BIBECHANA
Vol. 21, No. 3, December 2024, 290-299

ISSN 2091-0762 (Print), 2382-5340 (Online)
Journal homepage: http://nepjol.info/index.php/BIBECHANA
Publisher:Dept. of Phys., Mahendra Morang A. M. Campus (Tribhuvan University)Biratnagar

Characterization of the molecular interactions between Kaiso and CTCF using AlphaFold2 and molecular dynamics simulations

Bidhya Thapa[1,2], **Narayan Prasad Adhikari**[1,*]

[1] *Central Department of Physics, Tribhuvan University, Kirtipur, Kathmandu, Nepal*
[2] *Padma Kanya Multiple Campus, Tribhuvan University, Bagbazar, Kathmandu, Nepal*

[*]Corresponding author. Email: narayan.adhikari@cdp.tu.edu.np

Abstract

Zinc finger (ZF) protein CTCF is an architectural protein that plays a crucial role in global chromatin organization and remodeling via its interaction with several protein partners. The interaction between Kaiso and CTCF regulates the enhancer-blocking function of CTCF. Despite the important biological function of their interactions, structural characterization of the Kaiso-CTCF complex has yet to be carried out. In this work, we have employed molecular modeling and MD simulation to predict the complex between Kaiso and CTCF and investigate its structural features. We have employed Alphafold2 to predict the optimum complex between Kaiso and CTCF and MD simulation to explore the detailed dynamics of the interaction involved in complex formation and stabilization. We predicted the key residues and their interactions involved in the binding of Kaiso to CTCF. Our results show that the hydrophobic interaction between the inter-facial residues plays a significant role in forming the Kaiso-CTCF complex. In addition, several non-covalent interactions, such as hydrogen bonds and electrostatic and van der Waals interactions, stabilize the complex. The significant value of hydrophobic contact area and binding free energy signifies the stability of the predicted Kaiso-CTCF complex.

Keywords
Kaiso, CTCF, Protein-Protein Interactions, AlphaFold2, MD Simulation.

Article information
Manuscript received: August 16, 2024; September 5, 2024; Accepted: September 14, 2024
DOI https://doi.org/10.3126/bibechana.v21i3.68838

1 Introduction

Zinc finger (ZF) protein kaiso is a transcription factor that regulates various cellular functions, such as gene expression, cellular differentiation, development, and cancer progression [1–7]. Kaiso was first identified as the binding partner of armadillo repeat protein p120 catenin [3]. It contains three classical C2H2 ZF domains that are involved in DNA recognition and binding [4, 8]. The highly conserved hydrophobic N-terminal BTB/POZ domain is involved in interactions with other proteins

as well as in homodimerization [9]. Kaiso has dual specificity for DNA binding. It recognizes the doubly methylated CpG dinucleotides (mCpGmCpG) in the core site as well as the sequence-specific Kaiso binding site (KBS), having the nucleotide sequence TCCTGCNA, in DNA [4, 8]. Kaiso recruits nuclear co-repressor complex via its BTB/POZ domain to mediate the transcription repression in a methylation-dependent manner [9]. Depending on the context and cell type, it acts as a transcription repressor [8–10] or an activator [11]. Kaiso is implicated in regulating various genes involved in development and cancer. High levels of Kaiso expression are observed in multiple human cancers, such as lung [2], colon [12], breast [1], and prostate [13].

CCCTC- binding factor (CTCF) protein is a ubiquitous multi-functional transcription factor that binds with the thousands of tissue-specific and conserved sites spread throughout the genome [14–16]. CTCF was initially identified as the ZF protein binding with a CTC-rich sequence in the c-*myc* promoter [17]. It is a 727 amino acid long protein containing 11 ZF domains, two transcriptions repression domains, and N- and C-terminal unstructured loops [18]. Specifically, ZF 4-7 are directly involved in DNA binding, whereas other ZF domains enhance the stability of the CTCF-DNA complex [19]. The CTCF binding sites are primarily present in intergenic regions but also in the enhancer, promoter, and within the gene bodies [15,20]. It regulates gene expression both as a transcription repressor [21] as well as an activator [17], in different cellular contexts. Most importantly, it acts as an enhancer blocker by blocking the communication between an enhancer and target gene promoter and hence preventing transcriptional activation [22,23]. CTCF is an architectural protein that plays a crucial role in global chromatin organization and remodeling via its interaction with several protein partners [14]. In addition, it also participates in genomic imprinting, gametogenesis, and the early stages of mammalian development¸ Becaiter18use of its methylation-sensitive binding with DNA and transcription insulation activity, CTCF allows the accurate expression of the imprinted genes in the H19–Igf2 locus [23, 24].

Kaiso interacts with CTCF protein via the BTB/POZ domain and negatively regulates its enhancer blocker function [25]. Interaction between Kaiso and CTCF is physiologically significant because both proteins are observed to be co-expressed in many cells. Both Kaiso and CTCF are ubiquitously expressed proteins that are mainly present in the nucleus [3, 26]. The Kaiso consensus-sequence KBS is positioned close to the CTCF binding site on the human 5' -HS5 insulator at the -globin gene cluster. The close positioning of the KBS sequence significantly reduces the enhancer-blocking of the -globin insulator [25]. In addition, the Kaiso-CTCF interaction is highly specific, as CTCF does not interact with other POZ domain-containing proteins such as BCL-6, PLZF, and HIC-1 [25]. The Kaiso-CTCF interaction inhibits CTCF binding to its target sites, and it might be used to down-regulate the undesirable enhancer-blocking activity in specific cell types. Similarly, since the POZ domain of Kaiso is involved in recruiting the nuclear co-repressor complex to mediate transcription repression, this interaction might regulate the transcriptional activity of Kaiso. In addition, the POZ domain is involved in the homodimerization of Kaiso (Daniel et al., 1999). Interactions of Kaiso with CTCF could possibly affect its homodimerization and, thereby, other Kaiso-mediated activity [25].

