eMicrobe
https://www.sciltp.com/journals/eMicrobe

Article

Molecular Dynamics Simulation to Explore Functional Regions Involving Pho4 Interaction with Co-Activator and *PHO5* Promoter

Tanaporn Wangsanut [1], Nopawit Khamto [2] and Monsicha Pongpom [1,*]

[1] Department of Microbiology, Faculty of Medicine, Chiang Mai University, Chiang Mai 50200, Thailand
[2] Department of Biochemistry, Faculty of Medical Science, Naresuan University, Phitsanulok 65000, Thailand
* Correspondence: monsicha.p@cmu.ac.th

How To Cite: Wangsanut, T.; Khamto, N.; Pongpom, M. Molecular Dynamics Simulation to Explore Functional Regions Involving Pho4 Interaction with Co-Activator and *PHO5* Promoter. *eMicrobe* **2025**, *1*(1), 5. https://doi.org/10.53941/emicrobe.2025.100005.

Received: 20 June 2025
Revised: 8 September
Accepted: 15 September
Published: 24 September

Abstract: The PHO pathway is essential for the transcriptional response to inorganic phosphate starvation, and the core components are evolutionarily conserved across fungi. Pho4 and its co-activator Pho2 control the *PHO5* gene expression and serve as a paradigm for understanding chromatin remodelling and the evolution of combinatorial transcriptional control. Protein sequence alignment and structural modelling reveal that Pho4 homologs contain a large proportion of intrinsically disordered regions (IDRs). Using molecular dynamics (MD) simulations, we monitored the behaviors of the yeast *Saccharomyces cerevisiae* Pho4 protein (ScPho4), which includes the DNA Binding Domain (DBD), and the Pho2 Interaction Domain (P2ID), an IDR, in complex with the *PHO5* promoter (*ScPHO5*), both in the presence and absence of the Pho2 protein (ScPho2). The ScPho2-ScPho4 dimer formed a more stable complex at *ScPHO5*, as reflected by a lower RMSD and greater binding energy, compared to ScPho4 alone. This dimer also induced more structural changes in promoter DNA. Notably, the P2ID of ScPho4 adopted an open conformation when interacting with ScPho2, and a closed conformation in its absence. In human pathogenic yeast *Candida glabrata*, MD simulations revealed dynamic IDR movements in Pho4 homologs. Unlike other previous studies, our approach enables direct visualization and quantification of conformational changes in protein-DNA complex over simulation time. These findings highlight the value of MD simulation as a routine complement to experimental methods, advancing our knowledge of structure-motion-function.

Keywords: molecular dynamics simulation; Pho2; Pho4; yeasts; *Saccharomyces cerevisiae*; *Candida glabrata*

1. Introduction

Inorganic phosphate (Pi) is an essential nutrient for fungi, playing important roles in signal transduction, metabolism, and interactions with various hosts (such as animals, plants, and microbes [1]) and environments. To maintain phosphate homeostasis, fungi rely on the PHO pathway, a signaling system that regulates intracellular Pi levels. The PHO pathway is best studied in the model yeast *Saccharomyces cerevisiae* [2,3]. The basic helix-loop-helix (bHLH) transcription factor Pho4 (ScPho4) and the homeodomain protein Pho2 (ScPho2) coordinate the expression of the PHO regulon [3,4], which is involved in Pi acquisition, transportation, and storage under phosphate limitation. Under high Pi, the cyclin-dependent kinase complex Pho80–Pho85 phosphorylates ScPho4, retaining it in the cytoplasm. Under low Pi, Pho81 inhibits this kinase complex, leading to ScPho4 dephosphorylation, its nuclear localization, and activation of the PHO regulon (For details, see Figure 1). Inositol

pyrophosphate IP7 also plays a critical regulatory role. A recent study in the human pathogenic yeast *Cryptococcus neoformans* has shown that IP7 directly binds to Pho81, stabilizing its interaction with Pho80-Pho85 [5] and thereby activating *PHO* gene expression, irrespective of phosphate status [6].

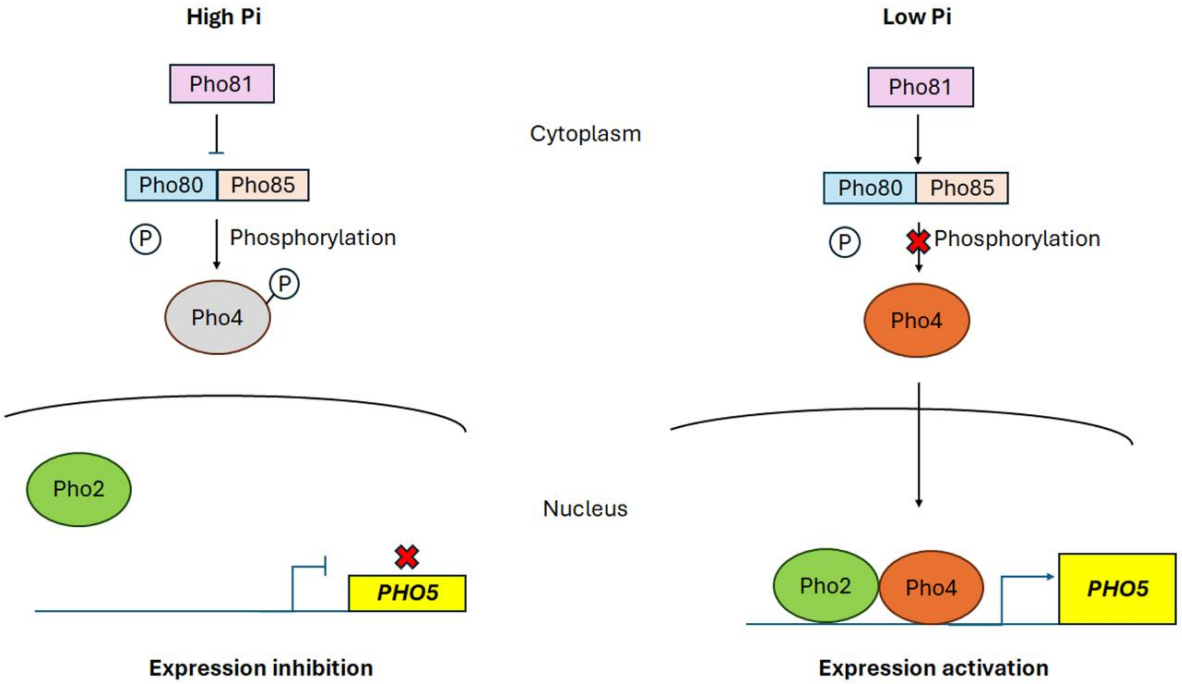

Figure 1. The PHO signaling pathway in *Saccharomyces cerevisiae*. Under high phosphate (Pi) conditions, Pho4 is phosphorylated by the Pho80–Pho85 kinase complex and retained in the cytoplasm. Under low Pi, Pho4 is unphosphorylated and translocated to the nucleus, where it interacts with Pho2 to activate the expression of *PHO* genes.

The PHO pathways have been investigated in other fungal species and revealed the extended roles of phosphate homeostasis to cellular transport of nutrients, carbohydrate and lipid metabolism, and response to various environmental stresses [2,7]. Specifically, while the mechanism of the phosphate starvation response in *S. cerevisiae* is well conserved in other fungi, the key transcriptional regulators can vary and account for the expansion of target genes. In the model fission yeast *Schizosaccharomyces pombe*, the phosphate starvation response is regulated by the Zn_2Cys_6 binuclear transcription factor Pho7, which is structurally different from ScPho4 and ScPho2. Pho7 regulates the expression of genes in response to multiple stress conditions by coordinating the proper expression of each stress-specific nutrient transport [8]. In human pathogenic yeasts *Candida glabrata* (*Nakaseomyces glabratus*) and *Candida albicans*, the expression of the *PHO* regulon under phosphate deprivation is less dependent on Pho2 [9–11]. The evolutionary analysis suggested that reduced Pho2 dependency could allow for rapid expansion of the Pho4 gene regulatory network [11]. For instance, in *C. glabrata*, the CgPho4 additionally targets the genes involved in carbon metabolism and the response to osmotic and oxidative stress [11]. In *C. albicans*, CaPho4 is also required for stress resistance and metal homeostasis [10,12]. In *C. neoformans*, CnPho4 solely regulates the PHO pathway during phosphate starvation, and additionally governs the expression of genes involved in nutrient transporters, carbohydrate, and lipid metabolism [13]. Together, the expansion of Pho4 target genes leads to an emerging paradigm that phosphate homeostasis is integrated with the regulation of other metabolisms and stress response pathways to facilitate pathogenic lifestyles of *Candida* spp. and *C. neoformans*.

