UNIVERSITY OF CALIFORNIA SAN DIEGO

Theoretical Model of HP1-STAT Interactions

A Thesis submitted in partial satisfaction of the requirements
for the degree Master of Science

in

Biology

by

Kangxin Xu

Committee in charge:

Professor Willis (Xiaowei) Li, Chair
Professor Cory Root, Co-Chair
Professor Barry J Grant

2022

The Thesis of Kangxin Xu is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2022

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

The FP7 WeNMR (project# 261572), H2020 West-Life (project# 675858), the EOSC-hub (project# 777536) and the EGI-ACE (project# 101017567) European e-Infrastructure projects are acknowledged for the use of their web portals, which make use of the EGI infrastructure with the dedicated support of CESNET-MCC, INFN-PADOVA-STACK, INFN-LNL-2, NCG-INGRID-PT, TW-NCHC, CESGA, IFCA-LCG2, UA-BITP, SURFsara and NIKHEF, and the additional support of the national GRID Initiatives of Belgium, France, Italy, Germany, the Netherlands, Poland, Portugal, Spain, UK, Taiwan and the US Open Science Grid.

I want to express thanks to the help from my advisor Professor Willis (Xiaowei) Li and committee members Professor Cory Root as well as Professor Barry J Grant, without whose help I could not complete this thesis. I want to express thanks to my family, especially to my mother, who constantly supported me when I was composing this thesis.

VITA

2020 Bachelor of Science in Biology, University of California San Diego

2022 Master of Science in Biology, University of California San Diego

ABSTRACT OF THE THESIS

Theoretical Model of HP1-STAT Interaction

by

Kangxin Xu

Master of Science in Biology

University of California San Diego, 2022

Professor Willis (Xiaowei) Li, Chair
Professor Cory Root, Co-Chair

HP1(Heterochromatin Protein 1) is a major component of heterochromatin, which is a highly condensed form of DNA playing an important role in multiple cellular activities including gene silencing. New research proposes that HP1 proteins could compartmentalize DNA into compacted chromatin by phase separation, which could be promoted by diverse HP1-binding proteins. STAT is a promising candidate. Besides its canonical role in JAK-STAT signaling pathway, previous research in our lab indicated that STAT contains HP1-

binding PXVXI motif. Unphosphorylated STAT could bind to HP1 and maintain the stability of heterochromatin while phosphorylated STAT disperses from HP1, resulting in heterochromatin disruption. Thus, I hypothesized that phosphorylation induced conformational change on STAT, switching it from an HP1-binding state to a DNA-binding state. In this paper, I constructed computational models among HP1α, STAT3 and DNA to examine the influence of phosphorylation on STAT3's binding affinity to both HP1α and DNA. During the preparation stage, I modified and constructed biomolecular structures for protein docking by Pymol and SWISS-MODEL. I imported the prepared biomolecular structures into HADDOCK and the web server provided potential binding complexes as output. I used the PRODIGY program to measure the Gibbs free energy and equilibrium constant of binding among unphosphorylated STAT3, HP1α and DNA as well as phosphorylated STAT3, HP1α and DNA. Compared to phosphorylated STAT3 homodimers, unphosphorylated STATA3 homodimers have higher binding affinity to HP1α and lower binding affinity to DNA. Although computational model has limitations and needs confirmation by further experiments in vitro, my results support the conclusion that phosphorylation drives STAT from HP1-binding to DNA-binding.

.

INTRODUCTION

1. Heterochromatin

Heterochromatin is a highly condensed structure of DNA whose compaction levels could vary between facultative heterochromatin and constitutive heterochromatin[1]. Facultative heterochromatin is normally regarded as transcriptionally silent non-repetitive chromatin which still conserves the potential to convert back to euchromatin, usually in response to developmental cues[2]. Constitutive heterochromatin normally consists of repetitive elements with a higher condensed level, usually serving structural functions[3]. The classical position effect variegation (PEV) in *Drosophila* is a typical example of constitutive heterochromatin, which shows that the conformational changes on chromatin could affect the transcription status of *white eye* gene[4]. Constitutive heterochromatin is known to be important for its role in transcription and gene silencing[5]. The stability of heterochromatin contributes to the inhibition of undesired recombination between repeat transposons as well as to the elimination of transcription of active transposons, and therefore maintains the integrity of the genome[1]. On the other hand, research has already identified tumor suppressors, for instance, the breast cancer type 1 susceptibility protein (BRCA1) and the retino-blastoma protein (RB), exerting their functions by promoting the formation of heterochromatin or maintaining the stability of constitutive heterochromatin, which suggests that heterochromatin could plays an essential role in tumor suppression [6-9]. Figure 1 provides an overview of the important findings that contribute to the development of the concept of heterochromatin[10].

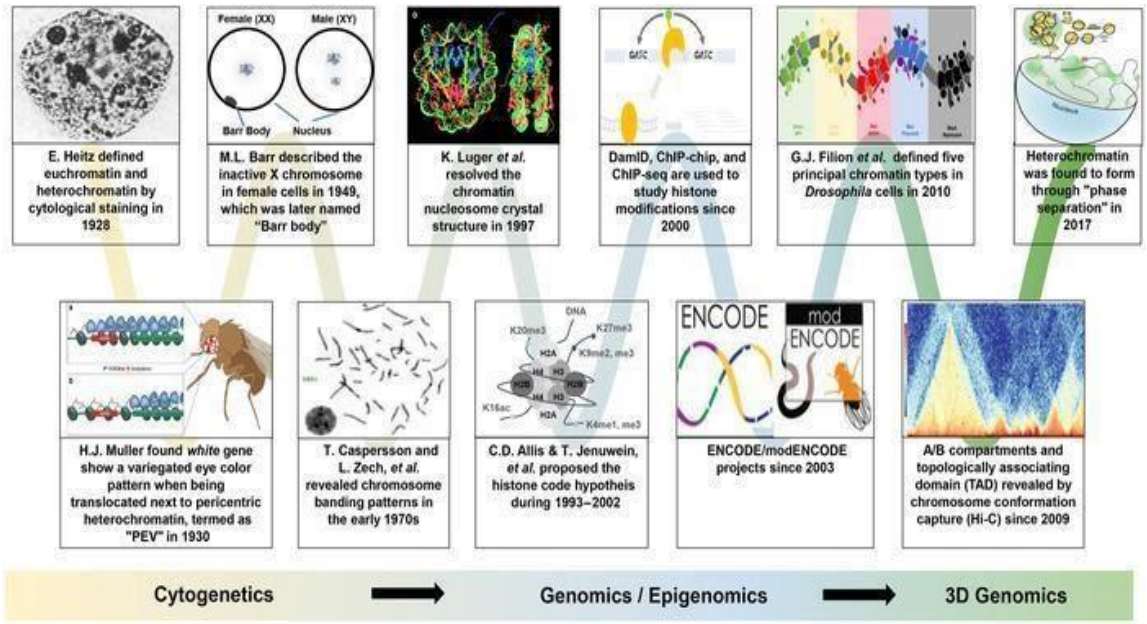

Figure 1. The Evolution of Concept of Heterochromatin

The significant findings and technical progress contribute to the evolution of heterochromatin

Figure 1 is from Liu, Jing et al. Annals of the New York Academy of Sciences vol. 1476,1 (2020): 59-77, with permission.

