



Thermodynamic Quantification of Multi-Sequence DNA Recognition by Prokaryotic Transcription Factor: A Molecular Dynamics Simulation Study

Soumi Das 

Independent Researcher, Kolkata 700032, West Bengal, India

Article History

Received: December 09, 2025

Accepted: April 17, 2026

Published: June 08, 2026

Abstract

DNA recognition by transcription factors is governed not only by binding affinity but also by DNA sequence-encoded modulation of protein conformational dynamics. Quantitative analysis of residue-specific conformational thermodynamics, based on side-chain torsion angle distributions derived from molecular dynamics simulations and principal component analysis, demonstrates that the symmetric Lac repressor protein (LacI) undergoes operator-specific conformational redistribution upon binding to the native lac operator O2 and the synthetic operator O-SymL. Binding to O2 is associated with enhanced conformational entropy in the N- and C-terminal regions, whereas binding to O-SymL promotes localized conformational stabilization. Polar and hydrophobic residues contribute differently to protein dynamics, with DNA sequence-dependent asymmetry governing these effects. Binding to either operator stabilizes residues at the protein–DNA interface; however, distal core subdomains continue to exhibit sequence-dependent remodeling of their conformational landscape. Long-range modulation of these distal regions induces asymmetry in monomer dynamics, resulting in the functional non-equivalence of structurally identical monomers and reflecting sequence-dependent allosteric propagation through the LacI scaffold. These findings support a thermodynamic framework showing how distinct DNA sequences remodel the conformational ensemble of a transcription factor, thereby modulating gene regulation beyond static binding affinity. This sequence-dependent asymmetric recognition provides a mechanistic basis for differential energetic and entropic coupling in gene regulatory systems. More broadly, these findings may inform the rational design of synthetic genetic circuits and the development of allosteric modulators targeting protein–nucleic acid and protein–protein interactions, including targets traditionally considered difficult to drug.

Keywords:

operator DNA sequences; lac repressor; protein–DNA recognition; asymmetric protein dynamics; conformational thermodynamics; molecular dynamics simulation

1. Introduction

Protein–DNA interactions (PDIs) play a central role in cellular function, governing gene regulation, DNA replication and repair, chromatin organization, and cell-cycle control. These interactions are essential for regulating transcription and maintaining genomic integrity. Dysregulation of protein–DNA interactions (PDIs) is a hallmark of numerous diseases, including cancer, neurodevelopmental disorders, and immune-related disorders. Although PDIs were historically considered “undruggable”

due to the absence of well-defined binding pockets and the predominance of electrostatic interactions, recent advances have challenged this view. Protein–DNA interactions (PDIs) are now recognized as promising therapeutic targets. They are governed by both direct base readout and indirect DNA shape recognition. These interactions are stabilized by electrostatic interactions (including salt bridges), hydrogen bonding, hydrophobic interactions, and π – π stacking interactions. Multiple therapeutic strategies targeting protein–DNA interactions (PDIs) can be broadly categorized into: (i) direct inhibition of DNA–

binding proteins; (ii) targeting DNA through sequence- or structure-specific recognition; and (iii) modulation of the protein–DNA interface via allosteric or indirect mechanisms, in which small molecules bind to distal sites on either the protein or DNA, thereby altering conformational dynamics and indirectly reducing binding affinity. Given the inherent flexibility of protein–DNA interfaces, indirect strategies can exploit transient, druggable conformational states, thereby offering a promising approach for achieving selective and effective therapeutic intervention [1,2].

Transcription factors (TFs) are essential regulatory proteins that modulate gene expression by binding to specific DNA sequences, thereby enabling the activation or repression of target genes. Elucidating the molecular principles governing TF discrimination among DNA sequences is fundamental to understanding the mechanisms of gene regulation. Although transcription factors (TFs) bind to multiple genomic loci, their regulatory outputs are highly sequence-dependent and cannot be explained solely by differences in binding affinity. Instead, specificity emerges from the coordinated interplay of multiple factors, including direct base contacts, DNA shape readout, conformational selection within dynamic protein–DNA ensembles, long-range allosteric communication, cooperative interactions with additional regulatory factors, sequence-encoded modulation of protein dynamics, and chromatin accessibility. However, the mechanistic basis by which these factors collectively translate into precise sequence-specific transcriptional regulation remains unresolved [3].

In bacteria, the most prevalent DNA-binding motif is the helix–turn–helix (HTH) motif. This structural motif typically engages inverted repeat (IR) DNA sequences, which form palindromic arrangements that enable transcription factors such as LacI, TetR, and IclR to bind as dimers or tetramers in a head-to-head configuration. In contrast, direct repeat (DR) sequences require a head-to-tail binding mode. This type of interaction is also observed in transcription factors such as Xis, Atox1, and BldC. Over the past decade, several curated databases, including JASPAR, TRANSFAC, UniPROBE, RegulonDB, and CollecTF, have facilitated the prediction and systematic analysis of transcription factor binding sites (TFBS). In parallel, a range of artificial intelligence (AI)-driven approaches have been developed to enhance the identification and modeling of TF–DNA interactions. DanQ integrates convolutional neural networks (CNNs) with recurrent neural networks (RNNs) to predict transcription factor binding sites (TFBS). DeepBind, DeepSEA, and DESSO are based on convolutional neural network (CNN) architectures. In addition, FCNSignal employs

fully convolutional networks (FCNs) for the prediction of transcription factor binding signals. Deep learning methods have demonstrated strong performance in the prediction of transcription factor binding sites (TFBS). Machine learning (ML) models can accurately predict and classify TFBS; however, traditional ML approaches remain comparatively underexplored [4]. The JASPAR 2026 database significantly expands its open-access repository of transcription factor (TF) binding profiles and integrates deep learning-based models to enhance the accuracy and scalability of TF–DNA binding predictions [5].

Recent advances underscore the importance of integrating physical principles with data-driven approaches to improve the understanding of transcription factor (TF)–DNA recognition. Physics-based machine learning frameworks incorporate explicit energetic descriptors, including electrostatic, van der Waals, and solvation contributions, into statistical models, thereby enabling both accurate prediction and mechanistic interpretability of binding interactions. These approaches demonstrate that binding specificity arises not solely from DNA sequence features but from the underlying energetic landscape of protein–DNA interactions. This framework enables physically interpretable feature attribution, identifies key amino acid–nucleotide base interactions, and overcomes the limitations of conventional black-box models [6]. Wang et al. (2026) demonstrated that sequence-based deep learning models, including convolutional neural networks (CNNs) and self-attention transformer architectures, can accurately predict transcription factor–DNA binding across both high- and low-affinity sites. These models outperform traditional statistical approaches in predictive accuracy. These models are capable of identifying low-affinity binding sites and providing mechanistic insights into transcription factor–DNA recognition, thereby contributing to a deeper understanding of gene regulation and facilitating the rational design of proteins [7].

DeepCAC combines convolutional neural networks (CNNs) with multi-head self-attention mechanisms to enable more effective modeling of transcription factor (TF)–DNA interactions [8]. Similarly, DeepCTF demonstrates that integrating DNA sequence information with structural features, such as minor groove width and helix twist, within an attention-based deep learning framework enhances the prediction of TF binding specificity [9]. Barissi et al. (2022) introduced DNAAffinity, a physics-based machine learning framework that predicts transcription factor binding affinities using DNA structural and mechanical properties derived from atomistic molecular dynamics simulations. This study highlights that binding affinity is governed not only by direct base contacts but also by indirect readout mechanisms arising from DNA

shape and associated energetic contributions [10]. Sedhom et al. (2022) developed a method to align major-groove hydrogen-bonding patterns across DNA sequences that bind the same protein, revealing conserved interaction signatures that are not captured by sequence-based analyses alone. Applying this approach to multiple binding sites of the Lac repressor and the λ CI repressor, they demonstrated that key hydrogen-bonding interactions remain conserved despite variations in DNA sequence. This study highlights the importance of direct readout information in protein–DNA specificity, with implications for protein design and the elucidation of evolutionary relationships [11]. Recent studies further demonstrate that the electrostatic properties of intrinsically disordered regions (IDRs) in transcription factors regulate efficient target search processes through a range of transient interactions [12,13]. Marklund et al. (2022) demonstrated that DNA-binding specificity is primarily governed by the efficiency of protein association with different DNA sequences. Their results indicate that modulating association rates, rather than dissociation rates, reduces the likelihood of proteins becoming kinetically trapped on non-target sequences during the search process. This suggests that binding kinetics, particularly association rates, play a dominant role in determining specificity in protein–DNA interactions [14].

The lactose repressor (LacI), a sequence-specific transcription factor that regulates the *lac* operon in *Escherichia coli*, serves as a canonical model system for investigating the relationship between sequence-specific DNA recognition and allosteric regulation, supported by extensive structural, genetic, and thermodynamic characterization.

While biophysical techniques such as nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography have elucidated the architectures of LacI–DNA complexes, these static structures do not fully capture the dynamic and energetic contributions that distinguish specific from non-specific DNA recognition. Within the cellular environment, LacI samples numerous non-specific DNA sites during its target search via mechanisms such as one-dimensional sliding and three-dimensional diffusion, before forming a stable complex at its cognate operator. Elucidating the structural and thermodynamic distinctions between these binding modes is therefore essential for understanding the basis of recognition specificity.

In the *lac* operon, the homotetrameric LacI repressor binds simultaneously to the primary operator O1 and auxiliary operators O2 and O3, thereby inducing DNA looping. This looping increases the local concentration of the repressor near the promoter, sterically hinders RNA polymerase access, and results in efficient transcriptional re-

pression. Deletion of auxiliary operators substantially reduces transcriptional repression, underscoring the critical role of cooperative interactions in LacI-mediated gene regulation.

LacI functions as a homotetramer (a dimer of dimers) composed of four identical 37.5 kDa monomers. Each monomer comprises an N-terminal helix–turn–helix (HTH) DNA-binding domain (residues 1–62), a central core domain (residues 63–329) that contains the inducer-binding site and dimerization interface, and a C-terminal tetramerization domain (residues 330–360). The core domain, which is subdivided into N- and C-terminal subdomains, is connected to the DNA-binding domain via hinge helices (residues 50–57). Figure 1a illustrates the architecture of the operator-bound tetrameric Lac repressor. Upon binding inducers such as allolactose or IPTG, LacI undergoes an allosteric conformational transition within the core domain that reduces operator DNA-binding affinity and thereby initiates transcription. The N-terminal helix–turn–helix (HTH) motif of the Lac repressor recognizes the major groove of operator DNA through sequence-specific hydrogen-bonding interactions with base pairs. The hinge helices insert into the minor groove, inducing an approximately 45° DNA bend via an induced-fit mechanism [15–37].

Fried et al. (2002) demonstrated that hydration plays a significant role in the interaction between the Lac repressor and operator DNA. Formation of the specific LacI–O1 repressor–operator complex is accompanied by the release of approximately 260 water molecules, resulting in a substantial entropy gain that contributes to high-affinity complex formation. This process also reduces the solvent-accessible surface area and stabilizes the protein–DNA interface. Interfacial water molecules trapped at the protein–DNA interface mediate hydrogen-bonding networks that complement direct base-specific hydrogen-bonding interactions. The thermodynamic basis of binding reflects a balance between enthalpic contributions arising from direct and water-mediated interactions and entropic effects associated with solvent reorganization [38]. Barr and van der Vaart (2012) demonstrated, through molecular dynamics simulations, that the bent DNA conformation in the Lac repressor headpiece–O1 operator complex arises primarily from specific protein–DNA interactions and hydration thermodynamics, rather than intrinsic DNA curvature alone [39].

Lüking et al. (2022) identified two distinct conformational states adopted by the Lac repressor during DNA diffusion: (i) a flexible “search” conformation characterized by disordered hinge regions that facilitates efficient one-dimensional scanning along relatively straight DNA, and (ii) a helical “recognition” conformation that enables

high-affinity binding to bent, sequence-specific operator DNA. During the search phase, the protein adopts a flexible conformation characterized by disordered hinge regions, enabling the Lac repressor to remain mobile and facilitating rapid one-dimensional sliding along non-specific, largely straight DNA while minimizing kinetic trapping. Upon encountering specific operator sequences, the hinge regions undergo a disorder-to-order transition, adopting α -helical conformations that insert into the minor groove. This helical “recognition” state stabilizes the Lac repressor–operator complex. Thus, the reversible transition between disordered (search) and ordered helical (recognition) conformations underlies the efficiency of target-site localization and high-affinity binding to the cognate operator [40]. Using all-atom molecular dynamics simulations, Liao et al. (2019) suggested that LacI preferentially binds to bent A-form DNA, stabilizing the protein structure while concomitantly reducing conformational entropy. Their results further indicate a modest increase in hydrogen-bonding interactions in specific LacI–

DNA complexes relative to non-specific DNA [41]. Enhanced sampling simulations indicate that cognate DNA sequences stabilize specific polar contacts, thereby lowering the free energy of the specific complex. At the same time, they preserve a relatively smooth energy landscape that enables rapid sliding along DNA without necessitating large conformational rearrangements [42,43]. Xu et al. (2015) investigated the discrimination between specific and non-specific DNA binding by the Lac repressor through an integration of molecular dynamics simulations and correlation network analysis. Their findings show that binding to specific operator DNA gives rise to a more centralized and densely connected interaction network, which enhances long-range communication between the DNA-binding domain and distal regions, thereby stabilizing allosteric signaling networks. In contrast, non-specific complexes exhibit a more dynamic and weakly coordinated interaction network, predominantly governed by transient electrostatic and hydrogen-bonding interactions [44].

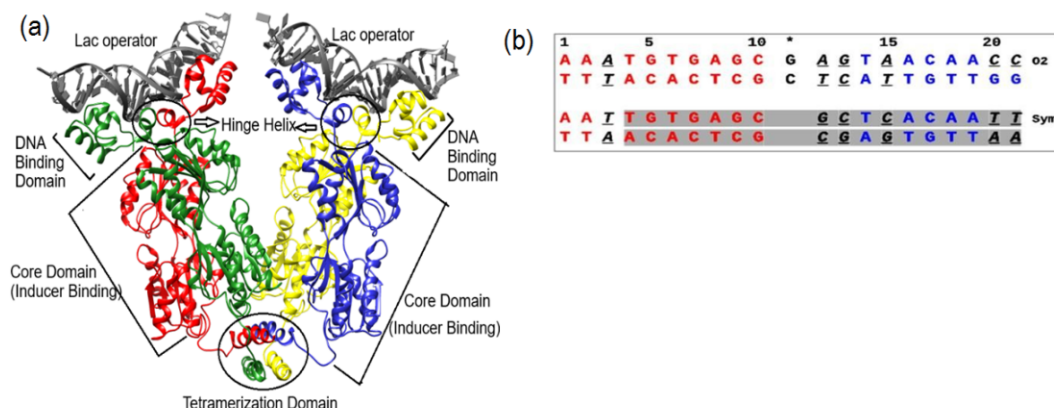


Figure 1: (a) Ribbon representation of the tetrameric Lac repressor bound to two DNA operator sites, based on the 4.80 Å resolution crystal structure (PDB ID: 1LBG), visualized using UCSF Chimera. (b) Nucleotide sequences of the naturally occurring pseudo-palindromic O2 operator and the synthetic symmetric (palindromic) SymL operator.

Sengupta et al. (2017) demonstrated that the high-free-energy transition state (TS) of Lac repressor–operator binding retains a substantial fraction of electrostatic interactions, while DNA phosphates remain hydrated and key aromatic and amide functional groups become progressively buried [45].

Computational and experimental studies show that local protein flexibility, electrostatic interactions, and solvent effects play key roles in modulating LacI dynamics. The hinge helix enhances non-specific DNA binding by strengthening electrostatic interactions and modulating counterion distribution, thereby reducing the dissociation rate [46,47].

Molecular dynamics simulations of full-length LacI under varying hydrostatic pressures indicate that elevated pressure inhibits the transition from the DNA-unbound “open” state to the DNA-bound “closed” conformation, an effect analogous to that induced by DNA or ligand binding. These findings suggest that hydration of specific residues, particularly in the hinge-helix region and at domain interfaces, plays a critical role in stabilizing distinct functional conformations. Water-mediated interactions act as key contributors to LacI functional allostery [48]. High-resolution nuclear magnetic resonance (NMR) spectroscopy has shown that the Lac repressor dimer exists in a dynamic equilibrium between two distinct confor-

mational states corresponding to the DNA-bound and inducer-bound forms. Binding of either effector shifts this equilibrium in a manner consistent with the Monod–Wyman–Changeux (MWC) allosteric model, in which redistribution of pre-existing conformational ensembles mediates transcriptional switching [49].