In filamentous fungal model *Neurospora crassa*, Nuc-1 is the ScPho4 homolog that controls phosphate acquisition [14]. Notably, the regulatory function of Nuc-1 is also expanded to carbohydrate metabolism and transport, consistent with *Candida* spp. and *C. neoformans* [15]. In another filamentous fungal model, *Aspergillus nidulans*, the PalcA protein is the ScPho4 homolog that regulates the PHO systems [16]. So far, there is currently no evidence that Nuc-1 or PalcA regulates the PHO pathways with other transcription factors. In methylotrophic yeast *Hansenula polymorpha*, Pho4 (HpPho4), but not the Pho2 homolog (HpPho2), is required for the expression of repressible acid phosphatase under low phosphate conditions [17]. These data from non-pathogenic fungi argue against the paradigm that the expansion of Pho4 target genes and the extended role of the PHO pathway are merely

to accommodate a pathogenic lifestyle. Thus, the evolution of PHO pathways and the network of transcriptional regulators remains to be elucidated.

In particular, the requirement of Pho2 in the regulation of *PHO* genes has sparked research into the evolutionary basis of combinatorial gene regulation. The transcriptional control of the *PHO5* repressible acid phosphatase gene by ScPho4 and ScPho2 complex is well characterized in *S. cerevisiae*. He et al. 2017 conducted the heterologous expression study by expressing *PHO4* and *PHO2* homologous genes from hemiascomycete yeast species in the *S. cerevisiae pho2*Δ*pho4*Δ strain background, and examining the ability of these homologous genes to induce acid phosphatase activity and to regulate target genes at genome-wide levels [11]. Importantly, the authors concluded that the observed target gene expansion is not consistent with the phylogenetic relationship, but rather is a property specific to Pho4 homologs with decreased Pho2-dependence. Kerwin and Wykoff (2010) proposed that CgPho4 induces phosphatase activity independently of CgPho2, likely due to its increased size (from 312 residues in *S. cerevisiae* to 533 residues in *C. glabrata*) [9]. Indeed, recent studies by He's laboratory group have shown that the differences in Pho2-dependence were primarily from the differences located in Intrinsically Disordered Regions (IDRs), which were varied in both number and sequences of amino acids [18]. The IDRs of ScPho4 and CgPho4 differently modulate the function of the Activation Domain (AD), DNA binding Domain (DBD), and Pho2 Interaction Domain (P2ID). Thus, the evolution of PHO regulators is mainly driven by IDR divergence.

IDRs are widely distributed across all kingdoms of life, playing important roles in numerous biological processes [19]. One of the key characteristics of IDRs is a lack of stable secondary and tertiary structures and existence in a collection of interconverting states known as an ensemble. The absence of a defined folded state challenges the studies that are based on conventional sequence-structure-function relationships of proteins. Moreover, the sequences of IDRs are poorly conserved, hindering the prediction of IDR conformation properties and their evolutionary conservation and variants. In fact, the structural ensembles of most IDRs cannot be easily characterized by conventional experimental methods due to their highly dynamic nature and conformational heterogeneity. Thus, our understanding of the functional roles and evolution of IDRs is still limited compared to that of structured proteins.

Molecular dynamics (MD) simulation is a powerful computational-based method that can accurately predict how the system of proteins or biomolecules moves over time by capturing behaviors of these molecules in full atomic detail [20]. MD simulations can provide a framework for experimental designs, which have dramatically impacted the studies in, but not limited to, structural biology, molecular biology, and drug discovery. Furthermore, MD simulation can be routinely applied to the studies of IDRs and has become an important complementary tool to predict IDRs structures and a sequence-ensembles-function relationship [21–23]. For the transcriptional regulation of the PHO pathway in fungi, a variety of experiments have been performed to provide the functions of Pho4 and the dependency of Pho2 coactivator. Nonetheless, the obtained data are generally limited in their three-dimensional and temporal resolution. To gain insights into the atomic-level motions of Pho2 and Pho4 during the binding at the *PHO5* promoter, an MD simulation was performed to assess structural changes upon protein-DNA complex formation. The complex stability, compactness of biomolecules, atomic contact number, and free binding energy were computationally quantified. Also, we applied the MD simulation method to predict the structural ensembles and functions of P2ID, one of the key Pho4 IDRs. Overall, our results complement previous experimental studies and demonstrate the versatility of MD simulation techniques in advancing research in both evolutionary biology and structural bioinformatics.

2. Materials and Methods

2.1. Sequence Analysis

DNA sequences for the *PHO5* promoter of *S. cerevisiae* were reported previously [4]. Protein sequences of Pho2, Pho4, and Pho4 homologs were identified using the yeast genome database (https://www.yeastgenome.org; accessed on 20 June 2025). Structures of each protein were generated using the AlphaFold3 web-based analysis tool [24]. The pLDDT scores were obtained from AlphaFold3, which was used to estimate the confidence of the model prediction for the position of each residue in a protein and to predict IDR. Protein structures were visualized and edited in the BIOVIA Discovery Studio 2024. Functional domains were drawn using the IBS 2.0 tool [25]. Protein sequence alignment was performed using the Clustal Omega program (https://www.ebi.ac.uk/Tools/msa/clustalo; accessed on 20 June 2025) and T-Coffee (https://tcoffee.crg.eu; accessed on 20 June 2025).

2.2. Modeling Structure of Pho4 Complex

The complex structures of biomolecules were predicted from full-length protein sequences, using the AlphaFold3 server as available on (https://alphafoldserver.com; accessed on 20 June 2025) and converted to the protein databank (pdb) file using the BIOVIA Discovery Studio 2024. The 35 base pairs of double-stranded DNA of the *PHO5* UAS1 region were added to the complex. Residues with low and very low pLDDT scores located outside the P2ID and bHLH domains were removed using the protein editing tool, the BIOVIA Discovery Studio 2024.

2.3. Molecular Dynamics Simulation

MD simulations were conducted on an HPE Apollo 6500 (Hewlett Packard Enterprise, Spring, TX, USA) based on AMD EPYC 7742 with NVIDIA A100 SXM4 GPU (San Francisco, CA, USA). All simulations were performed using the GROMACS 2022.4 package with AMBER ff99SB force field [26–29]. The gmx pdb2gmx command was used to convert the pdb file of the complex structure to molecule topology, a position restraint, and a post-processed structure file. The protein-DNA complex was placed in the center of a cubic box with 1.0 nm from the box edge, using the gmx editconf command. Solvent (water) was filled into a defined box, using the three-point water model TIP3P with gmx editconf. Sodium (Na^+) and chloride (Cl^-) ions were added to the system to neutralize the net charge and achieve a physiological salt concentration of 0.15 M (Table S1). The ions were introduced using the gmx genion command in GROMACS, following preparation with gmx grompp. The protein complex underwent energy minimization and two phases of equilibration: (i) the constant number of particles, volume, and temperature (NVT) ensemble and (ii) the constant number of particles, pressure, and temperature (NPT) ensemble for 100 ps each, using the gmx mdrun -v -deffnm em, gmx mdrun -v -deffnm nvt, and gmx mdrun -v -deffnm npt commands, respectively. The production run was conducted for 400 ns under periodic boundary conditions (PBC) in all three spatial directions (pbc = xyz). Long-range electrostatic interactions were treated using the particle mesh Ewald (PME) method (coulombtype = PME) with a real-space cutoff of 1.0 nm. Lennard-Jones (van der Waals) interactions were also treated with a cutoff of 1.0 nm (rvdw = 1.0). The system temperature was maintained at 310.15 K using the V-rescale (modified Berendsen) thermostat, with separate coupling groups for Protein_DNA and Water_Ion, and a coupling time constant (tau_t) of 0.1 ps. The pressure was maintained at 1 bar using the Parrinello-Rahman barostat (pcoupl = Parrinello-Rahman), with a coupling constant (tau_p) of 2.0 ps and compressibility of 4.5×10^{-5} bar^{-1}.

To analyze the dynamics of protein complexes, trajectory files were first concatenated (gmx trjcat command), and then all molecules were centered in the box (gmx trjconv with -pbc mol-center command). The gmx make_ndx command was used to set the individual group within the protein-DNA complex. Structural stability of the protein complex was measured using the gmx rms command, which calculates the root-mean-square deviation. The compactness (radius of gyration) and number of atomic contacts were determined using the gmx gyrate and gmx_midist with -on commands, respectively. The outputs were analyzed and used to generate a graphical visualization in Excel. The 3D atomic coordinates of the ScPho2-ScPho4-DNA complex at 400 ns for all three independent MD simulation replicates are available in Supplemental Data S2–S4.

