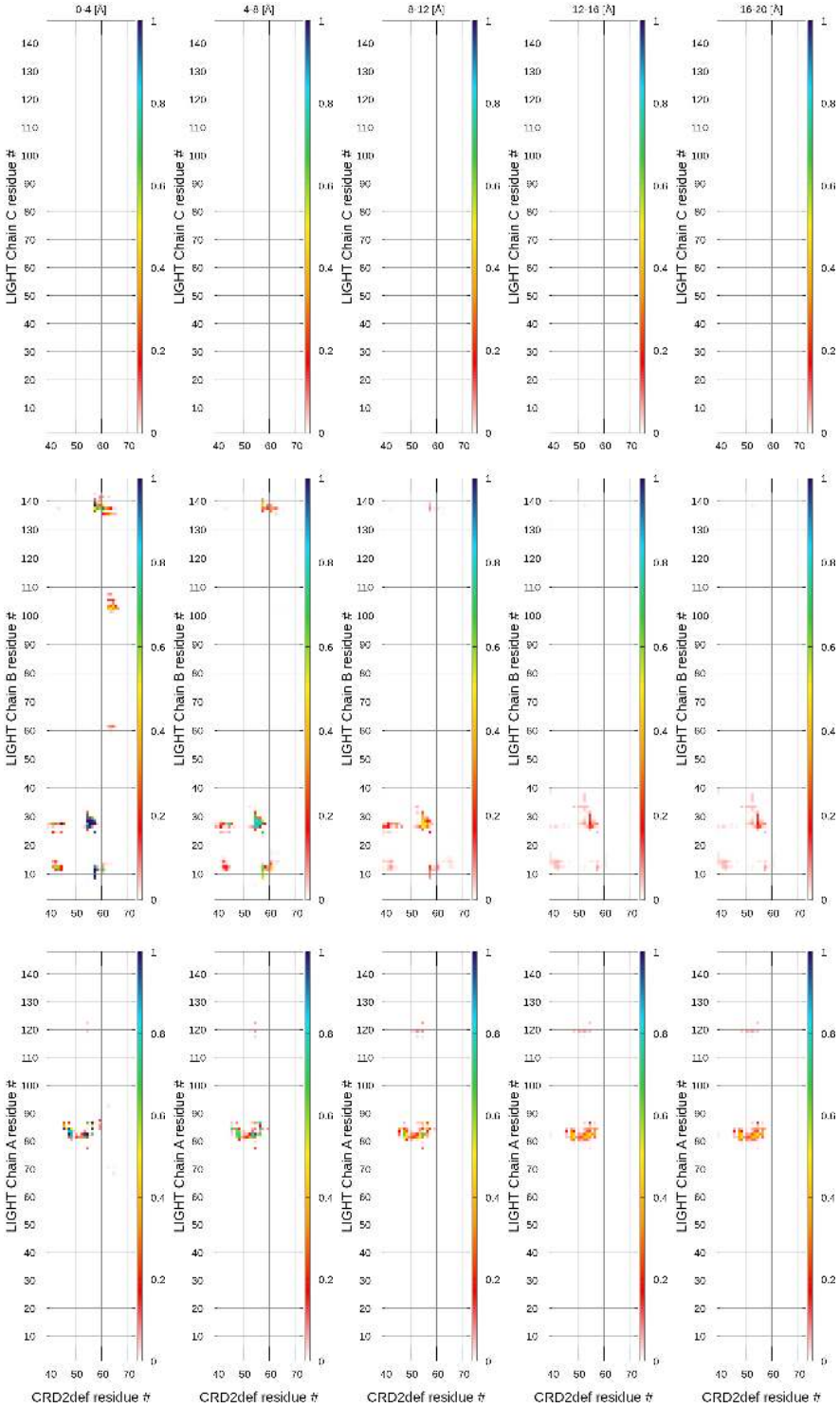


Figure S10: continued: Contact heatmap averaged over 25 SMD simulations for HVEM variant CRD2def with the LIGHT trimer. The values are normalized in the range of increasing extension (relative difference of the distance between centers of mass of HVEM variant and LIGHT trimer) averaged for 4 Å.

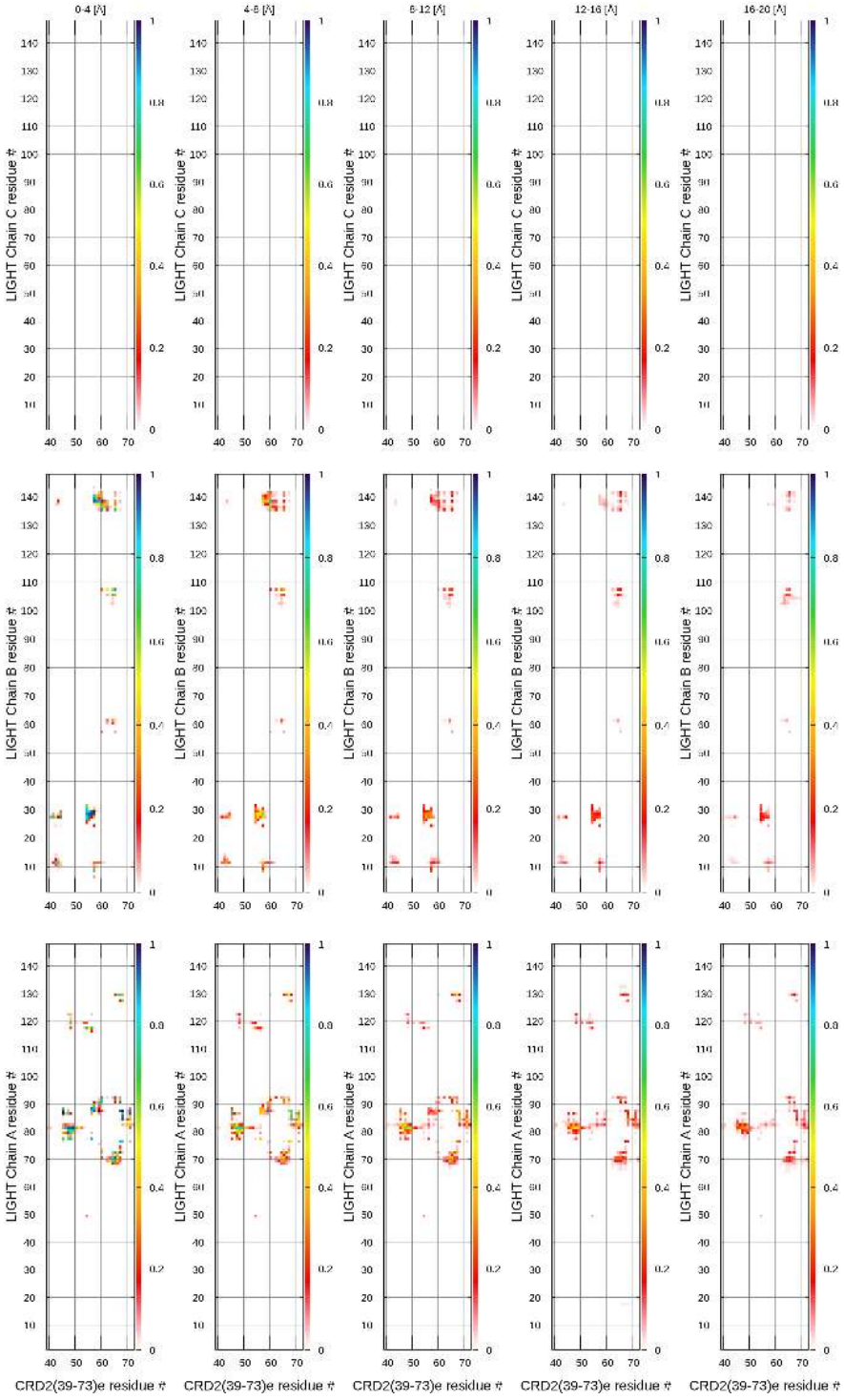


Figure S10: continued: Contact heatmap averaged over 25 SMD simulations for HVEM variant CRD2(39-73)e with the LIGHT trimer. The values are normalized in the range of increasing extension (relative difference of the distance between centers of mass of HVEM variant and LIGHT trimer) averaged for 4 Å.

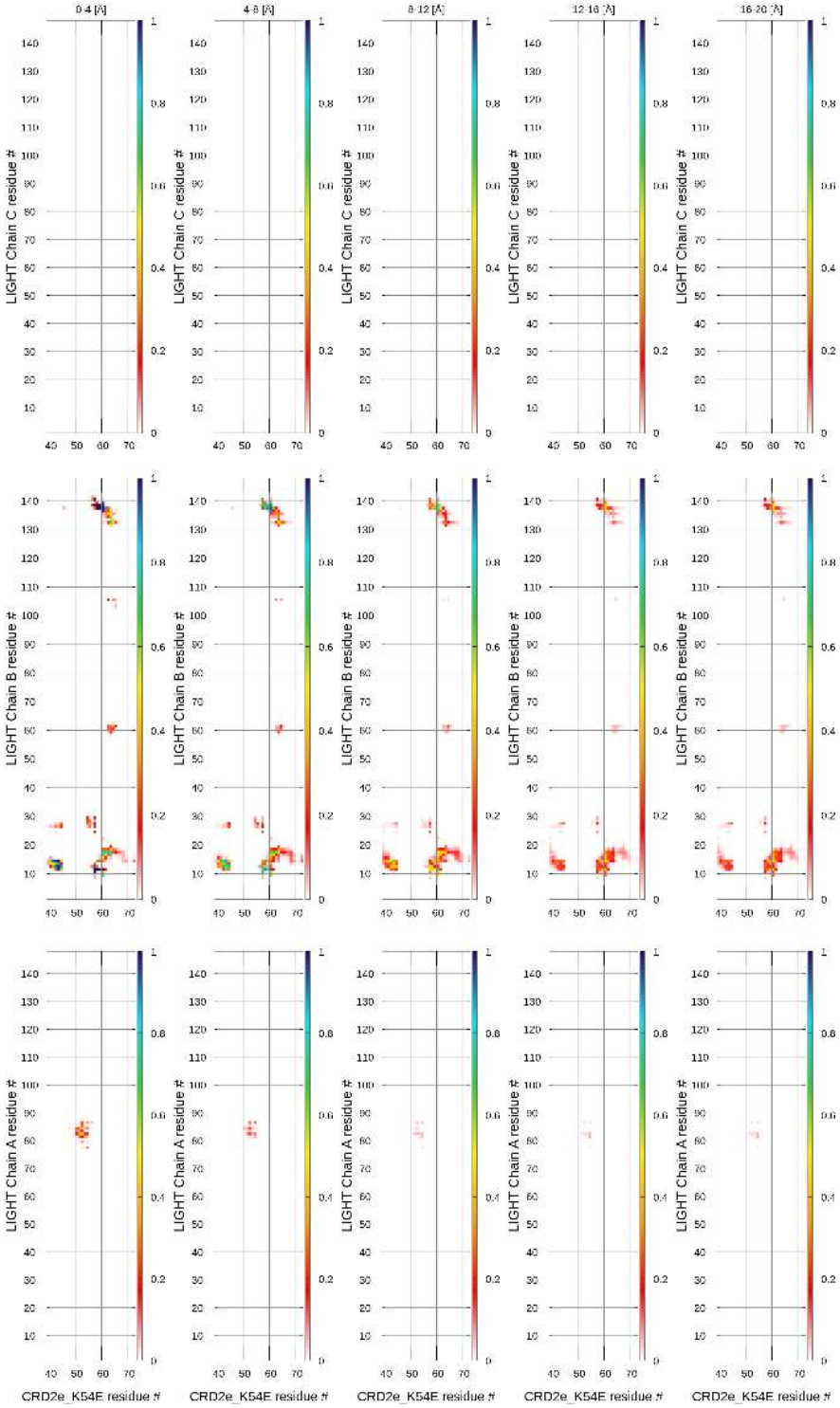


Figure S10: continued: Contact heatmap averaged over 25 SMD simulations for HVEM variant CRD2e_K54E with the LIGHT trimer. The values are normalized in the range of increasing extension (relative difference of the distance between centers of mass of HVEM variant and LIGHT trimer) averaged for 4 Å.

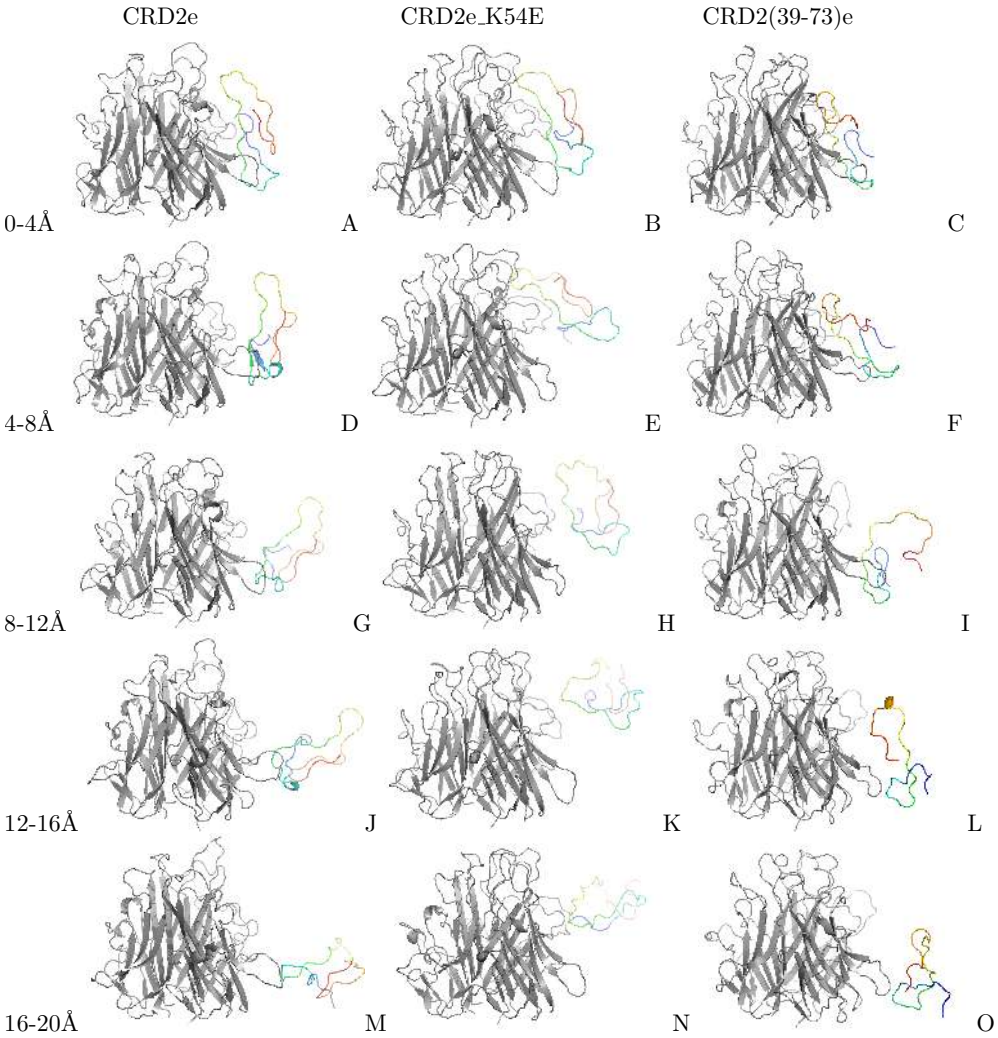


Figure S10: Representative structures upon extending (relative difference of the distance between centers of mass of HVEM variant and LIGHT trimer) averaged for 4 Å.