Machine learning approaches significantly improve the accurate prediction of how mutations affect LacI-mediated transcriptional repression, achieving higher performance than traditional methods [50]. OptoLacI variants were engineered by coupling the Lac repressor scaffold with a light-oxygen-voltage 2 (LOV2) photosensory domain, enabling direct light-dependent control of gene expression in the absence of IPTG. Optogenetic engineering of LacI enables precise light-mediated control of DNA-binding activity, highlighting its utility in synthetic biology applications that do not require chemical inducers [51].

Xu et al. (2022) showed that LacI-mediated DNA loops function as topological barriers whose stability is modulated by transcription-generated supercoiling. Their results demonstrate that torsional stress reduces the lifetime and stability of LacI-induced loops, thereby weakening their ability to impede RNA polymerase progression. These findings highlight a dynamic interplay between DNA topology and protein-mediated looping, illustrating how mechanical strain modulates transcriptional repression within the lac operon [52]. Kinetic analyses support a multistep binding pathway characterized by operator-dependent energetic barriers and partially ordered intermediate states. Nuclear magnetic resonance (NMR) spectroscopy, single-molecule spectroscopy, and molecular dynamics simulations collectively indicate that recognition between the Lac repressor and its operators does not conform to a classical lock-and-key model but instead proceeds through a dynamic multistep mechanism involving both direct and indirect base readout [53–57].

Despite these advances, the influence of operator sequence on distal LacI protein dynamics and intraprotein communication remains incompletely understood. In particular, it is unclear whether symmetric LacI dimers respond equivalently to different operators or whether sequence-specific binding induces asymmetric dynamics between the two monomers. Addressing this question requires residue-specific thermodynamic analyses capable of quantifying entropy redistribution and energetic coupling across the protein architecture.

The histogram-based method (HBM) provides a rigorous framework for dissecting the conformational thermodynamic basis of bimolecular function across diverse systems, including protein–peptide, protein–protein, and protein–nucleic acid complexes. This method enables the

prediction of functional ligand-binding sites in proteins as well as putative binding pockets for non-cognate ligands targeting nucleic acids, thereby providing structural and thermodynamic insights that can inform the rational design of protein- and nucleic acid-targeted therapeutics.

Protein–protein interactions, as exemplified by the MDM2–p53 and SAP–SLAM complexes, induce cooperative intraprotein conformational changes that propagate allosteric effects beyond the immediate binding interface. The histogram-based method (HBM) quantitatively characterizes these conformational rearrangements by analyzing residue-specific thermodynamic changes, thereby providing microscopic insight into the structural and thermodynamic basis of ligand binding and allostery. Studies of membrane-associated and pathogenic protein–protein interfaces further underscore the role of entropy redistribution in stabilizing transient yet functionally critical interactions.

Thermodynamic quantification using the histogram-based method (HBM) reveals how free energy and entropic contributions collectively govern the stability of biomolecular complexes and ligand recognition, highlighting the method’s ability to link microscopic conformational dynamics to functional outcomes. Beyond proteins, the histogram-based method (HBM) has been successfully extended to nucleic acid systems, where it has elucidated the thermodynamic basis of metal ion specificity in DNA cleavage by restriction endonucleases, protein-induced stabilization of Hoogsteen base pairs, ligand recognition by riboswitch aptamers, and the thermosensing behavior of riboswitch aptamers. Conformational thermodynamic analyses of ligand-binding studies of the fluoride riboswitch aptamer have identified potential hotspots for RNA-targeted antimicrobial drug design. In the histogram-based method (HBM), residue-specific torsional probability distributions are extracted from equilibrium molecular dynamics trajectories, enabling direct computation of conformational entropy and free energy based on principles of statistical mechanics [58–71]. This method has broad implications for understanding structure–function relationships in biomolecules by characterizing conformational stability and ordering under diverse perturbations, including temperature, pressure, solvent composition, pH, and mutations.

The histogram-based method (HBM) offers distinct advantages over conventional enhanced-sampling molecular dynamics (MD) approaches such as metadynamics and umbrella sampling, which rely on predefined collective variables (CVs) to accelerate rare conformational transitions. In protein–nucleic acid complexes, conformational dynamics arise from a wide range of microscopic variables, including torsion and pseudo-torsion

angles of nucleic acids, backbone dihedral angles and side-chain rotamers of amino acid residues, and solvent-mediated interactions. This complexity makes it challenging to define a minimal and unbiased set of collective variables. Incorrect or suboptimal selection of collective variables can introduce sampling bias, impede convergence, and reduce interpretability at the residue level. By eliminating reliance on predefined reaction coordinates, the histogram-based method (HBM) enables efficient and unbiased residue-specific characterization of conformational thermodynamics, capturing both free-energy and entropic contributions. Moreover, the histogram-based method (HBM) complements traditional energy decomposition approaches, such as MM/PBSA and MM/GBSA, by providing detailed residue-level insights into the mechanisms of molecular recognition within receptor–ligand complexes. These mechanistic details are often inaccessible to collective-variable-dependent enhanced-sampling methods [72–76].

Figure 1b illustrates the two target DNA sequences investigated in this study: (i) the naturally occurring pseudo-palindromic operator O2 and (ii) the synthetic operator O-SymL, a perfect palindrome derived from the left half-site of the native O1 operator. Complete twofold symmetry in O-SymL is achieved through deletion of the central G:C base pair. O-SymL is widely used in biophysical studies due to its exceptionally high affinity for the Lac repressor. In contrast, operator O2 exhibits sequence asymmetry between the left (consensus) and right (non-consensus) binding sites. Conserved bases are highlighted in red (left) and blue (right), whereas asymmetric regions in O2 are underlined and italicized (**Figure 1b**).

In this study, region-specific thermodynamic changes in the Lac repressor protein (holo) were quantified relative to the unbound (apo) state using HBM. Negative conformational free-energy changes (ΔG) and entropy contributions ($T\Delta S$) indicate energetic stabilization and increased conformational ordering, whereas positive ΔG and $T\Delta S$ values indicate destabilization and increased conformational disorder. Principal component analysis (PCA) is further employed to characterize the dominant motions of LacI bound to operators O-SymL and O2, thereby validating the trends observed using the histogram-based method (HBM).

Overall, this work provides direct thermodynamic evidence linking side-chain dynamics to sequence-dependent allostery in LacI. Transcriptional repression is therefore governed not only by static binding affinity but also by the dynamic redistribution of entropic and energetic contributions across the protein scaffold. By linking subtle DNA sequence variations to global allosteric reprogramming, this study advances mechanistic understanding of transcrip-

tion factor regulation and establishes a general framework for interpreting how sequence-encoded dynamics modulate function across diverse protein–DNA systems.

2. Methods

2.1. Modeling Strategy

The coordinates of the Lac repressor dimer bound to operator DNA were obtained from the Protein Data Bank (PDB ID: 1EFA) [77], containing the full-length Lac repressor. Chains A and B were retained for all simulations, whereas chain C (residues 46–331) was removed to reduce system size and maintain the symmetric dimeric structure of the repressor. The co-crystallized anti-inducer ONPF in the structure was excluded, as previous studies indicated that ONPF binding does not significantly perturb the ligand-binding pocket or the overall Lac repressor–DNA conformation. All crystallographic water molecules were removed from the system. The resulting processed structure was used as the model for the Lac–O-SymL repressor–operator complex.

Because a high-resolution X-ray or NMR structure for the full-length Lac–O2 complex (residues 2–329) is unavailable, a hybrid modeling strategy was employed. The N-terminal DNA-binding domain of the Lac repressor dimer from PDB entry 1EFA was structurally aligned with the corresponding domain of the NMR ensemble structure 2KEJ. The remaining DNA-binding region present in 2KEJ was subsequently removed to eliminate redundancy. Among the NMR ensemble models of 2KEJ, model 8, identified by the WHAT IF server (<https://swift.cmbi.umcn.nl/servers/html/index.html>, accessed on 1 September 2015) as the closest to the average NMR structure, was selected as the representative structure. The resulting LacI–O2 complex model was subjected to energy minimization to remove steric clashes introduced during modeling. Energy minimization was performed using 5,000 steps of steepest descent, followed by 10,000 steps of conjugate gradient optimization in UCSF Chimera [78]. In the modeled Lac–O2 complex, chain A interacts with the consensus half-site of the operator, whereas chain B binds the non-consensus half-site of the O2 operator.

The free Lac repressor (operator-unbound state) was generated from the same PDB entry (1EFA) by removing the DNA molecule.

2.1.1. Model Validation of Lac-O2 Repressor-Operator Complex

The structural integrity and reliability of the modeled LacI–O2 repressor–operator complex were evaluated using a multi-criteria assessment to ensure conformational

consistency, stereochemical quality, interface integrity, and structural stability prior to production simulations.

2.1.2. Structural Quality and Stereochemistry

The modeled complex was subjected to stereochemical validation using standard structural assessment tools, including PROCHECK and MolProbity. Ramachandran plot analysis confirmed that the majority of residues occupy favored regions, with no significant population in disallowed conformational space. Side-chain rotamer distributions were within acceptable limits, and no severe steric clashes were observed following energy minimization. Bond length and bond angle deviations were within standard tolerances for high-quality protein structures, indicating that the modeling and minimization procedures did not introduce significant geometric artifacts.

2.1.3. Structural Conservation Relative to Template

To ensure preservation of the canonical LacI fold, the modeled LacI–O2 complex was superimposed onto the reference LacI–O–SymL crystal structure (PDB ID: 1EFA). Backbone root-mean-square deviation (RMSD) analysis of the N-terminal DNA-binding domains yielded values below 0.5 Å, demonstrating retention of the helix–turn–helix architecture and proper positioning of the hinge helices. The core domains also exhibited minimal deviation, confirming that global structural features of the Lac repressor dimer were preserved during model construction. DNA bending and minor-groove geometry remained consistent with known LacI-induced deformation patterns, further supporting the structural and functional validity of the model [79].

2.1.4. Protein–DNA Interface Validation

The structural validity of the modeled LacI–O2 repressor–operator complex was further assessed by comparing hydrogen bonds, salt bridges, and van der Waals contacts at the protein–DNA recognition interface with those observed in the eighth conformer of the NMR ensemble (2KEJ). Key major-groove, base-specific contacts characteristic of LacI recognition were preserved, particularly those involving residues within the helix–turn–helix motif and the hinge-helix region. The asymmetric engagement of the two Lac repressor subunits, in which chain A binds the consensus half-site and chain B interacts with the non-consensus half-site, is structurally consistent with the sequence asymmetry of the O2 operator.

Hybrid modeling strategies are well established for the construction and refinement of biomolecular complexes [57,68].

Following the validated protocol ensures that the observed sequence-dependent conformational differences reflect intrinsic LacI–DNA dynamics rather than modeling artifacts.

2.2. Molecular Dynamics Simulation Protocol

All-atom molecular dynamics simulations were performed using GROMACS 5.1 for the repressor–operator complexes (LacI–O–SymL and LacI–O2), as well as for the unbound Lac repressor protein (Table 1) [80]. Proteins and DNA were parameterized using the AMBER ff99SB-ILDN and AMBER parmbsc1 force fields, respectively — a validated combination that improves protein side-chain dynamics and DNA backbone conformational accuracy.

The AMBER ff99SB-ILDN force field was employed for protein modeling due to its improved backbone and side-chain torsion parameters, which enhance structural accuracy, better reproduce experimental NMR observables, and ensure stability in long-timescale molecular dynamics simulations. DNA was modeled using the AMBER parmbsc1 force field, which corrects known artifacts of earlier DNA force fields, including excessive terminal base-pair fraying, an inaccurate BI/BII backbone equilibrium, and underestimation of DNA helical parameters. This force field enables stable microsecond-scale simulations of both canonical B-DNA and protein-bound DNA systems [41,68,81–85].

For histidine residues, the N ϵ atom was treated as a secondary amino nitrogen, whereas the N δ atom was considered an imino nitrogen. The systems were solvated using the TIP3P water model in a cubic box, maintaining a minimum distance of 1 nm between the solute complex and the box boundaries. The systems were subsequently neutralized by adding the required number of Na⁺ and Cl[−] ions. Periodic boundary conditions were applied in all directions. Energy minimization was performed using the steepest-descent algorithm to alleviate steric clashes. Each system was subsequently equilibrated for 0.5 ns in the NVT ensemble, followed by 1.0 ns in the NPT ensemble at 300 K and 1 atm. Temperature was controlled using the Berendsen thermostat, while pressure was maintained using the Parrinello–Rahman barostat. Lennard–Jones and short-range electrostatic interactions were truncated at 1.0 nm, while long-range electrostatic interactions were treated using the particle mesh Ewald (PME) method. All bonds involving hydrogen atoms were constrained using the LINCS algorithm. Following equilibration, production molecular dynamics simulations were performed for 500 ns with a 2-fs integration time step. Atomic coordinates were recorded every 10.0 ps, resulting in a total of 50,000 trajectory frames for subsequent analysis [86,87].

Table 1: Systems taken for molecular dynamics simulation.

System Number	System Taken for Analysis	Initial Structure	DNA Sequence Used for MD Simulation	Presence of Protein	Duration of MD Simulation	Number of Solvent	Number of Na ⁺	Force Field	Average RMSD in Å with Mean and SD (Within First Bracket)
1	Lactose repressor-Operator SymL	As in PDB ID 1EFA	d(TGTGAGCG CTCACAATT). d(AATTGTGAGC GCTCACA)	Yes	500 ns	51,750	38	Amber ff99SB-ILDN force field for protein and Amber ParmBSC1 force field for DNA	2.0 (0.02)
2	Lactose repressor-operator O2	Model structure	d(AAATGTGA GCGAGTAACAAC C).d(GGTTGTTA CTCGCTCACAT TT)	Yes	500 ns	55,844	50	Amber ff99SB-ILDN force field for protein and Amber ParmBSC1 force field for DNA	3.4 (0.01)
3	Lactose repressor dimer	As in PDB ID 1EFA without DNA	Not Applicable	Yes	500 ns	41,017	6	Amber ff99SB-ILDN force field for protein	5.4 (0.10)

The protein–DNA complex systems (LacI–O-SymL and LacI–O2) comprise 165,000 and 179,000 atoms, respectively. Three independent simulations were performed for each system using different randomized velocity seeds. Each trajectory was propagated for 500 ns under identical equilibration and production conditions. All reported analyses are based on ensemble averages across the three independent molecular dynamics simulation trajectories.

2.3. Analysis of Structure and Conformational Stability

Periodic boundary condition (PBC) artifacts often caused apparent spatial separation of the two DNA strands and the protein subunits during simulations. This issue was consistently observed across all systems and could not be fully resolved using standard trajectory rewrapping procedures, including the “PBC MOL” option in GROMACS or the wrapping functions available in VMD. To address this, each trajectory was converted to PDB format using the GROMACS utility *trjconv*, followed by post-processing with an in-house FORTRAN program to reconstruct continuous protein–DNA complexes for downstream analysis.

Root-mean-square deviation (RMSD) and root-mean-square fluctuation (RMSF) analyses were performed using GROMACS on the corrected trajectories. Protein–DNA hydrogen bonds were identified using a modified version of *pyrHBfind* [88], applying a donor–acceptor distance cutoff of 3.0 Å and a minimum hydrogen bond angle of 150°.

Side-chain χ_1 torsion angle distributions were analyzed using VMD. DNA conformational properties were characterized in terms of inter-base-pair and base-pair step parameters using the NUPARM software package [89].

Structural stability and convergence of the LacI–O-SymL complex were assessed using backbone root-mean-square deviation (RMSD) relative to the corresponding experimental crystal structure, along with root-mean-square fluctuation (RMSF) analysis of protein and DNA backbones. The backbone RMSD values stabilized within 2.0 Å for LacI–O-SymL and 4.0 Å for LacI–O2, respectively. The helix–turn–helix motifs remained stably inserted within the DNA major groove throughout the simulations, with no evidence of unfolding events. The preservation of key protein–DNA contacts, together with maintenance of canonical DNA helical parameters, confirms the structural stability of the LacI–O2 repressor–operator model.

2.4. Conformational Thermodynamics from Histogram-Based Methods (HBM)

The complete protocol of the histogram-based method (HBM) for calculating conformational thermodynamics

from equilibrated molecular dynamics trajectories has been described in previous publications [58–71]. Histograms of χ_1 torsional angles for each residue were generated over a 0°–360° range and discretized into 90 equal bins. These histograms represent the probability distribution of the system’s occupancy of specific conformational states. The corresponding free energy and conformational entropy are then computed using Boltzmann weighting and the Gibbs entropy formalism, respectively. The normalized probability distribution of any conformational variable, such as χ in complex states (i.e., operator bound) and free states (i.e., operator unbound) are represented by $H_i^{\text{complex}}(\chi)$ and $H_i^{\text{free}}(\chi)$, respectively.