2.4. Binding Free Energy

The molecular mechanics with generalized Born and surface area (MM-GBSA) was calculated to measure the binding energy between the protein-DNA complex, using the gmx_MMPBSA package [30]. After installation of the gmx_MMPBSA package, an input file was generated (gmx_MMPBSA --create_input gb command), and the temperature was set to 310.15 K. An index for individual groups within the protein-DNA complex was created using the gmx make_ndx command. The gmx_MMPBSA package calculated binding free energy using the following equation:

$$\Delta G_{bind} = \Delta E_{MM} - T\Delta S + \Delta G_{sol}$$

$$\Delta E_{MM} = \Delta E_{vdw} + \Delta E_{elect}$$

$$\Delta G_{sol} = \Delta G_{polar} + \Delta G_{nonpolar}$$

The outputs were analyzed at the indicated time duration (see results) and visualized using the gmx_MMPBSA_ana module and Excel.

3. Results

3.1. Functional Domains and Size of the Pho4 Homologs

The artificial intelligence-based protein structure model AlphaFold3 predicted that approximately 75% of ScPho4 and CgPho4 proteins were IDRs [18]. We therefore examined the proportion of IDRs within the Pho4 homologs from yeast, dimorphic fungi, and filamentous fungi. As shown in Figure 2A and B, the sizes of Pho4 proteins from other fungal species are significantly larger than ScPho4. Using AlphaFold3, the 3D structure of the size-increased region was mostly predicted with a lower predicted structure score (pIDDT < 50), suggesting that these residues are IDRs (Figure 2B). Linear protein sequence alignment identified three highly conserved regions across Pho4 homologs. First, the bHLH DBD showed the highest conservation (Figure 2A). The second highly conserved region was mapped to the known Pho80/85 post-translational recognition motifs ([S/T]PX[LMIF]) (Figure 2C top), suggesting that these Pho4 homologs could be targeted for Pho80/Pho85 phosphorylation. The third region, the KSNYQ motif, is well conserved only in dimorphic fungi and molds, while this motif is completely absent in yeast species (Figure 2C bottom). Importantly, the sequences of P2ID identified in ScPho4 were poorly conserved in dimorphic and filamentous fungi, and even between ScPho4 vs. CgPho4 (data not shown [18]).

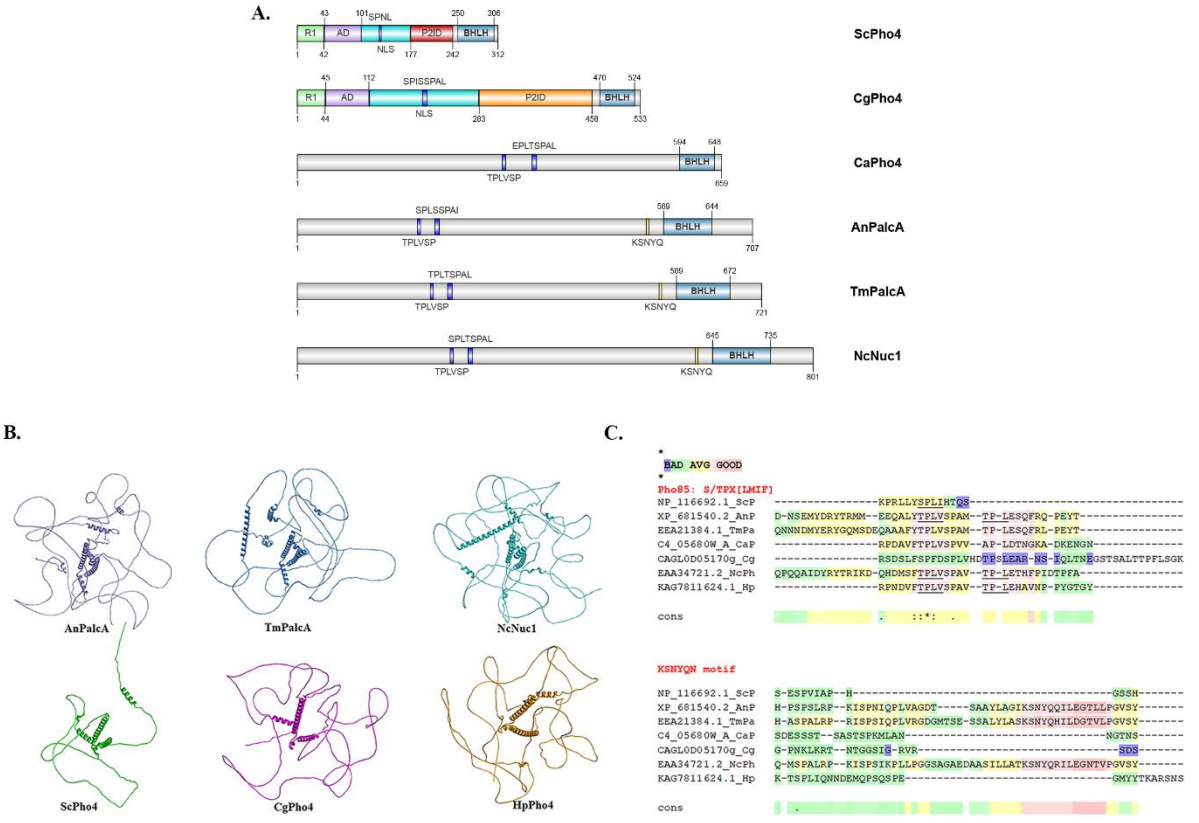

Figure 2. Structural basis of fungal Pho4 homologs. (**A**) Functional domains of Pho4 homologs in various fungal species. (**B**) 3D structures of Pho4 homologs predicted using the AlphaFold3 server. (**C**) Protein sequence alignment of fungal Pho4 homologs. The Pho85 phosphorylation motif (top) is conserved across fungal species, whereas the KSNYQ motif (bottom) is found only in filamentous and thermally dimorphic fungi. The T-Coffee multiple sequence alignment web server was used to identify conserved residues.

3.2. Molecular Dynamics of ScPho4 Complex at PHO5 Promoter

The requirement of ScPho2 in *PHO5* gene induction is experimentally well established, which provides a solid foundation for computational modeling. To set up biomolecules for initial analysis, protein sequences of ScPho2 and ScPho4 and DNA sequences of *PHO5* promoter (*ScPHO5*) were retrieved from NCBI and the yeast genome database and then subjected to structure prediction using the AlphaFold3 program [24]. The AlphaFold3 model predicted that ScPho4 interacted with DNA sequences 5′ AATTAGCACGTTTTCGCA 3′ while ScPho2

bound at the AT-rich region of *PHO5* promoter, consistent with binding sites determined by gel shift experiments (Figure 3; [4]). This information serves as a control and validates the accuracy of computational methods.

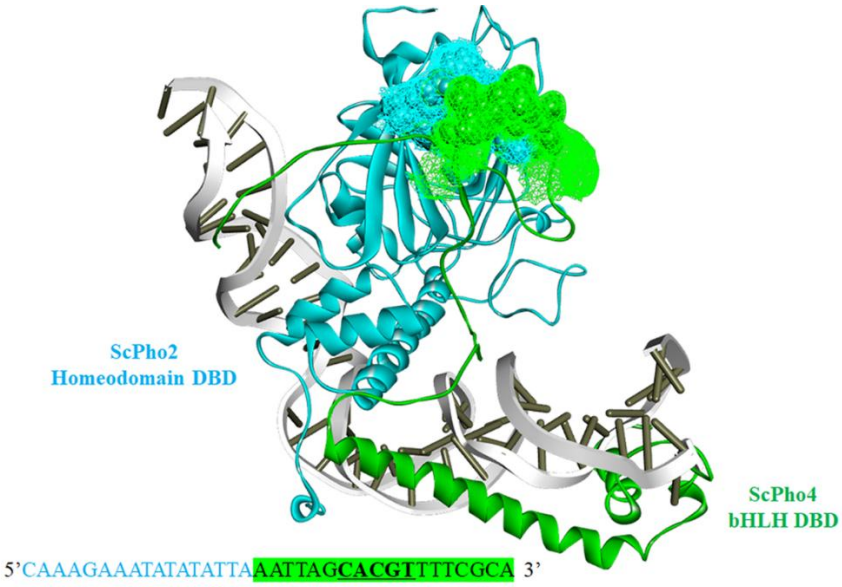

Figure 3. ScPho2–ScPho4 complex structure at the *PHO5* promoter. The protein-DNA interaction model was generated using AlphaFold3. Molecular dynamics (MD) simulation was performed with the Gromacs 2022.4 package. The structure shown was extracted at 400 ns of the MD simulation. The blurred green region represents the ScPho4 interaction domain (residues 199–247), while the blurred blue region indicates the ScPho2 interaction domain (residues 343–390), as previously identified by yeast two-hybrid experiments and mutational analyses. Key hydrogen bond interactions are highlighted in Figure 5. DBD: DNA-binding domain. The DNA sequence of the *ScPHO5* promoter is shown below the complex. Sequences bound by ScPho2 are highlighted in blue, and those bound by ScPho4 are highlighted in green. The consensus binding motif CACGT, recognized by Pho4, is shown in bold.