Defossez et al. (2005) have shown that the C-terminal region of CTCF binds with Kaiso. The amino-acid residues 640-724 in CTCF are reported to be involved in the interactions with the POZ domain of Kaiso. However, no experimentally solved structure is available for this complex because of the largely unstructured and flexible nature of this region. In addition, the specific binding region and key residues involved in the interactions are not identified. Identification of key residues involved in the interactions of the protein with its binding partner is essential to understand the structural features of the complex and their binding affinity. In this work, we used the neural network-based modeling method AlphaFold2 to predict the complex between CTCF and Kaiso. Furthermore, we optimized this interaction using MD simulation and assessed the structural stability of the predicted complex. We predicted the key residues involved in the formation of the complex. In addition, we investigated the dynamics of the interactions that are responsible for the binding of the POZ domain of the Kaiso with CTCF. Our results show that the -sheet consisting of two parallel -strand interacts with the C-terminal end of the POZ domain. As the POZ domain in Kaiso is highly hydrophobic and the -sheet in CTCF contains several hydrophobic residues at the binding site, hydrophobic interactions play a significant role in forming the Kaiso-CTCF complex. In addition, several hydrogen bonds, electrostatic, and van der Waals interactions stabilize the complex. The considerable value of binding free energy suggests that the predicted complex is stable.

2 Materials and Methods

2.1 Protein Structures and Molecular Modeling

The input sequences for the Kaiso and CTCF proteins were taken from the UniPort Knowledgebase (UniProtKB) [27]. The complete structure of the

Kaiso and CTCF, shown in Figures 1a and 1b, were modeled using the neural network (NN)-based Alphafold2 implemented in ColabFold [28–30]. ColabFold is an accelerated approach to structure prediction that combines the fast homology search of optimized multiple sequence alignment (MSAs) generation by MMseqs2 (Many-against-Many sequence searching) [30]. We used ColabFold v1.5.5: AlphaFold2 using MMseqs2. The C-terminal region of CTCF is reported to interact with the BTB/POZ domain of Kaiso [25]. Using this information, the amino acids 640-724 in the C-terminal domain were taken as input sequences for CTCF. Similarly, the residues 1-120 in the BTB/POZ domain were taken as the input sequence for Kaiso. The complex between Kaiso and CTCF was modeled using the ColabFold. We used the prediction model alphafold2_ptm for the monomer prediction, whereas the Kaiso-CTCF complex was predicted using the alphafold2_multimer_v3 prediction model. The number of recycles controls the number of times the prediction is repeatedly fed through the model. For monomeric prediction, the number of recycles is chosen to be 3 (num_recycles=3). However, we have used 20 recycles (num_recycles=20) for complex prediction using AlphaFold2_multimer_v3. The recyle_early_stop_tolerance, which specifies when to stop the recycling, was chosen to be 0.0 for monomer and 0.5 for multimer. In addition, we used a single random seed to initialize the prediction (num_seed =1). Similarly, to optimize the MSA, we set the max_msa parameter to auto and the msa_mod to mmseqs2_uniref_env. Similarly, we used the unpaired_paired pair_mod, which pairs sequences from the same species and unpaired MSA. The rank one model predicted from ColabFold was used as the input structure of the Kaiso-CTCF complex for molecular dynamics simulation and further analysis.

2.2 System Setup and Molecular Dynamics Simulations

The system input files for the MD simulation were prepared using the solution builder package of the CHARMM-GUI web server [31]. The system was solvated using TIP 3P water in a cubical box with a 10 Å padding around the complex and neutralized by adding 0.15M of NaCl, which resulted in the system containing 36,210 atoms in the box with dimensions of $73 \times 73 \times 73$ Å^3. All-atom MD simulations were performed with the NAMD 2.14 package [32] using CHARMM36m force field [33], an improved force field for folded and intrinsically disordered protein. Standard MD protocol was followed for the simulation as in our previous works. [34–36]. Briefly, the energy minimization of the system was carried out for 10,000 steps using the conjugate gradient and line search algorithm. The equilibration run was performed for 125 ps with a 1 fs integration time step at 300K. During the equilibration run, the protein's heavy atoms were harmonically restrained with a force constant of 1.0 kcal/mol/Å for the backbone and 0.1 kcal/mol/Å for the side chain. The particle Mesh Ewald (PME) method [37] was employed to calculate the long-range interactions. The cut-off distance of 12 Å taken for non-bonded interactions. The Nosé Hoover Langevin method with a piston period of 50 fs and a decay of 25 fs was used to control the pressure. Similarly, the temperature was controlled by employing the Langevin temperature coupling with a friction coefficient of 1ps^{-1}. The production run was propagated for 350 ns in NPT condition at 300 K and 1 atm pressure, taking a 2 fs time step.

2.3 System Setup and Molecular Dynamics Simulations

The MD simulation trajectories were analyzed using visual molecular dynamics (VMD) [38]. VMD was also used for visualization and image rendering purposes. The MM/GBSA binding free energy was estimated using NAMD. The hydrogen bond plugin in VMD was used to analyze the complex's inter-protein hydrogen bonding. The heavy atom distance cut-off of 3.5 Å and bond angle cut-off of 30° were used to calculate hydrogen bonds. Similarly, the solvent-accessible surface area (SASA) of the proteins and complex was estimated using VMD. The non-bonded interaction energies were calculated using the NAMD energy plugin in VMD.