The change in free energy associated with the microscopic conformational variable χ_1 upon binding of the repressor protein to operator DNA is defined by the following expression:

$$\Delta G_i^{\text{conf}}(\chi) = -k_B T \ln(H_{\text{max},i}^{\text{complex}}(\chi)/H_{\text{max},i}^{\text{free}}(\chi))$$

where ‘max’ indicates the peak value of the histogram and i represents each of the lac repressor protein residues.

The change in side chain conformational entropy for χ_1 due to binding of the repressor protein with operator DNA is expressed by the following formula:

$$T\Delta S_i^{\text{conf}}(\chi) = -k_B T \left[\sum_j H_i^{\text{complex}}(\chi) \ln H_i^{\text{complex}}(\chi) - \sum_j H_i^{\text{free}}(\chi) \ln H_i^{\text{free}}(\chi) \right]$$

where the sum is taken over all histogram bins j , and i represents each of the amino acid residues of the lac repressor protein. k_B and T indicate the Boltzmann constant and temperature, respectively.

The changes in conformational thermodynamics of Lac repressor chains A and B in the operator-bound state were computed relative to those of chains A and B in the operator-unbound state, respectively.

Conformational thermodynamic quantities were derived from the χ_1 torsional degree of freedom, which captures dominant rotameric transitions of both interface residues and distal residues far from the protein–DNA interface. The χ_1 torsion angle defines the primary rotameric state governing side-chain orientation relative to the backbone and directly modulates intermolecular contacts. Although higher-order side-chain torsion angles (χ_2 , χ_3 , χ_4) and backbone (φ/ψ) dihedrals contribute to finer conformational adjustments, their effects are generally smaller on the sampled timescales or indirectly coupled to χ_1 , consistent with restricted backbone flexibility of the DNA-binding domain.

All reported thermodynamic quantities (ΔG and $T\Delta S$) represent ensemble averages over three independent molecular dynamics trajectories per system, with uncertainties estimated as the standard error of the mean across replicas.

2.5. Principal Component Analysis (PCA) of Molecular Dynamics Simulation Trajectories

Principal component analysis, integrated with molecular dynamics simulations, transforms high-dimensional trajectory data into a reduced set of orthogonal modes capturing the most significant collective motions. This is achieved by constructing and diagonalizing the covariance matrix (C_{ij}) of atomic positional fluctuations, thereby projecting the dataset onto a lower-dimensional subspace that captures the key dynamical features of the protein–DNA complex during the simulation [90–93]. The dominant principal components typically correspond to slow, large-scale conformational transitions, indicating that functionally relevant motions are largely confined to a low-dimensional subspace in which a limited number of eigenvectors account for most of the conformational variance.

The elements C_{ij} in the covariance matrix are defined by $C_{ij} = \langle (x_i - \langle x_i \rangle) (x_j - \langle x_j \rangle) \rangle$ where $\langle x_i \rangle$, $\langle x_j \rangle$ are the mass-weighted mean Cartesian coordinates of the atoms present in the system, and x_i (or x_j) is the coordinate of the i th (or j th) atom of the system, and “ $\langle \rangle$ ” indicates ensemble average. Diagonalization of the covariance matrix yields a set of eigenvectors and corresponding eigenvalues, where the eigenvectors define the principal directions of correlated atomic displacements and the eigenvalues quantify the amplitude of fluctuations along these directions. The principal components (PCs), ranked in order of decreasing eigenvalues, thus capture the dominant motions of the system, with the first few PCs typically representing slow, large-scale collective movements associated with conformational transitions of the complex. The mass-weighted covariance matrix was constructed and diagonalized using the `gmxcovar` module of the GROMACS simulation package, and the resulting eigenvectors were further analyzed using `gmx ana eig`. To reduce computational cost while preserving essential global dynamics, only the $C\alpha$ atoms of the protein and the phosphate atoms of DNA were included in the analysis, as the dimensionality of the covariance matrix scales quadratically with the number of atoms. The dominant collective fluctuations involving multiple regions of the macromolecule correspond to large-scale “breathing” motions. The extreme projections along the first principal component (PC1) were identified to obtain representative conformations corresponding to

minimum and maximum amplitudes (hereafter referred to as PC1-first and PC1-last frames), thereby defining the endpoints of motion along this mode. Porcupine plots were generated using an in-house FORTRAN program, employing the endpoint structures (PDB format) to compute displacement vectors. These vectors represent both the direction and magnitude of atomic displacements associated with the principal mode of motion.

PyMOL, UCSF Chimera, and Visual Molecular Dynamics (VMD) were used for visual inspection of the trajectory files and for image generation. Graphs were generated using Gnuplot, Microsoft Excel, and Python scripts.

3. Results

Representative snapshots from molecular dynamics simulations of the unbound and bound LacI–O-SymL and LacI–O2 complexes are shown in Figure 2.

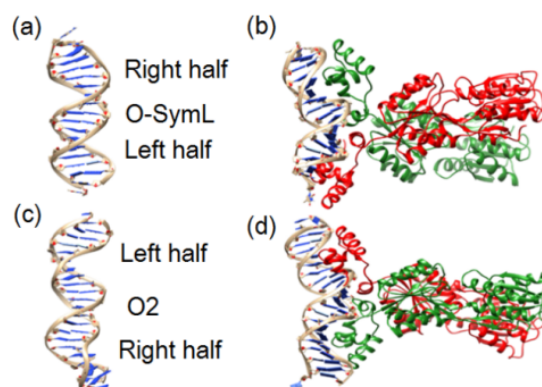


Figure 2: Representative average structures derived from molecular dynamics trajectories of the lac repressor in (a) the unbound O-SymL operator system, (b) the holo state corresponding to the lac repressor bound to O-SymL, (c) the unbound O2 operator system, and (d) the holo state corresponding to the lac repressor bound to O2, Reproduced from Ref. [79].

3.1. Assessment of MD Simulation Convergence

Structural stability and conformational variability of the Lac repressor were assessed by calculating the root-mean-square deviations (RMSDs) of $C\alpha$ atoms for the dimeric protein in both operator-bound and unbound states, relative to their initial structures (Figure S1). RMSD analysis shows that the DNA-bound systems (Lac–O-SymL and Lac–O2) are more stable than the free protein, indicating that DNA binding constrains the Lac repressor structure. All systems reached a plateau after ~100–150 ns, indicating that the simulations were long enough to capture equilibrated LacI dynamics. In contrast, simulations of the free

protein dimer show persistent fluctuations in RMSD over time, reflecting slower relaxation and increased conformational variability in the absence of DNA. This RMSD analysis demonstrates that DNA binding not only stabilizes the LacI conformation but also reveals operator-specific effects, with the synthetic symmetric sequence (O-SymL) inducing a more rigid protein conformation than the native O2 operator. These differences may underlie the variations in binding affinity and specificity observed experimentally. Based on these observations, snapshots from 250 ns to 500 ns were selected for further analyses of DNA-protein interactions, bending, groove dynamics, conformational thermodynamics, and essential dynamics, as all systems reached structural equilibrium by 250 ns.

3.2. Comparative Hydrogen Bonding Interaction Analysis of Lac Repressor Protein with O-Sym and O2 DNA Operators

Comparison of hydrogen-bond occupancies over 250–500 ns MD simulations, using a donor–acceptor distance cut-off ≤ 3.5 Å, angle threshold $\geq 120^\circ$, and considering interactions with $\geq 50\%$ occupancy, reveals pronounced asymmetry in Lac repressor–DNA recognition. Residue numbering follows that of Figure 1. Tables 2 and 3 present the hydrogen-bonding interactions between the Lac repressor and the different DNA operators analyzed in this study.

Despite the intrinsic symmetry of the Lac repressor dimer and the palindromic O-SymL operator, hydrogen-bond occupancy analysis (Table 2) reveals pronounced asymmetry in protein–DNA interactions. The two monomers do not contribute equivalently to DNA recognition, indicating that sequence symmetry does not necessarily

translate into a symmetric interaction landscape at the protein–DNA interface. In the O-SymL complex, Chain A exhibits a combination of base-specific and extensive backbone-mediated interactions, with a pronounced preference for phosphate contacts. Residues such as Ser16A, Thr19A, and Gln54A show highly persistent occupancies, forming a dense interaction network predominantly localized on strand E. Additional contributions from Tyr7A (75%), Asn25A (79%), and Lys-59A (69%) further stabilize the consensus half-site. Notably, Tyr-17A forms a highly persistent hydrogen bond with Gua7D (O6/N7; ~ 78 –100%), while Gln-18A establishes a stable base-specific interaction with Gua-5D (O6; $\sim 95\%$), underscoring their roles in sequence-specific recognition. In contrast, Chain B exhibits a more structured and diverse interaction profile, combining strong base-specific recognition with multiple backbone-mediated contacts. This is dominated by Arg22B, which forms highly stable bifurcated hydrogen bonds with Gua5E (O6/N7; ~ 99 –100%), supported by Tyr17B interactions with Gua7E (O6/N7; ~ 74 –100%), indicative of dominant major groove readout. Additionally, Chain B exhibits extensive backbone interactions, including Leu6B–Cyt13D (O2P; 100%), Ser16B–Gua5E (O2P/O5'; ~ 77 –100%), Thr19B–Gua5E (O2P; 100%), and unique contacts with Thy4E via Ser31B and Thr34B (O2P/O1P; ~ 98 –100%). Furthermore, Chain B engages both DNA strands through Asn50B–Cyt13D (O1P; $\sim 74\%$) and Gln54B–Thy14D (O1P; $\sim 97\%$), indicating a broader and more spatially distributed interaction interface. However, several interactions that would be expected to be symmetric—particularly those involving Gln18—are weakened or absent in chain B, reflecting a breakdown of ideal symmetry under dynamic conditions.

Table 2: Hydrogen bonds between Lac repressor and operator O-SymL, identified from the MD trajectory (250 ns–500 ns) using Amber-ff99SB-ILDN force field for protein and amber ParmBSC1 force field for DNA. For DNA, chain D corresponds to the first strand and chain E corresponds to the second strand.

Donor	Donor Atoms	Acceptor	Acceptor Atoms	% of Occurrence
Leu-6A	N	Cyt-13E	O2P	100
Tyr-7A	OH	Cyt-13E	N4	75
Ser-16A	OG	Gua-5D	O2P	100
Ser-16A	N	Gua-5D	O2P	100
Tyr-17A	OH	Gua-7D	O6	100
Tyr-17A	OH	Gua-7D	N7	78
Gln-18A	NE2	Gua-5D	O6	95
Thr-19A	OG1	Gua-5D	O2P	100
Ser-21A	OG	Thy-14E	O2P	100
Asn-25A	ND2	Thy-14E	O1P	79

Table 2: *Cont.*

Donor	Donor Atoms	Acceptor	Acceptor Atoms	% of Occurrence
Tyr-47A	OH	Cyt-13E	O2P	100
Asn-50A	N	Cyt-13E	O5'	98
Gln-54A	NE2	Thy-14E	O1P	100
Lys-59A	NZ	Cyt-15E	O1P	69
Leu-6B	N	Cyt-13D	O2P	100
Ser-16B	OG	Gua-5E	O2P	100
Ser-16B	N	Gua-5E	O2P	100
Ser-16B	OG	Gua-5E	O5'	77.29
Tyr-17B	OH	Gua-7E	N7	74
Tyr-17B	OH	Gua-7E	O6	100
Thr-19B	OG1	Gua-5E	O2P	100
Arg-22B	NH2	Gua-5E	O6	100
Arg-22B	NH1	Gua-5E	N7	99
Ser-31B	N	Thy-4E	O2P	100
Thr-34B	OG1	Thy-4E	O1P	98
Asn-50B	ND2	Cyt-13D	O1P	74
Gln-54B	NE2	Thy-14D	O1P	97

Table 3: Hydrogen bonds between the Lac repressor and operator O2, identified from the MD trajectory (250 ns–500 ns) using Amber-ff99SB-ILDN force field for protein and Amber ParmBSC1 force field for DNA. For DNA, chain D corresponds to the first strand and chain E corresponds to the second strand.

Donor	Donor Atoms	Acceptor	Acceptor Atoms	% of Occurrence
Leu-6A	N	Cyt-13E	O2P	100
Tyr-7A	OH	Cyt-13E	N4	98.50
Ser-16A	N	Gua-5D	O2P	100
Ser-16A	OG	Gua-5D	O2P	100
Tyr-17A	OH	Gua-7D	O6	100
Tyr-17A	OH	Gua-7D	N7	80.30
Thr-19A	OG1	Gua-5D	O2P	100
Ser-21A	OG	Thy-14E	O2P	88
Arg-22A	NH1	Gua-5D	N7	79.70
Arg-22A	NH2	Gua-5D	O6	86.26
Asn-25A	ND2	Thy-14E	O1P	81
Asn-25A	ND2	Thy-14E	O2P	77
Ser-31A	N	Thy-4D	O2P	100
Thr-34A	OG1	Thy-4D	O1P	99
Tyr-47A	OH	Cyt-13E	O2P	77
Asn-50A	N	Cyt-13E	O1P	94

Table 3: *Cont.*

Donor	Donor Atoms	Acceptor	Acceptor Atoms	% of Occurrence
Asn-50A	N	Cyt-13E	O5'	75
Gln-54A	NE2	Thy-14E	O1P	95
Lys-59A	NZ	Thy-14E	O1P	98
Gln-60A	N	Cyt-15E	O1P	97
Arg-118A	NH1	Gua-11D	O1P	65
Arg-118A	NH2	Gua-11D	O1P	54
Leu-6B	N	Ade-12D	O2P	99
Ser-16B	N	Gua-5E	O2P	100
Ser-16B	OG	Gua-5E	O2P	94
Tyr-17B	OH	Gua-13D	N7	64
Gln-18B	NE2	Thy-6E	O4	90
Thr-19B	OG1	Gua-5E	O2P	100
Arg-22B	NH1	Gua-5E	N7	100
Arg-22B	NH2	Gua-5E	O6	100
Asn-25B	ND2	Gua-13D	O2P	66
Ser-31B	N	Thy-4E	O2P	100
Ser-31B	OG	Thy-4E	O1P	74
Thr-34B	OG1	Thy-4E	O1P	93
Tyr-47B	OH	Ade-12D	O1P	100
Asn-50B	N	Ade-12D	O1P	98
Gln-54B	NE2	Gua-13D	O1P	89

In contrast, the O2 complex exhibits a heterogeneous and adaptive interaction network (Table 3), driven by intrinsic sequence asymmetry. In Chain A, the interaction pattern is dominated by a highly stabilized and functionally coherent network that combines strong phosphate backbone anchoring with selective base readout on strand E. Residues Ser16A, Thr19A, Ser31A, and Thr34A form near-persistent contacts with phosphate groups (O2P/O1P; ~99–100%), establishing a rigid electrostatic scaffold that stabilizes DNA binding. Superimposed on this backbone framework, Tyr7A (98.5%) contributes a strong base-specific contact with Cyt-13E (N4), while Tyr-17A maintains fully persistent interactions with Gua7D (O6/N7; ~80–100%), indicating conserved major groove recognition. Additionally, Arg22A provides bifurcated base readout through stable interactions with Gua-5D (N7/O6; ~79–86%), and Asn25A engages Thy14E via dual phosphate contacts (O1P/O2P; ~77–81%), further reinforcing half-site specificity.

Supporting stabilization is further provided by Gln54A (95%) and Lys-59A (98%) through interactions with Thy14E, as well as Gln60A (97%) engaging Cyt15E,

collectively defining a dense and highly cooperative interaction network. Notably, Arg118A contributes weaker yet persistent long-range phosphate contacts with Gua11D, suggesting auxiliary stabilization beyond the core recognition interface. On the other hand, Chain B exhibits a more distributed yet highly stable interaction architecture, integrating robust base recognition with extensive backbone engagement across both DNA strands. Arg22B serves as the principal specificity determinant, forming fully persistent bidentate hydrogen bonds with Gua5E (N7/O6; 100%). This interaction is strongly supported by Thr19B (100%) and Ser16B (94–100%) contacts with the phosphate backbone, thereby reinforcing recognition of strand E. Additional sequence readout is mediated by Gln18B, which interacts with Thy6E (O4; 90%), and Tyr17B, which forms a weaker yet persistent contact with Gua13D (N7; 64%), indicating asymmetry in major groove engagement relative to Chain A. Backbone stabilization is further extended through Ser31B and Thr34B interactions with Thy4E (O2P/O1P; 74–100%), while Tyr47B and Asn50B establish highly persistent phosphate contacts

with Ade12D (100% and 98%, respectively), suggesting a compensatory stabilization strategy on strand D. Moreover, Gln54B contributes additional phosphate recognition with Gua13D (89%), reinforcing cross-strand stabilization.