A recent study demonstrated that a P2ID in ScPho4 enhances the ScPho4 activity by mediating the interaction with ScPho2 partner while autoinhibiting the ScPho4 function when ScPho2 is not present [18]. To assess the structural dynamics of ScPho4 and its activity, the IDR P2ID and structured DBD (residues 180–312) of ScPho4 protein were subjected to MD simulation. Then, ScPho4-Sc*PHO5* complex in combination with and without ScPho2 was simulated for 400 ns. The root-mean-square-deviation (RMSD) and radius of gyration (Rg) were monitored to quantify structural stability and compactness of the biomolecular complex, respectively. As demonstrated in Figure 4A and B, all complexes reached equilibrium after 50 ns of simulation and converged in 400 ns of simulation. The RMSD profile of ScPho4 in the complex without ScPho2 exhibited higher RMSD values than the complex with ScPho2 throughout the simulations (Figure 4B).

At 400 ns, the ScPho4 P2ID formed an open structure that extends to interact with ScPho2 (Figure 5A left). Specifically, intermolecular hydrogen bonds were formed between ScPho2 Glu374–ScPho4 Thr210 and ScPho2 Gly375–ScPho4 Ser211 (Figure 5B). This intermolecular interaction between the ScPho2 interaction region (residues 343–390) and ScPho4 interaction region (residues 199–247) corresponds to the previous results from yeast-two-hybrid experiments and mutational analyses (Figures 3 and 5A; [3,31]). Conversely, the ScPho4 P2ID adopted a more closed structure when simulated without ScPho2. This closed structure was mediated by intramolecular bonds formed between the ScPho4 residues Thr210 and Ser211 (Figure 5C, region 1) and Gly216-Ala227, Val218-Val225, and Ala 220-Ser223 (Figure 5C, region 2). Thus, our MD simulation is consistent with the previous experimental data on dual effects of the P2ID, which drives ScPho2-dependent transactivation activity of ScPho4 [18].

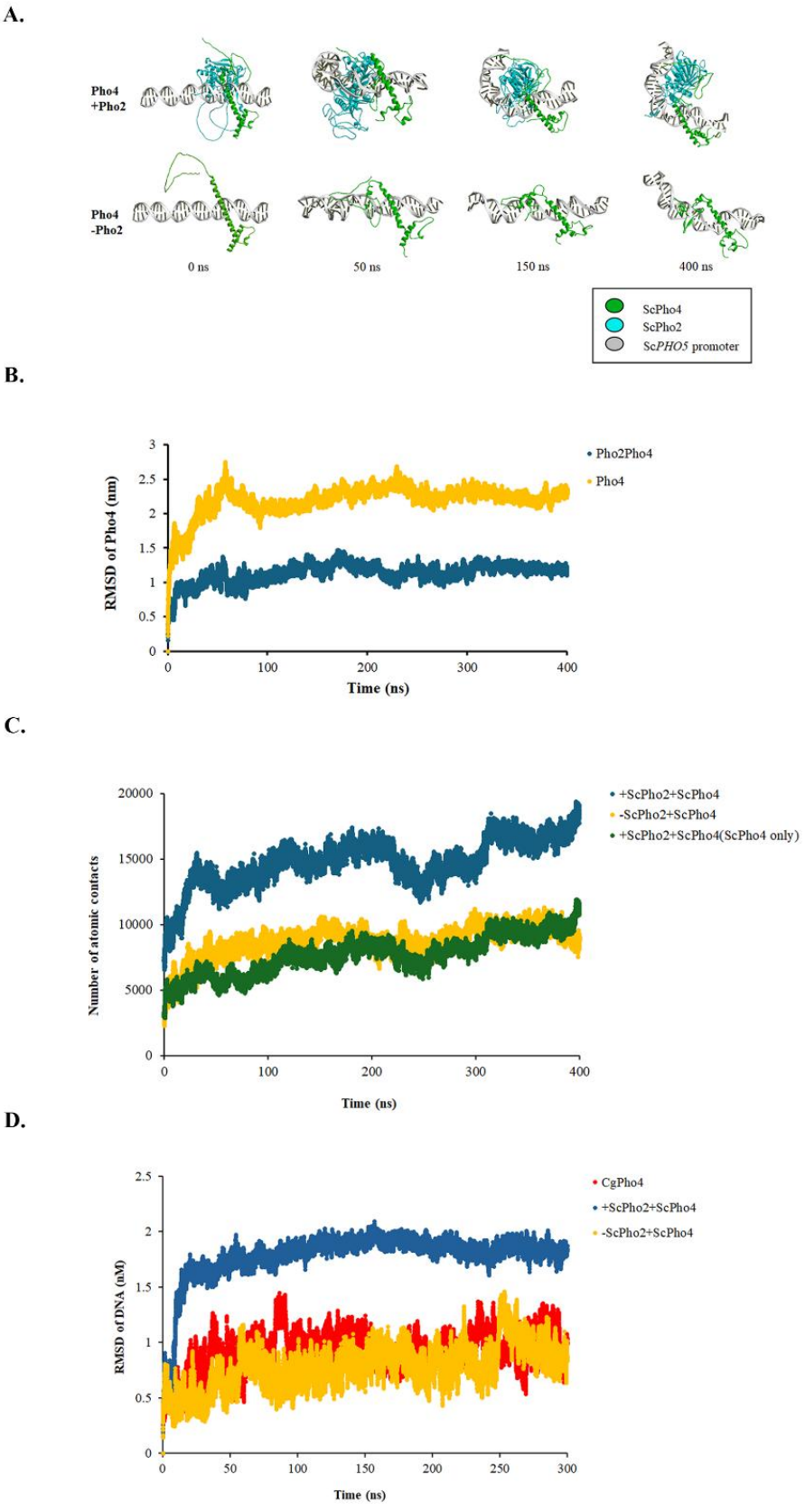

Figure 4. MD simulation of ScPho4–DNA complexes with or without ScPho2. Structural dynamics of ScPho2-ScPho4 complex or ScPho4 alone in association with *PHO5* DNA promoter were simulated for 400 ns using the Gromacs 2022.4 package. (**A**) Structural changes of the ScPho4-*PHO5* complex with and without ScPho2 are shown at 0, 50, 150, and 400 ns. (**B**) Root-mean-square deviation (RMSD) of ScPho4 was measured to assess the stability of each interaction complex. (**C**) The number of atom contacts between ScPho4 and DNA was calculated. (**D**) The radius of gyration (Rg) was determined to evaluate the compactness of the DNA (*PHO5* promoter) complexes.