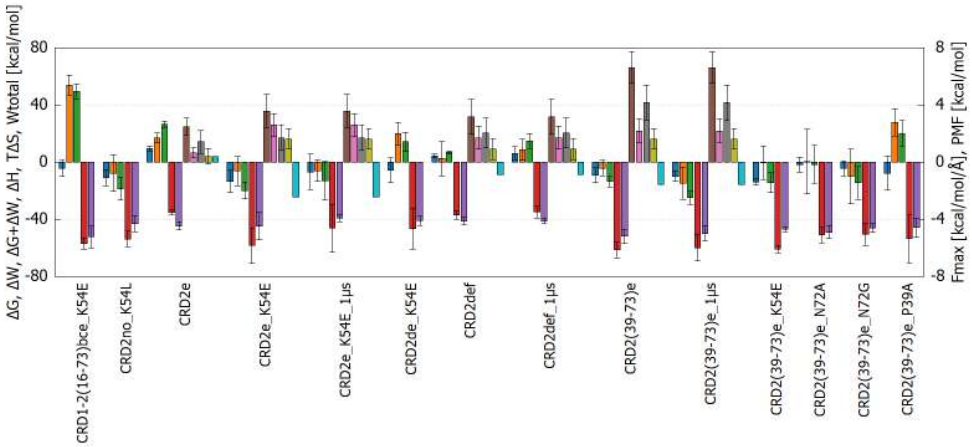


Figure S11: Change of: Binding free energy (ΔG, **blue**), internal work (ΔW, **orange**), summary of Gibbs free energy and internal work (ΔG+ΔW, **green**), enthalpy (ΔH, **red**), entropy (ΔS, **violet**) in 200 ns and 1000 ns long simulations ("1μs" postscript), total work (Wtotal, pulling speed 0.05 m/s **brown**, 0.01 m/s **pink**), force max (Fmax, pulling speed 0.05 m/s **gray**, 0.01 m/s **yellowish green**) and potential mean force (PMF, **cyan**) calculated on SMD basis of eleven selected systems with highest affinity to LIGHT. CRD2e_1μs, due to dissociating, is not shown.

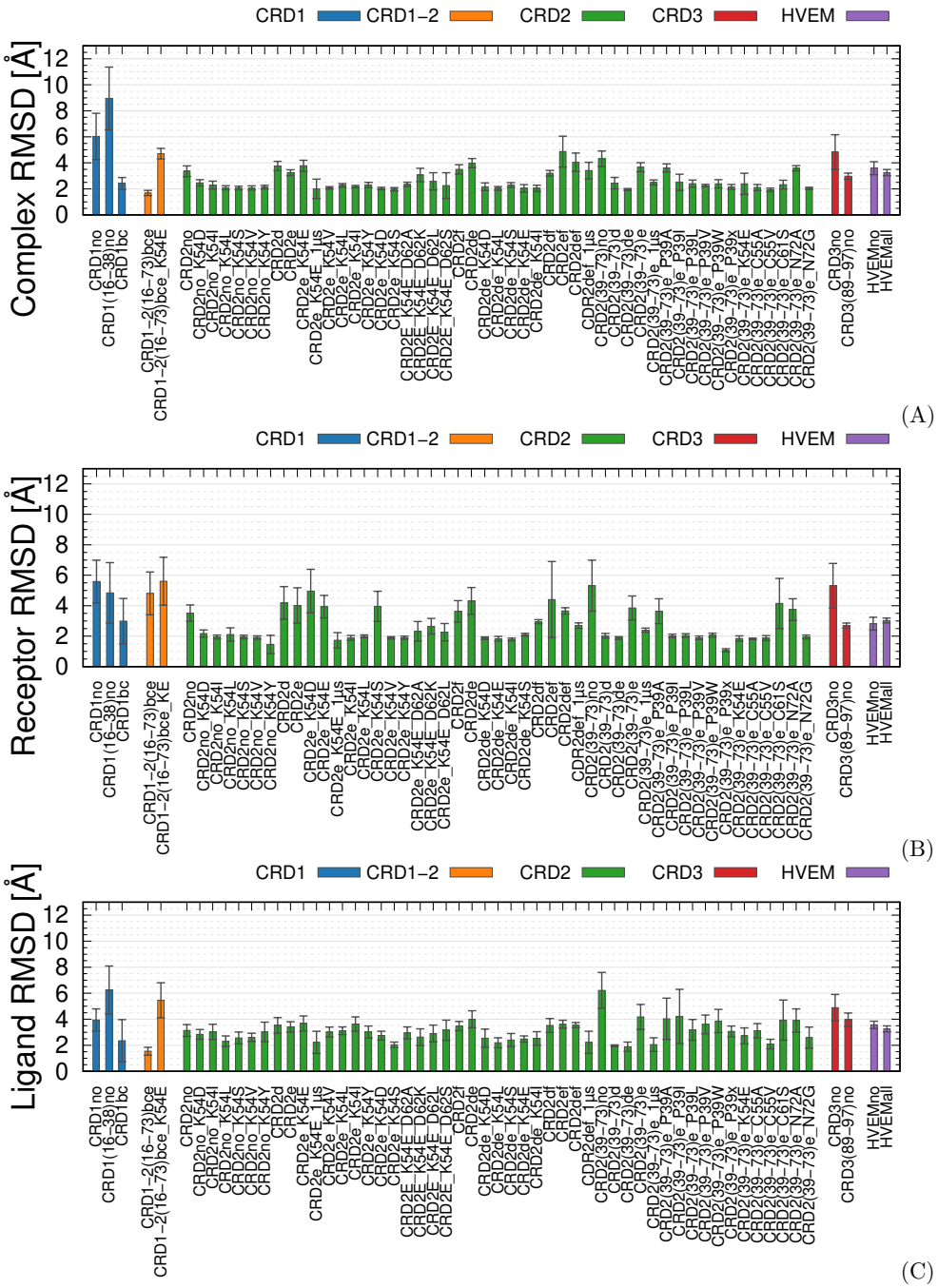


Figure S12: Bar plots of the stable complexes of the structural properties: (A) Complex CαRMSD, (B) Receptor CαRMSD, and (C) Ligand CαRMSD values averaged for 3 trajectories for the last 20 ns of simulation. For the RMSD calculation, an initial conformation adapted from the PDB file was used as a reference.

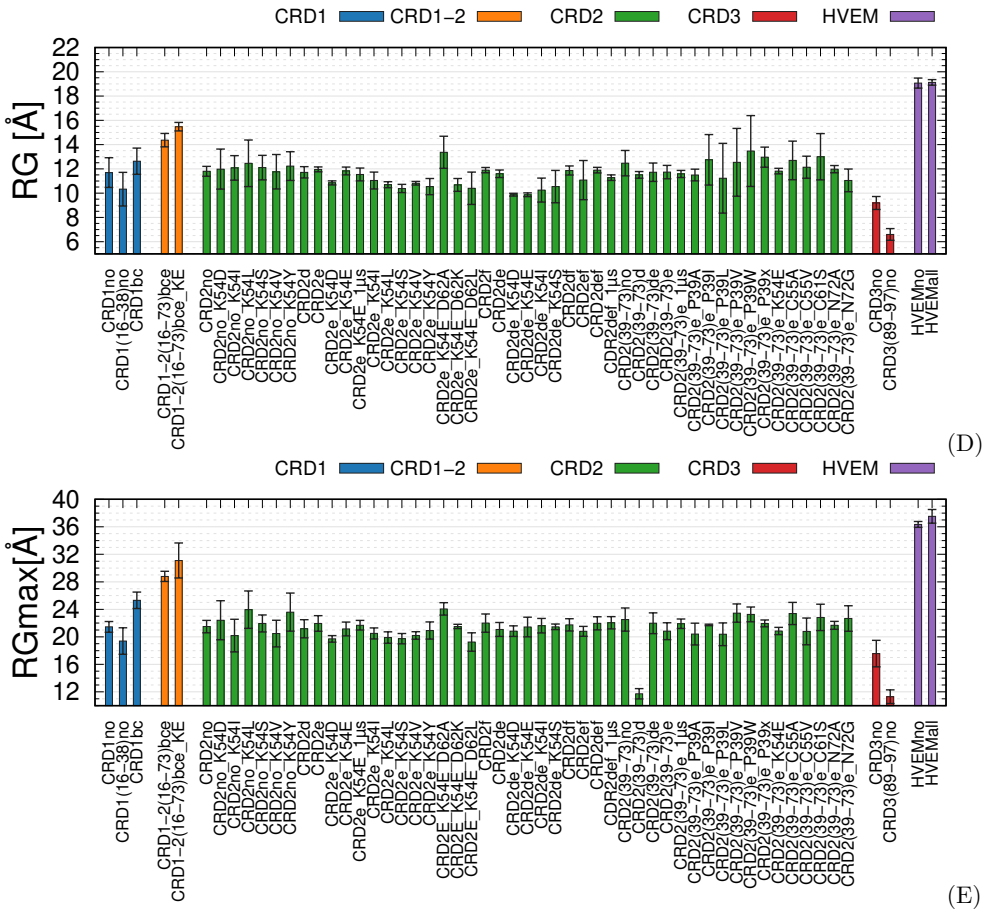


Figure S12: continued: Bar plots of the stable complexes of the structural properties: (D) RG and (E) RGmax values averaged for 3 trajectories for the last 20 ns of simulation. For the RMSD calculation, an initial conformation adapted from the PDB file was used as a reference.

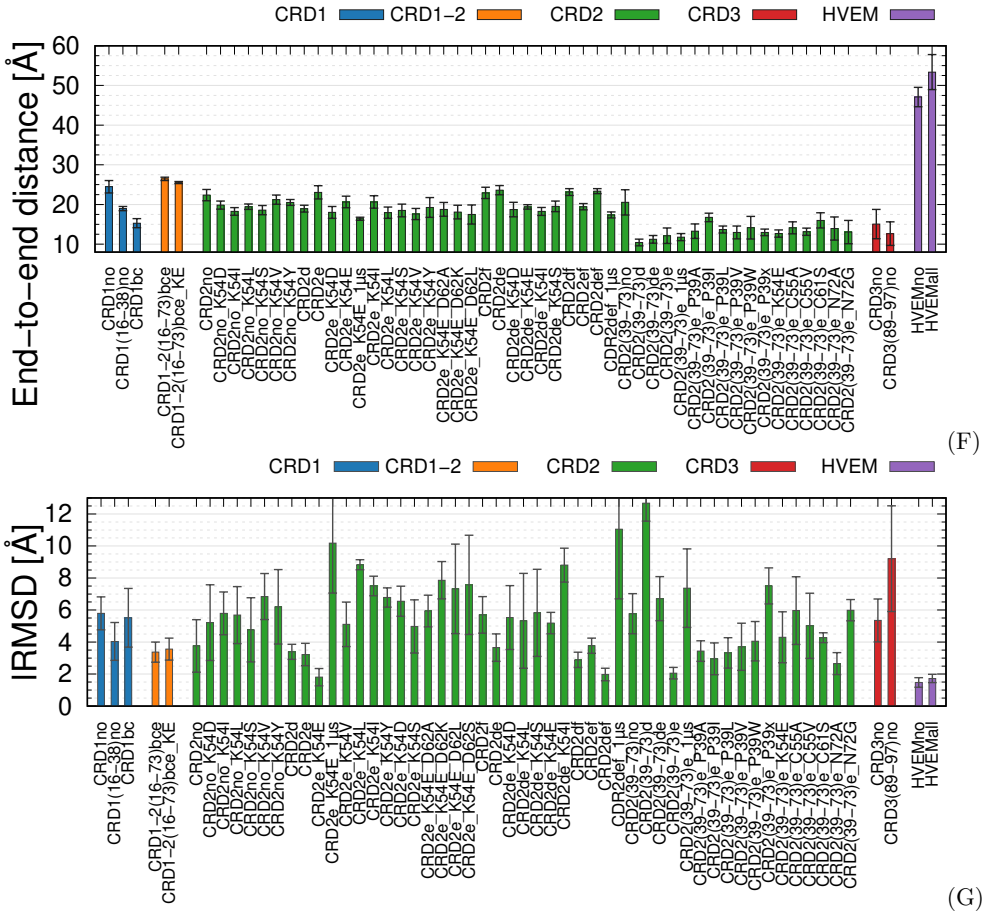


Figure S12: continued: Bar plots of the stable complexes of the structural properties: (F) End-to-end distance and (G) lRMSD values averaged for 3 trajectories for the last 20 ns of simulation. For the RMSD calculation, an initial conformation adapted from the PDB file was used as a reference.