3 Results and Discussion

We performed modeling and MD simulation to predict the complex between Kaiso and CTCF and investigated the structural stability and dynamics of the interactions in the Kaiso-CTCF complex. The optimum complex between Kaiso and CTCF was predicted using Alphafold2-multimer implemented in Google Collaboratory, ColabFold. The MD simulation was employed to assess the structural integrity of the predicted complex and identify the major residues involved in forming and stabilizing the Kaiso-CTCF complex. In addition, we studied the dynamics of the non-covalent interactions, such as hydrogen bonding, hydrophobic, electrostatic, and van der Waals interactions, during the simulation. We also estimated the hydrophobic contact area between Kaiso and CTCF as well as the binding free energy of the Kaiso-CTCF complex.

3.1 Modeling of the Kaiso-CTCF Complex

The amino acid residues 641-727 in the C-terminal region of CTCF were predicted to be involved in the interactions with the BTB/POZ domain of Kaiso [25]. Based on this information, we selected 641-727 residues from the canonical sequence of CTCF and used them to create a complex with the BTB/POZ domain of Kaiso. As shown in Figure 1a, the C-terminal region of CTCF is the unstructured loop. Solving the crystal structure of the protein-protein complex with large unstructured loops is experimentally challenging. In such conditions, computational techniques such as molecular modeling and MD simulations guide the structure prediction and characterization.

The CTCF residues 641-727 were predicted to be involved in the interactions with the BTB/POZ domain of Kaiso [25]. Based on this information, we selected 640-727 residues from the canonical sequence of CTCF and BTB/POZ domain residues 1-120 of Kaiso as the input sequence in the ColabFold v1.5.5: AlphaFold2 using MMseqs2 to predict the complex. The Kaiso-CTCF complex was predicted using the alphafold2_multimer_v3 prediction model. The Materials and Methods section explains the details of the parameters and conditions used for the complex prediction. We chose the complex with rank one, with a predicted local distance difference test (pLDDT) score of 75.5 and a predicted template modeling (pTM) score of 0.65. The pTM score greater than 0.5 indicates that the predicted complex is close to the true structure. Since the BTB/POZ domain of Kaiso has a stable structural domain (Figure 1b), the predicted lDDT of Kaiso residues in the range of 10-117 is greater than 80, whereas the lDDT of most of the CTCF residues is less than 50. However, the predicted lDDT score of the CTCF residues in the range 670-690, which are involved in binding with Kaiso, is greater than 60. Figure 1c shows the Kaiso-CTCF complex predicted by ColabFold. Since all-atom MD simulation is computationally expensive, we truncated the complex by taking the regions likely involved in interactions. Since only the residues 666-690 in CTCF are within 5 Å of Kaiso, they are most likely to interact with it. Therefore, we truncated the predicted complex by taking residues 662-692 for CTCF and 9-120 for Kaiso. As the BTB/POZ domain of Kaiso is a stable domain, we have taken the entire domain as an input structure to maintain its stability. Figure 1d shows the Kaiso-CTCF complex used for all-atom MD simulation. The predicted Kaiso-CTCF complex is stable throughout the simulation and features several interprotein interactions that stabilize it, as shown in Figure 2a.

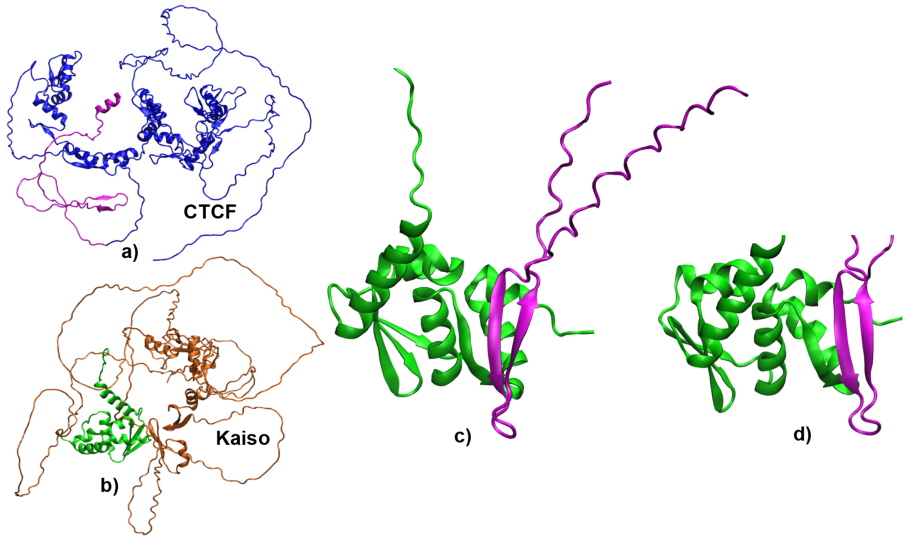

Figure 1: (a) The full-length structure of CTCF highlighting the residues 640-727 in the C-terminal domain in magenta. (b) The complete structure of the Kaiso showing the BTB/POZ domain in green. (c) Kaiso-CTCF complex structure predicted from ColabFold. (d) Structure of the Kaiso-CTCF complex used for the MD simulation.