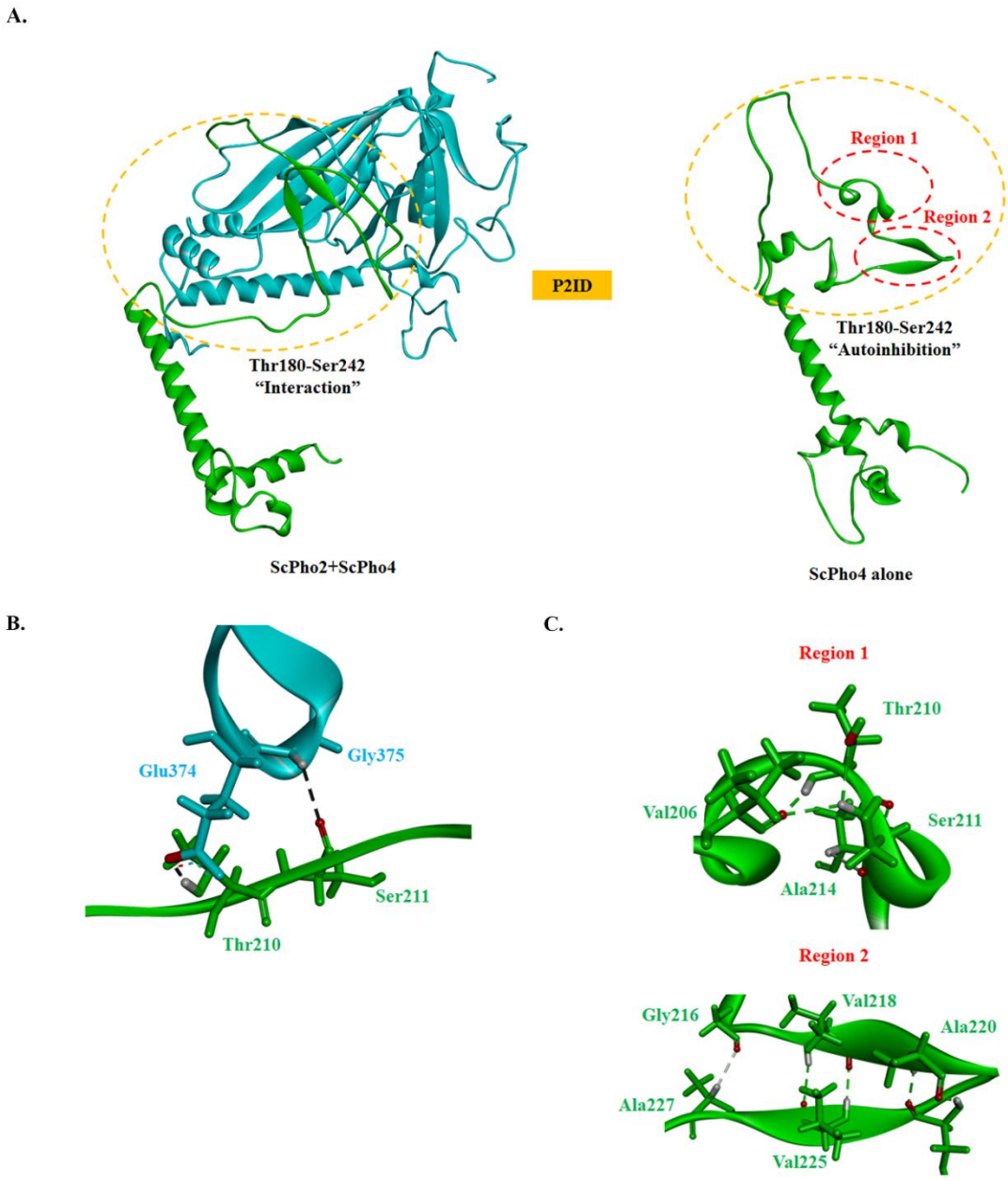

Figure 5. Structural dynamics of the ScPho4 P2ID with and without ScPho2. (**A**) Conformations of the ScPho4 P2ID with and without ScPho2 are shown. The ScPho4 P2ID residues Thr180-Ser242 are indicated with yellow circles, while regions forming intramolecular hydrogen bonds in the absence of ScPho2 are highlighted with red circles. (**B**) In the presence of ScPho2, ScPho4 residues Thr210 and Ser211 form intermolecular hydrogen bonds with ScPho2 residues Glu374 and Gly375, mediating the protein-protein interaction. (**C**) Key residues predicted to contribute to the autoinhibition of ScPho4 are illustrated. Intramolecular hydrogen bonds (highlighted in red circles in (**A**) are formed when ScPho4 is alone. The depicted structures were extracted at 400 ns of MD simulation.

Next, we monitored the number of atomic contacts between the protein-DNA complexes with and without ScPho2. The ScPho2-ScPho4 together showed the greatest number of atomic contacts with target DNA (Figure 4C). When considering the number of atomic contacts that ScPho4 contributes to the ScPho2-ScPho4-DNA complex, the number was comparable to the ScPho4-DNA complex (Figure 4C). This data suggested that interaction with ScPho2 provided a greater number of atomic contacts to the ScPho2-ScPho4-DNA complex.

When monitoring the structural changes during simulations, we noticed a drastic distortion of *PHO5* DNA in the presence of ScPho2-ScPho4 dimer (Figure 4A). Notably, structural changes in *PHO5* DNA are less prominent when ScPho4 is present alone (Figure 4C). In the absence of ScPho2, the RMSD profile of *PHO5* DNA showed lower values overall but with significant fluctuations throughout the simulation. Frequent fluctuations in the RMSD profile indicated that the DNA structure undergoes continual conformational changes, reflecting less

structural stability over the course of the simulation (Figure 4D). In contrast, when *PHO5* DNA is in complex with the ScPho2–ScPho4 dimer, the RMSD values were higher but fluctuated less, indicating that although the DNA adopts a new structural conformation, this state is more stable and consistent during the simulation (Figure 4D). Together, these results imply that ScPho4 requires ScPho2 as a co-activator not only to enhance DNA binding but also to promote and stabilize specific conformational changes in the target DNA sequence.

To determine the binding ability of ScPho4 to DNA, MM-GBSA calculations were carried out to quantify the relative binding free energy between the ScPho2-ScPho4 dimer vs. ScPho4 alone to *ScPHO5* DNA. As shown in Table 1, the ScPho4-ScPho2 dimer showed 1.6-fold greater binding energy than the ScPho4 alone to *ScPHO5* DNA, being −506.91 ± 116.88 vs. −326.54 ± 26.57, respectively. The ScPho4-ScPho2-DNA complex exhibited stronger van der Waals interactions and electrostatic interactions than the ScPho4-DNA complex, suggesting a stronger protein-DNA formation. Moreover, the ScPho4-DNA showed more favored polar solvation, suggesting that the solvent molecules favorably stabilized the electrostatic interactions, likely due to the greater solvent-exposed surface of ScPho4 alone without interaction with ScPho2. For non-polar solvation, the ScPho4-DNA complex exhibited less negative values than the ScPho2-ScPho4-DNA complex, suggesting a decrease in favorable non-polar interactions with solvent molecules. Overall, the binding free energy of ScPho2-ScPho4-DNA is more negative than that of ScPho4-DNA, indicating a more stable protein-DNA complex. Our data from MD simulation were consistent with experimental evidence that the function of ScPho4 is dependent on co-activator ScPho2 [9,11,18].

Table 1. Relative binding free energies of ScPho4–DNA complexes with and without ScPho2 determined by MM-GBSA[#].

Component	Free Energy (kcal/mol)	
	+ScPho2 + ScPho4	**-ScPho2 + ScPho4**
Rep1		
Gas phase		
Van der Waals (VDWAALS)	−373.76 ± 10.66	−177.23 ± 7.98
Electrostatic (EEL)	−19663.70 ± 180.42	−12556.44 ± 114.85
Solvent phase		
Polar solvation (EGB)	19485.20 ± 178.54	12434.83 ± 109.92
Non-polar solvation (ESURF)	−51.20 ± 0.72	−25.10 ± 0.82
Total	−603.46 ± 16.46	−323.94 ± 13.47
Rep2		
Gas phase		
Van der Waals (VDWAALS)	−204.90 ± 204.83	−181.44 ± 8.22
Electrostatic (EEL)	−16117.85 ± 155.82	−12452.01 ±149.01
Solvent phase		
Polar solvation (EGB)	15982.53 ± 153.33	12358.63 ± 147.01
Non-polar solvation (ESURF)	−36.76 ± 0.69	−26.55 ± 0.55
Total	−376.97 ± 211.09	−301.37 ± 14.59
Rep3		
Gas phase		
Van der Waals (VDWAALS)	−289.64 ± 7.84	−212.12 ± 9.65
Electrostatic (EEL)	−16357.00 ± 260.65	−12674.02 ±169.76
Solvent phase		
Polar solvation (EGB)	16144.87 ± 256.19	12562.29 ± 155.87
Non-polar solvation (ESURF)	−38.53 ± 0.66	−30.47 ± 0.81
Total	−540.31 ± 14.43	−354.32 ± 17.11
Average	−506.91 ± 116.88	−326.54 ± 26.57

[#] Binding free energies were calculated at a temperature of 310.15 K, using complex frames from start frame 39900 (399 ns) to end frame 40000 (400 ns). Three replicates were performed for each system. Values for each replicate and the average of three independent simulations are shown.

3.3. Molecular Dynamics of CgPho4 at ScPHO5 Promoter

The P2ID in CgPho4 has been previously identified by protein sequence alignment search with the P2ID of ScPho4 [18]. The CgPho4 P2ID sequences are presented at a longer length and showed less conservation with P2ID ScPho2 (Figure 2A; [18]). In a current study by He et al. (2025), the Pho4 chimeras were constructed by swapping the well-aligned region between ScPho4 and CgPho4, and the transcriptional activity of these fusion proteins was quantified by a dual fluorescence reporter assay at *ScPHO5* [18]. The authors discovered that the P2ID of CgPho4 did not have a dual role in co-activator interaction and autoinhibition as observed in the P2ID of ScPho4. In other words, the P2ID of CgPho4 allows the protein to activate transcription independently of Pho2.

To explore structural dynamics and flexibility of the CgPho4 P2ID during *ScPHO5* binding, the CgPho4 containing the P2ID and DBD portion was used in MD simulation. When monitoring 3D structures of P2ID over a 300 ns simulation time, we found that this region could fold into different ensembles (Figure 6A). The RMSD profile showed that CgPho4 underwent a drastic structural change and reached equilibrium at 25 ns and then remained stable thereafter (Figure 6A,B). Notably, CgPho4 could distort *ScPHO5* DNA to a higher degree than ScPho4 alone, especially in the first 150 ns (Figures 4A and 6A). The RMSD profile of CgPho4-DNA complex revealed that *ScPHO5* DNA adopted the most compacted structure at 90 ns (Figure 4D). Overall, MD simulation could be applied to predict conformational ensembles of disordered regions of Pho4 proteins.