Thus, the two subunits differ not only in interaction strength but also in residue utilization and strand preference. Gln60A forms a highly stable hydrogen bond with Cyt-15E (~97%), while Arg118A contributes additional phosphate backbone contacts in non-consensus regions. These interactions collectively extend the recognition interface beyond the helix–turn–helix (HTH) motif, incorporating residues such as Tyr47 and Asn50, thereby enhancing the adaptability of DNA binding. The absence of a central GC base pair in O-SymL and the intrinsic asymmetry of the O2 operator result in distinct binding modes and interaction distributions. While O-SymL exhibits partial symmetry breaking at the dynamic level, O2 displays a more flexible and distributed interaction network that accommodates sequence variability through compensatory stabilization mechanisms. Collectively, these findings demonstrate that Lac repressor–DNA recognition is governed by stable phosphate backbone anchoring, sequence-specific base readout, and dynamically redistributed auxiliary contacts. The results underscore the critical role of DNA sequence context in shaping protein–DNA interaction landscapes beyond static structural models, highlighting the importance of dynamic reorganization in molecular recognition.

3.3. Modulation of Protein Side-Chain Dynamics upon Operator DNA Binding

Figure S2 shows that DNA binding differentially affects C α fluctuations in the two Lac repressor subunits, indicating unequal chain stabilization upon operator binding. In the Lac-SymL complex, RMSF analysis of DNA indicates minimal fluctuations in bases forming hydrogen bonds with amino acids, while enhanced dynamics are observed in residues His29, Ser31, Ala32, Asn46, Gly58, Lys59, Gln60, Leu61, Arg101, Ser102, Gly103, Asn142, Gly236, Ile237, Gly310, Gln311, Ala312, Val313, and Lys314. Notably, the most pronounced increase in dynamicity upon binding of both SymL and O2 operators is observed in hinge, turn, and coil regions. In the unbound protein, loop regions exhibit high mobility, which is substantially reduced upon DNA binding, indicating stabilization induced by protein–DNA interactions. Both monomers insert the hydrophobic side chain of Leu56 into the DNA minor groove, resulting in the unstacking of two contiguous central CpG base pairs and the induction of a DNA kink. This van der Waals interaction re-

mains stable throughout the simulations. RMSF values for Leu56 are reduced in the DNA-bound state (SymL-bound: chain A, 1.029 Å; chain B, 1.0 Å; O2-bound: chain A, 0.943 Å; chain B, 0.848 Å) compared with the free protein (chain A, 1.57 Å; chain B, 1.62 Å), confirming stabilization of the kink region. Residues in the N- and C-terminal core domains of the Lac–O2 complex, particularly in chain A (residues ~75–230), exhibit enhanced fluctuations, whereas the Lac–SymL complex shows no significant changes in dynamics. In addition, asymmetric dynamic behavior in the operator-unbound protein is observed in residues 285–310, which are distal from the DNA-binding interface. These findings indicate that O2 binding induces a distinctly different dynamic behavior in chain A compared to SymL, suggesting that sequence-dependent differences in overall protein dynamics may play a critical role in coordinating the C-terminal domains during tetramerization.

To assess whether the observed asymmetric conformational dynamics of the Lac repressor are robust and not artifacts of force field selection, molecular dynamics simulations were repeated using an alternative protein force field (AMBER ff14SB), in addition to AMBER ff99SB-ILDN. Residue-wise C α fluctuations of Lac repressor chains A and B were systematically compared for both the O2- and O-SymL-bound complexes using two different force fields (**Figure S3**). The RMSF profiles of the O-SymL- and O2-bound Lac repressor complexes reveal clear operator-dependent modulation of protein dynamics, and this trend remains consistent across the force fields examined.

In the Lac–O–Sym repressor–operator complex, fluctuations are moderately distributed, with distinct peaks localized to the N-terminal DNA binding domain (residues 1–62) and a pronounced enhancement near the hinge region (residues 50–62), while the core domain displays comparatively restrained yet spatially heterogeneous motions. Both monomers exhibit closely comparable fluctuation amplitudes, suggesting a symmetric dynamic response upon binding to the symmetric operator. In contrast, the Lac–O2 repressor–operator complex displays elevated RMSF values within the DNA-binding domain, particularly in one monomer, indicating increased local flexibility and pronounced dynamic asymmetry. The hinge region of the Lac–O2 system remains dynamically prominent but exhibits less sharply defined, broader peaks, consistent with reduced conformational constraint relative to the Lac–O–Sym system. Beyond residue 100, the core domain of the Lac–O2 repressor–operator complex displays generally lower flexibility, with only a few localized peaks at specific positions, suggesting a sequence-

dependent redistribution of flexibility rather than uniform or global conformational stabilization.

Agreement between two independently parameterized protein force fields demonstrates that the observed conformational dynamics are robust, reproducible, and not artifacts of force-field choice. This concordance therefore reinforces the validity of the mechanistic interpretations derived from the molecular dynamics simulations.

3.4. Conformational Thermodynamics of Binding from Side-Chain χ_1 Dihedral Angle Histograms

To assess residue-level conformational fluctuations of the Lac repressor, histograms of the side-chain dihedral angle χ_1 were generated for all residues from MD trajec-

ries in both protein-bound and unbound states. Conformational analysis was performed over the 250–500 ns interval of each trajectory, during which χ_1 torsional angle distributions remained stationary and side-chain rotamer transitions converged to stable probability distributions.

Representative distributions (Figure 3a,b) display both multimodal and unimodal peaks, reflecting distinct rotameric states and variable side-chain flexibility across residues. Sharper unimodal peaks are observed for His29 in chain A (O2 complex), Asp149 in chain B (free O2 system), Asp219 in chain A (free O-Sym system), and Ser224 in chain A (free O-Sym system), indicating comparatively higher rigidity at these positions. In contrast, bimodal distributions, such as those for chain B Lys84 (free, O-Sym) and chain B Asp149 (O2 complex), reveal increased flexibility.

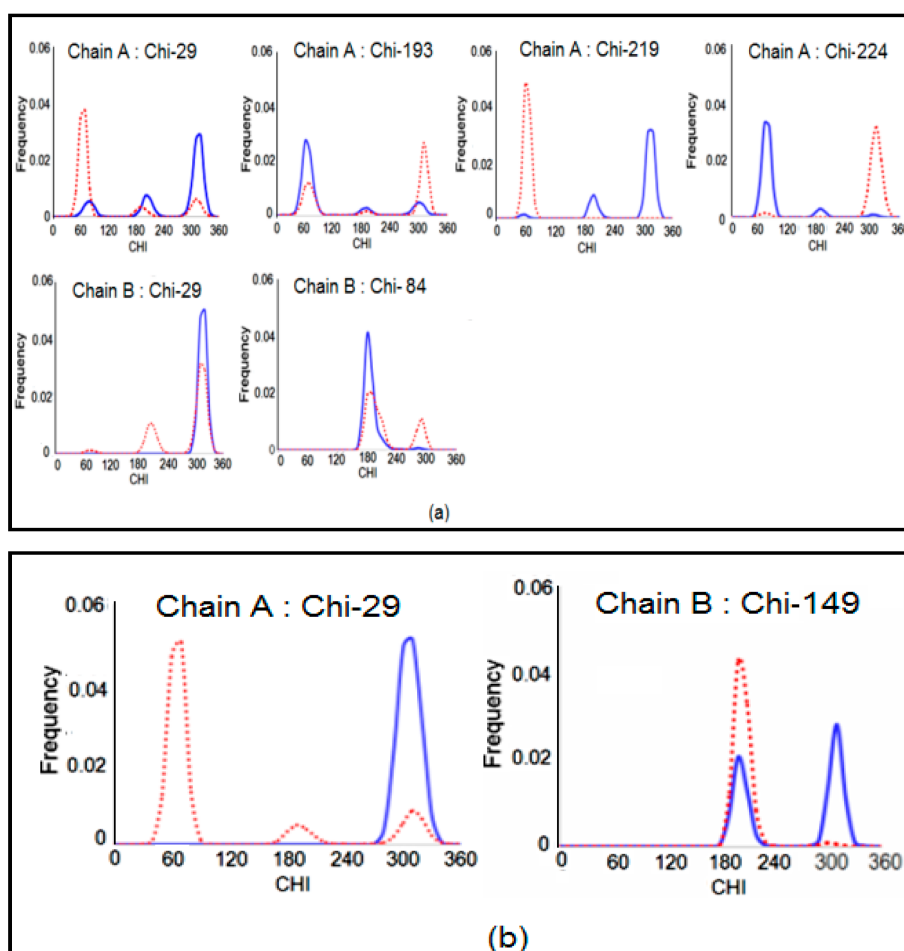


Figure 3: (a): Multimodal histogram distributions of the side-chain χ_1 torsion angles for selected residues of the Lac repressor dimer in the absence and presence of operator DNA (O-Sym). Chain A residues include His29, Ser193, Asp219, and Ser224, while chain B includes His29 and Lys84. The apo state is represented by red dotted lines, and the DNA-bound holo state by blue lines. (b) χ_1 side-chain torsion angle distributions highlighting distinct rotameric states in the Lac repressor bound to operator DNA (O2), shown for His29 in chain A and Asp149 in chain B (Reproduced from [79]).

The comparatively sharp and narrow peaks suggest constrained conformational dynamics, leading to a more rigid and well-ordered conformational ensemble. In contrast, broadened peaks are indicative of pronounced conformational heterogeneity and enhanced dynamic fluctuations. The presence of bimodal distributions further suggests interconversion between two distinct conformational states, indicating the coexistence of or dynamic exchange between discrete isomeric conformations separated by an energy barrier. In contrast, sharp unimodal peaks are consistent with the dominance of a single energetically favorable conformer, reflecting conformational stabilization and reduced stochastic fluctuations.

Overall, variations in peak number, height, and width demonstrate that Lac repressor exhibits sequence-specific, site-dependent differences in side-chain conformational fluctuations between free and operator-bound states.

Based on these observations, residue-wise and domain-wise changes in conformational free energy (ΔG) and entropy ($T\Delta S$) were computed using a histogram-based method (HBM), which sums χ_1 contributions across residues within defined regions (Figure S4a–d). Negative ΔG and $T\Delta S$ values correspond to conformational stabilization and increased order upon DNA binding, whereas positive values indicate destabilization and enhanced disorder. The results indicate that binding to both O-Sym and O2 induces conformational stabilization in specific residues, particularly within regions proximal to the DNA-binding interface.

Enhanced segmental motions are observed in the core domains (N- and C-terminal subdomains) of the Lac repressor in the O2-bound state compared with O-Sym, particularly near the protein–protein interface. These motions are more prominently reflected in conformational entropy changes (Figure S4b) than in free-energy changes (Figure S4a), highlighting the increased flexibility of distal regions induced by operator-specific binding. Table S1 quantitatively illustrates how O-Sym and O2 differentially modulate residue-specific conformational thermodynamics across secondary structural elements of the homodimeric Lac repressor. Collectively, these analyses demonstrate that side-chain conformational dynamics and thermodynamic responses of the Lac repressor are highly dependent on the bound DNA sequence, with O2 promoting increased distal flexibility, whereas O-Sym induces more localized stabilization.

Two helices ($\alpha 2$ helix of lac repressor (O-Sym bound) and $\alpha 4$ helix lac repressor (both O-Sym and O2 bound)) as well as two loops (loop3 of chain B lac repressor (O2 bound) $\Delta G = -7.46$ kJ/mol, $T\Delta S = -10.74$ kJ/mol and loop5 chain A lac repressor (operator O-Sym

bound) $\Delta G = -7.21$ kJ/mol, $T\Delta S = -2.23$ kJ/mol and (operator O2 bound) $\Delta G = -6.14$ kJ/mol, $T\Delta S = -10.33$ kJ/mol) of DNA binding domain undergo significant energetically stabilized and entropically ordered state upon binding to operators due to formation of an array of hydrogen bonding interaction. The $\alpha 2$ helix exhibits conformational free-energy changes of -6.40 kJ/mol and -10.83 kJ/mol for chains A and B of the Lac repressor in the O-SymL-bound state, respectively, with corresponding entropy contributions of -22.04 kJ/mol and -15.09 kJ/mol, respectively. For the $\alpha 4$ helix, conformational free-energy changes are -14.58 kJ/mol and -3.38 kJ/mol for chains A and B of the Lac repressor in the O-SymL-bound state, respectively, and -8.87 kJ/mol and -10.01 kJ/mol for chains A and B in the O2-bound state, respectively. The corresponding conformational entropy changes for the $\alpha 4$ helix are -11.91 kJ/mol and -23.55 kJ/mol for chains A and B in the O-SymL complex, and -19.17 kJ/mol and -19.71 kJ/mol for chains A and B in the O2 complex, respectively. Most N-terminal residues of both dimers in the Lac repressor–O-SymL complex (DNA-binding domain) show reductions in both free energy and conformational entropy, consistent with stable sequence-specific hydrogen-bonding interactions with DNA. In contrast, a subset of residues in the DNA-binding loop region (Pro3, Gln26, Ser31, and Ile48) exhibit increased conformational entropy in the operator-bound state. Some of the residues located in core domain (the N and C-terminal sub domains) of different chains of lac repressor (Asp149, Thr154 and Asn157 of loop12), (Asp219 of loop16), (237 Ile of loop17), (Val111, Leu114, Leu115 of α -Helix6), (163His, 165Asp, 171Val of α -Helix8), (Val192, Arg197, Tyr204 of α -Helix9), (Leu296, Gln298, Thr299, Ser300, Leu304 of α -Helix13), (Thr68 of β -Sheet 1), (Asn125 of β -Sheet 3) have enhanced fluctuation in terms of conformational entropy around the protein–protein interaction site (Figure S4b).

A pronounced increase in fluctuations is observed for N-terminal DNA-binding residues (Tyr7, Ser21, Val24, Met42, Ile48) in the O2-bound complex. Asymmetric motional freedom is particularly evident in the LacI–O2 complex. In the operator-bound state, increased conformational entropy is observed for numerous residues within the N- and C-terminal core domains of both protein chains, particularly at or near the protein–protein interaction interface. Despite being a homodimer with an identical primary sequence, the Lac repressor exhibits asymmetric motional freedom between its subunits. Significant disorder is observed in α -helix 6 (chain B of the Lac repressor in the O-Sym-bound state), α -helix 8 (chains A and B in both O-Sym- and O2-bound states), α -helix 9 (chain A in both O-Sym- and O2-bound states), and α -helix 13

(chain A in the O-Sym-bound state). Additional disorder is present in β -sheet 1 (chain B in both O-Sym- and O2-bound states), β -sheet 3 (chain B in both O-Sym- and O2-bound states), loop 12 (chain A in the O2-bound state), and loop 16 (chain A in both O-Sym- and O2-bound states), indicating region-specific conformational flexibility across different secondary structural elements. A significant reduction in the flexibility of DNA-contacting side chains is observed upon complex formation. Most N-terminal residues of both dimers in the Lac repressor–O-SymL complex exhibit decreased conformational entropy as a result of additional interactions with the DNA. Residue-wise conformational entropy changes indicate increased ordering of residues that form stable sequence-specific hydrogen bonds with DNA. Further investigation of domain motions and the essential dynamics of the protein subunits may provide additional insight into this behavior. A few residues within the DNA-binding region also exhibit increased conformational entropy in the operator-bound state. These observations suggest that certain residues located at or near the protein–protein and protein–DNA interfaces acquire enhanced mobility upon DNA binding. **Figure S4c** shows that the N-terminal DNA-binding domains of both chain A and chain B of the Lac repressor are stabilized and ordered upon binding to both O-Sym and O2 operators. However, the conformational free energy (ΔG) and entropy ($T\Delta S$) changes in the N-terminal DNA-binding domain of the O2-bound complex differ substantially between the two chains, with chain A showing $\Delta G = -26.72$ kJ/mol and $T\Delta S = -45.63$ kJ/mol, and chain B showing $\Delta G = -3.51$ kJ/mol and $T\Delta S = -20.40$ kJ/mol. In comparison, the O-Sym-bound complex exhibits larger magnitude changes, with chain A displaying $\Delta G = -37.0$ kJ/mol and $T\Delta S = -40.67$ kJ/mol, and chain B showing $\Delta G = -19.01$ kJ/mol and $T\Delta S = -57.44$ kJ/mol, indicating pronounced chain-dependent differences in conformational thermodynamics upon operator binding. **Figure S4d** shows the thermodynamics of conformational changes in the combined N- and C-terminal subdomains of chain A of the Lac repressor in the O2-bound state, which is destabilized ($\Delta G = 5.3$ kJ/mol) and disordered ($T\Delta S = 7.68$ kJ/mol). In contrast, chain B of the Lac repressor in the O2-bound complex is stabilized ($\Delta G = -25.77$ kJ/mol) and ordered ($T\Delta S = -18.21$ kJ/mol). On the other hand, the N- and C-terminal subdomains of chain A in the Lac repressor–O-SymL complex are destabilized ($\Delta G = 12.61$ kJ/mol) but exhibit increased ordering ($T\Delta S = -11.66$ kJ/mol). In contrast, the corresponding subdomains of chain B are marginally stable ($\Delta G = -2.07$ kJ/mol) yet significantly disordered ($T\Delta S = 29.48$ kJ/mol).