3.2 Structural stability of the complex

We estimated the root mean square deviation (RMSD) of the predicted Kaiso-CTCF complex to assess the structural integrity. The time evolution of the RMSD of the complex was calculated using the initial frame of the simulation as the reference structure. The protein backbone atoms were used to calculate the RMSD values. As shown in the

RMSD plot in Figure 2b, the Kaiso-CTCF complex undergoes structural reorganization up to 200 ns of the simulation and becomes stable. The average RMSD of the complex settles down to nearly 3.5 Å. Individually, the CTCF shows larger structural variation compared to the BTB/POZ domain of Kaiso, owing to its flexible structure compared to the stable BTB domain of Kaiso. We used the MD simulation trajectories after 200ns, where the system is properly stabilized, to estimate the average properties of the system. In addition, to analyze the flexibility of each residue in the complex, we calculated the root mean square fluctuation (RMSF) of individual residues using the backbone C atoms. As shown in Figures 2c and 2d, the RMSF of Kaiso residues is significantly stable compared to the residues in CTCF. As expected, the extremities of both proteins are highly flexible due to a lack of interactions with another protein. In Kaiso, except for the loop containing residues 62-66, most residues have RMSF of less than 2 Å. In CTCF, the residues in 668-676 and 680-689 that are involved in interaction with Kaiso have lower RMSF, indicating higher stability.

3.3 Major interactions in the Kaiso-CTCF complex

We investigated the various non-covalent interactions involved in forming and stabilizing the complex between Kaiso and CTCF. As shown in Figure 2a, the Kaiso-CTCF complex has several hydrophobic, hydrogen bonding, and ionic interactions. We explored the detailed dynamics of these interactions during the simulation.

Hydrophobic Interactions

The BTB/POZ domain of Kaiso and the -sheet in the binding site of CTCF contain multiple hydrophobic residues. Therefore, several inter-protein hydrophobic interactions are present at the binding interface of the Kaiso-CTCF complex. Figure 3a shows the representative snapshot of the MD simulation showing the hydrophobic interactions in the Kaiso-CTCF. Our results predict that the VAL673 and ILE684 in CTCF make hydrophobic interactions with VAL93 and LEU98 of Kaiso. Similarly, ILE90, LEU101, and ILE102 in Kaiso create hydrophobic grooves with VAL686, ILE671, and ALA669 in CTCF. In addition, ILE113 and LEU116 of Kasio interact with ALA669 of CTCF. Moreover, PHE112 in Kaiso makes hydrophobic contact with PRO667 and VAL668 in CTCF.

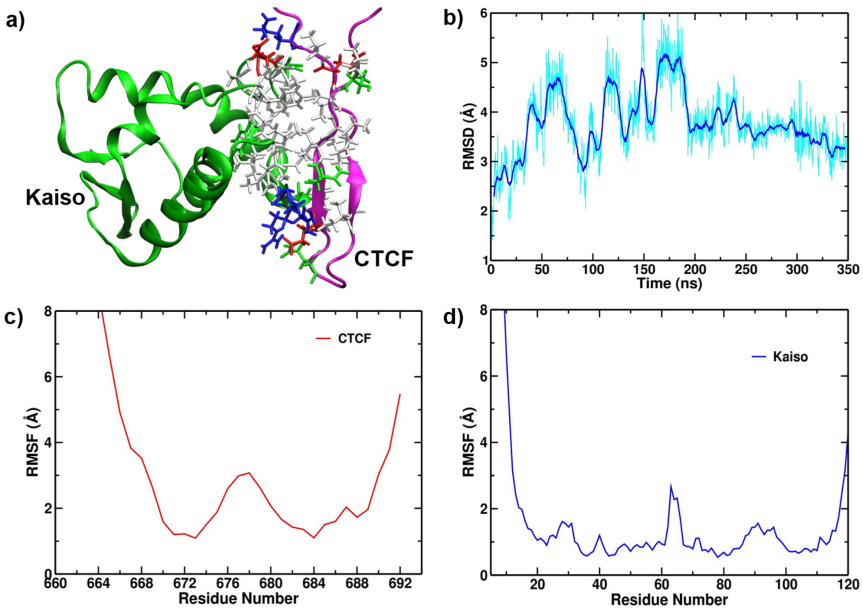

Figure 2: (a) Kasio-CTCF complex at the end of 350 ns MD simulation. (b) Time evolution of the RMSD measurement of the Kaiso-CTCF complex. RMSF measurements of the (c) CTCF and (d) CTCF residues in the Kaiso-CTCF complex during simulation.

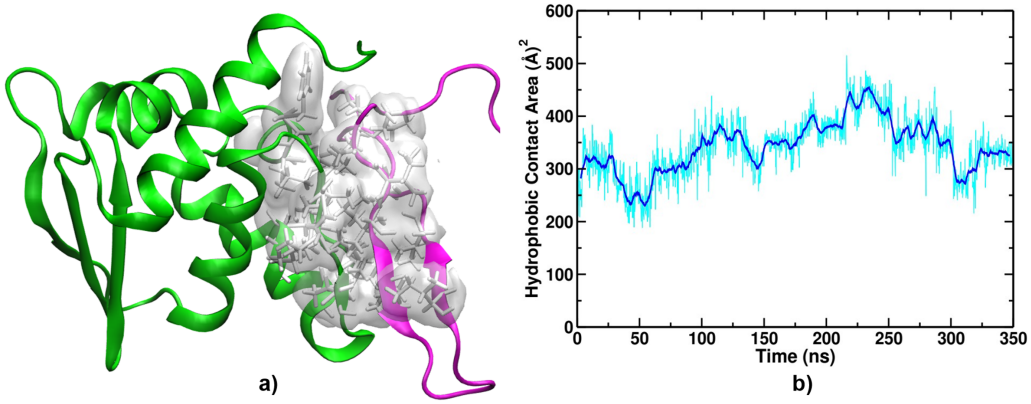

Figure 3: (a) Representative structure of the Kasio-CTCF complex during the simulation highlighting the hydrophobic interactions at the binding interface. (b) Time evolution of the hydrophobic contact area between the Kaiso and CTCF during the simulation.