2. The Assembly of Heterochromatin and HP1

The assembly of heterochromatin could be classified into three stages: initiation, spreading and maintenance[11]. Though the mechanism is not fully understood, the initiation stage is believed to start when some sequence-specific proteins or non-coding RNA (in yeast) bind to the nucleation centers, in order to recruit histone deacetylases (HDACs) and the SUV39 family histone H3K9 methyltransferases (HMTs) to produce deacetylated histone and methylated H3K9[12-14]. The methylated H3K9 would bind to another important component and a determinant of heterochromatin, the heterochromatin protein 1 (HP1)[15-16]. Abnormal behavior of HP1 is found in diverse human cancers including breast, colorectal, brain, ovarian, blood, thyroid, prostate, lung, bone, pancreatic, and liver cancer [17-23]. The structure of HP1 contains three disordered elements, the N-terminal extension (NTE), the Hinge region (HR) and the C-terminal

extension (CTE) as well as two folded domains; the chromodomain (CD) and the chromoshadow domain (CSD) [24]. HP1 paralogs share highly conserved sequences, especially among CD and CSD, which allows them to dimerize and form both homodimers as well as heterodimers[25]. Among all the dimerization, the CSD-mediated dimerization is mainly used in this paper. Commonly, the CSD-mediated dimerization takes place when the two CSD monomers interact with each other through their α-helices. This CSD-mediated dimerization enables HP1 to interact with multiple non-histone proteins with PXVXL, PXVXI, or relative motifs [26]. The right panel of Figure 2 shows the interaction between HP1 and PIWI which possess the PXVXL motif [27]. In the traditional view, HP1 forms static oligomeric networks so as to prevent transcription factors from accessing silenced genes located in the heterochromatin domain [28]. However, research has shown that HP1 transiently binds to heterochromatin and HP1's CD-H3K9me3 interaction is not sufficient for HP1's binding to heterochromatin [28]. By combining these findings, one possible explanation is that the binding of HP1 to heterochromatin is stabilized by multivalent interactions transiently. In this paper, I will focus specifically on one potential candidate, the unphosphorylated STAT (uSTAT).

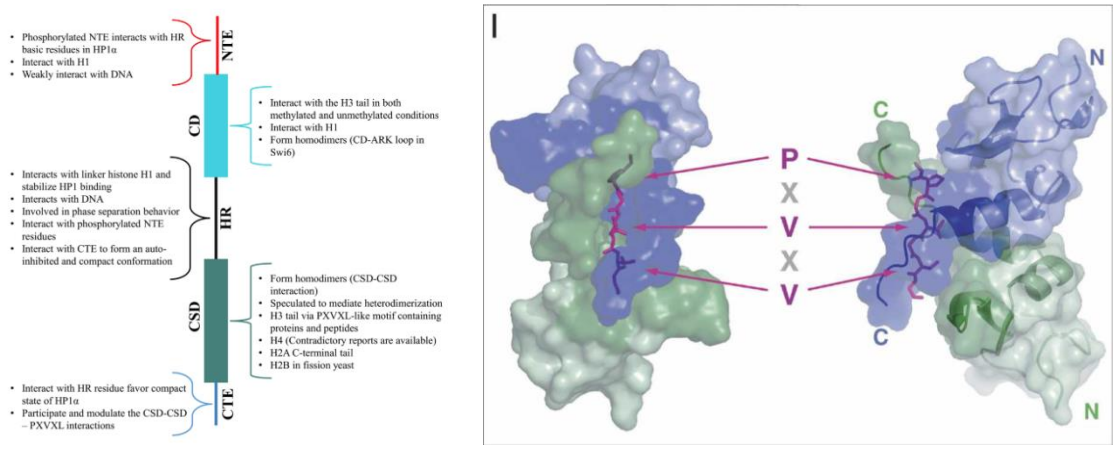

Figure 2. HP1 Structure and Interaction with PxVxI Motif

Summary of how HP1 dimerize and interact with chromatin. HP1 interact with PIWI with PXVXL motif
Panel left is from Kumar, Amarjeet, and Hidetoshi Kono. *Biophysical reviews* vol. 12,2 (2020): 387-400. Permission approved
Panel right is from Brower-Toland, Brent et al. *Genes & development* vol. 21,18 (2007): 2300-11. Permission approved

3. Liquid-liquid Phase Separation and Heterochromatin Formation

Recent research proposes a new possible explanation for the role of HP1 in heterochromatin formation[50]. Phosphorylation on the N-terminal extension of HP1α and binding of DNA trigger higher order oligomerization of HP1α which promotes the formation of phase-separated droplets[51]. Specific HP1α (shugoshin and LBR) could promote or reverse the phase separation[51]. Phase-separated droplets compartmentalize heterochromatin, and dynamically expose buried nucleosomal regions[52]. Hypothetically, the reshape of nucleosome core is correlated with histone modification and could increase multivalent interactions between nucleosomes, promoting phase separation[52]. The formation of phase-separated droplets could be separated into three steps as observed in DNA curtain assay experiments in vitro[53]. Before DNA condensation, HP1α would bind to distal points of DNA strands simultaneously. The bind HP1α would then capture lateral DNA fluctuations to start DNA compaction. The process would proceed by trapping more non-compacted DNA through HP1a-DNA and HP1a-HP1a interactions. Further experiments indicated that the stability of compaction is more susceptible to fluctuations in HP1 concentration than that in DNA level[53].However, how HP1 compact DNA in vivo is not completely understood. The experiment shows that binding with other proteins could affect the critical concentration of HP1 for condensate formation[53].

4. STAT and the JAK-STAT Pathway

Signal transducer and activator of transcription (STAT) proteins are latent cytoplasmic transcription factors which could be activated by diverse extracellular signaling pathways [29]. Normally, STAT proteins contain structural motifs including coil-coil, DNA binding domain, Src

homology as well as C-terminal transactivation domain [30]. In the canonical JAK-STAT signaling pathway, JAK tyrosine kinase is activated by extracellular bindings of cytokines or growth factors and could phosphorylate specific tyrosine on STAT proteins [31]. The phosphorylated STAT protein could dimerize via interaction between the SH2 domains and enter the nucleus via the importin α/β ternary complex [32]. After entering the nucleus, the dimerized STATs could bind to the target DNA sequence and regulate the transcription [32]. The JAK-STAT signaling pathway could thus accept an extracellular cue and translate it into a transcriptional response, which plays an important role in cell proliferation, differentiation, cell migration, as well as cell apoptosis and thus the abnormal activities of JAK/STAT signaling pathway could trigger multiple diseases including cancers [32].

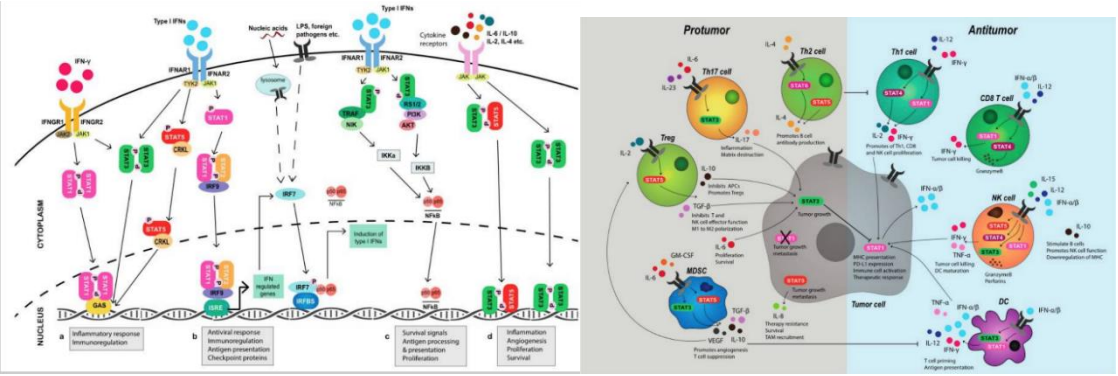

Figure 3. JAK-STAT Pathway and STATs Role in Oncogenesis

A simplified description of JAK-STAT signaling pathway
The contradicting role of STAT plays in oncogenesis
Panel left/right are from Owen, Katie L et al.*Cancers* vol. 11,12 2002. 12 Dec. 2019, Permission approved