Amino acids of lac repressor protein are classified into polar (Asn, Cys, Gln, Ser, Thr, and Tyr), hydrophobic

(Ala, Gly, Ile, Leu, Met, Phe, Pro, Trp, Val), acidic (Asp and Glu), and basic (Arg, Lys, and His) based on chemical nature. The contributions from four different amino acid groups to the changes in conformational thermodynamics of Lac repressor chains A and B in the operator-bound state, relative to chains A and B in the free state, are shown in **Figure S4e–h**.

Acidic amino acid residues contribute unfavorably in the O-Sym bound Lac repressor. Both chains of O-Sym bound Lac repressor show positive ΔG . Chain A of the Lac repressor in the O-SymL-bound complex is strongly destabilized, while chain B is marginally energetically destabilized. In both chains in the O-Sym bound Lac repressor, positive $T\Delta S$ indicates entropy gain, leading to overall destabilization. In the O2 bound Lac repressor, a clear asymmetry is observed. Chain A of the O2-bound complex remains weakly unfavorable with moderate positive ΔG and strongly positive $T\Delta S$, whereas chain B of the O2-bound complex exhibits negative ΔG with strongly negative $T\Delta S$. Basic amino acids induce marginal conformational destabilization and increased disorder in both chain A and chain B of the Lac repressor in the O-Sym-bound state, whereas they produce pronounced conformational destabilization and disorder in chain A of the Lac repressor in the O2-bound state. However, significant conformational stabilization and ordering of chain B in the Lac repressor–O2 complex is driven by contributions from both acidic and basic amino acids, indicating enthalpy-driven stabilization accompanied by conformational ordering. Hydrophobic amino acids contribute to significant destabilization and increased disorder in chain B of the Lac repressor in the O2-bound state, whereas chains A and B of the O-SymL-bound complex, as well as chain A of the O2-bound complex, are significantly stabilized and exhibit increased conformational ordering. Chain A of the Lac repressor in the O2-bound state, is stabilized (negative ΔG) with significant entropy loss (negative $T\Delta S$), indicating ordering upon binding, whereas chain B is destabilized (positive ΔG) despite a favorable entropy gain. This anti-correlated response suggests asymmetric burial and solvent exposure of hydrophobic patches, reflecting chain-specific packing differences. Polar amino acid residues from both Lac repressor chains A and B in the O-SymL- and O2-bound states contribute most significantly to conformational stability and ordering, suggesting that electrostatic interactions play a dominant role in the changes in conformational thermodynamics associated with repressor–operator complex formation. The change in conformational free energy for Lac repressor chains A and B in the O-SymL-bound state is -40.13 kJ/mol and -16.13 kJ/mol, respectively, whereas for the O2-bound system it is -23.67 kJ/mol and

−31.32 kJ/mol, respectively. These differences primarily arise from contributions of polar residues.

On the other hand, the conformational entropy cost for Lac repressor chains A and B in the O-SymL-bound state is −42.14 kJ/mol and −30.81 kJ/mol, respectively, while the change in conformational entropy for chains A and B in the O2-bound state is −49.01 kJ/mol and −31.23 kJ/mol, respectively, with these contributions primarily arising from polar residues. It is essential to mention that acidic and basic amino acids play the major role in the conformational stability and order of the Lac repressor chain B (operator O2 bound) [$\Delta G = -3.33$ kJ/mol, $T\Delta S = -13.57$ kJ/mol corresponding to acidic residues; $\Delta G = -9.45$ kJ/mol, $T\Delta S = -15.71$ kJ/mol corresponding to basic residues] instead of Lac repressor chain A (operator O2 bound) [$\Delta G = 3.72$ kJ/mol, $T\Delta S = 12.41$ kJ/mol corresponding to acidic residues; $\Delta G = 7.10$ kJ/mol, $T\Delta S = 11.31$ kJ/mol corresponding to basic residues] and chain A [$\Delta G = 13.51$ kJ/mol, $T\Delta S = 7.27$ kJ/mol corresponding to acidic residues; $\Delta G = 2.0$ kJ/mol, $T\Delta S = 2.24$ kJ/mol corresponding to basic residues] and chain B [$\Delta G = 0.81$ kJ/mol, $T\Delta S = 10.04$ kJ/mol corresponding to acidic residues; $\Delta G = 2.92$ kJ/mol, $T\Delta S = 0.77$ kJ/mol corresponding to basic residues] (operator O-Sym bound). In contrast, hydrophobic amino acids contribute to conformational stabilization and ordering of the Lac repressor. This effect is observed in chain A of the O-Sym-bound complex ($\Delta G = 0.45$ kJ/mol, $T\Delta S = -18.36$ kJ/mol), chain B of the O-Sym-bound complex ($\Delta G = -8.68$ kJ/mol, $T\Delta S = -9.06$ kJ/mol), and chain A of the O2-bound complex ($\Delta G = -8.59$ kJ/mol, $T\Delta S = -16.26$ kJ/mol). In contrast, chain B of the O2-bound complex exhibits destabilization, as indicated by positive values ($\Delta G = 14.80$ kJ/mol, $T\Delta S = 14.82$ kJ/mol), suggesting sequence- and chain-dependent modulation of conformational stability. Therefore, in addition to electrostatic interactions, van der Waals interactions also play a significant role in governing the thermodynamics of conformational changes in the Lac repressor protein in both O-SymL- and O2-bound states relative to the free Lac repressor.

3.5. Domain Motions and Conformational Transitions Revealed by PCA

Correlated motions govern energy transport and allosteric signaling in biomolecules. Since many functional processes are driven by slow, large-scale conformational rearrangements rather than local thermal fluctuations, the conformational transitions of the Lac repressor and its DNA-bound complexes were analyzed by projecting the molecular dynamics trajectories onto a two-dimensional space

defined by the first two principal components (PC1 and PC2), with PC1 capturing the maximum variance in the dataset.

Porcupine plots were used to visualize subdomain motions in the Lac–O–SymL and Lac–O2 protein–DNA complexes, as well as in the unbound Lac repressor (Figure 4). For the free Lac repressor, the first principal component is consistent with the RMSF analysis, revealing that the most pronounced motions occur in the helix–turn–helix motif of the DNA-binding domains in both monomers. In the Lac–O–SymL complex, the regulatory domain of monomer B exhibits greater overall motion than monomer A, with the largest displacements occurring in loops connecting α -helices and the DNA-binding domain. In both the Lac–O–SymL and Lac–O2 complexes, the dominant motions are primarily localized to loop and turn regions, whereas subunit A engages the left operator site and subunit B binds the right operator site.

Analysis of domain motions and essential dynamics reveals asymmetric behavior between the two monomers. Notably, binding to the O2 operator enhances fluctuations in both the N-terminal and C-terminal core domains, and the protein subunits show a tendency to move away from each other. In contrast, in the operator-unbound protein, major dynamics are largely confined to the DNA-binding domain, with minimal motion observed in other regions.

Essential dynamics extracted from the molecular dynamics trajectories highlight large-scale correlated motions in the protein–DNA complexes, filtering out recurrent fluctuations to reveal the dominant structural transitions. The intrinsic asymmetry of the two half-sites in the natural O2 operator, together with the absence of the central GC base pair in the symmetric synthetic SymL operator, suggests significant differences in the principal modes of motion of the dimeric Lac repressor upon binding to these DNA sequences.

In summary, porcupine plot analysis demonstrates that DNA sequence allosterically modulates protein dynamics, extending its influence to domains distal from the DNA-binding site. How these allosteric effects relate to gene regulation and how they are transmitted throughout the protein remains an open question, underscoring the complex interplay between DNA sequence and protein dynamics.

Analysis of domain motions and essential dynamics indicates that the two Lac repressor monomers exhibit asymmetric dynamics, strongly influenced by the bound DNA sequence. When bound to the natural O2 operator, fluctuations increase not only in the DNA-binding domain but also propagate to the N- and C-terminal regions of the core domain, with the subunits exhibiting a tendency to move slightly apart. This suggests that DNA binding in-

duces long-range dynamic effects, modulating structural regions distant from the protein–DNA interface. In contrast, the operator-free protein exhibits motions largely confined to the DNA-binding domain, with minimal fluctuations elsewhere in the structure, indicating that DNA binding is required to trigger extensive allosteric dynamics. By focusing on the dominant collective modes, it becomes evident that the asymmetry of the natural O2 operator, particularly the intrinsic differences between its two

half sites, induces distinct principal motions in each Lac repressor monomer. In contrast, binding to the symmetric synthetic operator SymL, which lacks the central GC base pair, produces a more uniform dynamic response across the protein. These findings demonstrate that subtle differences in DNA sequence can exert substantial effects on both the pattern and magnitude of protein motions, highlighting the critical role of sequence in shaping allosteric communication pathways.

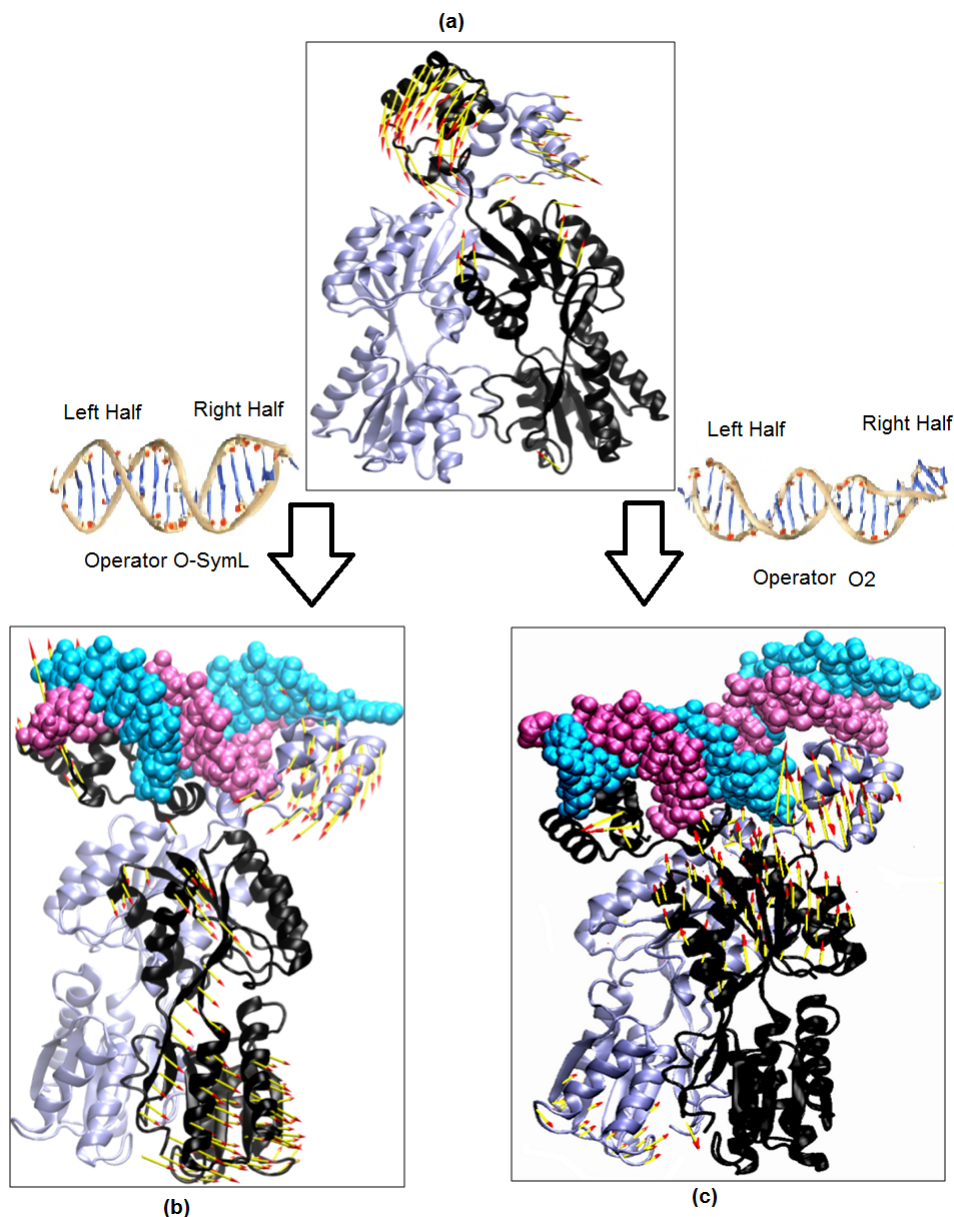


Figure 4: Porcupine plots representing the first principal component (PC1) of motion for (a) free Lac repressor protein, (b) operator SymL-bound Lac repressor protein, and (c) operator O2-bound Lac repressor protein. DNA chains are shown as Chain C (cyan) and Chain D (mauve), while protein chains are shown as Chain A (ice blue) and Chain B (black). Arrows indicate the direction and magnitude of collective motions. MD simulations were performed using the Amber ff99SB-ILDN force field for the protein and the Amber ParmBSC1 force field for DNA (Modified from author's preprint [79]).

Allostery refers to the process by which ligand binding at one site of a protein induces conformational or dynamic changes at a distal site, thereby modulating protein function. In this study, porcupine plot and essential dynamics analyses demonstrate that binding of DNA operators (O-SymL or O2) at the canonical DNA-binding site of the Lac repressor induces asymmetric motions and enhanced fluctuations in regions distant from the binding interface, including the regulatory domain and the N-terminal and C-terminal core domains. These changes are sequence-dependent, with different operators producing distinct patterns of subunit-specific dynamics. This provides robust evidence of allostery, as sequence-dependent perturbations at the DNA-binding interface are transmitted through the LacI dimer to modulate the dynamics of distal regions, resulting in non-equivalent behavior between the two otherwise symmetric monomers.

This sequence-dependent modulation of global dynamics provides mechanistic insight into the differential stabilization of gene regulatory networks, in which operators act as active allosteric effectors that modulate inter-domain coupling and conformational thermodynamics. This, in turn, influences transcriptional repression strength and the conformational plasticity of the LacI regulatory complex.

The porcupine plot derived from principal component analysis of the molecular dynamic trajectory of the operator-unbound Lac repressor protein reveals pronounced collective motions primarily localized within the N-terminal DNA-binding domains, suggesting exploration of multiple conformational substates consistent with a heterogeneous free-energy landscape. The dominant eigenvector (PC1) captures large-amplitude collective atomic fluctuations of the helix–turn–helix motifs, indicating substantial conformational flexibility in the absence of operator DNA. In contrast, the core regulatory domains exhibit comparatively smaller eigenvector components, indicating a relatively stable protein scaffold. The DNA-binding and core domains display coordinated, directionally consistent displacement patterns. Such anti-correlated motions are hallmarks of hinge-like or twisting dynamics. This intrinsic flexibility is mechanically consistent with facilitated target search, where conformational pre-equilibration reduces the energetic penalty associated with structural rearrangement upon operator binding.

4. Discussion

The present study demonstrates that the conformational thermodynamic landscape of the symmetric Lac repressor is reshaped in an asymmetric, operator-dependent manner,

highlighting the strong coupling between local side-chain fluctuations and long-range allosteric communication.

The symmetric O-SymL operator induces pronounced and balanced conformational stabilization, with reduced Lac repressor flexibility at the recognition interface, consistent with stable sequence-specific hydrogen-bonding networks. In contrast, the natural O2 operator elicits a heterogeneous response, with localized stabilization accompanied by increased conformational entropy in the N- and C-terminal core subdomains and regions near the protein–protein interface. Polar residues contribute to conformational stabilization via favorable electrostatic and hydrogen-bonding interactions, while acidic and basic residues introduce monomer-specific asymmetry in the LacI–O2 repressor–operator complex. Hydrophobic residues provide additional van der Waals stabilization in the LacI–O-SymL repressor–operator complex and selectively in one monomer of the O2-bound Lac repressor dimer, suggesting operator-dependent modulation of side-chain burial and the underlying enthalpy–entropy balance.

Principal component analyses support these thermodynamic findings, revealing operator-specific patterns of collective motion. In the LacI–O2 complex, the dominant eigenmodes exhibit amplified fluctuations that propagate into the distal N- and C-terminal core domains and accentuate asymmetric motions between the two subunits of the Lac repressor protein.

The dominant collective motion of the Lac repressor–O-SymL complex is anisotropic and asymmetric. One monomer exhibits larger displacement vectors, especially in the core and DNA-binding domains, while the partner monomer shows comparatively restricted mobility. This disparity demonstrates that the two subunits contribute unequally to the principal mode of motion, despite the dimer’s inherent structural symmetry. In contrast, the free Lac repressor shows fluctuations predominantly confined to the DNA-binding domain, particularly the helix–turn–helix motif, with minimal propagation into distal domains.