We calculated the hydrophobic contact area between Kaiso and CTCF to estimate the contribution of the hydrophobic interactions to the stability of the complex. The hydrophobic contact area is the hydrophobic surface buried at the binding interface between two proteins upon complex formation. The higher the contact area, the greater the interacting surface between the interacting partners, which leads to the greater stability of the complex [39]. We calculated the contact area from the SASA of individual proteins and complex using the following equation [39].

$$\text{Contact Area} = \frac{(S_{Kaiso}+S_{CTCF}-S_{Complex})}{2}$$

where S_{Kaiso}, S_{CTCF}, and $S_{complex}$ are the hydrophobic SASA of the Kaiso, CTCF, and Kaiso-CTCF complex.

Figure 3b shows the time evolution of the hydrophobic contact area between the Kaiso and CTCF. The average hydrophobic contact area for the last 150 ns of the simulation trajectories is 365 ± 54 Å^2. Previous studies have shown that the burial of 1 Å^2 of hydrophobic surface area at the binding interface contributes about -15 ± 1.2 cal/mol [39, 40] of free energy to stabilize the complex. Therefore, the 365 Å^2 of the hydrophobic contact area of the Kaiso-CTCF complex contributes about -5.5 kcal/mol to its stability. Therefore, hydrophobic interactions are crucial in the Kaiso-CTCF complex formation and stabilization.

Hydrogen Bonds Interactions

We analyzed the inter-protein hydrogen bond interactions between the Kaiso and CTCF using the criterion explained in the Methods section. The time evolution of the total number of hydrogen bonds between inter-facial residues in the Kaiso-CTCF complex is shown in Figure 4a. As the complex undergoes structural reorganization till 200 ns and becomes stable, we used the simulation trajectory of 200-350 ns to estimate the occupancy percentage of the hydrogen bonds. Figure 4b shows the occupancy percentage of major hydrogen bonds involved in forming the Kaiso-CTCF complex. As more than one pair of atoms are involved in forming hydrogen bonds in a given residue pair, the occupancy percentage included in Figure 4b is the aggregation of all hydrogen bonds formed within a given residue pair. Figure 4c shows the atomic details of the significant inter-protein hydrogen bonds in the Kaiso-CTCF complex. Since most of the residues in the binding interface are hydrophobic, the backbone atoms are involved in hydrogen bonding. Even with the polar residues, hydrogen bonds are formed between backbone atoms as the side chains are orientated away from the binding interface (Figure 4c). The backbone nitrogen of LYS689 in CTCF forms a hydrogen bond with the oxygen atom of GLU115 in Kaiso. This bond is formed around 125 ns and remains consistently stable after 150 ns, as shown in Figure 4d. Similarly, the hydrogen bond between GLY117 in Kaiso and LYS689 in CTCF becomes stable after 150 ns. In addition, the backbone nitrogen and oxygen of VAL93 in Kaiso form hydrogen bonds with the backbone atoms of GLN672 and GLU674 in CTCF, respectively. These bonds are formed at the start of the simulation and remain stable throughout (Figure 4d). In addition, the

ARG92(Kaiso)- GLU674(CTCF) is the only major hydrogen bond formed between the side chains of Kaiso and CTCF.