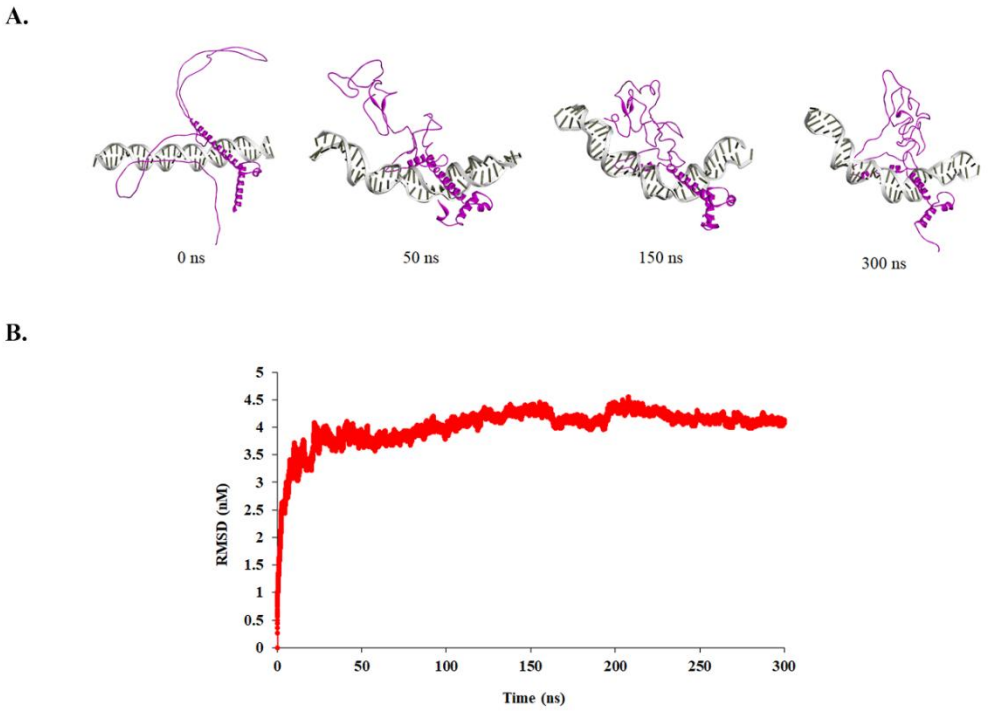

Figure 6. MD simulation of the CgPho4-DNA complex. Computational analysis of CgPho4 in association with Sc*PHO5* DNA promoter was performed using the Gromacs 2022.4 package for 300 ns. Only the P2ID and bHLH DBD domains of the CgPho4 were included in the simulation. (**A**) Structural changes of the CgPho4-*PHO5* complex are shown at 0, 50, 150, and 300 ns. (**B**) RMSD of CgPho4 was measured to assess the stability of the interaction complex.

4. Discussion

Though the "structure determines function" has long been a biology mantra, it has been known to molecular biologists that "structure moves to function". Biomolecule dynamics are involved with motions ranging from atomic vibrations to large-scale conformational changes. Here, we explored the structural dynamics of Pho4 during interaction with Pho2 at the *PHO5* promoter, using a computational-based method. First, we found that the binding of ScPho4 is more energetically favorable when forming a complex with ScPho2. Structural analysis demonstrated that the IDR within P2ID of ScPho4 is highly flexible and could form a structure that mediates interaction with ScPho2. IDRs are commonly found in proteins as part of linkers that separate functional domains [32,33] or spacers [34] that inhibit undesirable protein-protein interactions. For transcription factors, IDRs play crucial roles in recruiting cofactors, forming interactions with other transcription factors, mediating binding specificity, and forming molecular condensates [35–40]. Based on the experimental study in ScPho4 and CgPho4, IDRs aid in the process of transactivation, autoinhibition, and protein-protein interaction [18]. Our data from the MD simulation approach is consistent with this study [18]; we could capture conformational dynamics of the IDR-containing P2ID that formed an extended structure to interact with ScPho2, while forming a closed structure in the absence of ScPho2. Thus, our study highlighted the significance of using computational methods to complement experimental methods.

The exploration of IDRs has further underscored the importance of protein dynamics that contribute to their biological and biochemical functions. However, it is difficult to predict how the IDR sequences translate into a structural ensemble and functions [22]. In this study, AlphaFold3 predicted that Pho4 homologs contain a large

portion of the low and very low confidence regions (i.e., pLDDT values less than 70), which are predictive of protein disorder regions. As such, the majority of Pho4 3D structures are characterized as IDRs, comprising more than 75% of protein sequences. Despite the lack of a definite folded structure, the presence of IDR sequences in Pho4 *per se* could potentially predict its dependence on Pho2 binding. For example, in filamentous and dimorphic fungi, Pho4 homologs do not contain the P2ID sequences, and therefore the transactivation activity of these Pho4 homologs can be predicted to be less dependent on the Pho2 proteins. In contrast, the interaction region of Pho2 homologs is highly conserved across fungal species and predicted to form folded structures [41,42]. It is likely that the Pho2 and Pho4 homologs have evolved differently, with Pho2 having remained a requirement for co-transcription factor(s), while Pho4 has gained the ability to regulate target genes independently.

Moreover, DNA distortion was detected from our ScPho4 complex simulation. Importantly, nucleosome configurations at the *PHO5* promoter are highly dynamic, ranging from a maximally repressive, fully nucleosomal state to a maximally non-repressive, nucleosome-free state [43–45]. Nucleosomes at the *PHO5* promoter are continuously assembled, disassembled, and slide in an intrinsically stochastic manner [44]. Mechanical insights into the relationship between transcription factor-mediated nucleosome dynamics and *PHO5* gene regulation have been gained through both experimental and computational approaches [4,43,44,46]. ScPho4 is required for two structurally inseparable events: (i) promoting nucleosome disassembly and (ii) inducing transcriptional activation [44]. From computational models, ScPho4 has been shown to mediate "local remodeling activity" on *PHO5* promoter nucleosomes [46] and regulate transcriptional burst frequency at *PHO5* promoter by accelerating nucleosome removal [43]. In our MD simulation approach, the observed DNA distortion may represent a ScPho4-induced dynamic structural change in the *PHO5* promoter to facilitate transcription. We noted that we used only 35 bp of the UAS1 of the *PHO5* promoter [3,4]. MD simulation of Pho4 with a larger portion of the *PHO5* promoter region and other relevant chromatin cofactors [44] could potentially provide a more complete picture of how Pho4 binding influences chromatin organization and nucleosome structure.

The ability to predict structures from sequences could potentially provide valuable insights into their functional properties and support structure-based development of drugs and therapeutic agents. In particular, characterizing the conformational ensembles of IDRs (e.g., Pho4 homologs) generally requires the integration of experimental and simulation methods. For computational approaches, there has been substantial development and improvement in simulation and prediction models for studying IDRs. First, recent force field strategies have been developed to improve the accuracy of MD simulations of IDRs, including adjustments to dihedral parameters, the addition of grid-based energy correction map (CMAP) parameters, and refinements to protein–water interactions [23]. Second, force field benchmarking is recommended for studying IDRs to ensure accuracy and reliability in simulation results [47]. The study by Chan-Yao-Chong et al. 2023 showed that distinct conformational ensembles of IDRs could be generated depending on the force field and its associated water model used in MD simulations [41]. For characterizing specific IDRs, the authors suggested a comprehensive evaluation of the chosen force fields and water models by assessing both local and global conformations and combining these assessments into a final score [47]. The force field and solvent model with good performances against typical IDR-containing protein systems have been excellently reviewed somewhere else [23]. For instance, MD simulations were used to model the structural ensembles of the transactivation domains in p53 tumor suppressor, a transcription factor that suppresses tumor growth and is frequently mutated in human cancer. In fact, the transactivation domain 1 (aa 1–61) and 2 (aa 40–61) in p53 were best characterized by ff99SB-disp [48] and ff99SB-IDLN(TIP3P) [49], respectively.