Each Lac operator has a distinct base sequence and binding affinity. These differences alter protein–DNA interaction patterns, including hydrogen bonding, electrostatic interactions, and van der Waals contacts, which modulate local side-chain flexibility, inter-residue coupling, collective domain motions, and overall energetic stabilization.

Previous studies have shown that structural perturbations in DNA can propagate long-range signals and influence distal regulatory sites [94–97].

This study extends current understanding by demonstrating that DNA sequences act as active allosteric effectors of the Lac repressor through selective reweighting of

conformational entropy, rather than serving merely as passive binding substrates.

Asymmetry within the homo-dimeric LacI has been reported previously in structural studies, mutational analyses, and experimental thermodynamic investigations [33, 34].

Rastinejad et al. (1993) demonstrated that the intrinsic sequence asymmetry of the natural Lac operator governs an asymmetric pattern of contacts with the DNA-binding domains of the Lac repressor. The protein interacts symmetrically only when the intrinsic asymmetry at both half-sites is eliminated. These findings emphasize that deviations from a perfect palindromic operator sequence are functionally important in specifying LacI–DNA binding geometry in long-range regulatory contexts. Frank et al. (1997) demonstrated that operator sequence variation from consensus symmetric Lac operator to variant or non-specific sequences alters binding affinity by several orders of magnitude. This reduction is accompanied by non-additive free-energy penalties, altered enthalpy–entropy compensation, and reduced heat-capacity changes, indicating that DNA sequences actively modulate LacI conformational dynamics as well as binding thermodynamics. Flynn et al. (2003) proposed that the allosteric transition of LacI from the high-affinity operator-bound state to the low-affinity inducer-bound state is not a symmetric concerted motion, but instead involves asymmetric, multi-pathway conformational changes that propagate from the inducer-binding site to the DNA-binding domains via the core dimer interface [55].

The concept that DNA sequence modulates protein dynamics and cooperative behavior is well established in other repressor systems [98–104]. McGinnis et al. (2022) demonstrated that asymmetric DNA sequences can be recognized by the mycobacteriophage immunity repressor through cooperative binding of two distinct DNA-binding domains, each contributing essential and functionally distinct contacts. This study shows that regulatory specificity does not necessarily depend on symmetric dimerization, but can instead emerge from coordinated multi-domain recognition of non-palindromic operator sequences. Mazumder et al. (2012) established that the λ CI repressor binds cooperatively to adjacent operator sites in an orientation-dependent manner. In a subsequent study, Mazumder et al. (2017) further showed that DNA sequence can allosterically enhance distal protein–protein interactions by modulating conformational dynamics. Similar sequence-dependent dynamic effects have also been reported for both the λ Cro and Gal repressors. Using ultrafast fluorescence spectroscopy, Choudhury et al. (2016) investigated the λ Cro repressor, a transcriptional regulator that controls the lytic–lysogenic switch by

binding specific operator DNA sequences. Using ultrafast fluorescence spectroscopy, they demonstrated that the minor-groove binding domains of two operator DNA sequences exhibit pronounced sequence-dependent flexibility on the picosecond timescale, while the overall protein backbone dynamics remain largely unchanged. In the λ Cro repressor system, enhanced fast DNA motions, particularly in the OR3 complexes, were associated with differences in dynamic behavior and entropic contributions to binding specificity. Similarly, in the Gal repressor, distinct operator sequences differentially modulate protein motions, underscoring the importance of dynamic allostery in sequence recognition. Naiya et al. (2016) further proposed conformational selection as a key mechanism for multi-sequence recognition, wherein specificity arises from the preferential stabilization of pre-existing conformational states.

Recent computational and experimental studies further reinforce the emerging paradigm that DNA sequence intrinsically encodes regulatory information through modulation of protein dynamics and thermodynamics. Molecular dynamics simulations of various transcription factors, including SOC1, ZTA, TBR1, EGR/Sp1, and AP-1, consistently demonstrate that subtle variations in DNA sequence remodel hydrogen-bonding networks, electrostatic interactions, binding affinity, and the enthalpy–entropy balance [105–109]. Beyond individual transcription factors, Onyema et al. (2026) revealed that nucleosomes possess long-range allosteric networks connecting DNA and histone proteins. Specific DNA and histone residues act as allosteric hubs, enabling the propagation of local perturbations to modulate chromatin accessibility. Quantitative characterization of this allosteric behavior enables the identification of residues and domains that may serve as potential targets for therapeutic intervention [110].

Taken together, these studies strongly support the present findings that DNA sequences are not merely passive binding substrates but actively modulate protein conformational ensembles, binding thermodynamics, and long-range allosteric communication. The present results extend the ensemble-based model by providing residue-level thermodynamic quantification of sequence-dependent entropy redistribution in LacI, thereby establishing a direct link between operator DNA sequence and asymmetric modulation of the protein’s conformational free-energy landscape.

4.1. Significance and Broader Implications

The torsional angle–based thermodynamic framework using the histogram-based method is broadly applica-

ble to protein–DNA, protein–RNA, protein–protein, and protein–ligand systems in which functional regulation arises from redistribution of conformational entropy. Estimation of conformational thermodynamics using the histogram-based method provides a precise framework for identifying allosteric, druggable sites. Residues or regions exhibiting pronounced ligand- or DNA/RNA-induced entropy redistribution define dynamic regulatory nodes within the allosteric network. Such sites represent attractive targets for drug discovery because modulation of their conformational flexibility can reshape the global free-energy landscape. From this perspective, allosteric ligands function as thermodynamic effectors that reshape conformational entropy distributions rather than merely occupying cryptic binding pockets. Consequently, residue-level entropy mapping integrates dynamic free-energy considerations into the rational design of selective therapeutics.

Sequence-dependent redistribution of conformational ensembles emerges as a general principle underlying transcription factor specificity. DNA-mediated ensemble reweighting has been reported for p53, the catabolite activator protein, NF- κ B, and the glucocorticoid receptor, where distinct response elements stabilize alternative conformational substates despite comparable binding affinities. These findings indicate that regulatory discrimination is governed not solely by equilibrium binding affinity but also by DNA-encoded modulation of the protein conformational free-energy landscape [111–115].

These results further demonstrate that regulatory specificity can arise from entropy redistribution and residue-level dynamic asymmetry, rather than solely from static structural differences.

Beyond transcriptional regulation, residue-specific free-energy and conformational entropy analysis provides a transferable thermodynamic framework for genetic engineering and synthetic biology applications. In systems such as CRISPR–Cas9, target discrimination depends not only on binding affinity but also on protein flexibility and intrinsic dynamics that modulate mismatch tolerance and off-target activity. Quantification of torsional entropy changes between cognate and mismatched complexes enables identification of residues that govern target discrimination specificity, thereby guiding the rational engineering of higher specificity variants. A similar principle can be extended to engineered transcription factors and epigenetic reader proteins, where subtle shifts in conformational entropy modulate the recognition of DNA and histone modifications, thereby altering regulatory outcomes without substantial changes in binding free energy [116].

RNA-guided nucleases further exemplify entropy-driven allostery. In CRISPR–Cas13a, RNA binding propa-

gates long-range allosteric communication from the REC lobe to the HEPN nuclease domains, stabilizing catalytically competent conformations through reorganization of interdomain interactions [117]. RNA-mediated allosteric activation in CRISPR–Cas13a reflects an ensemble reorganization rather than discrete conformational transitions; therefore, the histogram-based method may enable residue-wise quantification of conformational entropy, recasting experimentally identified allosteric hotspots as sites of significant changes in torsional entropy. This approach provides a dynamic and thermodynamic framework for the rational engineering of RNA-targeting tools [117].

From a translational perspective, DNA-induced changes in thermodynamic parameters have important implications for drug discovery. Allosteric pockets may be transiently formed or preferentially stabilized within DNA-bound conformational ensembles, thereby enabling selective targeting of transcription factors that are otherwise difficult to modulate through orthosteric inhibition. Rather than directly blocking the large protein–DNA interface, small molecules can act by shifting the equilibrium between pre-existing conformational states sampled by the protein. By preferentially stabilizing specific conformational states, such ligands function as ensemble modulators that effectively reshape the protein free-energy landscape and thereby alter its functional behavior. This principle aligns with topology-guided redesign of allosteric LacI-family transcription factors, highlighting the mechanistic relevance of conformational thermodynamics as a quantitative framework for rational protein design [118]. In drug discovery, engineered protein interfaces are designed using diverse strategies, including AI-guided mutational scanning, interface residue screening, and machine learning-based affinity ranking, to develop high-affinity inhibitors and biologics that selectively modulate disease-relevant targets such as protein–protein and protein–DNA interactions [119].

4.2. Limitations and Future Directions

The present results indicate that DNA sequence alone drives asymmetric redistribution of conformational free energy and entropy across structurally identical subunits, manifested at the conformational ensemble level through coordinated side-chain and collective motions rather than discrete allosteric hubs. This distributed asymmetry supports an entropy-driven model of allostery. Although targeted mutational or in silico perturbation analyses could provide complementary validation, the conclusions are independent of such interventions, as all systems compris-

ing identical protein sequences were simulated under identical force-field parameters and simulation protocols.

MM/PBSA-based energy decomposition, dynamic network analysis, electrostatic potential calculations, and detailed hydration-shell or water-dynamics analyses, including residence time calculations, were not performed in the present study, as these approaches involve additional model-dependent assumptions and parameter choices such as dielectric constants, solvent-accessible surface definitions, and shell boundaries, which may introduce systematic bias. Similarly, correlation network analysis relies on user-defined thresholds that may influence the resulting connectivity patterns. In contrast, principal component analysis (PCA), based on direct diagonalization of the covariance matrix of atomic fluctuations, provides a parameter-independent description of dominant collective motions and enables an unbiased characterization of large-scale dynamic asymmetry.

Conformational thermodynamic quantities were computed exclusively from the χ_1 side-chain torsion angle using the histogram-based method, capturing the principal rotameric transitions that underlie side-chain conformational plasticity at both interfacial and distal residues. Accordingly, the conformational thermodynamics derived from the HBM quantify relative internal conformational changes between ligand-bound and ligand-free states, rather than absolute binding free energies of the full LacI–DNA repressor–operator complexes. Higher-order side-chain torsion angles (χ_2 , χ_3 , χ_4) and backbone dihedrals (ϕ/ψ) were not included in the analysis. Accordingly, the reported values represent χ_1 -specific entropy contributions rather than a full multidimensional estimate of conformational entropy. This reduction in the number of microscopic conformational degrees of freedom may underestimate the absolute entropy and neglect certain torsional cross-correlations. Nevertheless, the consistent application of the χ_1 -based approach across all systems, together with its agreement with essential dynamics analyses, supports its reliability for comparative assessment of sequence-dependent allosteric effects.

Future microsecond-scale molecular dynamics simulations of LacI–O1 and LacI–O3 complexes, incorporating all microscopic conformational degrees of freedom, would enable a more comprehensive thermodynamic characterization of the protein–DNA recognition landscape and provide deeper insight into sequence-dependent differences in global dynamics. In addition, targeted mutational analyses could further elucidate the role of protein dynamics in allostery-driven protein–protein communication. Systematic base substitutions in DNA would enable quantification of how base-specific energetic contributions modulate the redistribution of free energy and en-

trophy. Mutations that perturb protein–DNA interactions (PDIs) are key drivers of disease. Mutations in DNA-binding proteins or in their target sequences can disrupt interaction energetics, leading to aberrant gene expression, replication stress, and DNA damage. These alterations often result in transcriptional misregulation, aberrant signaling, and genomic instability. Disease-associated mutations frequently modify the thermodynamics or kinetics of protein–DNA interactions, producing phenotypic consequences without complete loss of molecular function [2].

5. Conclusions

In summary, this work shows that multi-sequence DNA recognition by LacI is governed by asymmetric redistribution of conformational free energy and entropy across structurally identical subunits. These dynamic disparities arise from variations in hydrogen-bonding patterns, DNA base-pair step deformations, and minor groove geometry, underscoring the role of DNA shape and intrinsic sequence asymmetry in modulating protein dynamics. The absence of a central GC base pair in O-SymL and the inherent half-site asymmetry of the O2 operator promote distinct recognition modes. Sequence-dependent differences in global protein dynamics upon DNA binding play a crucial role in modulating recognition and inter-domain communication between the C-terminal domains during Lac repressor tetramerization, thereby contributing to its functional regulation. Operator DNA acts as an active allosteric ligand, reshaping the Lac repressor conformational landscape and differentially stabilizing gene regulatory networks in a sequence-dependent manner. Distinct domain motions in the symmetric subunits of the Lac repressor, characterized through essential dynamics and residue-specific conformational thermodynamics, provide a quantitative framework for understanding entropy-driven allostery in transcription factors. This further establishes a general paradigm for sequence-encoded dynamic regulation. By defining the dynamic and thermodynamic basis of sequence-dependent allostery, this study provides a rational framework for engineering Lac repressor variants, with implications for gene regulation, synthetic biology, and guidance for drug discovery via selective modulation of protein–DNA interactions and allosteric pathways.

List of Abbreviations

CVs	Collective Variables
GROMACS	GROningen Machine for Chemical Simulations
HBM	Histogram Based Method

HTH	Helix Turn Helix
IPTG	Isopropyl β -D-1thiogalactopyranoside
LacI	Lactose Repressor Protein
LINCS	Linear Constraint Solver
MD	Molecular Dynamics
MM/PBSA	Molecular Mechanics Poisson-Boltzmann Surface Area
MM/GBSA	Molecular Mechanics Generalized Born Surface Area
NMR	Nuclear Magnetic Resonance
OptoLacI	Optogenetic Engineering of LacI
PBC	Periodic Boundary Condition
PCA	Principal Component Analysis
PDB	Protein Data Bank
PME	Particle Mesh Ewald
PYMOL	Python Molecular Graphics System
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
TFs	Transcription Factors
TS	Transition State
VMD	Visual Molecular Dynamics

Author Contributions

The author (Soumi Das) is solely responsible for the conceptualization, methodology, software, validation, formal analysis, funding acquisition, investigation, resources, data curation, writing---original draft preparation, writing---review and editing, visualization, and project administration. The author has reviewed and accepted the published version of the manuscript.

Availability of Data and Materials

Data supporting the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The author declares no conflicts of interest.

Funding

This work was supported by the Council of Scientific and Industrial Research (CSIR), Government of India, under the CSIR--NET fellowship scheme in Chemical Sciences (awarded to Soumi Das; File No: 31/002(0963)/2013-EMR-I).

Acknowledgments

The author sincerely acknowledges Siddhartha Roy (Retd.) and Dhananjay Bhattacharyya (Retd.) for insightful discussions. Computational facilities were provided by the Bose Institute, the Saha Institute of Nuclear Physics, and the Technical Research Center at the

S. N. Bose National Center for Basic Sciences, Kolkata, India. This work is the author's original contribution and is partly derived from the author's Ph.D. thesis [Das, S. (2023). "Computational Study Of Structure And Dynamics of Protein-DNA Complexes", University of Calcutta, <http://hdl.handle.net/10603/508120>, accessed on 1 September 2024] and a preprint [Das, S., Roy, S. and Bhattacharyya, D. (2024), "In-silico studies on structural and thermodynamics basis of interaction of lac repressor protein binding to different DNA operators." <https://doi.org/10.1101/2024.07.07.602393>, accessed on 1 September 2024] available on bioRxiv.

AI Declaration

The artificial intelligence tool ChatGPT was used solely for language refinement. All scientific concepts, data analyses, interpretations, and intellectual contributions were independently conceived, critically evaluated, and validated by the author. The author performed all substantive revisions, approved the final version of the manuscript, and assumes full responsibility for the content and scientific integrity of the work.

Supplementary Materials

Supplementary material associated with this article has been published online and is available at: <https://doi.org/10.69709/MolModC.2026.165319>.