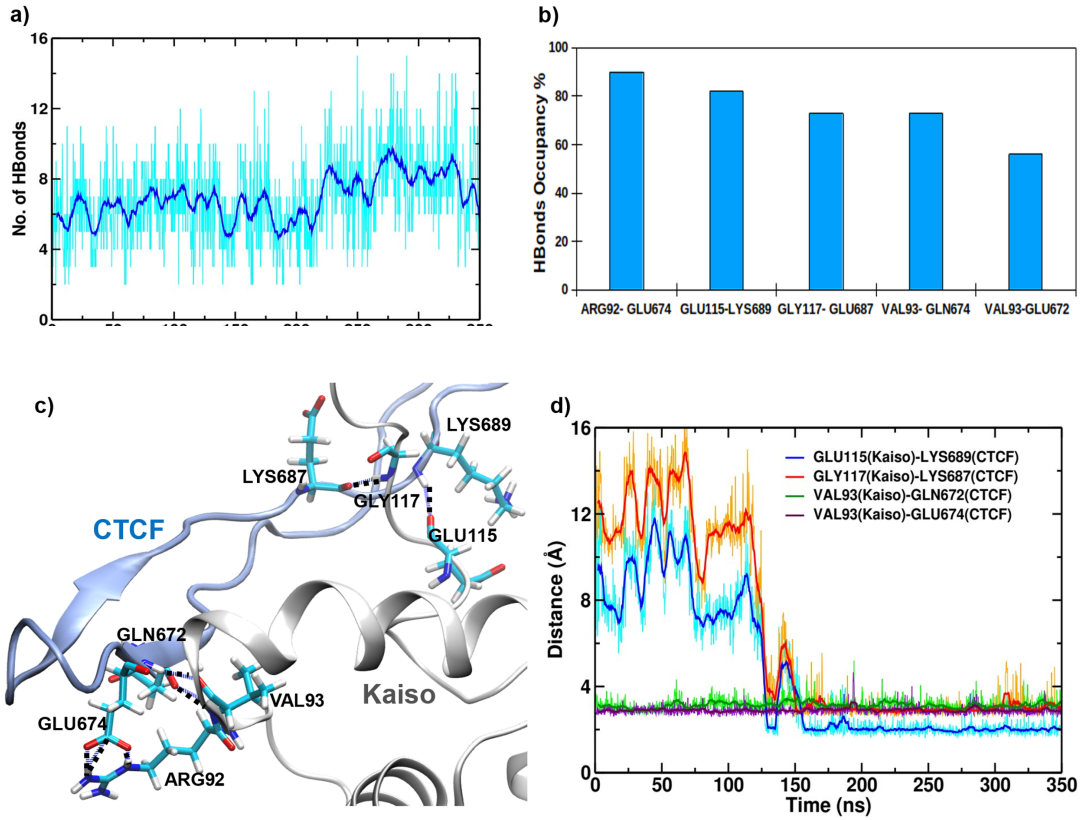

Figure 4: (a) Time evolution of the total inter-protein hydrogen bonds in the Kaiso-CTCF complex during the simulation. (b) Percentage occupancy of the major hydrogen bonds in the Kaiso-CTCF complex. (c) Details of the major inter-protein hydrogen bonds in the Kaiso-CTCF complex (shown in the licorice structure). The dotted lines represent the hydrogen bonds. (d) Time evolution of the distance between the heavy atoms forming major inter-protein hydrogen bonds in the Kaiso-CTCF complex.

Ionic Interactions

The ionic interactions result from the electrostatic attraction between the side chains of the charged residues at the binding interface between two proteins. The negatively charged side chain of GLU674 in CTCF makes ionic interaction with the positively charged side chains of ARG92 and ARG94 of Kaiso. Similarly, the GLU115 in Kaiso makes ionic interactions with the side chains of LYS663, LYS689, and LYS690 in CTCF. In the Kaiso-CTCF complex, a strong, consistent salt bridge is not formed between these charged residues, as their heavy atom pairs are beyond 3.2 Å during most of the simulation time. Nevertheless, the ARG92 and ARG94 in Kaiso make transient salt bridges with GLU674 in CTCF. These ionic interactions enhance the binding of these two proteins to form a stable complex. We further estimated the contribution of the non-bonded interaction in stabilizing the Kaiso-CTCF complex. We calculated the average values of non-bonded interaction energy using the last 150 ns of the simulation trajectory. The electrostatic and van der Waals interactions contribute about -199 ± 72 kcal/mol and -56 ± 10 kcal/mol to the total energy of the complex.

3.4 Binding Free Energy of the Complex

We estimated the binding free energy of the Kaiso-CTCF complex using molecular mechanics with a generalized Born and surface area solvation (MM/GBSA) approach [41–43]. Previous studies have shown that the MM/GBSA method has some limitations in calculating the absolute binding free energy of the complex and tends to overestimate the binding free energy [44, 45]. However, it is a computationally efficient way to calculate the relative binding energy of a complex. As in our previous works [36, 46, 47], we used NAMD to calculate the MM/GBSA energy using the simulation trajectories from the last 150 ns of the simulation, where the complex is properly stabilized. The binding free energy of the Kaiso-CTCF complex is -45.1 $\pm$ 5.2 kcal/mol. Even though the MM/GBSA approach overestimates the binding free energy [45], comparing this value with our previous works [46, 47] using the same method indicates a stable complex between Kaiso and CTCF.

4 Conclusion

In this work, we employed the neural network-based modeling method AlphaFold2 and MD simulation to predict the complex between the Kaiso and CTCF and investigated the structural features of the predicted Kaiso-CTCF complex. We predicted the specific binding region as well as the key residues and their role in forming and stabilizing the complex. In addition, we studied the atomic-level detailed dynamics of various inter-protein non-covalent interactions to understand the complex formation between the CTCF and Kaiso. Our results show that the residues 662-690 towards the end of the C-terminal end of CTCF, containing a -sheet with two parallel -strands, bind with the BTB/POZ domain of Kaiso. Our simulation results reveal that hydrophobic interactions between the inter-facial residues are crucial in forming the stable complex between Kaiso and CTCF. In addition, several hydrogen bonding and ionic interactions are essential for binding two proteins. Similarly, electrostatic and van der Waals interactions also contribute to complex stabilization. The significant value of hydrophobic contact area and binding free energy indicate the stability of the predicted complex. To the best of our knowledge, it is the first computational work in modeling the Kaiso-CTCF complex and investigating its structural features. Previous studies have reported that Kaiso-CTCF interaction negatively regulates CTCF insulator activity [25]; our first structural characterization outlines the molecular basis of the interaction of Kaiso with CTCF.

Acknowledgments

BT and NPA acknowledge support from the Research Coordination and Development Council (RCDC) of Tribhuvan University Grants number TU-NPAR-077/78-ERG 14. BT acknowledges the PhD Fellowship and Research Support Grant (Award number PhD-78/79-S&T-15) from the University Grants Commission (UGC), Nepal.

Author Contributions

BT performed the modeling and MD simulation, analyzed the data, and wrote the manuscript. NPA supervised the work and contributed to data analysis and interpretation.

Data Availability

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

References

[1] BI Bassey-Archibong et al. Kaiso depletion attenuates the growth and survival of triple negative breast cancer cells. *Cell death & disease*, 8(3):e2689–e2689, 2017.

[2] SD Dai et al. Cytoplasmic kaiso is associated with poor prognosis in non-small cell lung cancer. *BMC cancer*, 9:1–13, 2009.

[3] JM Daniel and AB Reynolds. The catenin p120ctn interacts with kaiso, a novel btb/poz domain zinc finger transcription factor. *Molecular and cellular biology*, 19(5):3614–3623, 1999.