Third, the development of deep learning models for large-scale prediction of IDR ensembles has recently increased. For example, ALBATROSS is a deep learning model that can directly predict the global dimensions of IDRs from their amino acid sequences. [22]. Another model, AlphaFold-Metainference, generates structural ensembles of IDRs by using AlphaFold's predicted distograms –inter-residue distance maps– as structural restraints in molecular dynamics (MD) simulations within the metainference framework. [21,50]. Notably, the computational cost of MD simulations can hinder the ability to benchmark appropriate force fields, especially for long-timescale simulations or large biomolecular complexes. The coarse-grained models, which represent groups of atoms as an interaction center or "bead", can reduce the computational cost of simulations, enabling the study of longer simulation timescales or more complex biological systems [51]. However, specific molecular motions or interactions may be missed due to the reduced resolution in the representation of protein complexes. Recently, cg2all, a novel deep learning model, has been developed to accurately reconstruct all-atom protein structures from coarse-grained representation models [52,53].

For future studies, the MD simulation strategies described above could be implemented to further characterize the IDRs of Pho4 homologs or predict the structural dynamics of the Pho4-*PHO5* complex within a larger system, such as the Pho4-nucleosome complex. By integrating simulations with experimental validation, such as site-

directed mutagenesis or cryo-EM, these approaches could provide insights into the conformational transitions that drive Pho4's regulatory function. This knowledge will deepen our Pho4 understanding from 'sequences-structure-function to 'structural motions-functions'. Furthermore, a comparative analysis of Pho4 homologs across different fungal species could reveal evolutionary trends in regulatory mechanisms, potentially informing new strategies for manipulating gene expression in synthetic biology applications and developing antifungal therapeutic agents.

5. Conclusions

Our study provides new insights into the dynamic behavior of Pho4 and its interaction with Pho2 at the *PHO5* promoter, underscoring the essential roles of intrinsically disordered regions in mediating complex. Through molecular dynamics simulations, we demonstrated that the ScPho2-ScPho4 complex forms a more stable and distinct structural assembly at the promoter compared to ScPho4 alone, resulting in specific conformational changes in both protein and DNA. Importantly, our work illustrates how MD simulations serve as a powerful tool not only for visualizing and quantifying protein–DNA interactions at atomic detail, but also for exploring the structural basis underlying the evolution of gene regulatory networks. Routine integration of such computational approaches with experimental studies will significantly advance our understanding of how both 3D structure and IDR properties contribute to transcriptional regulation across diverse fungal species.

Supplementary Materials

The additional data and information can be downloaded at: https://media.sciltp.com/articles/others/2509220950069856/Supplemental-Materials.zip. Table S1: Total number of counterions added to the system.

Author Contributions

T.W. and M.P. conceived and supervised the study; T.W., N.K., and M.P. performed the MD simulations; T.W., N.K., and M.P. analyzed data and generated the figures and tables; T.W., N.K., and M.P. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding

This research was financially supported by the Faculty of Medicine Research Fund, grant no. 138-2568 to T.W. and no. 046-2566 to M.P.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

The protein ID and sequences (Supplemental Data S1) as well as the 3D coordinate files (Supplemental Data S2–S4), used to generate figures in this manuscript, are available in the supporting materials.

Acknowledgments

We would like to thank ERAWAN HPC, Information Technology Service Center (ITSC) of Chiang Mai University, for their support in the high-performance computing resources.

Conflicts of Interest

The authors declare no competing interests.

References

1. Bhalla, K.; Qu, X.; Kretschmer, M.; et al. The phosphate language of fungi. *Trends Microbiol.* **2022**, *30*, 338–349. https://doi.org/10.1016/j.tim.2021.08.002.
2. Lev, S.; Djordjevic, J.T. Why is a functional PHO pathway required by fungal pathogens to disseminate within a phosphate-rich host: A paradox explained by alkaline pH-simulated nutrient deprivation and expanded PHO pathway function. *PLoS Pathog.* **2018**, *14*, e1007021. https://doi.org/10.1371/journal.ppat.1007021.

3. Barbaric, S.; Münsterkötter, M.; Goding, C.; et al. Cooperative Pho2-Pho4 interactions at the *PHO5* promoter are critical for binding of Pho4 to UASp1 and for efficient transactivation by Pho4 at UASp2. *Mol. Cell. Biol.* **1998**, *18*, 2629–2639. https://doi.org/10.1128/mcb.18.5.2629.
4. Barbarić, S.; Münsterkötter, M.; Svaren, J.; et al. The homeodomain protein Pho2 and the basic-helix-loop-helix protein Pho4 bind DNA cooperatively at the yeast *PHO5* promoter. *Nucleic Acids Res.* **1996**, *24*, 4479–4486. https://doi.org/10.1093/nar/24.22.4479.
5. Lee, Y.S.; Mulugu, S.; York, J.D.; et al. Regulation of a cyclin-CDK-CDK inhibitor complex by inositol pyrophosphates. *Science* **2007**, *316*, 109–112. https://doi.org/10.1126/science.1139080.
6. Desmarini, D.; Lev, S.; Furkert, D.; et al. IP(7)-SPX Domain interaction controls fungal virulence by stabilizing phosphate signaling machinery. *mBio* **2020**, *11*, e01920-20. https://doi.org/10.1128/mBio.01920-20.
7. Gomes-Vieira, A.L.; Wideman, J.G.; Paes-Vieira, L.; et al. Evolutionary conservation of a core fungal phosphate homeostasis pathway coupled to development in *Blastocladiella emersonii*. *Fungal Genet. Biol.* **2018**, *115*, 20–32. https://doi.org/10.1016/j.fgb.2018.04.004.
8. Carter-O'Connell, I.; Peel, M.T.; Wykoff, D.D.; et al. Genome-wide characterization of the phosphate starvation response in *Schizosaccharomyces pombe*. *BMC Genom.* **2012**, *13*, 697. https://doi.org/10.1186/1471-2164-13-697.
9. Kerwin, C.L.; Wykoff, D.D. *Candida glabrata PHO4* is necessary and sufficient for Pho2-independent transcription of phosphate starvation genes. *Genetics* **2009**, *182*, 471–479. https://doi.org/10.1534/genetics.109.101063.
10. Ikeh, M.A.; Kastora, S.L.; Day, A.M.; et al. Pho4 mediates phosphate acquisition in *Candida albicans* and is vital for stress resistance and metal homeostasis. *Mol. Biol. Cell* **2016**, *27*, 2784–2801. https://doi.org/10.1091/mbc.E16-05-0266.
11. He, B.Z.; Zhou, X.; O'Shea, E.K. Evolution of reduced co-activator dependence led to target expansion of a starvation response pathway. *Elife* **2017**, *6*, e25157. https://doi.org/10.7554/eLife.25157.
12. Ikeh, M.; Ahmed, Y.; Quinn, J. Phosphate acquisition and virulence in human fungal pathogens. *Microorganisms* **2017**, *5*, 48. https://doi.org/10.3390/microorganisms5030048.
13. Toh-e, A.; Ohkusu, M.; Li, H.M.; et al. Identification of genes involved in the phosphate metabolism in *Cryptococcus neoformans*. *Fungal Genet. Biol.* **2015**, *80*, 19–30. https://doi.org/10.1016/j.fgb.2015.04.019.
14. Kang, S.; Metzenberg, R.L. Molecular analysis of nuc-1+, a gene controlling phosphorus acquisition in *Neurospora crassa*. *Mol. Cell. Biol.* **1990**, *10*, 5839–5848. https://doi.org/10.1128/mcb.10.11.5839-5848.1990.
15. Gras, D.E.; Persinoti, G.F.; Peres, N.T.; et al. Transcriptional profiling of *Neurospora crassa Δmak-2* reveals that mitogen-activated protein kinase MAK-2 participates in the phosphate signaling pathway. *Fungal Genet. Biol.* **2013**, *60*, 140–149. https://doi.org/10.1016/j.fgb.2013.05.007.
16. Wu, D.; Dou, X.; Hashmi, S.B.; et al. The Pho80-like cyclin *of Aspergillus nidulans* regulates development independently of its role in phosphate acquisition. *J. Biol. Chem.* **2004**, *279*, 37693–37703. https://doi.org/10.1074/jbc.M403853200.
17. Zhou, Y.; Yuikawa, N.; Nakatsuka, H.; et al. Core regulatory components of the PHO pathway are conserved in the methylotrophic yeast *Hansenula polymorpha*. *Curr. Genet.* **2016**, *62*, 595–605. https://doi.org/10.1007/s00294-016-0565-7.
18. Snyder, L.F.; O'Brien, E.M.; Zhao, J.; et al. Divergence in a eukaryotic transcription factor's co-TF dependence involves multiple intrinsically disordered regions affecting activation and autoinhibition. *Nat. Commun.* **2025**, *16*, 5340. https://doi.org/10.1038/s41467-025-59244-w.
19. Tesei, G.; Trolle, A.I.; Jonsson, N.; et al. Conformational ensembles of the human intrinsically disordered proteome. *Nature* **2024**, *626*, 897–904. https://doi.org/10.1038/s41586-023-07004-5.
20. Hollingsworth, S.A.; Dror, R.O. Molecular dynamics simulation for All. *Neuron* **2018**, *99*, 1129–1143. https://doi.org/10.1016/j.neuron.2018.08.011.
21. Brotzakis, Z.F.; Zhang, S.; Murtada, M.H.; et al. AlphaFold prediction of structural ensembles of disordered proteins. *Nat. Commun.* **2025**, *16*, 1632. https://doi.org/10.1038/s41467-025-56572-9.
22. Lotthammer, J.M.; Ginell, G.M.; Griffith, D.; et al. Direct prediction of intrinsically disordered protein conformational properties from sequence. *Nat. Methods* **2024**, *21*, 465–476. https://doi.org/10.1038/s41592-023-02159-5.
23. Mu, J.; Liu, H.; Zhang, J.; et al. Recent Force field strategies for intrinsically disordered proteins. *J. Chem. Inf. Model* **2021**, *61*, 1037–1047. https://doi.org/10.1021/acs.jcim.0c01175.
24. Abramson, J.; Adler, J.; Dunger, J.; et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **2024**, *630*, 493–500. https://doi.org/10.1038/s41586-024-07487-w.
25. Xie, Y.; Li, H.; Luo, X.; et al. IBS 2.0: An upgraded illustrator for the visualization of biological sequences. *Nucleic Acids Res.* **2022**, *50*, W420–W426. https://doi.org/10.1093/nar/gkac373.
26. Abraham, M.J.; Murtola, T.; Schulz, R.; et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1*, 19–25.
27. Case, D.A.; Aktulga, H.M.; Belfon, K.; et al. AmberTools. *J. Chem. Inf. Model* **2023**, *63*, 6183–6191. https://doi.org/10.1021/acs.jcim.3c01153.