5. The *Drosophila* Model: uSTAT92E Could Stabilize Heterochromatin by Binding HP1

Previous research in our lab has found that overactivated JAK could disrupt the formation of heterochromatin in a *Drosophila* leukemia model [33-35]. In the canonical JAK-STAT signaling pathway, JAK would phosphorylate STAT, which induces transcriptional regulation on specific target genes. Therefore, the overactivated JAK could disrupt the heterochromatin

through phosphorylation of STAT. To further investigate the role of STAT in heterochromatin formation and stability maintenance, further experiments were performed, and *Drosophila* was used as the organism model [36]. *Drosophila* was a perfect fit for this experiment. On the one hand, *Drosophila* is the only genome whose constitutive heterochromatin sequences are assembled into the genome map, and position-effect variegation (PEV) could be used to measure the state of heterochromatic gene silencing [37]. On the other hand, *Drosophila* possesses only one JAK and one STAT, which eliminates the redundancies of paralogs. The experiment revealed that STAT92E could colocalize with HP1, and the loss of STAT92E would trigger delocalization of HP1 which impaired the formation and stability of heterochromatin [36]. It was observed that decrease of STAT92E and overactivation of JAK would both induce HP1 delocalization and overall loss of H3mK9, resulting in the disruption of heterochromatin, which seem to be contradictory since JAK activates STAT through the JAK-STAT pathway [36]. Also, it was found that cells with overexpression of STAT92E would exhibit higher levels of heterochromatin than wild type cells [36]. Based on these findings, a hypothesis was proposed that it was unphosphorylated STAT92E (uSTAT92E) that actually maintains the stability of heterochromatin [36]. It was identified that STAT92E protein had a perfect and an imperfect HP1-binding sequence motif PxVxL(x donates any a.a.)[36]. Further immunofluorescence experiments proved that only unphosphorylated STAT92E protein would physically associate with HP1; phosphorylated STAT92E spread more evenly in the nucleus and did not possess colocalization with HP1 [36]. Treatment with pervanadate (H2O2/vanadate) was used to increase the phosphorylation level of STAT92E [36]. The results confirmed that phosphorylation would compromise the association between STAT92E and HP1, triggering HP1 dispersal and destabilizing heterochromatin [36]. To test whether the effect of STAT92E on heterochromatin

involved transcription of other proteins, cycloheximide was used to block protein synthesis and the result suggested that synthesis of other proteins were not involved in this process [36].

6. The Mammalian Cell: The Non-canonical Role of Human STAT Proteins

The experiments in the *Drosophila* model demonstrated several essential points. Traditionally, it is believed that phosphorylation is required for STAT protein to enter the nucleus, while in the experiments, it was found that unphosphorylated STAT92E would colocalize with HP1 in the nucleus and it has the PV HP1 binding motif. Moreover, phosphorylation of STAT92E would trigger dispersal of HP1 which impairs the formation and stability of the heterochromatin. Therefore, it seems that in *Drosophila*, STAT92E and heterochromatin are important to the maintenance of genome integrity. However, further experiments are necessary to see whether the unphosphorylated STAT in mammalian cells also has the same non-canonical functions. Thus, our lab performed experiments in human cells and mouse xenograft models to explore the interaction between STAT5A and HP1α (also known as "chromobox protein homolog 5," or CBX5) [38]. It is known that STAT5A has the HP1-binding motif PxVxI which is located from a.a 465 to 469 [38]. To test whether STAT5A is physically associated with HP1, point mutation was introduced to the HP1 binding motif and the result of ChIP showed that only STAT5A with intact HP1-binding motif would be co-immunoprecipitated with HP1α [38]. Also, the result found that the mutations that interrupted the H3K9me binding, HP1α dimerization, and interaction with the PxVxL motif of HP1 would also decrease the affinity between STAT5A and HP1α [38]. Further experiments of point mutation confirmed that only unphosphorylated STAT5A with intact HP1-binding would physically interact with HP1 [38].

Similar to uSTAT92E in *Drosophila*, uSTAT5A could advance the formation of heterochromatin through binding to HP1 [38]. Analysis with fluorescence recovery after photobleaching (FRAP) indicated that unphosphorylated STAT5A stabilized heterochromatin by inhibiting the mobility of HP1, consistent with the competition model which suggested that the binding between HP1 and heterochromatin is *transiently* stabilized by available binding partners [38]. In the mouse xenograft model, it was found that the overexpression of unphosphorylated STAT5A could suppress tumor development and this ability was strongly correlated with its binding affinity with HP1 [38]. The experiment on human colon cancer cells confirmed the same conclusion [38]. Moreover, it was found that genes downregulated by overexpression of $STAT5A^{Y694F}$ (mutated STAT5A which could not be phosphorylated) or overexpression of HP1α were largely overlapped and a majority of which were tumor-promoter genes [38]. By using the cBio Cancer Genomics Portal, it was found that the underexpression of STAT5 and HP1α correlate with poor prognosis in certain human cancers[38].

Finally, similar experiments were done for human STAT3 and it was found that unphosphorylated STAT3 also promotes heterochromatin formation and functions as a tumor suppressor in mouse xenograft models[54]

7. A Hypothetical Model

Though the mechanism is still unclear, the research about STAT92E, STAT3 and STAT5A seems to suggest that besides the canonical JAK-STAT signaling pathway, STAT could also be involved in gene transcription by heterochromatin maintenance. Recent preliminary data in our lab indicates that unphosphorylated STAT might also participate in the initiation stage of heterochromatin formation, in addition to its maintenance. These findings

provide new insight into the role of STAT. Based on this, a hypothetical model could be proposed that unphosphorylated STAT could both initiate and maintain the stability of heterochromatin by binding to HP1. During the initiation stage, unphosphorylated STAT compete with histone to bind to HP1α. Furthermore, unphosphorylated STAT could bind to the HP1 already associated with heterochromatin and form a complex to promote the formation of phase-separated droplets. Phosphorylation could trigger conformational change which makes STAT switch from the HP1-binding state to the DNA-binding state. To construct a theoretical model, three essential components are involved: HP1, uSTAT and DNA.

8. Thermodynamics

Since the binding between HP1 and heterochromatin is transiently stabilized by STAT, it is assumed that the binding among DNA, HP1 and STAT is non-covalent and thus reversible and follows thermodynamic principles. According to the Gibbs energy of reaction, the change of free energy (ΔG) is equal to change in enthalpy minus temperature (T) times the change in entropy (ΔS)[40].

$$\Delta G=\Delta H-T\Delta S \quad (1)$$

When the value of ΔG is negative, the reaction is spontaneous and proceeds in the forward direction. When the value of ΔG is zero, the reaction reaches equilibrium. When the value of ΔG is positive, the reaction would proceed in the reverse direction. For the binding of the model in this paper, the reaction could be summarized as follows:

$$HP1 + DNA + uSTAT \leftrightharpoons HP1\text{-}uSTAT + DNA \leftrightharpoons HP1\text{-}uSTAT\text{-}DNA$$

After Phosphorylation

HP1 + DNA + pSTAT ⇋HP1+pSTAT-DNA

The equilibrium constant of a chemical reaction is the value of the reaction quotient when the rate of forward reaction is equal to the rate of reverse reaction. And the equilibrium constant could be calculated as follows:

$K_d = \frac{[A]^x[B]^y}{[A_xB_y]}$ for $A_xB_y \leftrightharpoons xA + yB$

And thus when there is more free HP1 and uSTAT in the nucleus, the uSTAT3 would bind to HP1 and stabilize its binding on heterochromatin until the reaction reaches equilibrium. Furthermore, it suggests that the reaction could be pushed to the forward direction by increasing one of the reactants. If there is more uSTAT in the nucleus, more HP1 could be stabilized on the heterochromatin.