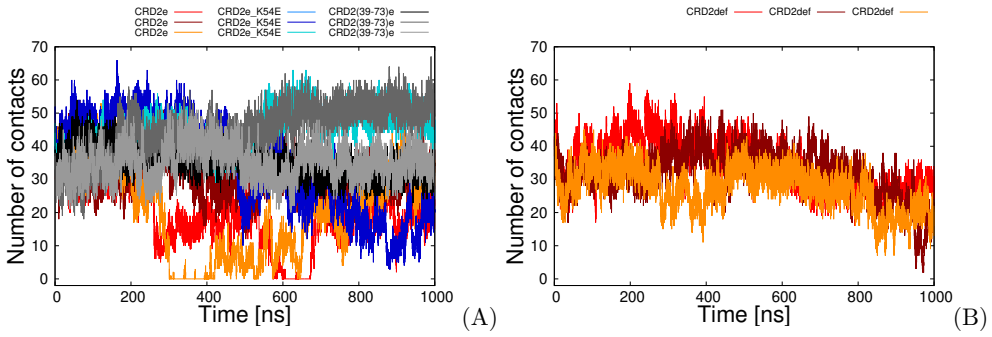
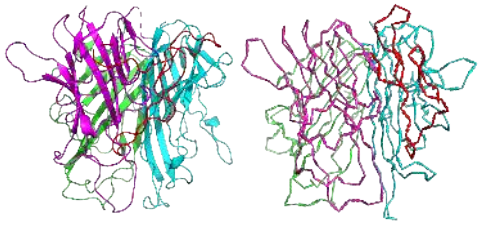
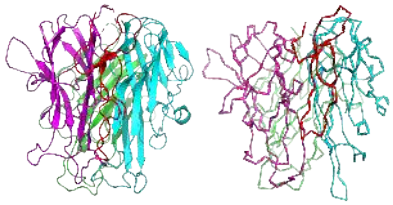


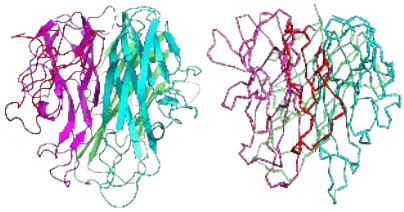
Figure S13: Number of native contacts between HVEM variant and LIGHT trimer as a function of time for 1μs simulations (A): CRD2e (red), CRD2e_K54E (blue), and CRD2(39-73)e (grey); (B): CRD2def (red). Different trajectories are depicted as various color tone.



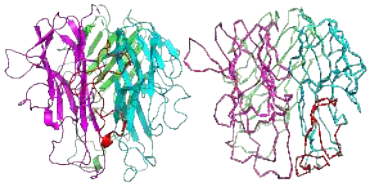
(A)



(B)



(C)



(D)

Figure S14: Comparison between the all-atom dominant structure (left, cartoon representation) with the UNRES-dock cluster with lowest RMSD to respective all-atom structure (right, ribbon representation). Each LIGHT chain is colored green, cyan, magenta and peptide is colored in red for: A) CRD2e (RMSD 4.80Å), B) CRD2def (RMSD 4.39Å), C) CRD2_K54E (RMSD 6.30Å) and D) CRD2(39-73)e (RMSD 4.95Å).

3.3 A Practical Guide to All-Atom and Coarse-Grained Molecular Dynamics Simulations Using Amber and Gromacs: A Case Study of Disulfide-Bond Impact on the Intrinsically Disordered Amyloid Beta

Intrinsically disordered proteins present substantial challenges to conventional experimental techniques due to their pronounced conformational fluctuations and transient structural elements. To address this, we present comprehensive computational protocols for investigating IDPs at various resolutions utilizing the Amber and Gromacs packages.

We focused our case study on the monomeric amyloid-β42 (Aβ42), examining the peptide with and without engineered disulfide bonds (L17C/L34C) to elucidate the structural impact of these constraints on the initial stages of aggregation. To ensure robust sampling of the conformational space, I employed a multi-resolution approach. Specifically, I implemented all-atom simulations using the Amber ff19SB force field coupled with the OPC water model, which is optimized to balance protein–protein and protein–water interactions. For the coarse-grained components, I utilized the Martini 3 force field in Gromacs, incorporating specific scaling of solute–solvent interactions to better represent disordered states, and the SIRAH force field within the Amber ecosystem to evaluate its efficacy against standard all-atom benchmarks. Furthermore, to treat the disulfide bonds dynamically and allow for bond formation and rupture under stress, I applied a novel protocol based on finite distance restraints. This approach provides a significantly more realistic representation of bond behavior compared to conventional static covalent modeling.

In my analysis of the structural ensembles, I found that the choice of force field critically dictates the sampling of the conformational landscape. The average radius of gyration (R_g) for the wild-type peptide was 1.29 ± 0.14 nm in the Amber ff19SB simulations. The Martini 3 model closely mirrored this compactness (average R_g of 1.32 ± 0.01 nm) while sampling a broader energy landscape, effectively capturing the extended nature of the IDP. In contrast,

the SIRAH model produced significantly more compact structures (R_g of 1.13 ± 0.11 nm), suggesting a tendency to overstabilize collapsed states. Across all methods, the introduction of the L17C/L34C disulfide bond consistently generated more compact structures; this effect was most pronounced in the Martini simulations, where the average R_g decreased by approximately 10%.

My analysis of the secondary structure content revealed a paradoxical but highly significant finding. While cross-linking amyloid fibrils via L17C/L34C mutations is experimentally known to expedite aggregation, I observed that the addition of the disulfide bond actually decreased the overall β-strand content by approximately 50% in the Amber ff19SB and SIRAH simulations. However, through hierarchical clustering and RMSD analysis against known fibril structures (such as 2BEG and 2M4J), I demonstrated that the disulfide-bonded variant exhibits enhanced structural similarity to fibrillar loop conformations. Although our simulations were strictly restricted to monomeric sampling, this significant shift of the monomeric conformational ensemble toward fibril-like topologies lowers the entropic barrier required to form aggregation-prone seed structures. This elevated propensity at the monomeric level bridges the gap between single-peptide dynamics and macroscopic behavior, effectively explaining the accelerated aggregation rates observed *in vitro* despite the reduction in native secondary structure.

To rigorously validate these models against experimental observables, I compared calculated chemical shifts with experimental data, finding high agreement across all methods (with correlation coefficients for C_α atoms ranging from 0.981 to 0.992). This confirms that despite the varying degrees of compactness observed, all three force fields provide a fundamentally reasonable representation of the IDP structural ensemble. However, this accuracy comes at vastly different computational costs. In terms of efficiency, coarse-grained models offer a monumental acceleration: the SIRAH approach demonstrated an approximate 60-fold speed-up compared to the all-atom Amber ff19SB simulations when executed on a low-end GPU. Furthermore, the Martini 3 model executed in Gromacs proved even faster, operating roughly 15 times faster than SIRAH in Amber when using comparable CPU hardware. This massive acceleration makes coarse-grained models particularly advantageous for the extended

microsecond-level sampling required to fully characterize the highly fluctuating equilibrium states of IDPs, even if they sacrifice some fine atomistic details, such as precise secondary structure recognition, due to their simplified representations.

We conclude that while Martini 3 offers superior sampling of extended states, the all-atom Amber ff19SB force field provides essential atomistic detail crucial for identifying specific secondary structure elements and transient contacts. This work establishes that integrating the speed of coarse-grained models with the accuracy of all-atom force fields yields the most reliable characterization of IDPs. We provide these adaptable protocols to the broader community to facilitate multiscale investigations into other disordered and structurally complex systems.

3.4 Visualising the Energy Landscape for a Molecular Dynamics Trajectory

Understanding how structural constraints, such as disulfide bonds, and intrinsic disorder shape the conformational space of proteins requires robust analytical frameworks. To address this, we developed an open-source software package–MDDG, which transforms molecular dynamics trajectories into disconnectivity graphs to visualize the sampled energy landscape topology. By applying Savitzky–Golay smoothing to per-frame energy traces, we identified proxy minima and transition states without the need for computationally expensive geometry optimizations. This approach enables the rapid analysis of conformational hierarchies directly from simulation data, making it highly applicable to the diverse biomolecular systems and force fields explored throughout this thesis.

We processed energy time series, potential energy for NVE/NVT ensembles or enthalpy for NPT, using Savitzky–Golay filters to suppress high-frequency thermal fluctuations while preserving slower inter-basin transitions. Local extrema in the smoothed traces served as proxies for stationary points: interior minima satisfied $y_s < y_{s-1}$ and $y_s < y_{s+1}$, with boundary points assessed via cubic polynomial fits. Transition state proxies were assigned to the maxima between consecutive minima. Output files detailed proxy energies, connections, and frame mappings for disconnectivity graph generation using established visualization tools. Key parameters included window length $2m+1$ and polynomial degree $p = 2$, tuned to match conformational timescales via the approximate cutoff $t_c \approx 1/f_c$, where f_c derives from the filter bandwidth. Edge effects were handled by interpolation to avoid artifacts. This workflow successfully processed 10^4–10^5 frames in seconds, remaining ensemble-agnostic and fully compatible with both the all-atom and coarse-grained models utilized in my research.

To directly align this methodological advancement with the primary focus of this thesis, the impact of dynamic disulfide bonds on protein folding and stability, I focused extensively on the hydrophobic seed protein (PDB: 1HYP). This system was simulated using the UNRES coarse-grained force field at 320 K over ∼489 ns (with frames saved every 4.89 ps). Because UNRES

treats dynamic disulfides via a bimodal switching potential rather than fixed covalent topology, the model can explore alternative connectivity arrangements within a single continuous trajectory. By applying Savitzky–Golay smoothing (window length 21, polynomial order 2, effective cutoff 96 ps), I resolved 18,511 proxy minima spanning 162 distinct disulfide topologies. The 15 most frequent disulfide patterns accounted for 68.6% of the sampled minima. The resulting disconnectivity graphs, colored by disulfide connectivities (identified via greedy matching of cysteine side-chain distances below 5.5 Å, excluding adjacent pairs), revealed pronounced funnels governed by conserved bonding networks. Branches in the graph signified large-scale reshuffling events across barriers, explicitly highlighting dynamic disulfide exchange as the dominant organizational principle shaping the protein's energy landscape.

Building upon the investigations of IDPs presented in the previous chapter, I also applied MDDG to the Aβ_{42} monomer (simulated for 200 ns using CHARMM36m). The generated graphs accurately depicted highly disordered, shallow funnels characterized by sparse β-sheet content and rare D23–K28 salt bridges, perfectly capturing the frustrated energy landscape typical of IDPs. To further validate the versatility of the package on complex macromolecular assemblies, we analyzed all-atom simulations of the FGF-2–heparin$_{dp6}$ complex (PDB: 1BFC). AMBER (ff14SB/GLYCAM) exhibited two major funnels with low glycosaminoglycan RMSDs ($\sim$1–4 Å), while CHARMM36m showed three funnels with broader structural exploration. Quenching proxy minima to their true local minima yielded mean RMSDs of 1.24 Å, affirming the high structural fidelity of the proxy states.

In conclusion, MDDG provides an efficient, interpretive summary of MD-sampled landscapes, capturing structural hierarchies without exhaustive stationary point searches. Crucially for this thesis, the 1HYP analysis demonstrates the tool's power to elucidate disulfide-mediated conformational control, successfully separating complex trajectories into distinct structural subfunnels driven by covalent reshuffling. These broader applications confirm the method's versatility across the varying resolutions and interaction potentials utilized in modern structural biology.

Chapter 4

Conclusion

The research presented in this dissertation aimed to redefine the role of disulfide bonds from static structural reinforcements to dynamic, context-dependent regulators of protein biophysics. Through the use of multiscale computational methods, including various all-atom and coarse-grained models, these studies extended the current understanding of how disulfide bonds influence protein structure, stability, mechanics, and function. By investigating four complementary biophysical problems, namely the redox-dependent mechanics of Ribonuclease A, the rational design of HVEM–LIGHT peptide inhibitors, the conformational dynamics of the intrinsically disordered Aβ42 peptide, and the topology of energy landscapes, this work demonstrated that the formation, rupture, and dynamic treatment of disulfide bonds are critical determinants of protein behaviour and pathology.

A central outcome of this thesis is the demonstration that disulfide bonds cannot be interpreted adequately within a purely static framework. Their contribution depends not only on their presence in the native structure, but also on redox conditions, mechanical perturbation, conformational context, and the broader energy landscape in which the protein operates. For this reason, the implementation and validation of computational strategies capable of approximating disulfide-bond dynamics constituted an important methodological advance of this work. Across the studied systems, explicitly moving beyond permanently fixed covalent representations provided a more realistic description of how disulfide bonds regulate protein behaviour.

In the case of Ribonuclease A, the investigation of mechanical stability under varying redox conditions showed that the role of disulfide bonds in resisting unfolding is strictly environment-dependent. By applying a dynamic disulfide-bond potential in the UNRES coarse-grained model and implementing finite-distance restraints in Amber all-atom simulations, the physical

rupture of these bridges under mechanical stress was successfully mimicked, with good agreement with experimental observations published later by an independent research group. The results revealed that in highly oxidising environments, the rupture of specific disulfide bonds correlates directly with force peaks in the extension profile, thereby providing additional protection against unfolding. Under weakly oxidising or reductive conditions, this correlation disappears, indicating that the mechanical integrity of the protein is compromised long before the bonds themselves are physically broken. This finding confirms that representing disulfide bonds as permanent covalent constraints is an inaccurate approximation across a broad range of environmental conditions. The analysis further identified a hierarchy of stability, with Cys65–Cys72 acting as the most stable bond and a key contributor to the structural core, whereas Cys40–Cys95 was found to be the most susceptible to mechanical stress. This distinction suggests that individual disulfide bonds may be preserved for different purposes, including long-term structural support and the maintenance of enzymatic function under transient oxidative stress. The accuracy of the findings describing disulfide-bond stability in RNase A was recently supported by an independent experimental group, demonstrating how these computational predictions help interpret macroscopic topological observations[185].

These observations naturally extend to systems in which conformational adaptability, rather than maximal rigidity, determines function. In the study of HVEM-based peptide inhibitors targeting the HVEM–LIGHT complex, a central goal was to maximise inhibitory effectiveness while minimising pharmaceutical complexity by reducing both peptide size and the number of disulfide bonds required for activity. The results showed that this problem cannot be approached by assuming that increased covalent stabilisation necessarily improves binding. On the contrary, a clear flexibility paradox emerged: for selected HVEM-derived fragments, the absence of specific disulfide bonds, as in the CRD2no variant, enhanced binding by allowing the peptide to open and adjust its geometry during complex formation. This increased conformational freedom facilitated a more effective induced-fit interaction with the LIGHT trimer. Consequently, optimised candidates such as the truncated CRD2(39–73)e and the K54E mutant retained only a single strategically placed disulfide bond while achieving superior selectivity and mechanical interlocking. These results provide a more

nuanced design principle for peptide engineering, showing that the role of disulfide bonds is not simply to rigidify ligands, but to tune the balance between structural preorganisation and adaptive flexibility. This part of the work also enabled a detailed comparison of methods used to estimate binding affinities in equilibrium and transition-state regimes, including MM-GBSA-based approaches, steered molecular dynamics at different pulling speeds, and umbrella sampling as an approximation of the equilibrium limit. Taken together, these analyses showed that different computational approximations capture different aspects of the binding process, and that physiologically meaningful interpretation requires integrating rather than isolating these perspectives. Recent experimental observations using surface binding techniques and cellular assays successfully confirmed the high affinity of these designed HVEM-variants[186, 187], proving that this computational methodology can accurately predict functional peptide behaviour.