28. Maier, J.A.; Martinez, C.; Kasavajhala, K.; et al. ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713. https://doi.org/10.1021/acs.jctc.5b00255.
29. Hornak, V.; Abel, R.; Okur, A.; et al. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **2006**, *65*, 712–725. https://doi.org/10.1002/prot.21123.
30. Valdés-Tresanco, M.S.; Valdés-Tresanco, M.E.; Valiente, P.A.; et al. gmx_MMPBSA: A New tool to perform end-state free energy calculations with GROMACS. *J. Chem. Theory Comput.* **2021**, *17*, 6281–6291. https://doi.org/10.1021/acs.jctc.1c00645.
31. Bhoite, L.T.; Allen, J.M.; Garcia, E.; et al. Mutations in the pho2 (bas2) transcription factor that differentially affect activation with its partner proteins bas1, pho4, and swi5. *J. Biol. Chem.* **2002**, *277*, 37612–37618. https://doi.org/10.1074/jbc.M206125200.
32. Santner, A.A.; Croy, C.H.; Vasanwala, F.H.; et al. Sweeping away protein aggregation with entropic bristles: Intrinsically disordered protein fusions enhance soluble expression. *Biochemistry* **2012**, *51*, 7250–7262. https://doi.org/10.1021/bi300653m.
33. Li, M.; Cao, H.; Lai, L.; et al. Disordered linkers in multidomain allosteric proteins: Entropic effect to favor the open state or enhanced local concentration to favor the closed state? *Protein Sci.* **2018**, *27*, 1600–1610. https://doi.org/10.1002/pro.3475.
34. Jamecna, D.; Polidori, J.; Mesmin, B.; et al. An intrinsically disordered region in OSBP acts as an entropic barrier to control protein dynamics and orientation at membrane contact sites. *Dev. Cell* **2019**, *49*, 220–234. https://doi.org/10.1016/j.devcel.2019.02.021.
35. Brodsky, S.; Jana, T.; Mittelman, K.; et al. Intrinsically disordered regions direct transcription factor in vivo binding specificity. *Mol. Cell* **2020**, *79*, 459–471. https://doi.org/10.1016/j.molcel.2020.05.032.
36. Ferrie, J.J.; Karr, J.P.; Tjian, R.; et al. "Structure"-function relationships in eukaryotic transcription factors: The role of intrinsically disordered regions in gene regulation. *Mol. Cell* **2022**, *82*, 3970–3984. https://doi.org/10.1016/j.molcel.2022.09.021.
37. Ferrie, J.J.; Karr, J.P.; Graham, T.G.W.; et al. p300 is an obligate integrator of combinatorial transcription factor inputs. *Mol. Cell* **2024**, *84*, 234–243. https://doi.org/10.1016/j.molcel.2023.12.004.
38. Ji, D.; Shao, C.; Yu, J.; et al. FOXA1 forms biomolecular condensates that unpack condensed chromatin to function as a pioneer factor. *Mol. Cell* **2024**, *84*, 244–260. https://doi.org/10.1016/j.molcel.2023.11.020.
39. Krois, A.S.; Dyson, H.J.; Wright, P.E. Long-range regulation of p53 DNA binding by its intrinsically disordered N-terminal transactivation domain. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, e11302–e11310. https://doi.org/10.1073/pnas.1814051115.
40. Mindel, V.; Brodsky, S.; Cohen, A.; et al. Intrinsically disordered regions of the Msn2 transcription factor encode multiple functions using interwoven sequence grammars. *Nucleic Acids Res.* **2024**, *52*, 2260–2272. https://doi.org/10.1093/nar/gkad1191.
41. Wangsanut, T.; Tobin, J.M.; Rolfes, R.J. Functional mapping of transcription factor Grf10 that regulates adenine-responsive and filamentation genes in *Candida albicans*. *mSphere* **2018**, *3*. https://doi.org/10.1128/mSphere.00467-18.
42. Pongpom, M.; Khamto, N.; Sukantamala, P.; et al. Identification of homeobox transcription factors in a dimorphic fungus *Talaromyces marneffei* and protein-protein interaction prediction of RfeB. *J. Fungi* **2024**, *10*, 687. https://doi.org/10.3390/jof10100687.
43. Brown, C.R.; Mao, C.; Falkovskaia, E.; et al. Linking stochastic fluctuations in chromatin structure and gene expression. *PLoS Biol.* **2013**, *11*, e1001621. https://doi.org/10.1371/journal.pbio.1001621.
44. Korber, P.; Barbaric, S. The yeast *PHO5* promoter: From single locus to systems biology of a paradigm for gene regulation through chromatin. *Nucleic Acids Res.* **2014**, *42*, 10888–10902. https://doi.org/10.1093/nar/gku784.
45. Wolff, M.R.; Schmid, A.; Korber, P.; et al. Effective dynamics of nucleosome configurations at the yeast *PHO5* promoter. *Elife* **2021**, *10*, e58394. https://doi.org/10.7554/eLife.58394.
46. Kharerin, H.; Bhat, P.J.; Marko, J.F.; et al. Role of transcription factor-mediated nucleosome disassembly in *PHO5* gene expression. *Sci. Rep.* **2016**, *6*, 20319. https://doi.org/10.1038/srep20319.
47. Chan-Yao-Chong, M.; Chan, J.; Kono, H. Benchmarking of force fields to characterize the intrinsically disordered R2-FUS-LC region. *Sci. Rep.* **2023**, *13*, 14226. https://doi.org/10.1038/s41598-023-40801-6.
48. Liu, X.; Chen, J. Residual structures and transient long-range interactions of p53 transactivation domain: Assessment of explicit solvent protein force fields. *J. Chem. Theory Comput.* **2019**, *15*, 4708–4720. https://doi.org/10.1021/acs.jctc.9b00397.
49. Ouyang, Y.; Zhao, L.; Zhang, Z. Characterization of the structural ensembles of p53 TAD2 by molecular dynamics simulations with different force fields. *Phys. Chem. Chem. Phys.* **2018**, *20*, 8676–8684. https://doi.org/10.1039/c8cp00067k.

50. Bonomi, M.; Camilloni, C.; Cavalli, A.; et al. Metainference: A Bayesian inference method for heterogeneous systems. *Sci. Adv.* **2016**, *2*, e1501177. https://doi.org/10.1126/sciadv.1501177.
51. Kmiecik, S.; Gront, D.; Kolinski, M.; et al. Coarse-grained protein models and their applications. *Chem. Rev.* **2016**, *116*, 7898–7936.
52. Pang, Y.T.; Yang, L.; Gumbart, J.C. From simple to complex: Reconstructing all-atom structures from coarse-grained models using cg2all. *Structure* **2024**, *32*, 5–7. https://doi.org/10.1016/j.str.2023.12.004.
53. Heo, L.; Feig, M. One bead per residue can describe all-atom protein structures. *Structure* **2024**, *32*, 97–111. https://doi.org/10.1016/j.str.2023.10.013.