9. Residues for the Interface

In this study, STAT3 would be used to construct the theoretical model since its crystallized structures of phosphorylated state and unphosphorylated state have already been collected by Nkansah et al. [41]. The crystallized structure also suggests that phosphorylation is not necessary for STAT3 binding to DNA, which partly supports the hypothesis [41]. STAT3 is first identified as a DNA-binding protein that can be activated by interleukin 6 [42]. STAT3 could be phosphorylated at Tyrosine 705 and the constitutive activation of STAT3 is observed in many human cancers including breast and head and neck carcinomas, lymphomas, leukemias, melanomas, and pancreatic, lung, ovarian, and brain tumors, which also partly confirmed that unphosphorylated STAT3 might contribute to maintain the stability of heterochromatin [43-44]. STAT3 is predominantly localized in the nucleus, which might imply higher frequency of

interactions with HP1 and DNA [45]. Since the affinity to HP1 could be influenced by conformational change, the structure of both unphosphorylated STAT3 and phosphorylated STAT3 used in the theoretical model would be homodimers. Considering the homology among STAT3, STAT5A and STAT92E, the HP1-binding motif is identified and will be used to determine the interface between uSTAT3 and HP1 during binding. Previous experiments with STAT5A also suggest that dimerized HP1 exhibits higher binding affinity with uSTATA and thus the structure of homodimers of HP1 would be used to construct the theoretical model [38]. It is already found that an intact chromo shadow domain is required for the HP1 interacting with PXVXI motif [46]. Research on HP1 interacting with other peptides has identified several amino acids that might be directly involved in the binding. Experiments done by Thiru et al on binding between HP1 and CAF-1 has found that Ala-125, Phe-163, Arg-167, Leu-168, Trp-170, Thr-126, Leu-135 and Leu-146 might physically associate with PXVXL motif, and thus those amino acids would be labelled as active residues that would be part of the interface [46].

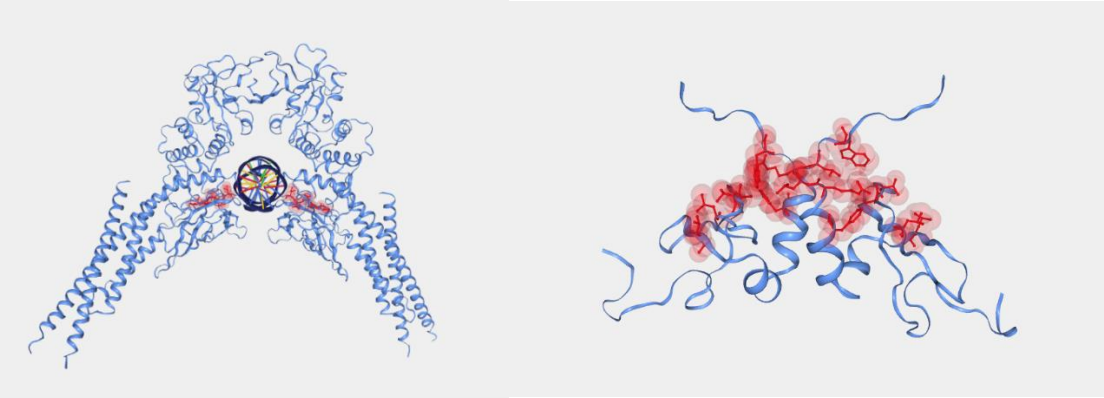

Figure 4. Active Residues on uSTAT3 and HP1

Red part highlights the potential active residues that might involve in the interaction. The left panel represent uSTAT3 with DNA and the right panel represent HP1.

Thus, the paper aims to explore:

- If phosphorylation would change the binding affinity between STAT3 and HP1.
- What is responsible for the change in binding affinity

METHODS & MATERIALS

Protein structures

The structure of proteins used in modeling were downloaded from the PDB Structure database on the NCBI website. The structure of unphosphorylated STAT3 with double stranded DNA (PDB number 4E38)[41] and the structure of phosphorylated STAT3 double stranded DNA (PDB number 1BG1)[4] were derived from the crystallized structure produced by Nkansah et al. The structure of HP1 homodimers (PDB number 1S4Z) came from the crystallized structure produced by Thiru et al[46]. The original structure of HP1 was a complex with CAF-1 binding to HP1, and the CAF-1 was removed through Pymol[37].

SWISS-MODEL

SWISS-MODEL was used in the experiment to generate theoretical HP1α structure[48]. The protein sequence of HP1α was obtained via accession number NP_001120794.1. The template used HP1β homodimers (PDB number 1S4Z). The original template is heterotrimers; pymol was used to remove the CAF factors.

HADDOCK

HADDOCK (High Ambiguity Driven protein-protein DOCKing) is used in experiments for modeling[47,55-56]. Protein structure (PDB number 4E38 and PDB number 1BG1) was trimmed with PDB tools to remove water molecules, merge peptides and renumber residues to fit in the format of HADDOCK. Protein structures were uploaded to the HADDOCK web server and active residues were selected according to the PXVXI binding motif. All parameters were kept as default values, except setting the number of structures for rigid body docking to 10000, the number of structures for semi-flexible refinement to 400, the number of structures for the final

refinement to 400 and the number of structures to analyze to 400. The docking results were downloaded for analysis.

HADDOCK perform protein docking through three stages: rigid-body energy minimization stage (it0), semi flexible simulated annealing stage (it1) and refinement in explicit solvent stage (itw). During rigid-body energy minimization stage, the system generated several thousands of models in a short period of time. During the semi flexible simulated annealing stage, more flexibility is allowed at the interface of rigid-body, side-chain as well as backbone + side chain. At the refinement in explicit solvent stage, the models are put in explicit solvent to refine residue-residue contacts, mainly electrostatic, at the interface.

Haddock uses van der Waals (E_{vdw}), electrostatic (E_{elec}), desolvation (E_{desol}) and restraint violation (E_{air}) energies as well as buried surface area (BSA) to calculate grades of predictive models. Van der Waals energy refers to the potential energy possessed by the Van der Waal force. Electrostatic energy refers to the total work done by external agents in bringing the system of charge to its current configuration. Desolvation energy refers to the energy electrostatic and van der Waals energy loss of the interaction between protein and solvent due to binding. The restraints violation energy checks if restraints (distance, dihedral angle, RDC, etc) used in the model are completely respected. And the calculation equations are listed:

$$*HADDOCK_{score}\text{-it0} = 0.01\ E_{vdw} + 1.0\ E_{elec} + 1.0\ E_{desol} + 0.01\ E_{air} - 0.01\ BSA$$

$$* HADDOCK_{score}\text{-it1} = 1.0\ E_{vdw} + 1.0\ E_{elec} + 1.0\ E_{desol} + 0.1\ E_{air} - 0.01\ BSA$$

$$* HADDOCK_{score}\text{-water} = 1.0\ E_{vdw} + 0.2\ E_{elec} + 1.0\ E_{desol} + 0.1\ E_{air}$$

Generally, the more negative the energy, the more reliable the model proposed.