The role of disulfide bonds proved to be even more subtle in the context of intrinsically disordered proteins, where there is no single native fold to stabilise. Using the Aβ42 peptide as a model system, this research investigated how a single engineered disulfide bond, L17C/L34C, redirects the conformational ensemble of a disordered monomer. The results provided a mechanistic explanation for experimentally observed differences in aggregation behaviour. It was shown that the introduction of the disulfide bond paradoxically decreases the overall β-sheet content of the monomeric peptide while simultaneously increasing its tendency to adopt fibril-like conformations. This indicates that the additional covalent constraint narrows the conformational landscape and promotes states that are structurally primed for toxic aggregation. In this setting, the disulfide bond does not act as a simple stabiliser, but rather as a conformational filter that reshapes the ensemble toward amyloidogenic pathways. Furthermore, the comparison of results obtained with Amber ff19SB, Martini 3, and SIRAH showed that different levels of representation are informative in different ways. While all-atom models provided the most reliable description of semi-stable and transient structural states, coarse-grained models such as Martini 3, when combined with appropriate water scaling, proved more suitable for sampling the broad and shallow energy landscapes characteristic of true intrinsic disorder.

Beyond the system-specific findings, this thesis also advanced the methodological framework available for studying disulfide-rich and conformationally heterogeneous proteins. An important objective was to reduce the practical difficulty of carrying out such simulations and to improve the interpretability of the resulting data. The published "Practical Guide" yielded both a highly transferable simulation protocol, now broadly adopted across diverse computational disciplines [188–203], and impactful biological insights that subsequent researchers have utilized to further explore systems like amyloid β [204–206]. In parallel, the development of the MDDG software introduced a topological language for analysing the organisation of protein energy landscapes. By mapping molecular dynamics trajectories into hierarchical graphs, this approach made it possible to relate metastable basins and transition barriers to particular disulfide connectivities. In this way, the energy-landscape analysis showed that dynamic disulfide reshuffling can function as a primary organising principle of the folding funnel, rather than as a secondary structural detail. This represents an important conceptual extension of the thesis, because it connects local covalent rearrangements with global features of conformational organisation.

Taken together, the results of this dissertation show that disulfide bonds are not passive structural annotations, but active biophysical elements whose role emerges only when their dynamic, environmental, and topological context is taken into account. Their influence extends across mechanical resistance, molecular recognition, conformational selection, aggregation propensity, and energy-landscape organisation. The multiscale approaches and analytical strategies developed here provide a framework for future studies of disulfide-rich biomolecules, while also showing that each system must be considered individually and that the specific role of each disulfide bond requires case-by-case verification. In this sense, the principal advancement of this dissertation lies not only in the results obtained for the individual systems studied, but also in establishing a more realistic and transferable way of thinking about disulfide bonds in molecular biophysics.

References

[1] Thornton, J. M. Disulphide bridges in globular proteins. *Journal of Molecular Biology* **1981**, *151*, 261–287.

[2] Wetzel, R.; Perry, L. J.; Baase, W. A.; Becktel, W. J. Disulfide bonds and thermal stability in T4 lysozyme. *Proceedings of the National Academy of Sciences* **1988**, *85*, 401–405.

[3] Chinchio, M.; Czaplewski, C.; Liwo, A.; Ołdziej, S.; Scheraga, H. A. Dynamic formation and breaking of disulfide bonds in molecular dynamics simulations with the UNRES force field. *Journal of Chemical Theory and Computation* **2007**, *3*, 1236–1248.

[4] Smardz, P.; Sieradzan, A. K.; Krupa, P. Mechanical Stability of Ribonuclease A Heavily Depends on the Redox Environment. *Journal of Physical Chemistry B* **2022**, *126*, 6240–6249.

[5] Ciura, P.; Smardz, P.; Spodzieja, M.; Sieradzan, A. K.; Krupa, P. Multilayered Computational Framework for Designing Peptide Inhibitors of HVEM-LIGHT Interaction. *Journal of Physical Chemistry B* **2024**, *128*, 6770–6785.

[6] Smardz, P.; Anila, M. M.; Rogowski, P.; Li, M. S.; Różycki, B.; Krupa, P. A Practical Guide to All-Atom and Coarse-Grained Molecular Dynamics Simulations Using Amber and Gromacs: A Case Study of Disulfide-Bond Impact on the Intrinsically Disordered Amyloid Beta. *International Journal of Molecular Sciences* **2024**, *25*, 6698.

[7] Neuman, V.; Wesołowski, P. A.; Bojarski, K. K.; Dewan, D.; Schäffler, M.; Smardz, P.; Wales, D. J. Visualizing the energy landscape for a molecular dynamics trajectory. *The Journal of chemical physics* **2026**, *164*, 44112.

[8] Rose, G. D.; Geselowitz, A. R.; Lesser, G. J.; Lee, R. H.; Zehfus, M. H. Hydrophobicity of Amino Acid Residues in Globular Proteins. *Science* **1985**, *229*, 834–838.

[9] Wright, P. E.; Dyson, H. J. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *Journal of Molecular Biology* **1999**, *293*, 321–331.

[10] Kjaer, L. F.; Ielasi, F. S.; Winbolt, T.; Delaforge, E.; Tengo, M.; Bessa, L. M.; Pérez, L. M.; Erba, E. B.; Bouvignies, G.; Palencia, A.; Jensen, M. R. Hierarchical folding-upon-binding of an intrinsically disordered protein. *Nature Communications 2025 16:1* **2025**, *16*, 11346–.

[11] Sisk, T. R.; Robustelli, P. Folding-upon-binding pathways of an intrinsically disordered protein from a deep Markov state model. *Proceedings of the National Academy of Sciences of the United States of America* **2024**, *121*.

[12] Chothia, C. Hydrophobic bonding and accessible surface area in proteins. *Nature 1974 248:5446* **1974**, *248*, 338–339.

[13] Scheraga, H. A.; Nemethy, G.; Steinberg, I. Z. The Contribution of Hydrophobic Bonds to the Thermal Stability of Protein Conformations. *Journal of Biological Chemistry* **1962**, *237*, 2506–2508.

[14] Sturtevant, J. M. Heat capacity and entropy changes in processes involving proteins (thermodynamics/hydrophobic effect/vibrational modes/conformations). *Biochemistry* **1977**, *74*, 2236–2240.

[15] Anfinsen, C. B. Some observations on the basic principles of design in protein molecules. *Comparative Biochemistry and Physiology* **1962**, *4*, 229–240.

[16] Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **1973**, *181*, 223–230.

[17] Gambardella, G.; Notari, S.; Cavaterra, D.; Iavarone, F.; Castagnola, M.; Bocedi, A.; Ricci, G. The Anfinsen Dogma: Intriguing Details Sixty-Five Years Later. *International Journal of Molecular Sciences* **2022**, *23*, 7759.

[18] Frauenfelder, H.; Sligar, S. G.; Wolynes, P. G. The Energy Landscapes and Motions of Proteins. *Science* **1991**, *254*, 1598–1603.

[19] Leopold, P. E.; Montal, M.; Onuchic, J. N. Protein folding funnels: A kinetic approach to the sequence-structure relationship. *Biophysics* **1992**, *89*, 8721–8725.

[20] Abkevich, V. I.; Gutin, A. M.; Shakhnovich, E. I. Free energy landscape for protein folding kinetics: Intermediates, traps, and multiple pathways in theory and lattice model simulations. *The Journal of Chemical Physics* **1994**, *101*, 6052–6062.

[21] Bryngelson, J. D.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. G. Funnels, pathways, and the energy landscape of protein folding: A synthesis. *Proteins: Structure, Function, and Bioinformatics* **1995**, *21*, 167–195.

[22] Monterroso, B.; Margolin, W.; Boersma, A. J.; Rivas, G.; Poolman, B.; Zorrilla, S. Macromolecular Crowding, Phase Separation, and Homeostasis in the Orchestration of Bacterial Cellular Functions. *Chemical Reviews* **2024**, *124*, 1899–1949.

[23] Boyer, P. D. Spectrophotometric Study of the Reaction of Protein Sulfhydryl Groups with Organic Mercurials. *Journal of the American Chemical Society* **1954**, *76*, 4331–4337.

[24] Anfinsen, C.; Haber, E. Studies on the Reduction and Re-formation of Protein Disulfide Bonds. *Journal of Biological Chemistry* **1961**, *236*, 1361–1363.

[25] Bechtel, T. J.; Weerapana, E. From structure to redox: The diverse functional roles of disulfides and implications in disease. *PROTEOMICS* **2017**, *17*, 1600391.

[26] Astbury, W. X-Ray Studies of Protein Structure. *Nature* **1936**, *137*, 803–805.

[27] Saunders, M.; Wishnia, A.; Kirkwood, J. G. The Nuclear Magnetic Resonance Spectrum of Ribonuclease. *Journal of the American Chemical Society* **1957**, *79*, 3289–3290.

[28] Adrian, M.; Dubochet, J.; Lepault, J.; McDowall, A. W. Cryo-electron microscopy of viruses. *Nature 1984 308:5954* **1984**, *308*, 32–36.

[29] Crystallography: Protein data bank. *Nature New Biology* **1971**, *233*, 223–223.

[30] Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Research* **2000**, *28*, 235–242.

[31] Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature 2021 596:7873* **2021**, *596*, 583–589.

[32] Ahdritz, G. et al. OpenFold: retraining AlphaFold2 yields new insights into its learning mechanisms and capacity for generalization. *Nature methods* **2024**, *21*, 1514.

[33] Zheng, W.; Wuyun, Q.; Li, Y.; Liu, Q.; Zhou, X.; Peng, C.; Zhu, Y.; Freddolino, L.; Zhang, Y. Deep-learning-based single-domain and multidomain protein structure prediction with D-I-TASSER. *Nature Biotechnology 2025 44:4* **2025**, *44*, 641–653.

[34] Selsted, M. E.; Harwig, S. S. Determination of the disulfide array in the human defensin HNP-2. *Journal of Biological Chemistry* **1989**, *264*, 4003–4007.

[35] Shin, D. H.; Lee, J. Y.; Hwang, K. Y.; Kim, K. K.; Suh, S. W. High-resolution crystal structure of the non-specific lipid-transfer protein from maize seedlings. *Structure* **1995**, *3*, 189–199.

[36] McIntosh, M.; Cruz, L. J.; Hunkapiller, M. W.; Gray, W. R.; Olivera, B. M. Isolation and structure of a peptide toxin from the marine snail Conus magus. *Archives of Biochemistry and Biophysics* **1982**, *218*, 329–334.

[37] Sela, M.; White, F. H.; Anfinsen, C. The reductive cleavage of disulfide bonds and its application to problems of protein structure. *Biochimica et biophysica acta* **1959**, *31 2*, 417–26.

[38] Epstein, C.; Anfinsen, C. The reversible reduction of disulfide bonds in trypsin and ribonuclease coupled to carboxymethyl cellulose. *The Journal of biological chemistry* **1962**, *237*, 2175–9.

[39] Wiedemann, C.; Kumar, A.; Láng, A.; Ohlenschläger, O. Cysteines and Disulfide Bonds as Structure-Forming Units: Insights From Different Domains of Life and the Potential for Characterization by NMR. *Frontiers in Chemistry* **2020**, *8*.

[40] Ramanujam, V.; Shen, Y.; Ying, J.; Mobli, M. Residual Dipolar Couplings for Resolving Cysteine Bridges in Disulfide-Rich Peptides. *Frontiers in Chemistry* **2020**, *7*.

[41] Schwartze, T. A.; Morosky, S. A.; Rosato, T. L.; Henrickson, A.; Lin, G.; Hinck, C.; Taylor, A. B.; Olsen, S. K.; Calero, G.; Demeler, B.; Roman, B. L.; Hinck, A. Molecular Basis of Interchain Disulfide Bond Formation in BMP-9 and BMP-10. *Journal of molecular biology* **2025**, *437*, 168935 – 168935.

[42] Markus, G. Electrolytic reduction of the disulfide bonds of insulin. *The Journal of biological chemistry* **1964**, *239*, 4163–70.

[43] Awoonor-Williams, E.; Golosov, A. A.; Hornak, V. Benchmarking In Silico Tools for Cysteine pKa Prediction. *Journal of Chemical Information and Modeling* **2023**, *63*, 2170–2180.

[44] Manteca, A.; Álvaro Alonso-Caballero; Fertin, M.; Poly, S.; Sancho, D. D.; Perez-Jimenez, R. The influence of disulfide bonds on the mechanical stability of proteins is context dependent. *Journal of Biological Chemistry* **2017**, *292*, 13374–13380.

[45] Villanueva, M.; Vostal, L. E.; Cohen, D. N.; Biesbrock, D.; Kuwaye, E. P.; Driver, S. G.; Hart, K. M. Differential effects of disulfide bond formation in TEM-1 versus CTX-M-9 -lactamase. *Protein Science* **2024**, *33*, e4816.

[46] Daopin, S.; Piez, K. A.; Ogawa, Y.; Davies, D. R. Crystal Structure of Transforming Growth Factor-2: an Unusual Fold for the Superfamily. *Science* **1992**, *257*, 369–373.

[47] Wen, Y.; Jiang, N.; Wang, Z.; Xiao, Y. Versatile whey acidic protein four-disulfide core domain proteins: biology and role in diseases. *Frontiers in Cell and Developmental Biology* **2024**, *12*, 1459129.

[48] Betz, S. F. Disulfide bonds and the stability of globular proteins. *Protein Science* **1993**, *2*, 1551–1558.

[49] Vaz, D. C.; Rodrigues, J. R.; Sebald, W.; Dobson, C. M.; ; Rui; Brito, M. M. Enthalpic and entropic contributions mediate the role of disulfide bonds on the conformational stability of Interleukin-4. *Protein Science* **2006**, *15*, 33–44.