References

- [1] Mishra, R.; Ahmad, S. Protein-DNA Interactions in Disease and Drug Discovery. *Chem. Commun.* **2026**, 62, 3702–3729. [CrossRef]
- [2] Genz, L.R.; Nair, S.; Sweeney, A.; Topf, M. Drug Targeting of Protein-Nucleic Acid Interactions. *Curr. Opin. Struct. Biol.* **2025**, 95, 103165. [CrossRef]
- [3] Das, S. Decoding the Mechanistic Landscape of Protein-DeoxyriboNucleic Acid Recognition: A Comprehensive Review. *ES Chem. Sustain.* **2025**, 4, 1610. [CrossRef]
- [4] Farias, A.B.; Martinez, G.S.; Galán-Vásquez, E.; Nicolás, M.F.; Pérez-Rueda, E. Predicting Bacterial Transcription Factor Binding Sites through Machine Learning and Structural Characterization Based on DNA Duplex Stability. *Brief. Bioinform.* **2024**, 25, bbae581. [CrossRef]
- [5] Baydar, D.O.; Rauluseviciute, I.; Aronsen, D.R.; Blanc-Mathieu, R.; Bonthuis, I.; De Beukelaer, H.; Ferenc, K.; Jegou, A.; Kumar, V.; Lemma, R.B.; et al. JASPAR 2026: Expansion of Transcription Factor Binding Profiles and Integration of Deep Learning Models. *Nucleic Acids Res.* **2026**, 54, D184–D193. [CrossRef]
- [6] Masri, C.A.; Yu, J. Combining Physics-Based Protein-DNA Energetics with Machine Learning

- to Predict Interpretable Transcription Factor-DNA Binding. *J. Chem. Inf. Model.* **2025**, *65*, 11804–11817. [CrossRef] [PubMed]
- [7] Wang, Y.; Li, J.; Chiu, T.P.; Xin, B.; Rohs, R. Sequence-Based Modeling of Low-Affinity Transcription Factor-DNA Binding through Deep Learning. *NAR Genom. Bioinform.* **2026**, *8*, lqag027. [CrossRef]
- [8] Zhang, J.; Liu, B.; Wu, J.; Wang, Z.; Li, J. DeepCAC: A Deep Learning Approach on DNA Transcription Factors Classification Based on Multi-Head Self-Attention and Concatenate Convolutional Neural Network. *BMC Bioinform.* **2023**, *24*, 345. [CrossRef]
- [9] Tariq, S.; Amin, A. DeepCTF: Transcription Factor Binding Specificity Prediction Using DNA Sequence Plus Shape in an Attention-Based Deep Learning Model. *SIViP* **2024**, *18*, 5239–5251. [CrossRef]
- [10] Barissi, S.; Sala, A.; Wiczór, M.; Battistini, F.; Orozco, M. DNAffinity: A Machine-Learning Approach to Predict DNA Binding Affinities of Transcription Factors. *Nucleic Acids Res.* **2022**, *50*, 9105–9114. [CrossRef] [PubMed]
- [11] Sedhom, J.; Kinser, J.; Solomon, L.A. Alignment of Major-Groove Hydrogen Bond Arrays Uncovers Shared Information between Different DNA Sequences That Bind the Same Protein. *NAR Genom. Bioinform.* **2022**, *4*, lqac101. [CrossRef]
- [12] Sakong, S.; Fierz, B.; Suter, D.M. Electrostatic Properties of Disordered Regions Control Transcription Factor Search and Pioneer Activity. *Nat. Commun.* **2026**, *17*, 2512. [CrossRef]
- [13] Jonas, F.; Navon, Y.; Barkai, N. Intrinsically Disordered Regions as Facilitators of the Transcription Factor Target Search. *Nat. Rev. Genet.* **2025**, *26*, 424–435. [CrossRef]
- [14] Marklund, E.; Mao, G.; Yuan, J.; Zikrin, S.; Abdurakhmanov, E.; Deindl, S.; Elf, J. Sequence Specificity in DNA Binding Is Mainly Governed by Association. *Science* **2022**, *375*, 442–445. [CrossRef] [PubMed]
- [15] Wilson, C.J.; Zhan, H.; Swint-Kruse, L.; Matthews, K.S. The Lactose Repressor System: Paradigms for Regulation, Allosteric Behavior and Protein Folding. *Cell. Mol. Life Sci.* **2007**, *64*, 3–16. [CrossRef] [PubMed]
- [16] Swint-Kruse, L.; Matthews, K.S. Allostery in the LacI/GalR Family: Variations on a Theme. *Curr. Opin. Microbiol.* **2009**, *12*, 129–137. [CrossRef] [PubMed]
- [17] Kalodimos, C.G.; Bonvin, A.M.; Salinas, R.K.; Wechselberger, R.; Boelens, R.; Kaptein, R. Plasticity in Protein-DNA Recognition: Lac Repressor Interacts with Its Natural Operator O1 through Alternative Conformations of Its DNA-Binding Domain. *EMBO J.* **2002**, *21*, 2866–2876. [CrossRef]
- [18] Kalodimos, C.G.; Biris, N.; Bonvin, A.M.; Levandoski, M.M.; Guennuegues, M.; Boelens, R.; Kaptein, R. Structure and Flexibility Adaptation in Nonspecific and Specific Protein-DNA Complexes. *Science* **2004**, *305*, 386–389. [CrossRef]
- [19] Kalodimos, C.G.; Boelens, R.; Kaptein, R. Toward an Integrated Model of Protein-DNA Recognition as Inferred from NMR Studies on the Lac Repressor System. *Chem. Rev.* **2004**, *104*, 3567–3586. [CrossRef]
- [20] Romanuka, J.; Folkers, G.E.; Biris, N.; Tishchenko, E.; Wienk, H.; Bonvin, A.M.; Kaptein, R.; Boelens, R. Specificity and Affinity of Lac Repressor for the Auxiliary Operators O2 and O3 Are Explained by the Structures of Their Protein-DNA Complexes. *J. Mol. Biol.* **2009**, *390*, 478–489. [CrossRef]
- [21] Bell, C.E.; Lewis, M. A Closer View of the Conformation of the Lac Repressor Bound to Operator. *Nat. Struct. Biol.* **2000**, *7*, 209–214. [CrossRef]
- [22] Daber, R.; Stayrook, S.; Rosenberg, A.; Lewis, M. Structural Analysis of Lac Repressor Bound to Allosteric Effectors. *J. Mol. Biol.* **2007**, *370*, 609–619. [CrossRef]
- [23] Lewis, M. The Lac Repressor. *C. R. Biol.* **2005**, *328*, 521–548. [CrossRef] [PubMed]
- [24] Lewis, M.; Chang, G.; Horton, N.C.; Kercher, M.A.; Pace, H.C.; Schumacher, M.A.; Brennan, R.G.; Lu, P. Crystal Structure of the Lactose Operon Repressor and Its Complexes with DNA and Inducer. *Science* **1996**, *271*, 1247–1254. [CrossRef]
- [25] Bell, C.E.; Lewis, M. Crystallographic Analysis of Lac Repressor Bound to Natural Operator O1. *J. Mol. Biol.* **2001**, *312*, 921–926. [CrossRef] [PubMed]
- [26] Sadler, J.R.; Sasmor, H.; Betz, J.L. A Perfectly Symmetric Lac Operator Binds the Lac Repressor Very Tightly. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 6785–6789. [CrossRef]
- [27] Spronk, C.A.; Bonvin, A.M.; Radha, P.K.; Melacini, G.; Boelens, R.; Kaptein, R. The Solution Structure of Lac Repressor Headpiece 62 Complexed to a Symmetrical Lac Operator. *Structure* **1999**, *7*, 1483–1492. [CrossRef]
- [28] Kalodimos, C.G.; Folkers, G.E.; Boelens, R.; Kaptein, R. Strong DNA Binding by Covalently Linked Dimeric Lac Headpiece: Evidence for the Crucial Role of the Hinge Helices. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 6039–6044. [CrossRef]
- [29] Markiewicz, P.; Kleina, L.G.; Cruz, C.; Ehret, S.; Miller, J.H. Genetic Studies of the Lac Repressor. XIV. Analysis of 4000 Altered Escherichia coli Lac Repressors Reveals Essential and Non-Essential Residues, as Well as “Spacers” Which Do Not Require a Specific Sequence. *J. Mol. Biol.* **1994**, *240*, 421–433. [CrossRef] [PubMed]
- [30] Nelson, B.D.; Manoel, C.; Traxler, B. Insertion Mutagenesis of the Lac Repressor and Its Implications for Structure-Function Analysis. *J. Bacteriol.* **1997**, *179*, 3721–3728. [CrossRef]
- [31] Milk, L.; Daber, R.; Lewis, M. Functional Rules for Lac Repressor-Operator Associations and Implications for Protein-DNA Interactions. *Protein Sci.* **2010**, *19*, 1162–1172. [CrossRef]
- [32] Betz, J.L.; Sasmor, H.M.; Buck, F.; Insley, M.Y.; Caruthers, M.H. Base Substitution Mutants of the

- Lac Operator: In Vivo and In Vitro Affinities for Lac Repressor. *Gene* **1986**, *50*, 123–132. [[CrossRef](#)] [[PubMed](#)]
- [33] Rastinejad, F.; Artz, P.; Lu, P. Origin of the Asymmetrical Contact between Lac Repressor and Lac Operator DNA. *J. Mol. Biol.* **1993**, *233*, 389–399. [[CrossRef](#)] [[PubMed](#)]
- [34] Frank, D.E.; Saecker, R.M.; Bond, J.P.; Capp, M.W.; Tsodikov, O.V.; Melcher, S.E.; Levandoski, M.M.; Record, M. Thermodynamics of the Interactions of Lac Repressor with Variants of the Symmetric Lac Operator: Effects of Converting a Consensus Site to a Non-Specific Site. *J. Mol. Biol.* **1997**, *267*, 1186–1206. [[CrossRef](#)]
- [35] Mossing, M.C.; Record, M.J. Thermodynamic Origins of Specificity in the Lac Repressor-Operator Interaction. Adaptability in the Recognition of Mutant Operator Sites. *J. Mol. Biol.* **1985**, *186*, 295–305. [[CrossRef](#)]
- [36] Whitson, P.A.; Olson, J.S.; Matthews, K.S. Thermodynamic Analysis of the Lactose Repressor-Operator DNA Interaction. *Biochemistry* **1986**, *25*, 3852–3858. [[CrossRef](#)] [[PubMed](#)]
- [37] Daber, R.; Sochor, M.A.; Lewis, M. Thermodynamic Analysis of Mutant Lac Repressors. *J. Mol. Biol.* **2011**, *409*, 76–87. [[CrossRef](#)]
- [38] Fried, M.G.; Stickley, D.F.; Smirnakis, K.V.; Adams, C.; MacDonald, D.; Lu, P. Role of Hydration in the Binding of Lac Repressor to DNA. *J. Biol. Chem.* **2002**, *277*, 50676–50682. [[CrossRef](#)]
- [39] Barr, D.; Van der Vaart, A. The Natural DNA Bending Angle in the Lac Repressor Headpiece–O1 Operator Complex Is Determined by Protein–DNA Contacts and Water Release. *Phys. Chem. Chem. Phys.* **2012**, *14*, 2070–2077. [[CrossRef](#)]
- [40] Lüking, M.; Elf, J.; Levy, Y. Conformational Change of Transcription Factors from Search to Specific Binding: A Lac Repressor Case Study. *J. Phys. Chem. B* **2022**, *126*, 9971–9984. [[CrossRef](#)]
- [41] Liao, Q.; Lüking, M.; Krüger, D.M.; Deindl, S.; Elf, J.; Kasson, P.M.; Kamerlin, S.L. Long Time-Scale Atomistic Simulations of the Structure and Dynamics of Transcription Factor-DNA Recognition. *J. Phys. Chem. B* **2019**, *123*, 3576–3590. [[CrossRef](#)]
- [42] Furini, S.; Domene, C. DNA Recognition Process of the Lactose Repressor Protein Studied via Metadynamics and Umbrella Sampling Simulations. *J. Phys. Chem. B* **2014**, *118*, 13059–13065. [[CrossRef](#)]
- [43] Furini, S.; Barbini, P.; Domene, C. DNA-Recognition Process Described by MD Simulations of the Lactose Repressor Protein on a Specific and a Non-Specific DNA Sequence. *Nucleic Acids Res.* **2013**, *41*, 3963–3972. [[CrossRef](#)]
- [44] Xu, L.; Ye, W.; Jiang, C.; Yang, J.; Zhang, J.; Feng, Y.; Luo, R.; Chen, H.F. Recognition Mechanism between Lac Repressor and DNA with Correlation Network Analysis. *J. Phys. Chem. B* **2015**, *119*, 2844–2856. [[CrossRef](#)]
- [45] Sengupta, R.; Capp, M.W.; Shkel, I.A.; Record, M.J. The Mechanism and High-Free-Energy Transition State of Lac Repressor-Lac Operator Interaction. *Nucleic Acids Res.* **2017**, *45*, 12671–12680. [[CrossRef](#)]
- [46] Sun, L.; Tabaka, M.; Hou, S.; Li, L.; Burdzy, K.; Aksimentiev, A.; Maffeo, C.; Zhang, X.; Holyst, R. The Hinge Region Strengthens the Nonspecific Interaction between Lac-Repressor and DNA: A Computer Simulation Study. *PLoS ONE* **2016**, *11*, e0152002. [[CrossRef](#)] [[PubMed](#)]
- [47] Seckfort, D.; Pettitt, B.M. Price of Disorder in the Lac Repressor Hinge Helix. *Biopolymers* **2019**, *110*, e23239. [[CrossRef](#)] [[PubMed](#)]
- [48] Kariyawasam, N.L.; Ploetz, E.A.; Swint-Kruse, L.; Smith, P.E. Simulated Pressure Changes in LacI Suggest a Link between Hydration and Functional Conformational Changes. *Biophys. Chem.* **2024**, *304*, 107126. [[CrossRef](#)] [[PubMed](#)]
- [49] Romanuka, J.; Folkers, G.E.; Gnida, M.; Kovačič, L.; Wienk, H.; Kaptein, R.; Boelens, R. Genetic Switching by the Lac Repressor Is Based on Two-State Monod-Wyman-Changeux Allosterity. *Proc. Natl. Acad. Sci. USA* **2023**, *120*, e2311240120. [[CrossRef](#)]
- [50] Garruss, A.S.; Collins, K.M.; Church, G.M. Deep Representation Learning Improves Prediction of LacI-Mediated Transcriptional Repression. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2022838118. [[CrossRef](#)]
- [51] Liu, M.; Li, Z.; Huang, J.; Yan, J.; Zhao, G.; Zhang, Y. OptoLacI: Optogenetically Engineered Lactose Operon Repressor LacI Responsive to Light Instead of IPTG. *Nucleic Acids Res.* **2024**, *52*, 8003–8016. [[CrossRef](#)]
- [52] Xu, W.; Yan, Y.; Artsimovitch, I.; Dunlap, D.; Finzi, L. Positive Supercoiling Favors Transcription Elongation through Lac Repressor-Mediated DNA Loops. *Nucleic Acids Res.* **2022**, *50*, 2826–2835. [[CrossRef](#)]
- [53] Goodson, K.A.; Wang, Z.; Haeusler, A.R.; Kahn, J.D.; English, D.S. LacI-DNA-IPTG Loops: Equilibria among Conformations by Single-Molecule FRET. *J. Phys. Chem. B* **2013**, *117*, 4713–4722. [[CrossRef](#)] [[PubMed](#)]
- [54] Glasgow, A.; Hobbs, H.T.; Perry, Z.R.; Wells, M.L.; Marqusee, S.; Kortemme, T. Ligand-Specific Changes in Conformational Flexibility Mediate Long-Range Allosterity in the Lac Repressor. *Nat. Commun.* **2023**, *14*, 1179. [[CrossRef](#)] [[PubMed](#)]
- [55] Flynn, T.C.; Swint-Kruse, L.; Kong, Y.; Booth, C.; Matthews, K.S.; Ma, J. Allosteric Transition Pathways in the Lactose Repressor Protein Core Domains: Asymmetric Motions in a Homodimer. *Protein Sci.* **2003**, *12*, 2523–2541. [[CrossRef](#)] [[PubMed](#)]
- [56] Yonetani, Y.; Kono, H. Dissociation Free-Energy Profiles of Specific and Nonspecific DNA-Protein Complexes. *J. Phys. Chem. B* **2013**, *117*, 7535–7545. [[CrossRef](#)]
- [57] Villa, E.; Balaeff, A.; Schulten, K. Structural Dynamics of the Lac Repressor-DNA Complex Revealed by