PRODIGY

PRODIGY (PROtein binDIng enerGY prediction) was used to predict the binding affinity[58-59]. The outputs of HADDOCK were imported to PRODIGY and the complexes were labeled as “A” and “B” bands while the Gibbs free energy and dissociation constants were calculated.

RESULTS

Phosphorylation induces shift of DNA binding domain on STAT3

To construct a plausible model of how HP1 binds to uSTAT3, I first examined the conformational change on STAT3 induced by phosphorylation. The 3D structure of phosphorylated STAT3 and unphosphorylated STAT3 was downloaded from the ncbi Structure database and analyzed with Pymol. According to the experimental design of Nkansah et al, STAT3 was labeled with spectral variants of GFP, replacing the N-terminal 126 amino acids. The analysis of the sequence provided several points. First of all, phosphorylation of Tyrosine 705 induced a shift of the DNA binding domain on STAT3. The DNA binding domain on unphosphorylated STAT3 ranges from a.a 321 to a.a 484, while the DNA binding domain on phosphorylated STAT3 ranges from a.a 417 to a.a 610. Noticeably, both pSTAT3 and uSTAT3 interact with DNA at a.a 331-332 or a.a 340 or 343-344 or a.a 423 or a.a 465-466 or a.a 469. Based on the sequence and homology between STAT3 and STAT5A, the HP1-binding motif is identified from a.a 460 to a.a 464. It is believed that protein structures could interact via a similar architectural motif even if the overall structures are different. Thus, theoretically, the association between HP1 and unphosphorylated STAT3 shall also involve the PVVVI HP1 binding motif. The homodimer of HP1 is also downloaded from the ncbi Structure database via the PDB number 1S4Z. The original structure is a complex in which HP1 binds to a peptide with the PXVXL motif. The complex was trimmed with Pymol and only the homodimers of HP1 would be used for modeling. Both pdb files were trimmed with pdbtools to fit in the format of Haddock and were imported to Haddock for docking. Based on the interactions between CAF-1 and HP1β, Ala-125, Phe-163, Arg-167, Leu-168, Trp-170, Thr-126, Leu-135 and Leu-146 in HP1 were labeled as active residues which might participate in binding.

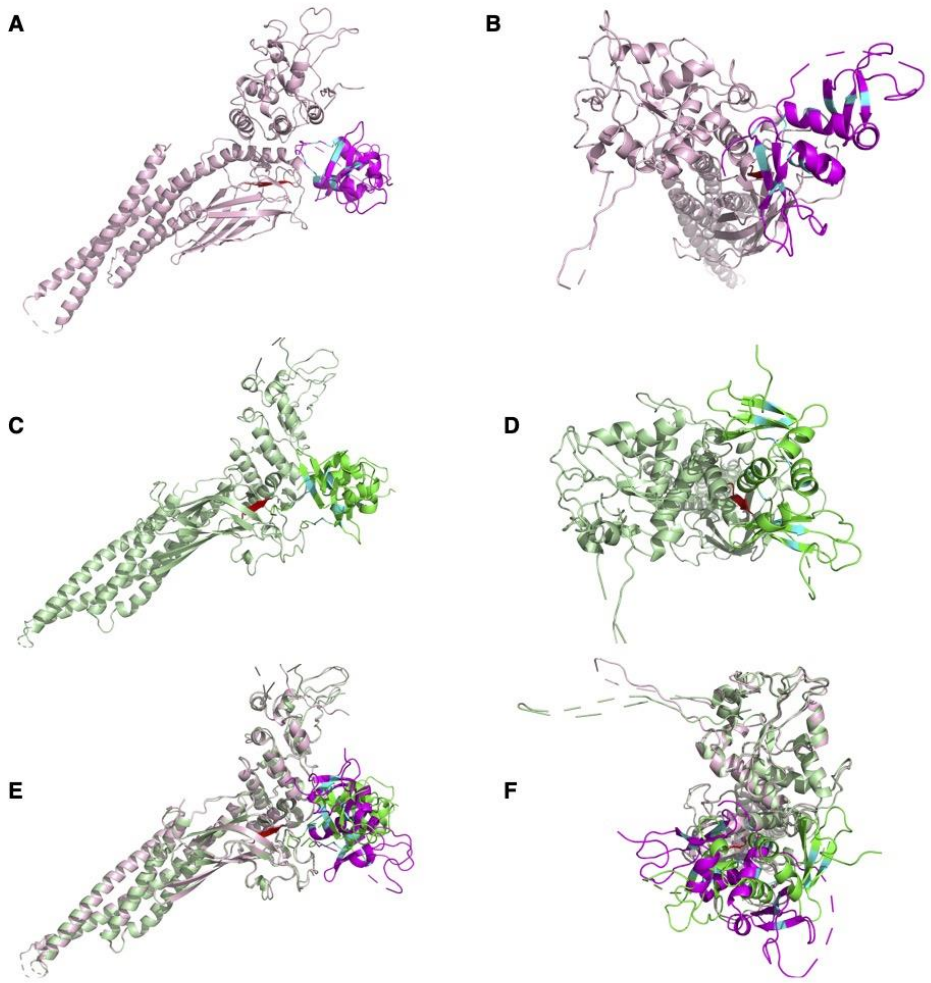

Figure 5. Predictive Model of HP1-STAT3(monomer)

Light pink refers to uSTAT3, dark pink refers to HP1 dimers bind to uSTAT3. Light green refers to pSTAT3. Dark green refers to HP1 dimers binds to pSTAT3.

Phosphorylation Might Influence HP1-STAT3 Binding via Electrostatic Interaction.

I performed protein docking between HP1 and STAT-DNA complex to see if DNA would affect the interaction between HP1 and STAT3. By comparing HP1-uSTAT3 and HP1-pSTAT3 models, it is obvious that the conformational change of STAT would influence the

orientation of HP1 binding to STAT (Figure 6). Noticeably, the electrostatic energy of uSTAT3-HP1 is surprisingly lower (more negative) than that of pSTAT3-HP1. According to Coulomb's law, the electrostatic energy could be calculated via the equation:

$$U = \sum_{i=1}^{N_A} \sum_{j=1}^{N_B} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}$$

N_A and N_B represent the number of point charges in the two biomolecules. From the equation, the electrostatic energy is proportional to the charge and inversely proportional to the distance between two charges. Thus, it could be implied that compared to pSTAT3-HP1, uSTAT3-HP1 has either larger charges or smaller distance during the binding. Besides electrostatic interactions, non-bonded interactions also include van der Waals interactions, which could be calculated with Lennard-Jones 12-6 functions:

$$\upsilon(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]$$

The collision diameters refers to the separation for which the energy is zero and refers to the well depth. Interestingly, for uSTAT3-HP1 and DNA-uSTAT3-HP1, the van der Waals energy did not change much, but the value of electrostatic energy was largely decreased. Thus, it might imply that the binding of DNA does not influence the distance of two molecules but shifts the overall charges which result in the different orientation of HP1. The same phenomenon was observed between DNA-uSTAT3-HP1 and DNA-uSTAT3-2HP1.