[50] Banach, M.; Kalinowska, B.; Konieczny, L.; Roterman, I. Role of Disulfide Bonds in Stabilizing the Conformation of Selected Enzymes—An Approach Based on Divergence Entropy Applied to the Structure of Hydrophobic Core in Proteins. *Entropy 2016, Vol. 18, Page 67* **2016**, *18*, 67.

[51] Shivaprasad, S.; Wetzel, R. An intersheet packing interaction in A fibrils mapped by disulfide cross-linking. *Biochemistry* **2004**, *43*, 15310–15317.

[52] Góngora-Benítez, M.; Tulla-Puche, J.; Albericio, F. Multifaceted roles of disulfide bonds. peptides as therapeutics. *Chemical Reviews* **2014**, *114*, 901–926.

[53] Holzwarth, G.; Doty, P. The Ultraviolet Circular Dichroism of Polypeptides. *Journal of the American Chemical Society* **1965**, *87*, 218–228.

[54] Binnig, G.; Quate, C. F.; Gerber, C. Atomic force microscope. *Physical Review Letters* **1986**, *56*, 930–933.

[55] Ashkin, A.; Dziedzic, J. M. Optical Trapping and Manipulation of Viruses and Bacteria. *Science* **1987**, *235*, 1517–1520.

[56] Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *Journal of Molecular Biology* **1976**, *103*, 227–249.

[57] Levy, A.; Slama, V.; Guilbert, S.; Antalík, A.; Johnson, S. K.; Frisari, G.; Rothlisberger, U. Modeling chemical reactivity in complex systems: Insights from hybrid QM/MM MD simulations. *Journal of Catalysis* **2026**, *453*, 116520.

[58] Fock, V. Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems. *Zeitschrift für Physik* **1930**, *61*, 126–148.

[59] Pople, J. A.; Nesbet, R. K. Self-Consistent Orbitals for Radicals. *The Journal of Chemical Physics* **1954**, *22*, 571–572.

[60] Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Physical Review* **1934**, *46*, 618.

[61] Barr, T. L.; Davidson, E. R. Nature of the Configuration-Interaction Method in Ab Initio Calculations. I. Ne Ground State. *Physical Review A* **1970**, *1*, 644.

[62] Krotscheck, E.; Kmmel, H.; Zabolitzky, J. G. Coupled-cluster many-body theory in a correlated basis. *Physical Review A* **1980**, *22*, 1243.

[63] McWeeny, R.; Sutcliffe, B. T. Fundamentals of Self-Consistent-Field (SCF), Hartree-Fock (HF), Multi-Configuration (MC)SCF and Configuration Interaction (CI) schemes. *Computer Physics Reports* **1985**, *2*, 219–278.

[64] Townsend, J.; Kirkland, J. K.; Vogiatzis, K. D. *Mathematical Physics in Theoretical Chemistry*; Elsevier, 2019; pp 63–117.

[65] Davidson, E. R.; Feller, D. Basis Set Selection for Molecular Calculations. *Chemical Reviews* **1986**, *86*, 681–696.

[66] Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. *Physical Review* **1964**, *136*, B864.

[67] Kohn, W.; Sham, L. J. Self-consistent equations including exchange and correlation effects. *Physical Review* **1965**, *140*, A1133.

[68] Samal, P.; Harbola, M. K. Local-density approximation for the exchange energy functional in excited-state density functional theory. *Journal of Physics B: Atomic, Molecular and Optical Physics* **2005**, *38*, 3765.

[69] Becke, A. D. A new mixing of Hartree–Fock and local density-functional theories. *The Journal of Chemical Physics* **1993**, *98*, 1372–1377.

[70] Dewar, M. J.; Thiel, W. Ground states of molecules. 38. The MNDO method. Approximations and parameters. *Journal of the American Chemical Society* **1977**, *99*, 4899–4907.

[71] Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. AMI: A New General Purpose Quantum Mechanical Molecular Model1. *J. Am. Chem. Soc* **1985**, *107*, 3902–3909.

[72] Stewart, J. J. Optimization of parameters for semiempirical methods I. Method. *Journal of Computational Chemistry* **1989**, *10*, 209–220.

[73] Stewart, J. J. Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *Journal of Molecular Modeling* **2007**, *13*, 1173–1213.

[74] Stewart, J. J. Optimization of parameters for semiempirical methods VI: More modifications to the NDDO approximations and re-optimization of parameters. *Journal of Molecular Modeling* **2013**, *19*, 1–32.

[75] Riniker, S. Fixed-Charge Atomistic Force Fields for Molecular Dynamics Simulations in the Condensed Phase: An Overview. *Journal of Chemical Information and Modeling* **2018**, *58*, 565–578.

[76] Levitt, M.; Lifson, S. Refinement of protein conformations using a macromolecular energy minimization procedure. *Journal of Molecular Biology* **1969**, *46*, 269–279.

[77] Steinbach, P. J.; Brooks, B. R. New spherical-cutoff methods for long-range forces in macromolecular simulation. *Journal of Computational Chemistry* **1994**, *15*, 667–683.

[78] Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* **1993**, *98*, 10089–10092.

[79] Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *The Journal of Chemical Physics* **1995**, *103*, 8577–8593.

[80] Kutzner, C.; Miletić, V.; Karen, .; Rodríguez, P.; Rampp, M.; Hummer, . G.; Groot, B. L. D.; Grubmüller, . H. Scaling of the GROMACS Molecular Dynamics Code to 65k CPU Cores on an HPC Cluster. *Journal of Computational Chemistry* **2025**, *46*, 70059.

[81] Weiner, S. J.; Kollman, P. A.; Singh, U. C.; Case, D. A.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. A New Force Field for Molecular Mechanical Simulation of Nucleic Acids and Proteins. *Journal of the American Chemical Society* **1984**, *106*, 765–784.

[82] MacKerell, A. D. et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *Journal of Physical Chemistry B* **1998**, *102*, 3586–3616.

[83] Jones, J. E. On the determination of molecular fields.—II. From the equation of state of a gas. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **1924**, *106*, 463–477.

[84] Laming, R. XLV. On the primary forces of electricity. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **1838**, *13*, 333–339.

[85] Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D. CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *Journal of Computational Chemistry* **1983**, *4*, 187–217.

[86] Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; Vries, A. H. D. The MARTINI force field: Coarse grained model for biomolecular simulations. *Journal of Physical Chemistry B* **2007**, *111*, 7812–7824.

[87] Fukuda, I.; Nakamura, H. Non-Ewald methods for evaluating the electrostatic interactions of charge systems: similarity and difference. *Biophysical Reviews 2023 14:6* **2023**, *14*, 1315–1340.

[88] Berendsen, H. J.; Postma, J. P.; van Gunsteren, W. F.; Hermans, J. Interaction models for water in relation to protein hydration. Intermolecular forces: proceedings of the fourteenth Jerusalem symposium on quantum chemistry and biochemistry held in jerusalem, israel, april 13–16, 1981. 1981; pp 331–342.

[89] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of chemical physics* **1983**, *79*, 926–935.

[90] Mark, P.; Nilsson, L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *The Journal of Physical Chemistry A* **2001**, *105*, 9954–9960.

[91] Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate- DNA helices. *Journal of the American Chemical Society* **1998**, *120*, 9401–9409.

[92] Massova, I.; Kollman, P. A. Computational alanine scanning to probe protein- protein interactions: a novel approach to evaluate binding free energies. *Journal of the American Chemical Society* **1999**, *121*, 8133–8143.

[93] Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opinion on Drug Discovery* **2015**, *10*, 449–461.

[94] Sun, H.; Duan, L.; Chen, F.; Liu, H.; Wang, Z.; Pan, P.; Zhu, F.; Zhang, J. Z.; Hou, T. Assessing the performance of MM/PBSA and MM/GBSA methods. 7. Entropy effects on the performance of end-point binding free energy calculation approaches. *Physical chemistry chemical physics : PCCP* **2018**, *20*, 14450–14460.

[95] Tuccinardi, T. What is the current value of MM/PBSA and MM/GBSA methods in drug discovery? *Expert Opinion on Drug Discovery* **2021**, *16*, 1233–1237.

[96] Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. A geometric approach to macromolecule-ligand interactions. *Journal of molecular biology* **1982**, *161*, 269–288.

[97] DesJarlais, R. L.; Kuntz, I. D.; Sheridan, R. P.; Venkataraghavan, R.; Dixon, J. S. Docking flexible ligands to macromolecular receptors by molecular shape. *Journal of medicinal chemistry* **1986**, *29*, 2149–2153.

[98] Evans, R. et al. Protein complex prediction with AlphaFold-Multimer. *bioRxiv* **2022**, 2021.10.04.463034.

[99] Johansson-Åkhe, I.; Wallner, B. Improving peptide-protein docking with AlphaFold-Multimer using forced sampling. *Frontiers in Bioinformatics* **2022**, *2*, 959160.

[100] Duin, A. C. V.; Dasgupta, S.; Lorant, F.; Goddard, W. A. ReaxFF: A reactive force field for hydrocarbons. *Journal of Physical Chemistry A* **2001**, *105*, 9396–9409.

[101] Senn, H. M.; Thiel, W.; Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angewandte Chemie International Edition* **2009**, *48*, 1198–1229.

[102] Leven, I.; Hao, H.; Tan, S.; Guan, X.; Penrod, K. A.; Akbarian, D.; Evangelisti, B.; Hossain, M. J.; Islam, M. M.; Koski, J. P.; Moore, S.; Aktulga, H. M.; Duin, A. C. V.; Head-Gordon, T. Recent Advances for Improving the Accuracy, Transferability, and Efficiency of Reactive Force Fields. *Journal of Chemical Theory and Computation* **2021**, *17*, 3237–3251.

[103] Warshel, A.; Weiss, R. M. An empirical valence bond approach for comparing reactions in solutions and in enzymes. *Journal of the American Chemical Society* **1980**, *102*, 6218–6226.

[104] Sit, C. S.; McKay, R. T.; Hill, C.; Ross, R. P.; Vederas, J. C. The 3D structure of Thuricin CD, a two-component bacteriocin with cysteine sulfur to -carbon cross-links. *Journal of the American Chemical Society* **2011**, *133*, 7680–7683.

[105] Levitt, M.; Warshel, A. Computer simulation of protein folding. *Nature 1975 253:5494* **1975**, *253*, 694–698.

[106] Kmiecik, S.; Gront, D.; Kolinski, M.; Wieteska, L.; Dawid, A. E.; Kolinski, A. Coarse-Grained Protein Models and Their Applications. *Chemical Reviews* **2016**, *116*, 7898–7936.

[107] McCammon, J. A.; Gelin, B. R.; Karplus, M. Dynamics of folded proteins. *Nature 1977 267:5612* **1977**, *267*, 585–590.

[108] Singh, B.; Martínez-Noa, Y.; Perez, A. How Well Do Molecular Dynamics Force Fields Model Peptides: A Systematic Benchmark across Diverse Folding Behaviors. *The Journal of Physical Chemistry B* **2026**, *130*, 4344–4357.

[109] Bayly, C. I.; Merz, K. M.; Ferguson, D. M.; Cornell, W. D.; Fox, T.; Caldwell, J. W.; Kollman, P. A.; Cieplak, P.; Gould, I. R.; Spellmeyer, D. C. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *Journal of the American Chemical Society* **1995**, *117*, 5179–5197.

[110] Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *Journal of molecular biology* **1963**, *7*, 95–99.

[111] Weiner, P. K.; Kollman, P. A. AMBER: Assisted model building with energy refinement. A general program for modeling molecules and their interactions. *Journal of Computational Chemistry* **1981**, *2*, 287–303.

[112] Fischer, A. L. M.; Tichy, A.; Kokot, J.; Hoerschinger, V. J.; Wild, R. F.; Riccabona, J. R.; Loeffler, J. R.; Waibl, F.; Quoika, P. K.; Gschwandtner, P.; Forli, S.; Ward, A. B.; Liedl, K. R.; Zacharias, M.; Fernández-Quintero, M. L. The Role of Force Fields and Water Models in Protein Folding and Unfolding Dynamics. *Journal of Chemical Theory and Computation* **2024**, *20*, 2321–2333.

[113] Song, D.; Wang, W.; Ye, W.; Ji, D.; Luo, R.; Chen, H. ff14IDPs force field improving the conformation sampling of intrinsically disordered proteins. *Chemical Biology Drug Design* **2017**, *89*.

[114] Maier, J.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *Journal of chemical theory and computation* **2015**, *11 8*, 3696–713.

[115] Tian, C.; Kasavajhala, K.; Belfon, K.; Raguette, L. E.; Huang, H.; Migues, A. N.; Bickel, J. D.; Wang, Y.; Pincay, J.; Wu, Q.; Simmerling, C. ff19SB: Amino-acid specific protein backbone parameters trained against quantum mechanics energy surfaces in solution. *Journal of chemical theory and computation* **2019**,

[116] Suresh, A.; Schweitzer-Stenner, R.; Urbanc, B. Amber ff24EXP-GA, Based on Empirical Ramachandran Distributions of Glycine and Alanine Residues in Water. *Journal of Chemical Theory and Computation* **2025**, *21*, 2515 – 2534.

[117] Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. Building water models: A different approach. *Journal of Physical Chemistry Letters* **2014**, *5*, 3863–3871.

[118] Ingólfsson, H. I.; Lopez, C. A.; Uusitalo, J. J.; de Jong, D. H.; Gopal, S. M.; Periole, X.; Marrink, S. J. The power of coarse graining in biomolecular simulations. *Wiley Interdisciplinary Reviews. Computational Molecular Science* **2013**, *4*, 225.

[119] Liwo, A.; Ołdziej, S.; Pincus, M. R.; Wawak, R. J.; Rackovsky, S.; Scheraga, H. A. A United-Residue Force Field for off-Lattice Protein-Structure Simulations. I. Functional Forms and Parameters of Long-Range Side-Chain Interaction Potentials from Protein Crystal Data. *Journal of Computational Chemistry* **1997**, *18*, 849–873.

[120] Liwo, A.; Czaplewski, C.; Pillardy, J.; Scheraga, H. A. Cumulant-Based Expressions for the Multibody Terms for the Correlation between Local and Electrostatic Interactions in the United-Residue Force Field. *Journal of Chemical Physics* **2001**, *115*, 2323–2347.

[121] Liwo, A.; He, Y.; Scheraga, H. A. Coarse-Grained Force Field: General Folding Theory. *Physical Chemistry Chemical Physics* **2011**, *13*, 16890–16901.