[4] JM Daniel et al. The p120ctn-binding partner kaiso is a bi-modal dna-binding protein that recognizes both a sequence-specific consensus and methylated cpg dinucleotides. *Nucleic acids research*, 30(13):2911–2919, 2002.

[5] J Jen and YC Wang. Zinc finger proteins in cancer progression. *Journal of biomedical science*, 23(1):53, 2016.

[6] J Jones et al. Nuclear kaiso indicates aggressive prostate cancers and promotes migration and invasiveness of prostate cancer cells. *The American journal of pathology*, 181(5):1836–1846, 2012.

[7] CC Pierre et al. Dancing from bottoms up–roles of the poz-zf transcription factor kaiso in cancer. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, 1871(1):64–74, 2019.

[8] A Prokhortchouk et al. The p120 catenin partner kaiso is a dna methylation-dependent transcriptional repressor. *Genes & development*, 15(13):1613–1618, 2001.

[9] HG Yoon et al. N-cor mediates dna methylation-dependent repression through a methyl cpg binding protein kaiso. *Molecular cell*, 12(3):723–734, 2003.

[10] NS Donaldson et al. Kaiso represses the cell cycle gene cyclin d1 via sequence-specific and methyl-cpg-dependent mechanisms. *PloS one*, 7(11):e50398, 2012.

[11] M. Rodova, K.F. Kelly, M. VanSaun, J.M. Daniel, and M.J. Werle. Regulation of the rapsyn promoter by kaiso and -catenin. *Molecular and Cellular Biology*, 24(16):7188–7196, 2004.

[12] E.C. Lopes, E. Valls, M.E. Figueroa, A. Mazur, F.G. Meng, G. Chiosis, P.W. Laird, W. Schüler, S.M. Lipkin, S.B. Baylin, et al. Kaiso contributes to dna methylation-dependent silencing of tumor suppressor genes in colon cancer cell lines. *Cancer Research*, 68(18):7258–7263, 2008.

[13] H. Wang, G. Liu, D. Shen, H. Ye, J. Huang, H. Jiang, and J. Lu. Kaiso, a transcriptional repressor, promotes cell migration and invasion of prostate cancer cells through regulation of mir-31 expression. *Oncotarget*, 7(5):5677–5691, 2016.

[14] S.J.B. Holwerda and W. de Laat. Ctcf: the protein, the binding partners, the binding sites and their chromatin loops. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1620):20120369, 2013.

[15] Y. Shen, F. Yue, D.F. McCleary, Z. Ye, L. Edsall, S. Kuan, U. Wagner, J. Dixon, L. Lee, V.V. Lobanenkov, et al. A map of the cis-regulatory sequences in the mouse genome. *Nature*, 488(7409):116–120, 2012.

[16] H. Wang, M.T. Maurano, H. Qu, K.E. Varley, J. Gertz, F. Pauli, K. Lee, T. Canfield, M. Weaver, R. Sandstrom, et al. Widespread plasticity in ctcf occupancy linked to dna methylation. *Genome Research*, 22(9):1680–1688, 2012.

[17] E.M. Klenova, R.H. Nicolas, H.F. Paterson, A.F. Carne, C.M. Heath, G.H. Goodwin, P.E. Neiman, and V.V. Lobanenkov. Ctcf, a conserved nuclear factor required for optimal transcriptional activity of the chicken c-myc gene, is an 11-zn-finger protein differentially expressed in multiple forms. *Molecular and Cellular Biology*, 13(10):7091–7101, 1993.

[18] M.M. Franco, A.R. Prickett, and R.J. Oakey. The role of ccctc-binding factor (ctcf) in genomic imprinting, development, and reproduction. *Biology of Reproduction*, 91(5):125, 1–9, 2014.

[19] M. Renda, I. Baglivo, B. Burgess-Beusse, S. Esposito, R. Fattorusso, G. Felsenfeld, and P.V. Pedone. Critical dna binding interactions of the insulator protein ctcf: a small number of zinc fingers mediate strong binding, and a single finger-dna interaction controls binding at imprinted loci. *Journal of Biological Chemistry*, 282(46):33336–33345, 2007.

[20] T.H. Kim, Z.K. Abdullaev, A.D. Smith, K.A. Ching, D.I. Loukinov, R.D. Green, M.Q. Zhang, V.V. Lobanenkov, and B. Ren. Analysis of the vertebrate insulator protein ctcf-binding sites in the human genome. *Cell*, 128(6):1231–1245, 2007.

[21] V. Lobanenkov, R.H. Nicolas, H. Paterson, P. de Jong, G.H. Goodwin, V.V. Lobanenkov, G.P. Lomonossoff, and P.E. Neiman. A novel sequence-specific dna binding protein which interacts with three regularly spaced direct repeats of the ccctc-motif in the 5-flanking sequence of the chicken c-myc gene. *Oncogene*, 5(12):1743–1753, 1990.

[22] A.C. Bell, A.G. West, and G. Felsenfeld. The protein ctcf is required for the enhancer blocking activity of vertebrate insulators. *Cell*, 98(3):387–396, 1999.

[23] A.T. Hark, C.J. Schoenherr, D.J. Katz, R.S. Ingram, J.M. Levorse, and S.M. Tilghman. Ctcf mediates methylation-sensitive enhancer-blocking activity at the h19/igf2 locus. *Nature*, 405(6785):486–489, 2000.

[24] A.C. Bell and G. Felsenfeld. Methylation of a ctcf-dependent boundary controls imprinted expression of the igf2 gene. *Nature*, 405(6785):482–485, 2000.