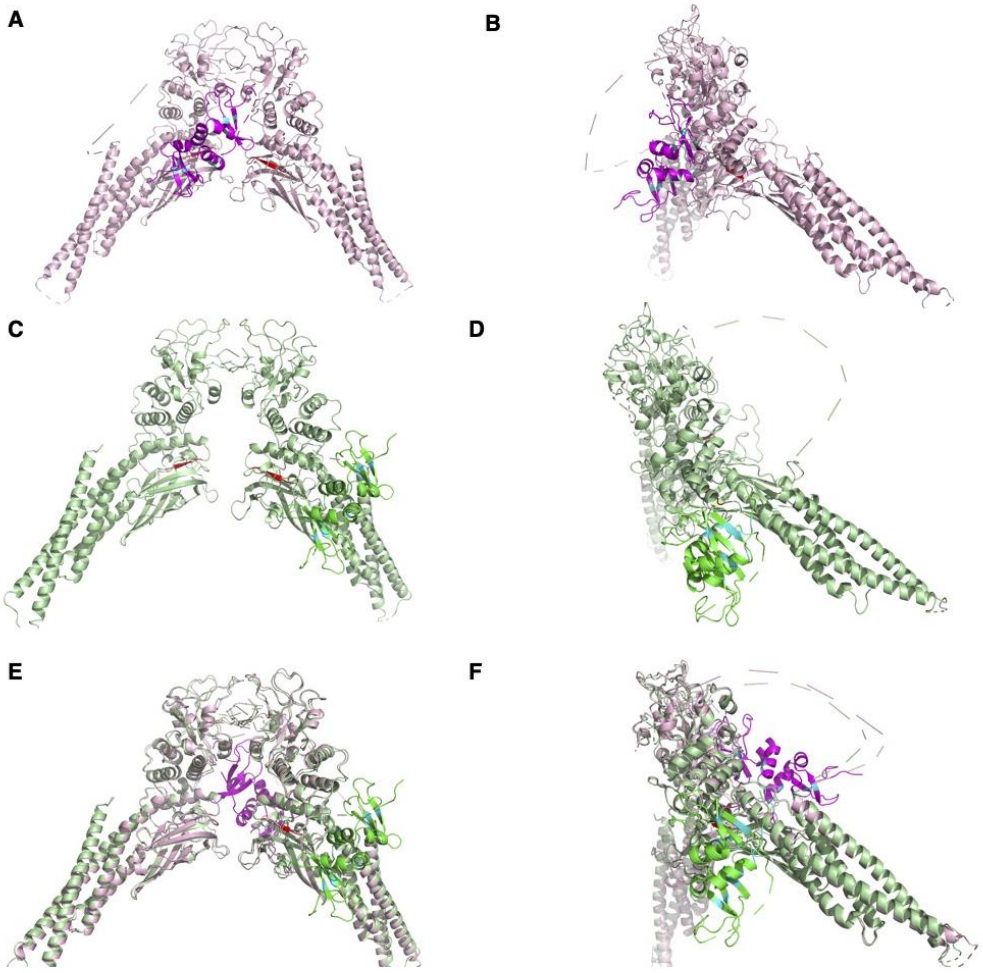

Figure 6. HP1-STAT3(homodimer) Model

Light pink refers to uSTAT3, dark pink refers to HP1 dimers bind to uSTAT3. Light green refers to pSTAT3. Dark green refers to HP1 dimers binds to pSTAT3.

Phosphorylation might affect the Binding Affinity of STAT3.

Normally, the more negative the grade, the better the model. Nonetheless, Haddock score could only be used to evaluate the quality of a model within the same run; it could not be used to compare the binding affinity of protein-protein interactions. Thus, I used PRODIGY (PROtein binDIng enerGY prediction) to measure the free energy and Kd of binding complexes. From the

data, it is shown that in monomer state, phosphorylated STAT3 has lower ΔG and Kd compared to unphosphorylated STAT3 when binding to HP1. According to the Dissociation Equation:

$$K_d = \frac{[A][B]}{[AB]}$$

Thus, K_d [AB] = [A][B]. The ΔG for both uSTAT3-HP1 and pSTAT3 1 were negative, which suggests that the equreaction could take place in the forward (association) direction spontaneously. On the other hand, since pSTAT3 has lower K_d than uSTAT3, it suggests that under the same circumstance at equilibrium state pSTAT1-HP1 will be less dissociated than uSTAT3-HP1 at 37 Celsius degrees.

Since the dimerization of STAT is reversible, STAT could also exist as homodimers. Thus, I also perform protein docking between HP1 and STAT3 in the homodimer state (Figure 6). The superimposition of two complexes shows that the conformational change induced by phosphorylation would largely affect the orientation of HP1 binding to STAT3 (Figure 6, panel E, F). The data show that in homodimer state, uSTAT3 exhibits lower ΔG value (more negative) and Kd than pSTAT3. Lower ΔG value shows that under the same circumstance the binding between uSTAT3 and HP1 are more likely to happen. Lower K_d indicates that at equilibrium state there should be a higher level of uSTAT3-HP1 complexes than pSTAT3-HP1.

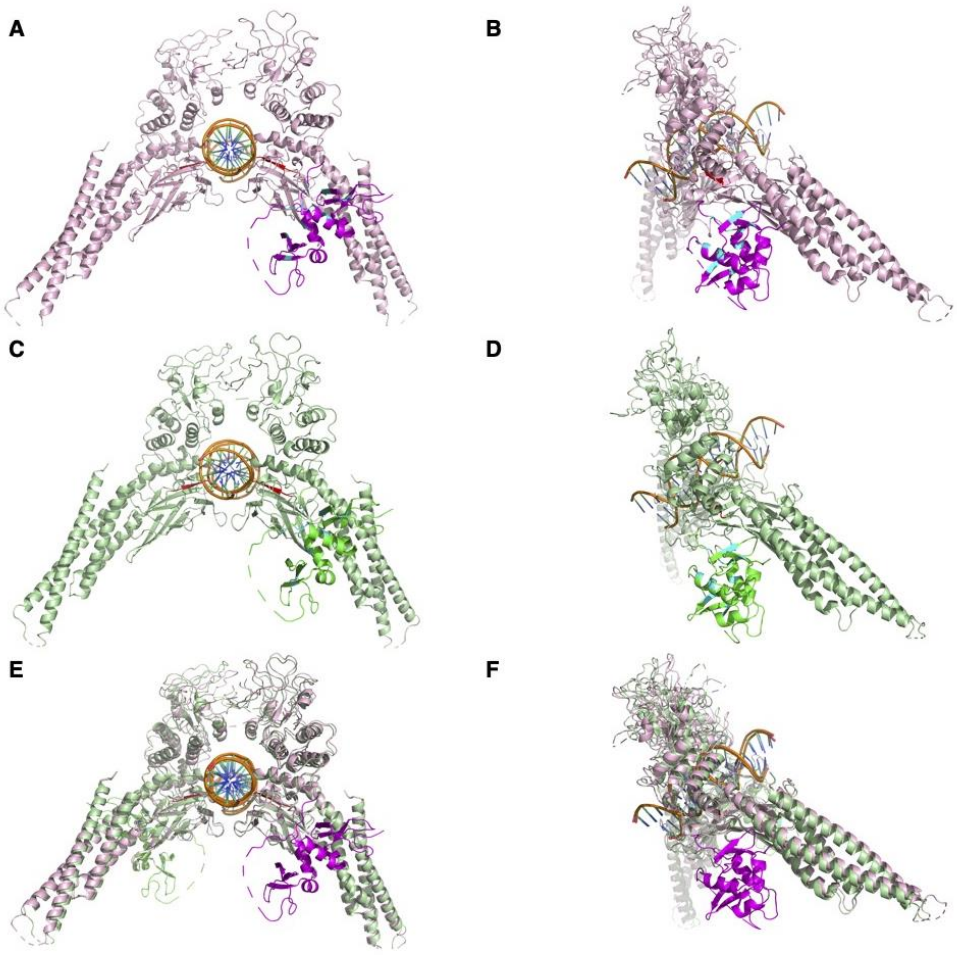

Figure 7. Predictive Model of DNA-STAT3-HP1

Light pink refers to uSTAT3, dark pink refers to HP1 dimers bind to uSTAT3. Light green refers to pSTAT3. Dark green refers to HP1 dimers binds to pSTAT3.