[122] Liwo, A.; Khalili, M.; Czaplewski, C.; Kalinowski, S.; Ołdziej, S.; Wachucik, K.; Scheraga, H. A. Modification and Optimization of the United-Residue (UNRES) Potential Energy Function for Canonical Simulations. I. Temperature Dependence of the Effective Energy Function and Tests of the Optimization Method. *Journal of Physical Chemistry B* **2007**, *111*, 260–285.

[123] Liwo, A.; Czaplewski, C.; Ołdziej, S.; Rojas, A. V.; Kaźmierkiewicz, R.; Makowski, M.; Murarka, R.; Scheraga, H. A. *Coarse-Graining of Condensed Phase and Biomolecular Systems*; 2008; pp 107–122.

[124] Kozłowska, U.; Liwo, A.; Scheraga, H. A.; Maisuradze, G. Determination of Side-Chain-Rotamer and Side-Chain and Backbone Virtual-Bond-Stretching Potentials of Mean Force from AM1 Energy Surfaces of Terminally-Blocked Amino-Acid Residues, for Coarse-Grained Simulations of Experimental Measurements. *Proceedings of the National Academy of Sciences of the United States of America* **2014**, *111*, 18243–18248, Verify page range; it conflicts with another PNAS entry in this list.

[125] Kozłowska, U.; Liwo, A.; Scheraga, H. A. Determination of Side-Chain-Rotamer and Side-Chain and Backbone Virtual-Bond-Stretching Potentials of Mean Force from AM1 Energy Surfaces of Terminally-Blocked Amino-Acid Residues, for Coarse-Grained Simulations of Protein Structure and Folding. I. The Method. *Journal of Computational Chemistry* **2010**, *31*, 1143–1153.

[126] Sieradzan, A.; Hansmann, U. H. E.; Scheraga, H. A.; Liwo, A. Extension of UNRES Force Field to Treat Polypeptide Chains with D-Amino Acid Residues. *Journal of Chemical Theory and Computation* **2012**, *8*, 4746–4757.

[127] Sieradzan, A. K. Introduction of Periodic Boundary Conditions into UNRES Force Field. *Journal of Computational Chemistry* **2015**, Incomplete in the source text; verify volume, pages, and publication status.

[128] Khalili, M.; Liwo, A.; Jagielska, A.; Scheraga, H. A. Molecular Dynamics with the United-Residue Model of Polypeptide Chains. II. Langevin and Berendsen-Bath Dynamics and Tests on Model Alpha-Helical Systems. *Journal of Physical Chemistry B* **2005**, *109*, 13798–13810.

[129] Kirkwood, J. G. Statistical Mechanics of Fluid Mixtures. *Journal of Chemical Physics* **1935**, *3*, 300–313.

[130] Nanias, M.; Chinchio, M.; Ołdziej, S.; Czaplewski, C.; Scheraga, H. A. Protein Structure Prediction with the UNRES Force-Field Using Replica-Exchange Monte Carlo-with-Minimization; Comparison with MCM, CSA, and CFMC. *Journal of Computational Chemistry* **2005**, *26*, 1472–1486.

[131] Duane, S.; Kennedy, A. D.; Pendleton, B. J.; Roweth, D. Hybrid Monte Carlo. *Physics Letters B* **1987**, *195*, 216–222.

[132] Lee, J.; Liwo, A.; Scheraga, H. A. Energy-Based de Novo Protein Folding by Conformational Space Annealing and an off-Lattice United-Residue Force Field: Application to the 10-55 Fragment of Staphylococcal Protein A and to Apo Calbindin D9K. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 2025–2030.

[133] Sieradzan, A. K.; Niadzvedtski, A.; Scheraga, H. A.; Liwo, A. Revised Backbone-Virtual-Bond-Angle Potentials to Treat the L- and D-Amino Acid Residues in the Coarse-Grained United Residue (UNRES) Force Field. *Journal of Chemical Theory and Computation* **2014**, *10*, 2194–2203.

[134] Maisuradze, G. G.; Zhou, R.; Liwo, A.; Xiao, Y.; Scheraga, H. A. Effects of Mutation, Truncation, and Temperature on the Folding Kinetics of a WW Domain. *Journal of Molecular Biology* **2012**, *420*, 350–365.

[135] Zhou, R.; Maisuradze, G. G.; Suñol, D.; Todorovski, T.; Macias, M. J.; Xiao, Y.; Scheraga, H. A.; Czaplewski, C.; Liwo, A. Folding Kinetics of WW Domains with the United Residue Force Field for Bridging Microscopic Motions and Experimental Measurements. *Proceedings of the National Academy of Sciences of the United States of America* **2014**, *111*, 18243–18248, Verify page range; it conflicts with another PNAS entry in this list.

[136] Rojas, A.; Liwo, A.; Browne, D.; Scheraga, H. A. Mechanism of Fiber Assembly: Treatment of Aβ Peptide Aggregation with a Coarse-Grained United-Residue Force Field. *Journal of Molecular Biology* **2010**, *404*, 537–552.

[137] He, Y.; Liwo, A.; Weinstein, H.; Scheraga, H. A. PDZ Binding to the BAR Domain of PICK1 Is Elucidated by Coarse-Grained Molecular Dynamics. *Journal of Molecular Biology* **2011**, *405*, 298–314.

[138] Gołaś, E.; Maisuradze, G.; Senet, P.; Ołdziej, S.; Czaplewski, C.; Scheraga, H. A.; Liwo, A. Simulation of the Opening and Closing of Hsp70 Chaperones by Coarse-Grained Molecular Dynamics. *Journal of Chemical Theory and Computation* **2012**, *8*, 1750–1764.

[139] Gołaś, E. I.; Czaplewski, C.; Scheraga, H. A.; Liwo, A. Common Functionally Important Motions of the Nucleotide-Binding Domain of Hsp70. *Proteins* **2015**, *83*, 282–299.

[140] Lipska, A. G.; Sieradzan, A. K.; Krupa, P.; Mozolewska, M. A.; D'Auria, S.; Liwo, A. Studies of Conformational Changes of an Arginine-Binding Protein from Thermotoga maritima in the Presence and Absence of Ligand via Molecular Dynamics Simulations with the Coarse-Grained UNRES Force Field. *Journal of Molecular Modeling* **2015**, *21*, 64.

[141] He, Y.; Maciejczyk, M.; Ołdziej, S.; Scheraga, H. A.; Liwo, A. Mean-Field Interactions between Nucleic-Acid-Base Dipoles Can Drive the Formation of a Double Helix. *Physical Review Letters* **2013**, *110*, 098101.

[142] Liwo, A. et al. A Unified Coarse-Grained Model of Biological Macromolecules Based on Mean-Field Multipole-Multipole Interactions. *Journal of Molecular Modeling* **2014**, *20*, 2306.

[143] Sieradzan, A. K.; Sans-Duñó, J.; Lubecka, E. A.; Czaplewski, C.; Lipska, A. G.; Leszczyński, H.; Ocetkiewicz, K. M.; Proficz, J.; Czarnul, P.; Krawczyk, H.; Liwo, A. Optimization of parallel implementation of UNRES package for coarse-grained simulations to treat large proteins. *Journal of computational chemistry* **2023**, *44*, 602–625.

[144] Wesołowski, P. A.; Sieradzan, A. K.; Winnicki, M. J.; Morgan, J. W.; Wales, D. J. Energy landscapes for proteins described by the UNRES coarse-grained potential. *Biophysical Chemistry* **2023**, *303*, 107107.

[145] Darré, L.; Machado, M. R.; Brandner, A. F.; González, H. C.; Ferreira, S.; Pantano, S. SIRAH: A structurally unbiased coarse-grained force field for proteins with aqueous solvation and long-range electrostatics. *Journal of Chemical Theory and Computation* **2015**, *11*, 723–739.

[146] Klein, F.; Soñora, M.; Santos, L. H.; Frigini, E. N.; Ballesteros-Casallas, A.; Machado, M. R.; Pantano, S. The SIRAH force field: A suite for simulations of complex biological systems at the coarse-grained and multiscale levels. *Journal of structural biology* **2023**, *215*.

[147] Marrink, S. J.; Vries, A. H. D.; Mark, A. E. Coarse Grained Model for Semiquantitative Lipid Simulations. *Journal of Physical Chemistry B* **2004**, *108*, 750–760.

[148] Souza, P. C. et al. Martini 3: a general purpose force field for coarse-grained molecular dynamics. *Nature Methods 2021 18:4* **2021**, *18*, 382–388.

[149] Wang, L.; Brasnett, C.; Borges-Araújo, L.; Souza, P. C.; Marrink, S. J. Martini3-IDP: improved Martini 3 force field for disordered proteins. *Nature Communications 2025 16:1* **2025**, *16*, 2874–.

[150] Frenkel, D. *Understanding Molecular Simulation: From Algorithms to Applications*; Academic Press, 2002; pp 63–108.

[151] Verlet, L. Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Physical Review* **1967**, *159*, 98–103.

[152] Leach, A. R. *Molecular Modelling: Principles and Applications*; Pearson Education, 2001; pp 353–409.

[153] Swope, W. C. A Computer Simulation Method for the Calculation of Equilibrium Constants for the Formation of Physical Clusters of Molecules: Application to Small Water Clusters. *Journal of Chemical Physics* **1982**, *76*, 637, Verify end page.

[154] Scully, J. L.; Hermans, J. Multiple Time Steps: Limits on the Speedup of Molecular Dynamics Simulations of Aqueous Systems. *Molecular Simulation* **1993**, *11*, 67–77.

[155] Liwo, A.; Pillardy, J.; Czaplewski, C. UNRES: A United-Residue Force Field for Energy-Based Prediction of Protein Structure—Origin and Significance of Multibody Terms. Proceedings of the Fourth Annual International Conference on Computational Molecular Biology (RECOMB 2000). 2000; pp 193–200.

[156] Berendsen, H. J.; Postma, J. P.; Gunsteren, W. F. V.; Dinola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* **1984**, *81*, 3684–3690.

[157] Rosta, E.; Buchete, N.-V.; Hummer, G. Thermostat Artifacts in Replica Exchange Molecular Dynamics Simulations. *Journal of Chemical Theory and Computation* **2009**, *5*, 1393–1399.

[158] Adelman, S. A. Generalized Langevin Theory for Many-Body Problems in Chemical Dynamics: General Formulation and the Equivalent Harmonic Chain Representation. *Journal of Chemical Physics* **1979**, *71*, 4471, Verify end page.

[159] Berendsen, H. J. C. *Simulating the Physical World: Hierarchical Modeling from Quantum Mechanics to Fluid Dynamics*; Cambridge University Press, 2007; pp 139–210.

[160] Parrinello, M.; Rahman, A. Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. *Physical Review Letters* **1980**, *45*, 1196.

[161] Jensen, M. R.; Zweckstetter, M.; Huang, J. R.; Blackledge, M. Exploring free-energy landscapes of intrinsically disordered proteins at atomic resolution using NMR spectroscopy. *Chemical Reviews* **2014**, *114*, 6632–6660.

[162] Hansmann, U. H.; Okamoto, Y. Prediction of peptide conformation by multicanonical algorithm: New approach to the multiple-minima problem. *Journal of Computational Chemistry* **1993**, *14*, 1333–1338.

[163] Nanias, M.; Czaplewski, C.; Scheraga, H. A. Replica Exchange and Multicanonical Algorithms with the coarse-grained UNRES force field. *Journal of chemical theory and computation* **2006**, *2*, 513–528.

[164] Hamelberg, D.; Mongan, J.; McCammon, J. A. Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. *The Journal of chemical physics* **2004**, *120*, 11919–11929.

[165] Barducci, A.; Bussi, G.; Parrinello, M. Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Physical Review Letters* **2008**, *100*, 020603.

[166] Young, M. R.; Pande, V. S. Multiplexed-replica exchange molecular dynamics method for protein folding simulation. *Biophysical journal* **2003**, *84*, 775–786.

[167] Czaplewski, C.; Kalinowski, S.; Liwo, A.; Scheraga, H. A. Application of Multiplexed Replica Exchange Molecular Dynamics to the UNRES Force Field: Tests with and + Proteins. *Journal of chemical theory and computation* **2009**, *5*, 627.

[168] Zwanzig, R. W. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *The Journal of Chemical Physics* **1954**, *22*, 1420–1426.

[169] van Gunsteren, W. F.; Berendsen, H. J. Thermodynamic cycle integration by computer simulation as a tool for obtaining free energy differences in molecular chemistry. *Journal of Computer-Aided Molecular Design* **1987**, *1*, 171–176.

[170] Lu, H.; Isralewitz, B.; Krammer, A.; Vogel, V.; Schulten, K. Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation. *Biophysical Journal* **1998**, *75*, 662.

[171] Merkel, R.; Nassoy, P.; Leung, A.; Ritchie, K.; Evans, E. Energy landscapes of receptor–ligand bonds explored with dynamic force spectroscopy. *Nature 1999 397:6714* **1999**, *397*, 50–53.

[172] Jarzynski, C. Nonequilibrium Equality for Free Energy Differences. *Physical Review Letters* **1997**, *78*, 2690.

[173] Torrie, G. M.; Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics* **1977**, *23*, 187–199.

[174] Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *Journal of Computational Chemistry* **1992**, *13*, 1011–1021.

[175] You, W.; Tang, Z.; Chang, C. E. A. Potential Mean Force from Umbrella Sampling Simulations: What Can We Learn and What Is Missed? *Journal of chemical theory and computation* **2019**, *15*, 2433.

[176] Kabsch, W. A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A* **1976**, *32*, 922–923.

[177] Flory, P. J. The Configuration of Real Polymer Chains. *The Journal of Chemical Physics* **1949**, *17*, 303–310.

[178] Kabsch, W.; Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983**, *22*, 2577–2637.

[179] Case, D. A.; Dyson, H. J.; Wright, P. E. Use of chemical shifts and coupling constants in nuclear magnetic resonance structural studies on peptides and proteins. *Methods in enzymology* **1994**, *239*, 392–416.

[180] Wales, D. J. *Atomic clusters and nanoparticles. Agregats atomiques et nanoparticules: Les Houches Session LXXIII 2–28 July 2000*; Springer, 2002; pp 437–507.

[181] Wales, D. J. PATHSAMPLE: A driver for OPTIM to create stationary point databases using discrete path sampling and perform kinetic analysis. 2025; https://www-wales.ch.cam.ac.uk/PATHSAMPLE/, Accessed: 2026-05-15.