- a Multiscale Simulation. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 6783–6788. [CrossRef]
- [58] Das, A.; Chakrabarti, J.; Ghosh, M. Conformational Contribution to Thermodynamics of Binding in Protein-Peptide Complexes through Microscopic Simulation. *Biophys. J.* **2013**, *104*, 1274–1284. [CrossRef]
- [59] Das, A.; Chakrabarti, J.; Ghosh, M. Conformational Thermodynamics of Metal-Ion Binding to a Protein. *Chem. Phys. Lett.* **2013**, *581*, 91–95. [CrossRef]
- [60] Sikdar, S.; Chakrabarti, J.; Ghosh, M. A Microscopic Insight from Conformational Thermodynamics to Functional Ligand Binding in Proteins. *Mol. Biosyst.* **2014**, *10*, 3280–3289. [CrossRef]
- [61] Samanta, S.; Mukherjee, S. Microscopic Insight into Thermodynamics of Conformational Changes of SAP-SLAM Complex in Signal Transduction Cascade. *J. Chem. Phys.* **2017**, *146*, 165103. [CrossRef]
- [62] Samanta, S.; Mukherjee, S. Co-Operative Intra-Protein Structural Response Due to Protein-Protein Complexation Revealed through Thermodynamic Quantification: Study of MDM2-p53 Binding. *J. Comput.-Aided Mol. Des.* **2017**, *31*, 891–903. [CrossRef]
- [63] Maganti, L.; Ghosh, M.; Chakrabarti, J. Molecular Dynamics Studies on Conformational Thermodynamics of Orail-Calmodulin Complex. *J. Biomol. Struct. Dyn.* **2018**, *36*, 3411–3419. [CrossRef] [PubMed]
- [64] Maganti, L.; Dutta, S.; Ghosh, M.; Chakrabarti, J. Allosteric in Orail Binding to Calmodulin Revealed from Conformational Thermodynamics. *J. Biomol. Struct. Dyn.* **2019**, *37*, 493–502. [CrossRef]
- [65] Moulick, A.G.; Chakrabarti, J. Fluctuation-Dominated Ligand Binding in Molten Globule Protein. *J. Chem. Inf. Model.* **2023**, *63*, 5583–5591. [CrossRef]
- [66] Mondal, M.; Chakrabarti, J.; Ghosh, M. Molecular Dynamics Simulations on Interaction between Bacterial Proteins: Implication on Pathogenic Activities. *Proteins* **2018**, *86*, 370–378. [CrossRef]
- [67] Dutta, S.; Ghosh, M.; Chakrabarti, J. In-Silico Studies on Conformational Stability of Flagellin-Receptor Complexes. *J. Biomol. Struct. Dyn.* **2020**, *38*, 2240–2252. [CrossRef] [PubMed]
- [68] Mandal, S.C.; Maganti, L.; Mondal, M.; Chakrabarti, J. Microscopic Insight to Specificity of Metal Ion Co-factor in DNA Cleavage by Restriction Endonuclease EcoRV. *Biopolymers* **2020**, *111*, e23396. [CrossRef]
- [69] Kole, K.; Gupta, A.M.; Chakrabarti, J. Conformational Stability and Order of Hoogsteen Base Pair Induced by Protein Binding. *Biophys. Chem.* **2023**, *301*, 107079. [CrossRef] [PubMed]
- [70] Das, S. Decoding the Effect of Temperatures on Conformational Stability and Order of Ligand Unbound Thermosensing Adenine Riboswitch Using Molecular Dynamics Simulation. *J. Biomol. Struct. Dyn.* **2025**, *43*, 11042–11055. [CrossRef]
- [71] Das, S. Unraveling Conformational Thermodynamics of Ligand Binding to Fluoride Riboswitch Aptamer: Implications for Therapeutic Design. *Mol. Model. Connect* **2025**, *2*, 0012. [CrossRef]
- [72] Shen, W.; Zhou, T.; Shi, X. Enhanced Sampling in Molecular Dynamics Simulations and Their Latest Applications—A Review. *Nano Res.* **2023**, *16*, 13474–13497. [CrossRef]
- [73] Mehdi, S.; Smith, Z.; Herron, L.; Zou, Z.; Tiwary, P. Enhanced Sampling with Machine Learning. *Annu. Rev. Phys. Chem.* **2024**, *75*, 347–370. [CrossRef]
- [74] Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J.Z.H.; Hou, T. End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. *Chem. Rev.* **2019**, *119*, 9478–9508. [CrossRef]
- [75] Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA Methods to Estimate Ligand-Binding Affinities. *Expert Opin. Drug Discov.* **2015**, *10*, 449–461. [CrossRef]
- [76] Jiang, D.; Du, H.; Zhao, H.; Deng, Y.; Wu, Z.; Wang, J.; Zeng, Y.; Zhang, H.; Wang, X.; Wang, E.; et al. Assessing the Performance of MM/PBSA and MM/GBSA Methods. 10. Prediction Reliability of Binding Affinities and Binding Poses for RNA-Ligand Complexes. *Phys. Chem. Chem. Phys.* **2024**, *26*, 10323–10335. [CrossRef]
- [77] 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 Res.* **2000**, *28*, 235–242. [CrossRef] [PubMed]
- [78] Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera--A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. [CrossRef] [PubMed]
- [79] Das, S.; Roy, S.; Bhattacharyya, D. In-Silico Studies on Structural and Thermodynamics Basis of Interaction of Lac Repressor Protein Binding to Different DNA Operators. *bioRxiv* **2024**. [CrossRef]
- [80] Abraham, M.J.; Murtola, T.; Schulz, R.; Pall, S.; Smith, J.C.; Hess, B.; Lindahl, E. Gromacs: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. [CrossRef]
- [81] Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J.L.; Dror, R.O.; Shaw, D.E. Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. *Proteins* **2010**, *78*, 1950–1958. [CrossRef]
- [82] Ivani, I.; Dans, P.D.; Noy, A.; Pérez, A.; Faustino, I.; Hospital, A.; Walther, J.; Andrio, P.; Goñi, R.; Balaceanu, A.; et al. Parmbsc1: A Refined Force Field for DNA Simulations. *Nat. Methods* **2016**, *13*, 55–58. [CrossRef]
- [83] Maier, J.A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K.E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713. [CrossRef]

- [84] Wang, X.; Xiong, D.; Zhang, Y.; Zhai, J.; Gu, Y.C.; He, X. The Evolution of the Amber Additive Protein Force Field: History, Current Status, and Future. *J. Chem. Phys.* **2025**, *162*, 030901. [[CrossRef](#)]
- [85] Liebl, K.; Zacharias, M. The Development of Nucleic Acids Force Fields: From an Unchallenged Past to a Competitive Future. *Biophys. J.* **2023**, *122*, 2841–2851. [[CrossRef](#)]
- [86] Darden, T.; Perera, L.; Li, L.; Pedersen, L. New Tricks for Modelers from the Crystallography Toolkit: The Particle Mesh Ewald Algorithm and Its Use in Nucleic Acid Simulations. *Structure* **1999**, *7*, R55–R60. [[CrossRef](#)]
- [87] Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52*, 7182–7190. [[CrossRef](#)]
- [88] Mukherjee, S.; Bansal, M.; Bhattacharyya, D. Conformational Specificity of Non-Canonical Base Pairs and Higher Order Structures in Nucleic Acids: Crystal Structure Database Analysis. *J. Comput.-Aided Mol. Des.* **2006**, *20*, 629–645. [[CrossRef](#)]
- [89] Bansal, M.; Bhattacharyya, D.; Ravi, B. NUPARM and NUCGEN: Software for Analysis and Generation of Sequence Dependent Nucleic Acid Structures. *Comput. Appl. Biosci.* **1995**, *11*, 281–287. [[Cross-Ref](#)]
- [90] Abdi, H.; Williams, L.J. Principal Component Analysis. *WIREs Comput. Stat.* **2010**, *2*, 433–459. [[Cross-Ref](#)]
- [91] David, C.C.; Jacobs, D.J. Principal Component Analysis: A Method for Determining the Essential Dynamics of Proteins. *Methods Mol. Biol.* **2014**, *1084*, 193–226. [[CrossRef](#)] [[PubMed](#)]
- [92] Moradi, S.; Nowroozi, A.; AryaeiNezhad, M.; Jalali, P.; Khosravi, R.; Shahlaei, M. A Review on Description Dynamics and Conformational Changes of Proteins Using Combination of Principal Component Analysis and Molecular Dynamics Simulation. *Comput. Biol. Med.* **2024**, *183*, 109245. [[CrossRef](#)]
- [93] Fei, J.; Ha, T. Watching DNA Breathe One Molecule at a Time. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 17173–17174. [[CrossRef](#)]
- [94] Segers, M.; Voorspoels, A.; Sakaue, T.; Carlon, E. Mechanisms of DNA-Mediated Allostery. *Phys. Rev. Lett.* **2023**, *131*, 238402. [[CrossRef](#)] [[PubMed](#)]
- [95] Kim, S.; Broströmer, E.; Xing, D.; Jin, J.; Chong, S.; Ge, H.; Wang, S.; Gu, C.; Yang, L.; Gao, Y.Q.; et al. Probing Allostery through DNA. *Science* **2013**, *339*, 816–819. [[CrossRef](#)]
- [96] Traverso, J.; Manoranjan, V.; Bishop, A.; Rasmussen, K.Ø.; Voulgarakis, N.K. Allostery through Protein-Induced DNA Bubbles. *Sci. Rep.* **2015**, *5*, 9037. [[CrossRef](#)] [[PubMed](#)]
- [97] Rosenblum, G.; Elad, N.; Rozenberg, H.; Wiggers, F.; Jungwirth, J.; Hofmann, H. Allostery through DNA Drives Phenotype Switching. *Nat. Commun.* **2021**, *12*, 2967. [[CrossRef](#)] [[PubMed](#)]
- [98] Mondol, T.; Batabyal, S.; Mazumder, A.; Roy, S.; Pal, S.K. Recognition of Different DNA Sequences by a DNA-Binding Protein Alters Protein Dynamics Differentially. *FEBS Lett.* **2012**, *586*, 258–262. [[CrossRef](#)]
- [99] Mazumder, A.; Bandyopadhyay, S.; Dhar, A.; Lewis, D.E.; Deb, S.; Dey, S.; Chakrabarti, P.; Roy, S. A Genetic Network That Balances Two Outcomes Utilizes Asymmetric Recognition of Operator Sites. *Biophys. J.* **2012**, *102*, 1580–1589. [[CrossRef](#)]
- [100] Mazumder, A.; Batabyal, S.; Mondal, M.; Mondol, T.; Choudhury, S.; Ghosh, R.; Chatterjee, T.; Bhattacharyya, D.; Pal, S.K.; Roy, S. Specific DNA Sequences Allosterically Enhance Protein-Protein Interaction in a Transcription Factor through Modulation of Protein Dynamics: Implications for Specificity of Gene Regulation. *Phys. Chem. Chem. Phys.* **2017**, *19*, 14781–14792. [[CrossRef](#)]
- [101] Choudhury, S.; Naiya, G.; Singh, P.; Lemmens, P.; Roy, S.; Pal, S.K. Modulation of Ultrafast Conformational Dynamics in Allosteric Interaction of Gal Repressor Protein with Different Operator DNA Sequences. *ChemBioChem* **2016**, *17*, 605–613. [[Cross-Ref](#)]
- [102] Choudhury, S.; Ghosh, B.; Singh, P.; Ghosh, R.; Roy, S.; Pal, S.K. Ultrafast Differential Flexibility of Cro-Protein Binding Domains of Two Operator DNAs with Different Sequences. *Phys. Chem. Chem. Phys.* **2016**, *18*, 17983–17990. [[CrossRef](#)]
- [103] Naiya, G.; Raha, P.; Mondal, M.K.; Pal, U.; Saha, R.; Chaudhuri, S.; Batabyal, S.; Pal, S.K.; Bhattacharyya, D.; Maiti, N.C.; et al. Conformational Selection Underpins Recognition of Multiple DNA Sequences by Proteins and Consequent Functional Actions. *Phys. Chem. Chem. Phys.* **2016**, *18*, 21618–21628. [[CrossRef](#)] [[PubMed](#)]
- [104] McGinnis, R.J.; Brambley, C.A.; Stamey, B.; Green, W.C.; Gragg, K.N.; Cafferty, E.R.; Terwilliger, T.C.; Hammel, M.; Hollis, T.J.; Miller, J.M.; et al. A Monomeric Mycobacteriophage Immunity Repressor Utilizes Two Domains to Recognize an Asymmetric DNA Sequence. *Nat. Commun.* **2022**, *13*, 4105. [[CrossRef](#)]
- [105] Kaur, S.; Sisodia, R.; Pareek, M.; Sri, T.; Chattopadhyay, S.; Yaduvanshi, P.; Asthana, S.; Madhuranakam, C.; Singh, A. Molecular Dynamics and Evolutionary Conservation of FUL-SOC1 Regulatory Interactions in Polyploid Brassica juncea. *Funct. Integr. Genom.* **2026**, *26*, 67. [[CrossRef](#)] [[PubMed](#)]
- [106] Duraisamy, B.; Pramanik, D. Influence of DNA Sequences on the Thermodynamic and Structural Stability of the ZTA Transcription Factor–DNA Complex: An All-Atom Molecular Dynamics Study. *J. Phys. Chem. B* **2025**, *129*, 4282–4297. [[CrossRef](#)]
- [107] Hartman, R.; Blane, A.; Kubrom, F.; Mayet, R.; Penedo, J.C.; Fanucchi, S. Molecular Insights into the Binding Dynamics of Transcription Factor TBR1 to T-Box DNA Sequences. *J. Mol. Biol.* **2026**, *438*, 169543. [[CrossRef](#)]

- [108] Patra, P.; Gao, Y.Q. Structural and Thermodynamic Insights into Competitive Binding of the EGR1 and Sp1 Zinc Finger Class of Transcription Factors with a GC-Rich DNA Motif. *J. Chem. Inf. Model.* **2025**, *65*, 11988–12005. [[CrossRef](#)]
- [109] Patra, P.; Gao, Y.Q. Structural and Dynamical Aspect of DNA Motif Sequence Specific Binding of AP-1 Transcription Factor. *J. Chem. Phys.* **2024**, *160*, 115103. [[CrossRef](#)] [[PubMed](#)]
- [110] Onyema, A.C.; Dikeocha, C.; Patel, R.; Moussa, J.; Loverde, S.M. Mapping Allosteric Communication in the Nucleosome with Conditional Activity. *J. Chem. Inf. Model.* **2026**, *66*, 2889–2899. [[CrossRef](#)] [[PubMed](#)]
- [111] Ning, S.; Zeng, C.; Wang, H.; Zhang, J.; Xue, Y.; Zhao, Y. The Allosteric Regulation of the DNA-Binding Domain of p53 by the Intrinsically Disordered C-Terminal Domain. *Pharmaceuticals* **2026**, *19*, 124. [[CrossRef](#)] [[PubMed](#)]
- [112] Lawson, C.L.; Swigon, D.; Murakami, K.S.; Darst, S.A.; Berman, H.M.; Ebright, R.H. Catabolite Activator Protein: DNA Binding and Transcription Activation. *Curr. Opin. Struct. Biol.* **2004**, *14*, 10–20. [[CrossRef](#)]
- [113] Prabhakant, A.; Panigrahi, A.; Krishnan, M. Allosteric Response of DNA Recognition Helices of Catabolite Activator Protein to cAMP and DNA Binding. *J. Chem. Inf. Model.* **2020**, *60*, 6366–6376. [[CrossRef](#)]
- [114] Ashkenazi, S.; Plotnikov, A.; Bahat, A.; Ben-Zeev, E.; Warszawski, S.; Dikstein, R. A Novel Allosteric Mechanism of NF- κ B Dimerization and DNA Binding Targeted by an Anti-Inflammatory Drug. *Mol. Cell. Biol.* **2016**, *36*, 1237–1247. [[CrossRef](#)]
- [115] Watson, L.C.; Kuchenbecker, K.M.; Schiller, B.J.; Gross, J.D.; Pufall, M.A.; Yamamoto, K.R. The Glucocorticoid Receptor Dimer Interface Allosterically Transmits Sequence-Specific DNA Signals. *Nat. Struct. Mol. Biol.* **2013**, *20*, 876–883. [[CrossRef](#)]
- [116] Yao, W.; Hu, X.; Wang, X. Crossing Epigenetic Frontiers: The Intersection of Novel Histone Modifications and Diseases. *Signal Transduct. Target. Ther.* **2024**, *9*, 232. [[CrossRef](#)] [[PubMed](#)]
- [117] Sinha, S.; Vargas, A.M.M.; Arantes, P.R.; Patel, A.; O’Connell, M.R.; Palermo, G. Unveiling the RNA-Mediated Allosteric Activation Discloses Functional Hotspots in CRISPR-Cas13a. *Nucleic Acids Res.* **2023**, *52*, 906–920. [[CrossRef](#)] [[PubMed](#)]
- [118] Hersey, A.N.; Kay, V.E.; Lee, S.; Realff, M.J.; Wilson, C.J. Engineering Allosteric Transcription Factors Guided by the LacI Topology. *Cell Syst.* **2023**, *14*, 645–655. [[CrossRef](#)]
- [119] Ridha, F.; Harini, K.; Shanmugam, N.R.S.; Nikam, R.; Gromiha, M.M. Computational Design of Protein Complexes: Influence of Binding Affinity. *Chem. Commun.* **2026**, *62*, 752–762. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The views expressed in this article are those of the author(s) and do not necessarily reflect the views of the publisher or editors. The publisher and editors assume no responsibility for any injury or damage resulting from the use of information contained herein.