[25] P.-A. Defossez, K.F. Kelly, G.J. Filion, R. Perez-Torrado, F. Magdinier, H. Menoni, C.L. Nordgaard, J.M. Daniel, and E. Gilson. The human enhancer blocker ctc-binding factor interacts with the transcription factor kaiso. *Journal of Biological Chemistry*, 280(52):43017–43023, 2005.

[26] G.N. Filippova, S. Fagerlie, E.M. Klenova, C. Myers, Y. Dehner, G.H. Goodwin, P.E. Neiman, S.J. Collins, and V.V. Lobanenkov. An exceptionally conserved transcriptional repressor, ctcf, employs different combinations of zinc fingers to bind diverged promoter

sequences of avian and mammalian c-myc oncogenes. *Molecular and Cellular Biology*, 16(6):2802–2813, 1996.

[27] The UniProt Consortium. Uniprot: the universal protein knowledgebase in 2021. *Nucleic acids research*, 49(D1):D480–D489, 2021.

[28] Milot Mirdita, Martin Steinegger, and Johannes Söding. Mmseqs2 desktop and local web server app for fast, interactive sequence searches. *Bioinformatics*, 35(16):2856–2858, 2019.

[29] Richard Evans, Michael O'Neill, Alexander Pritzel, et al. Protein complex prediction with alphafold-multimer. *bioRxiv*, page 2021.10.04.463034, 2021.

[30] Milot Mirdita, Konstantin Schütze, Yu Moriwaki, et al. Colabfold: making protein folding accessible to all. *Nature methods*, 19(6):679–682, 2022.

[31] Jooyoung Lee, Xiaolin Cheng, Jason M Swails, et al. Charmm-gui input generator for namd, gromacs, amber, openmm, and charmm/openmm simulations using the charmm36 additive force field. *Biophysical journal*, 110(3):641a, 2016.

[32] James C Phillips, Rosemary Braun, Wei Wang, et al. Scalable molecular dynamics with namd. *Journal of computational chemistry*, 26(16):1781–1802, 2005.

[33] Jing Huang and Alexander D MacKerell. Charmm36 all-atom additive protein force field: Validation based on comparison to nmr data. *Journal of computational chemistry*, 34(25):2135–2145, 2013.

[34] Rajendra Koirala, Bidhya Thapa, Shyam P Khanal, Jhulan Powrel, Rajendra P Adhikari, and Narayan P Adhikari. Binding of sars-cov-2/sars-cov spike protein with human ace2 receptor. *Journal of Physics Communications*, 5(3):035010, 2021.

[35] Bidhya Thapa and Narayan P Adhikari. Molecular interactions of zinc finger protein kaiso with hemimethylated dna. *Journal of Nepal Physical Society*, 9(1):1–10, 2023.

[36] Bidhya Thapa, Narayan P Adhikari, Purushottam B Tiwari, and Chapagain. Prem P. A 5-flanking c/g pair at the core region enhances the recognition and binding of kaiso to methylated dna. *Journal of Chemical Information and Modeling*, 63(7):2095–2103, 2022.

[37] Ulrich Essmann, Lalith Perera, Max L Berkowitz, et al. A smooth particle mesh ewald method. *The Journal of chemical physics*, 103(19):8577–8593, 1995.

[38] William Humphrey, Andrew Dalke, and Klaus Schulten. Vmd: visual molecular dynamics. *Journal of molecular graphics*, 14(1):33–38, 1996.

[39] Xue Zou, Wei Ma, Ilia A Solov'yov, et al. Recognition of methylated dna through methyl-cpg binding domain proteins. *Nucleic acids research*, 40(6):2747–2758, 2012.

[40] Barbara Vallone, Adriana E Miele, Pier Carlo Vecchini, et al. Free energy of burying hydrophobic residues in the interface between protein subunits. *Proceedings of the National Academy of Sciences*, 95(11):6103–6107, 1998.

[41] Tingjun Hou, Junmei Wang, Youyong Li, et al. Assessing the performance of the mm/pbsa and mm/gbsa methods. 1. the accuracy of binding free energy calculations based on molecular dynamics simulations. *Journal of chemical information and modeling*, 51(1):69–82, 2011.

[42] Hao Sun, Yu Li, Shengnan Tian, et al. Assessing the performance of mm/pbsa and mm/gbsa methods. 4. accuracies of mm/pbsa and mm/gbsa methodologies evaluated by various simulation protocols using pdbbind data set. *Physical Chemistry Chemical Physics*, 16(31):16719–16729, 2014.

[43] E Wang, Hao Sun, J Wang, et al. End-point binding free energy calculation with mm/pbsa and mm/gbsa: strategies and applications in drug design. *Chemical reviews*, 119(16):9478–9508, 2019.

[44] Tiziano Tuccinardi. What is the current value of mm/pbsa and mm/gbsa methods in drug discovery? *Expert opinion on drug discovery*, 16(11):1233–1237, 2021.

[45] Chethan Mulakala and Vellarkad N Viswanadhan. Could mm-gbsa be accurate enough for calculation of absolute protein/ligand binding free energies? *Journal of Molecular Graphics and Modelling*, 46:41–51, 2013.

[46] Lokendra S Dhami, Prabin Dahal, Bidhya Thapa, Narayan Gautam, Nurapati Pantha, Rameshwar Adhikari, and Narayan P Adhikari. Insights from in silico study of receptor energetics of sars-cov-2 variants. *Physical Chemistry Chemical Physics*, 2024.

[47] Bidhya Thapa, N P Adhikari, Rakesh P Koirala, et al. Investigating molecular interactions between kaiso and nuclear co-repressor using molecular simulations. *AIP Advances*, 14(6), 2024.