Previous research indicates that phosphorylation is not required for STAT3 to bind to DNA[41], which means unphosphorylated STAT3 could also bind to DNA. Therefore, I perform protein docking between the STAT3-DNA complex and HP1 to see if binding to DNA would affect the way STAT3 interacts with HP1 (Figure 7). Noticeably, after binding to DNA, HP1 would access uSTAT in a different orientation (Figure 7 A, B), while the binding between HP1 and pSTAT remains the same. The data shows that HP1-uSTAT3-DNA has lower ΔG and K_d

value than HP1-pSTAT3-DNA, which confirms that the binding between HP1 and uSTAT3-DNA complexes is more stable than the binding between HP1 and pSTAT3-DNA complexes. Moreover, I used PreDBA to calculate the binding affinity of STAT to DNA. The ΔG of uSTAT3-DNA is -6.15 kcal/mol, while the ΔG of pSTAT3-DNA is -10.46 kcal/mol. Thus, phosphorylation largely increases the binding affinity of STAT3 to DNA.

Table 1. Gibbs Free Energy and Disociation Constant of Complexes

Parameter/Model	ΔG (kcal mol-1)	Kd (M) at 37.0 ℃
uSTAT3(monomer)-HP1	-9.5	2.0E-07
pSTAT3(monomer)-HP1	-11.0	1.7E-08
uSTAT3-HP1	-12.9	8.7E-10
pSTAT3-HP1	-11.7	5.4E-09
DNA-uSTAT3-HP1	-13.4	3.5E-10
DNA-pSTAT3-HP1	-12.4	1.8E-09
DNA-uSTAT3-2HP1	-13.0	7.0E-10
DNA-pSTAT3-2HP1	-10.6	3.6E-08
DNA-uSTAT3	-6.15	4.64E-05
DNA-pSTAT3	-10.46	4.26E-08

STAT homodimers have two PXVXI HP1-binding motifs, and only one is used during each binding between STAT-DNA complexes and HP1. To explore if the first binding HP1 would affect the binding of the second HP1, protein docking was performed between DNA-STAT3-HP1 complexes and HP1 (Figure 8). The data shows that interaction with second HP1 has higher ΔG and K_d value than interaction with the first HP1, which implies that the association with second HP1 might be less preferable.

Though the difference is small, HP1-uSTAT3 (homodimers) and HP1-uSTAT3-DNA have lower ΔG values (more free energy) than their pSTAT3 counterparts respectively, which are consistent with the observation that phosphorylation trigger dispersal of STAT from HP1. Nonetheless, in monomer state, pSTAT3 exhibited higher affinity with HP1 than uSTAT3. One potential explanation for this could be that in monomer state the SH2 domain is exposed and the SH2 domain of pSTAT3 exhibits higher affinity to a wide range of proteins. I also used PreDBA to determine how phosphorylation would affect STATs' affinity with DNA, and the result suggests that phosphorylation could largely increase STATs' affinity with DNA.

Thus, the computational modelling suggest that unphosphorylated STAT3 has higher binding affinity with HP1 while phosphorylated STAT3 has higher binding affinity with DNA. Those changes in binding affinity might correlate with changes in electrostatic interactions and conformational change. To determine if electrostatic change induced by phosphorylation is responsible for the change of binding affinity, I add phosphate to the Tyr705 on uSTAT3-HP1a complexes and re-calculate the binding affinity of complexes(Table2). The binding affinity of uSTAT3 did not change significantly after adding the phosphate (only HP1-uSTAT3 complex showed slight increase in ΔG), which may suggest that phosphorylation toggle STAT from HP1-binding state to DNA-binding state through inducing conformational change on STAT.

Table 2 Gibbs Free Energy and Kd of USTAT3-HP1 after Adding Phosphate

Complexes Add Phosphate	ΔG (kcal mol-1)	Kd (M) at 37.0 ℃
uSTAT3-HP1a(monomer)	-9.5	2.0E-07
uSTAT3-HP1a	-12.8	8.8E-10
DNA-uSTAT3-HP1a	-14.2	1.0E-10
DNA-uSTAT3-2HP1a	-13.0	7.0E-10

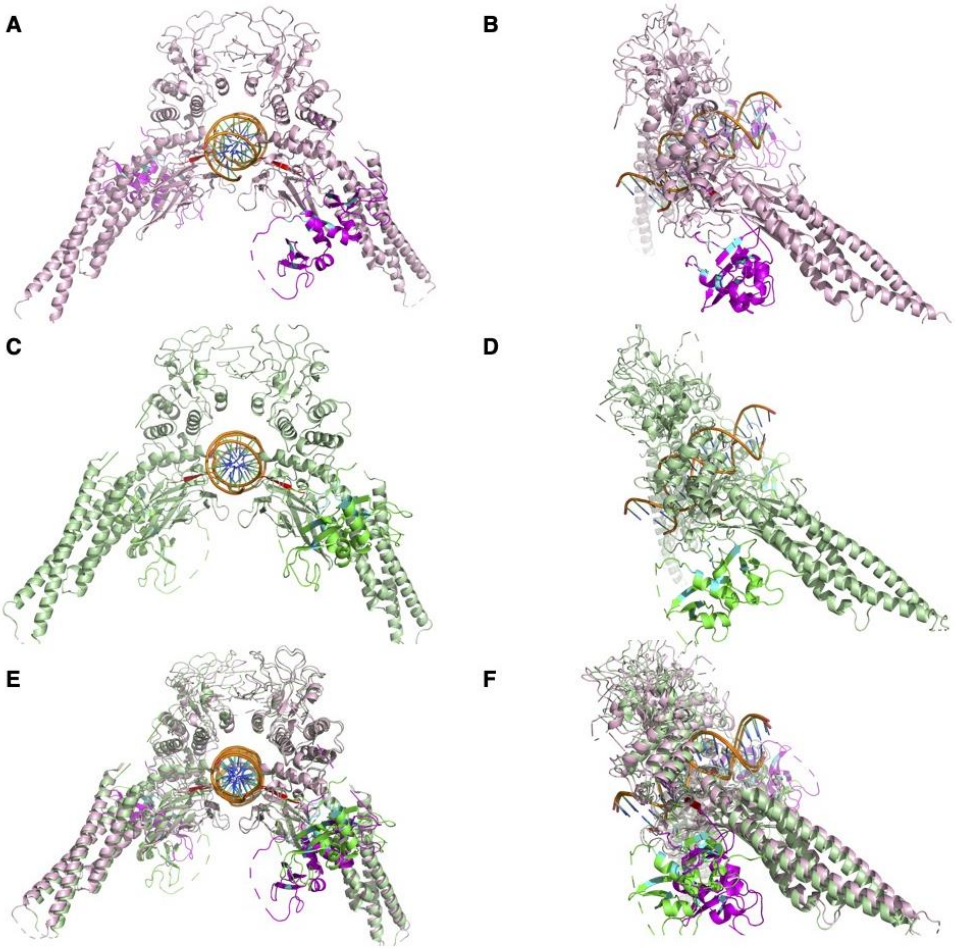

Figure 8. Predictive Model of DNA-STAT3-2HP1(on both PxVxI motif)

Light pink refers to uSTAT3, dark pink refers to HP1 dimers bind to uSTAT3. Light green refers to pSTAT3. Dark green refers to HP1 dimers binds to pSTAT3.

DISCUSSION

Protein Docking Provides Insight Into STAT-HP1 Interaction.