[182] Wales, D. J. OPTIM: A program for optimizing geometries and calculating reaction pathways. 2025; https://www-wales.ch.cam.ac.uk/OPTIM/, Accessed: 2026-05-15.

[183] Forst, W. *Theory of unimolecular reactions*; Elsevier, 2012.

[184] Krupa, P.; Sieradzan, A. K.; Mozolewska, M. A.; Li, H.; Liwo, A.; Scheraga, H. A. Dynamics of Disulfide-Bond Disruption and Formation in the Thermal Unfolding of Ribonuclease A. *Journal of Chemical Theory and Computation* **2017**, *13*, 5721–5730.

[185] Lee, J.; Im, D.; Liu, Y.; Fang, J.; Tian, X.; Kim, M.; Zhang, W. B.; Seo, J. Distinguishing Protein Chemical Topologies Using Supercharging Ion Mobility Spectrometry-Mass Spectrometry. *Angewandte Chemie International Edition* **2023**, *62*, e202314980.

[186] Ciura, P.; Gumpelmair, S.; Rodziewicz-Motowidło, S.; Steinberger, P.; Spodzieja, M.; Sieradzan, A. K. Evaluation of LIGHT-derived peptides to disrupt the HVEM/LIGHT immune checkpoint. *Scientific Reports 2025 15:1* **2025**, *15*, 40095–.

[187] Ciura, P.; Gumpelmair, S.; Sikorska, E.; Rodziewicz-Motowidło, S.; Steinberger, P.; Spodzieja, M.; Sieradzan, A. K. HVEM-derived peptides as inhibitors of HVEM/LIGHT complex formation. *Bioorganic Chemistry* **2025**, *166*, 109183.

[188] Xu, M.; Ruan, M.; Zhu, W.; Xu, J.; Lin, L.; Li, W.; Zhu, W. Integrative analysis of a novel immunogenic PANoptosis-related gene signature in diffuse large B-cell lymphoma for prognostication and therapeutic decision-making. *Scientific Reports 2024 14:1* **2024**, *14*, 30370–.

[189] Adekoya, G. J.; Adekoya, O. C.; Muloiwa, M.; Sadiku, E. R.; Kupolati, W. K.; Hamam, Y. Computational strategies for designing PEDOT-Based electrodes in electronic applications. *Advanced Composites and Hybrid Materials 2026 9:2* **2026**, *9*, 125–.

[190] Xu, M.; Xu, L.; Zhu, W.; Ruan, M.; Xu, J.; Qu, Y.; Wang, L.; Zhu, W. Deciphering the causal relationships between plasma metabolites, peripheral cells, inflammatory factors and DLBCL for discovering novel therapeutic targets. *Discover Oncology* **2025**, *16*, 1968–.

[191] Al-Ghorbani, M.; Gouda, M. A.; Alharbi, O.; Al-Mutairi, K. A.; Ramu, R. Synthesis and Biological Evaluation of New N-(4-Acetylphenyl)-2-cyanoacetamide Derivatives with Docking, and Molecular Dynamics Insights. *Russian Journal of General Chemistry* **2025**, *95*, 155–172.

[192] Sheikh, M. R.; Utsha, M. H.; Tasnim, J.; Alam, M. R.; Afrin, U.; Chakraborty, A.; Rahaman, M. M.; Ali, S. M. E.; Sayed, M. A. Systematic In-silico Analysis of Fisetin-Proteins Interactions Revealing the PTGS2 as a Potential Therapeutic Target. **2025**,

[193] Zou, X.; Ai, X.; Peng, P.; Zhong, L.; Ren, M.; Liang, H.; Wang, Y. 3d-Qsar Studies of Quinolines as Antimicrobial Agents and Molecular Dynamics Simulation. **2024**,

[194] Xi, Z.; Kuo, S.-T.; Cong, X.; Yan, X.; Russell, D. H. Revealing Anisotropic Growth of Liraglutide Oligomers by Native Ion Mobility Mass Spectrometry and Molecular Dynamics Simulation. *ACS Central Science* **2025**, *11*, 1154–1165.

[195] Ugarte La Torre, D.; Sugita, Y. CGBack: Diffusion Model for Backmapping Large-Scale and Complex Coarse-Grained Molecular Systems. *Journal of Chemical Information and Modeling* **2025**, *65*, 9974–9986, PMID: 40962313.

[196] El Mesky, M.; Zgueni, H.; bouadid, I.; Tasnim, J.; Rhazi, Y.; Hasan, M. M.; Tanghourte, M.; Qabouche, A.; Chalkha, M.; Saleh, N.; Chebabe, D.; Mabrouk, E. H.; Eddouks, M. Synthesis of spiro-lactam hydrazones by clay catalysis: toxicity, antioxidant, hypolipidemic and In silico assessments. *RSC Adv.* **2026**, *16*, 22949–22968.

[197] Zhou, W.; Zheng, Y.; Liu, Y.; Liu, J.; Liu, Y.; Niu, Y.; Yu, Y.; Zhang, X.; Zhang, Y.; Yu, C. ATP7A-fibulin-4 complex delivers copper in the Golgi to activate LOX in renal fibrosis. *JCI Insight* **2026**, *11*.

[198] Chen, J.; Qiu, S.; Ju, C.; Wang, D.; Fan, F.; Huang, J. Arg128*-Mediated Dual-Substrate Recognition and Dynamic Transport Mechanisms in (R)--Transaminase: Computational Insights and Mutational Profiling Guided Rational Engineering. *Biotechnology Journal* **2025**, *20*, e70032.

[199] Alabbas, A. Integrating Multi-Domain Approach for Identification of Neo Anti-DHPS Inhibitors Against Pathogenic Stenotrophomonas maltophilia. *Biology* **2025**, *14*.

[200] Kurt, B.; Kurt, F. AI-Powered Structural and Co-Expression Analysis of Potato (Solanum tuberosum) StABCG25 Transporters Under Drought: A Combined AlphaFold, WGCNA, and MD Approach. *Biology* **2025**, *14*.

[201] Abdullahi, M.; Nuhu, A. M.; Sa'adu, S. K.; Hamza, S. A.; Yakubu, Y.; da Costa, R. A. Structure-based virtual screening, quantum chemical calculations, and ADMET-based scaffold hopping of dual inhibitors targeting respiratory syncytial and influenza a viruses. *In Silico Research in Biomedicine* **2026**, *2*, 100156.

[202] Moulick, A. G.; Erhardt, S.; Wenzel, W.; Kozlowska, M. Structural dynamics in the CENP-A nucleosome impacted by protein-protein interactions with centromere protein N. *bioRxiv* **2025**,

[203] Gavari-Bami, S.; Bayramzadeh, H.; Yoosefian, M. Molecular dynamics-informed design of functionalized graphene oxide nanocarriers for enhanced romidepsin delivery in solid tumor therapy. *Advanced Journal of Nanochemistry and Medicine* **2025**, *1*, 1–16.

[204] Tolstova, A. P.; Adzhubei, A. A.; Strelkova, M. A.; Makarov, A. A.; Mitkevich, V. A. Survey of the A-peptide structural diversity: molecular dynamics approaches. *Biophysical Reviews 2024 16:6* **2024**, *16*, 701–722.

[205] Ghorbaninia, M.; Doroudgar, S.; Ganjalikhany, M. R. Delving into the crucial role of the initial structure in the dynamic and self-assembly of amyloid beta. *Biochemical and Biophysical Research Communications* **2025**, *758*, 151652.

[206] Boopathi, S.; Garduño-Juárez, R. A Small Molecule Impedes the A1–42 Tetramer Neurotoxicity by Preserving Membrane Integrity: Microsecond Multiscale Simulations. *ACS Chemical Neuroscience* **2024**, *15*, 3496–3512, PMID: 39292558.

Authorship statements

.1 Mechanical Stability of Ribonuclease A Heavily Depends on the Redox Environment

Mgr inż. Pamela Smardz
Instytut Fizyki Polskiej Akademii Nauk
al. Lotników 32/46, 02-668 Warszawa
psmardz@ifpan.edu.pl

Warszawa, 05.01.2023

Oświadczenie

Oświadczam, że w publikacji:

Tytuł: Mechanical Stability of Ribonuclease A Heavily Depends on the Redox Environment

Autorzy: Pamela Smardz, Adam K Sieradzan, Paweł Krupa

Opublikowanej w: The Journal of Physical Chemistry B 2022, 126 (33), 6240-6249

Moim udziałem było: wykonanie pełnoatomowych symulacji dynamiki molekularnej wraz z ich analizą oraz analiza symulacji gruboziarnistych, spisanie i opracowanie graficzne otrzymanych przeze mnie wyników w formie wstępnej wersji manuskryptu publikacji naukowej, a następnie wykonanie poprawek edytorskich w manuskrypcie.

P. Smardz

Dr hab. Adam K. Sieradzan, prof. UG
Wydział Chemii Uniwersytetu Gdańskiego
Wita Stwosza 63, 80-308 Gdańsk
adam.sieradzan@ug.edu.pl

Gdańsk, 09.01.2023

Oświadczenie

Oświadczam, że w publikacji:

Tytuł: Mechanical Stability of Ribonuclease A Heavily Depends on the Redox Environment

Autorzy: Pamela Smardz, Adam K Sieradzan, Paweł Krupa

Opublikowanej w: The Journal of Physical Chemistry B 2022, 126 (33), 6240-6249

Moim udziałem było: wykonanie gruboziarnistych symulacji sterowanej dynamiki molekularnej, redagowanie manuskryptu publikacji naukowej oraz współpełnienie obowiązków autora korespondencyjnego.

Sieradzan

dr Paweł Krupa
Instytut Fizyki Polskiej Akademii Nauk
al. Lotników 32/46, 02-668 Warszawa
pkrupa@ifpan.edu.pl

Warszawa, 05.01.2023

Oświadczenie

Oświadczam, że w publikacji:

Tytuł: Mechanical Stability of Ribonuclease A Heavily Depends on the Redox Environment

Autorzy: Pamela Smardz, Adam K Sieradzan, Paweł Krupa

Opublikowanej w: The Journal of Physical Chemistry B 2022, 126 (33), 6240-6249

Moim udziałem było: zaplanowanie badań naukowych oraz nadzór nad ich wykonaniem i analizą, wykonanie pełnoatomowych symulacji sterowanej dynamiki molekularnej wraz z ich analizą, redagowanie manuskryptu publikacji naukowej oraz współpełnienie obowiązków autora korespondencyjnego.



.2 Multilayered Computational Framework for Designing Peptide Inhibitors of HVEM–LIGHT Interaction

mgr Piotr Ciura
Faculty of Chemistry
University of Gdańsk
Wita Stwosza 63
80-308 Gdańsk

Gdynia, 15.01.2026

Statement

I hereby declare that in the publication: *Multilayered Computational Framework for Designing Peptide Inhibitors of HVEM–LIGHT Interaction*, Piotr Ciura, Pamela Smardz, Marta Spodzieja, Adam K. Sieradzan, Paweł Krupa, The Journal of Physical Chemistry B **2024**, 128 (28), 6770–6785, my contribution included the design of peptide inhibitors targeting the HVEM–LIGHT interaction, performing molecular dynamics simulations of the complexes, computation of thermodynamic properties, analysis of simulation trajectories with a particular focus on intermolecular interactions, visualization of results, and co-writing the main text of the manuscript.





Mgr inż. Pamela Smardz

Division of Theoretical Physics (ON-5.3),

Institute of Physics, Polish Academy of Sciences,

Al. Lotników 32/46, 02-668, Warsaw, Poland

Email: psmardz@ifpan.edu.pl

STATEMENT

I hereby declare that in the publication:

Multilayered Computational Framework for Designing Peptide Inhibitors of HVEM–LIGHT Interaction,

Piotr Ciura, Pamela Smardz, Marta Spodzieja, Adam K. Sieradzan, Paweł Krupa,

The Journal of Physical Chemistry B 2024, 128 (28), 6770–6785,

my contribution included the design and execution of molecular dynamics simulations of individual inhibitors, analysis of structural properties with a particular emphasis on disulfide bond behavior, running enhanced-sampling simulations, visualization of results, and co-writing the main text of the manuscript.

Date: 16.01.2026

Signature: P. Smardz

Uniwersytet Gdański

IN MARI VIA CUA

CHEMIA UG

Gdańsk, 17.12.2025 rok

dr hab. Marta Spodzieja
Katedra Chemii Biomedycznej
Wydział Chemii
Uniwersytet Gdański
ul. Wita Stwosza 63
80-309 Gdańsk, Poland

OŚWIADCZENIE

Niniejszym oświadczam, że w publikacji pt. *Multilayered computational framework for designing peptide inhibitors of HVEM-LIGHT interaction*, Piotr Ciura, Pamela Smardz, Marta Spodzieja, Adam K. Sieradzan, Pawel Krupa, The Journal of Physical Chemistry B, 2024, 128 (28), 6770–6785, moja rola obejmowała konsultacje naukowe oraz konceptualizację prowadzonych badań.

Marta Spodzieja

Marta Spodzieja

Uniwersytet Gdański

dr. hab. Adam Sieradzan, prof. UG
Katedra Chemii Teoretycznej
Wydział Chemii UG

10.04.2026r. Gdańsk

OŚWIADCZENIE

Niniejszym oświadczam, że w publikacji: Piotr Ciura, Pamela Smardz, Marta Spodzieja, Adam K. Sieradzan, Pawel Krupa, *Multilayered Computational Framework for Designing Peptide Inhibitors of HVEM-LIGHT Interaction*, The Journal of Physical Chemistry B 2024, 128 (28), 6770–6785 moja rola obejmowała: konsultacje naukowe, konceptualizację badań, prowadzenie części symulacji dynamik molekularnych, projektowanie skryptów, pozyskanie środków finansowych oraz współpisanie ostatecznej wersji manuskryptu.

Sieradzan

Adam K. Sieradzan



Dr. hab. Paweł Krupa

Division of Theoretical Physics (ON-5.3),
Institute of Physics, Polish Academy of Sciences,
Al. Lotników 32/46, 02-668, Warsaw, Poland
Email: pkrupa@ifpan.edu.pl

STATEMENT

I declare that I am the author of the following publication:

Multilayered Computational Framework for Designing Peptide Inhibitors of HVEM-LIGHT Interaction, Piotr Ciura, Pamela Smardz, Marta Spodzieja, Adam K. Sieradzan, and Pawel Krupa, The Journal of Physical Chemistry B 2024 128 (28), 6770-6785, DOI: 10.1021/acs.jpcb.4c02255

My contribution involved supervising the project from the disulfide-bond analysis perspective, including assistance with conducting enhanced-sampling simulations and their subsequent analysis.

Date: 18.12.2025

Signature: Paweł Krupa

.3 A Practical Guide to All-Atom and Coarse-Grained Molecular Dynamics Simulations Using Amber and Gromacs: A Case Study of Disulfide-Bond Impact on the Intrinsically Disordered Amyloid Beta

INSTITUTE OF PHYSICS
POLISH ACADEMY OF SCIENCES

mgr inż. Pamela Smardz
Division of Theoretical Physics (ON-5),
Institute of Physics, Polish Academy of Sciences,
Al. Lotników 32/46, 02-668, Warsaw, Poland
Email: psmardz@ifpan.edu.pl

STATEMENT

I declare that I am a co-author of the following publication:

- Smardz, P., Anila, M. M., Rogowski, P., Li, M. S., Różycki, B., & Krupa, P. (2024). A Practical Guide to All-Atom and Coarse-Grained Molecular Dynamics Simulations Using Amber and Gromacs: A Case Study of Disulfide-Bond Impact on the Intrinsically Disordered Amyloid Beta. International Journal of Molecular Sciences, 25(12), 6698. https://doi.org/10.3390/ijms25126698

My contribution involved performing the all-atom as well as the SIRAH coarse-grained simulations of amyloid beta, analyzing and interpreting the simulation trajectories, preparing all figures except for Figures 5, 9, and A5, contributing to the construction of the tables, and participating in the manuscript writing.

Signature : P. Sm

Date : 22.10.2025



mgr. Midhun Mohan Anila
Laboratory of Biological Physics (SL-4),
Institute of Physics, Polish Academy of Sciences,
Al. Lotników 32/46, 02-668, Warsaw, Poland
Email: midhun@ifpan.edu.pl

STATEMENT

I declare that im a co-author of the following publication:

- Smardz, P., Anila, M. M., Rogowski, P., Li, M. S., Różycki, B., & Krupa, P. (2024). A Practical Guide to All-Atom and Coarse-Grained Molecular Dynamics Simulations Using Amber and Gromacs: A Case Study of Disulfide-Bond Impact on the Intrinsically Disordered Amyloid Beta. International Journal of Molecular Sciences, 25(12), 6698. https://doi.org/10.3390/ijms25126698

My contributions to this publication include (i) performing the Martini coarse-grained simulations for amyloid beta, (ii) analyzing and interpreting the results from the simulation trajectories, (iii) preparing Figures 5, 9, and A5, constructing Table 2 (showing the structural dimensions of amyloid beta), (iv) participating in the construction of Table 9 (summarizing the computational times for different simulation models), (v) contributed to the manuscript writing.

Signature : 

Date : January 20, 2026

Dr. Pawel Rogowski

Department of Molecular Plant Physiology,

Faculty of Biology

Miecznikowa 1, 02-096

University of Warsaw

Email: p.rogowski3@uw.edu.pl

STATEMENT

I declare that I have co-authored the following publication:

- Smardz, P., Anila, M. M., Rogowski, P., Li, M. S., Różycki, B., & Krupa, P. (2024). A Practical Guide to All-Atom and Coarse-Grained Molecular Dynamics Simulations Using Amber and Gromacs: A Case Study of Disulfide-Bond Impact on the Intrinsically Disordered Amyloid Beta. International Journal of Molecular Sciences, 25(12), 6698. https://doi.org/10.3390/ijms25126698

My contributions to this work included assisting in the analysis and interpretation of the results from the Martini simulation trajectories of amyloid beta and contributing to the manuscript writing.

Signature : Paweł Rogowski

Date : 22.10.2025



Prof. dr hab. Mai Suan Li

Division of Theoretical Physics (ON-5),

Institute of Physics, Polish Academy of Sciences,

Al. Lotników 32/46, 02-668, Warsaw, Poland

Email: masli@ifpan.edu.pl

STATEMENT

I declare that I have co-authored the following publication:

- Smardz, P., Anila, M. M., Rogowski, P., Li, M. S., Różycki, B., & Krupa, P. (2024). A Practical Guide to All-Atom and Coarse-Grained Molecular Dynamics Simulations Using Amber and Gromacs: A Case Study of Disulfide-Bond Impact on the Intrinsically Disordered Amyloid Beta. International Journal of Molecular Sciences, 25(12), 6698. https://doi.org/10.3390/ijms25126698

My contribution in this publication was involved in the conceptualization, validation of results, and and writing of the original draft.

Warsaw , October 23, 2025





Dr. hab. Bartosz Rozycki
Associate Professor,
Theoretical Biophysics Group Leader,
Division of Theoretical Physics,
Institute of Physics, Polish Academy of Sciences,
Al. Lotników 32/46, 02-668, Warsaw, Poland
Email: rozycki@ifpan.edu.pl

STATEMENT

I declare that I am a co-author of the following publication:

- Smardz, P., Anila, M. M., Rogowski, P., Li, M. S., Różycki, B., & Krupa, P. (2024). A Practical Guide to All-Atom and Coarse-Grained Molecular Dynamics Simulations Using Amber and Gromacs: A Case Study of Disulfide-Bond Impact on the Intrinsically Disordered Amyloid Beta. International Journal of Molecular Sciences, 25(12), 6698. https://doi.org/10.3390/ijms25126698

My contributions to this publication include (i) conceptualization of the research, (ii) co-development of scripts for setting up, carrying out and analysing the molecular dynamics simulations with the Martini coarse-grained model, (iii) supervision of the research tasks involving the Martini simulations, (iv) validation of the Martini simulations, (v) providing access to computation resources, (vi) funding acquisition, (vii) project administration, (viii) writing — original draft preparation, (ix) writing — review and editing.

Signature : Bartosz Różycki

Date : January 12, 2026



Dr. hab. Paweł Krupa,
Division of Theoretical Physics (ON-5.3),
Institute of Physics, Polish Academy of Sciences,
Al. Lotników 32/46, 02-668, Warsaw, Poland
Email: pkrupa@ifpan.edu.pl

STATEMENT

I declare that I am the corresponding author of the following publication:

- Smardz, P., Anila, M. M., Rogowski, P., Li, M. S., Różycki, B., & Krupa, P. (2024). A Practical Guide to All-Atom and Coarse-Grained Molecular Dynamics Simulations Using Amber and Gromacs: A Case Study of Disulfide-Bond Impact on the Intrinsically Disordered Amyloid Beta. International Journal of Molecular Sciences, 25(12), 6698. https://doi.org/10.3390/ijms25126698

My contribution to this publication was conceptualization of both the all-atom and SIRAH coarse-grained simulations, software development, validation, formal analysis, providing resources, writing of the original draft, reviewing and editing, supervision, project administration, funding acquisition, and correspondence with the editorial office.

Signature: Paweł Krupa

Date: 25.10.2025

.4 Visualising the Energy Landscape for a Molecular Dynamics Trajectory

Pamela Smardz
Instytut Fizyki PAN
psmardz@ifpan.edu.pl

Warszawa, 04.05.2026

OŚWIADCZENIE

Oświadczam, że jestem współautorką następującej publikacji:

Tytuł: *Visualizing the energy landscape for a molecular dynamics trajectory*

Autorzy: Vilmos Neuman, Patryk A. Wesołowski, Krzysztof K. Bojarski, Diksha Dewan, Moritz Schäffler, Pamela Smardz, David J. Wales

Opublikowanej w: J. Chem. Phys. 164, 044112 (2026) https://doi.org/10.1063/5.0310206

Mój udział polegał na: udziale w planowaniu badań, przygotowaniu układu z wiązaniami disulfidowymi, wykonaniu symulacji dynamiki molekularnej białka RCSB PDB 1HYP wraz z analizą otrzymanych wyników, współudziale w opracowaniu metody analizy trajektorii, a także udziale w opracowaniu i spisywaniu wyników do manuskryptu publikacji naukowej.

P. Smardz

Name: Vilmos Neuman
Adress: Department of Chemistry, Physical and Theoretical Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QZ, U.K.
Email: vilmos.neuman@st-hildas.ox.ac.uk

STATEMENT

I declare that I am a co-author of the following publication:

Title: *Visualizing the energy landscape for a molecular dynamics trajectory*

Authors: *Vilmos Neuman, Patryk A. Wesołowski, Krzysztof K. Bojarski, Diksha Dewan, Moritz Schäffler, Pamela Smardz, David J. Wales*

Published in: *J. Chem. Phys. 164, 044112 (2026) https://doi.org/10.1063/5.0310206*

My contribution was

Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal).

Signature: 

Date: 04/05/2026

Name: Patryk Adam Wesołowski
Adress: Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, United Kingdom
Email: paw61@cam.ac.uk

STATEMENT

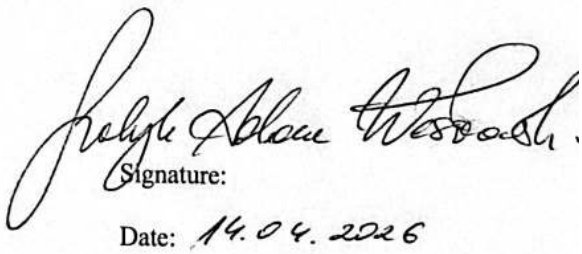
I declare that I am a co-author of the following publication:

Title: *Visualizing the energy landscape for a molecular dynamics trajectory*

Authors: *Vilmos Neuman, Patryk A. Wesołowski, Krzysztof K. Bojarski, Diksha Dewan, Moritz Schäffler, Pamela Smardz, David J. Wales*

Published in: *J. Chem. Phys. 164, 044112 (2026) https://doi.org/10.1063/5.0310206*

My contribution was Conceptualization (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Supervision (equal); Validation (equal); Writing – original draft (equal); Writing – review & editing (equal).



Signature:

Date: 14.04.2026

Krzysztof Bojarski
Gabriela Narutowicza 11/12, Gdańsk
krzysztof.bojarski@pg.edu.pl

STATEMENT

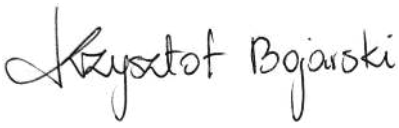
I declare that I am a co-author of the following publication:

Title: *Visualizing the energy landscape for a molecular dynamics trajectory*

Authors: *Vilmos Neuman, Patryk A. Wesołowski, Krzysztof K. Bojarski, Diksha Dewan, Moritz Schäffler, Pamela Smardz, David J. Wales*

Published in: *J. Chem. Phys. 164, 044112 (2026) https://doi.org/10.1063/5.0310206*

My contribution was data curation (equal); methodology (supporting); validation (equal).

Signature: 

Date: 04.05.2026

Name: Diksha Dewan
Address: University of Cambridge
Email: dd598@cam.ac.uk

STATEMENT

I declare that I am a co-author of the following publication:

Title: *Visualizing the energy landscape for a molecular dynamics trajectory*

Authors: *Vilmos Neuman, Patryk A. Wesołowski, Krzysztof K. Bojarski, Diksha Dewan, Moritz Schäffler, Pamela Smardz, David J. Wales*

Published in: *J. Chem. Phys. 164, 044112 (2026) https://doi.org/10.1063/5.0310206*

My contribution: Data curation (equal); Methodology (supporting); Validation (equal)

Signature: Diksha Dewan

Date: 13 April 2026

Name: Moritz Schäffler
Adress: Eichendorff Str. 37, 50825 Cologne, Germany
Email: mo.schaeffler@fz-juelich.de

STATEMENT

I declare that I am a co-author of the following publication:

Title: *Visualizing the energy landscape for a molecular dynamics trajectory*

Authors: *Vilmos Neuman, Patryk A. Wesołowski, Krzysztof K. Bojarski, Diksha Dewan, Moritz Schäffler, Pamela Smardz, David J. Wales*

Published in: *J. Chem. Phys. 164, 044112 (2026) https://doi.org/10.1063/5.0310206*

My contribution was:
Data curation (equal); Methodology (supporting); Validation (equal)

Signature: 

Date: 4th of May 2026

Name DAVID WALES
Adress Dept of chemistry, Cambridge
Email dw34@cam.ac.uk

STATEMENT

I declare that I am a co-author of the following publication:

Title: *Visualizing the energy landscape for a molecular dynamics trajectory*

Authors: *Vilmos Neuman, Patryk A. Wesołowski, Krzysztof K. Bojarski, Diksha Dewan, Moritz Schäffler, Pamela Smardz, David J. Wales*

Published in: *J. Chem. Phys. 164, 044112 (2026) https://doi.org/10.1063/5.0310206*

My contribution was

Conceptualization (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Supervision (equal); Validation (equal); Writing – original draft (equal); Writing – review & editing (equal).

Signature:

Date: 13/4/26

Dostępy do artykułów:

1. Mechanical Stability of Ribonuclease A Heavily Depends on the Redox Environment
Published as part of The Journal of Physical Chemistry virtual special issue "Protein Folding and Dynamics_ An Overview on the Occasion of Harold Scheraga's 100th Birthday".
Pamela Smardz, Adam K. Sieradzan,* and Paweł Krupa*
https://doi.org/10.1021/acs.jpcb.2c04718
J. Phys. Chem. B 2022, 126, 6240−6249

2. Multilayered Computational Framework for Designing Peptide Inhibitors of HVEM-LIGHT Interaction
Piotr Ciura, Pamela Smardz, Marta Spodzieja, Adam K. Sieradzan,* and Pawel Krupa
https://doi.org/10.1021/acs.jpcb.4c02255
J. Phys. Chem. B 2024, 128, 6770−6785

3. A Practical Guide to All-Atom and Coarse-Grained Molecular Dynamics Simulations Using Amber and Gromacs: A Case Study of Disulfide-Bond Impact on the Intrinsically Disordered Amyloid Beta
Pamela Smardz, Midhun Mohan Anila , Paweł Rogowski , Mai Suan Li, Bartosz Ró˙zycki and Pawel Krupa **
Int. J. Mol. Sci. **2024**, *25*, 6698, https://doi.org/10.3390/ijms25126698

4. Visualizing the energy landscape for a molecular dynamics trajectory

Vilmos Neuman, Patryk A. Wesołowski, Krzysztof K. Bojarski, Diksha Dewan, Moritz Schäffler, Pamela Smardz, and David J. Wales

J. Chem. Phys. **164**, 044112 (2026); doi: 10.1063/5.0310206 **164**, 044112-1