Protein-protein interactions play an important role in multiple biochemical activities. The interactome of interacting proteins in human cells is over 650000, while the dynamic of assembly as well as the time and localization of components could add to the complexity. On the other hand, crystallizing protein complexes and NMR are not easy tasks to perform and currently, only a few protein-protein complexes have been crystallized. Thus, to improve the structural understanding of protein-protein complexes, multiple computational tactics have been developed, including protein-protein docking. Protein-protein docking mainly serves to predict the structure of protein complexes based on their unbound components. From previous research in our lab, it is observed that uSTAT could promote the formation of heterochromatin and maintain the stability of heterochromatin through interacting with HP1. Therefore, protein docking between HP1 and STAT could help to unveil the details of interaction as well as to explore the mechanism behind it.

Computational Modelling Provides New Insight into HP1-STAT3 Binding.

Currently, there are several protein-protein docking softwares available and in this paper I used Haddock due to its broad applicability as well as its allowance of ambiguous restraints. Haddock could accept multiple biomolecular inputs including protein, ligand, nucleic acid, peptide, glycan, protein-ligand complex and protein-nucleic acid complex. Moreover, Haddock could convert the active/passive residues to ambiguous interaction restraints and allow conformational changes on both side chains and backbones during the modeling [47]. From previous experiments in our lab, point mutation has been used to confirm that PXVXI motif of

uSTAT3 is required for the binding. For experiments performed by Thiru et al, active amino acids interacting with the PXVXL motif have been identified on HP1 proteins.

Because dimerization is a reversible process, the monomer structure and homodimers structure of STAT3 were used in the protein docking respectively. As for HP1, only homodimers were used since previous research has proved that dimerization is required for HP1 to interact with PXVXI motif. By computational modeling, I found that phosphorylation increase STAT3's binding affinity with DNA while increase STAT3's binding affinity with HP1α. Such change is correlated with conformational change since only homodimer STAT3 exhibit such phenomenon. Moreover, phosphorylation changes the orientation that STAT3 binds to HP1α, mainly through affecting electrostatic force. Recent research indicates that HP1α with phosphorylated NTE/hinge tend to expose histone and promote the formation of phase separation. Thus, during the formation of heterochromatin, uSTAT3 might compete with histone to bind to HP1α and promote the phase separation. However, currently there is no structure of HP1α with phosphorylated NTE/hinge in the database, and therefore further experiments are needed.

AlphaFold: Development of Artificial Intelligence Sheds Light on Protein Structure.

The accuracy of computational prediction of protein structures has prevented large-scale structural bioinformatics, but this situation may not last for long. AlphaFold, a newly developed computational method in 2021, is able to construct protein three-dimensional structures simply with protein sequences[57]. In the CASP14(an independent mechanism for the assessment of methods of protein structure modeling)[57], AlphaFold has outcompeted all other computational modelling techniques in accuracy of protein structure prediction. Despite all its strengths, AlphaFold is not used in this experiment for multiple reasons. First and foremost, AlphaFold

could not tell difference between different state of the same protein: it could only produce the structure that are most likely to appear in the PDB file. Furthermore, AlphaFold currently could only predict protein structure, it could not predict protein-protein interactions. Hopefully, in the near future AlphaFold could be taught to perform protein-protein interaction.

Future Direction

Overall, this paper provides a predictive model about interactions between STAT3 and HP1, yet there are limitations. Previous experiments in our lab have shown that uSTAT5A and Ustat3 could interact with HP1. In the STAT5A paper, it is HP1α that interacts with uSTAT5A, yet there is no complete HP1α structure available (the one that was used in the model was constructed by SwissModel with the HP1β-CSD template). On the other hand, though previous research showed that intact CSD is sufficient for HP1 binding to PXVXI motif, there is evidence indicating that NLE is also important to determine the function of HP1α. Thus, a full copy of HP1α could improve the accuracy of protein docking. Furthermore, the HP1-uSTAT3-DNA exhibited large restraints violation energy, which suggests that the ambiguous restraints generated to instruct docking were hard to satisfy. Several factors might account for that. The PXVXI motif on STAT3 is buried, which is hard for HP1 to access. The active residues on HP1 are determined by interactions with CAF-1, but the details remain unclear (distance between atoms, distortion angle etc), and therefore further experiments are required to measure those parameters. Also, though the data showed that pSTAT3 exhibited higher affinity with DNA than uSTAT3, the ΔG value was smaller than pSTAT3-HP1. One possible explanation is that the ds DNA in the binding is not the DNA motif specifically enriched in STAT association. Thus, EMSA experiments are required to find the DNA STAT exhibit high affinity to and remeasure

the ΔG value. The new concept phase separation also sheds light on the interaction between STAT and HP1. We could test in vivo whether STAT3 could lower the critical concentration of HP1a via protein titration.

SUPPLEMENTARY MATERIAL

Besides models, HADDOCK also produced a series of parameters for each set of binding complexes respectively. The charts below are the van der Waals energy, electrostatic energy, desolvation energy, restraint's violation energy and buried surface area the system used to calculate the Haddock score. The system uses Haddock score to determine which model is the most credible.

Supplementary Table 3. HADDOCK Output for STAT3(monomer)-HP1

Parameter/Model	uSTAT3(monomer)-HP1	pSTAT3(monomer)-HP1
Haddock Score	-91.7 ±5.0	-94.1 ± 9.9
Van der Waals energy	-54.9 ± 6.1	-60.5 ± 1.3
Electrostatic energy	-286.9 ± 24.4	-222.1 ± 57.1
Desolvation energy	-6.1 ± 4.1	-5.5 ± 1.9
Restraint's violation energy	266.8 ± 18.1	163.5 ± 30.5
Buried Surface Area	1918.8 ± 70.5	1895.2 ± 94.0

Supplementary Table 4 HADDOCK Output for STAT-HP1

Parameter/Model	**uSTAT3-HP1**	**pSTAT3-HP1**
Haddock Score	-105.0 ± 10.7	-103.5 ± 5.7
Van der Waals energy	-70.4 ± 5.4	-66.4 ± 7.8
Electrostatic energy	-448.6 ± 58.1	-385.6 ± 54.0
Desolvation energy	1.1 ± 5.0	18.3 ± 8.1
Restraint's violation energy	540.3 ± 111.3	217.4 ± 56.2
Buried Surface Area	2749.7 ± 109.8	2309.6 ± 174.7

Supplementary Table 5. HADDOCK Output for HP1-STAT3 Complexes

Parameter/Model	DNA-uSTAT3-HP1	DNA-pSTAT3-HP1
Haddock Score	-74.7 ± 3.9	-64.3 ± 7.0
Van der Waals energy	-100.4 ± 10.2	-77.9 ± 6.5
Electrostatic energy	-59.5 ± 33.4	-63.9 ± 34.7
Desolvation energy	-0.3 ± 2.8	-3.7 ± 4.8
Restraint's violation energy	380.0 ± 32.6	300.5 ± 37.3
Buried Surface Area	2505.1 ± 254.2	2056.3 ± 120.9

Supplementary Table 6. Output for 2HP1-STAT3-DNA Complexes

Parameter/Model	DNA-uSTAT3-2HP1	DNA-pSTAT3-2HP1
Haddock Score	-105.6 ± 10.3	-115.5 ± 5.5
Van der Waals energy	-100.5 ± 7.7	-100.6 ± 8.8
Electrostatic energy	-139.3 ± 11.6	-235.8 ± 43.2
Desolvation energy	-1.1 ± 2.2	11.4 ± 6.3
Restraint's violation energy	237.9 ± 34.7	208.9 ± 96.4
Buried Surface Area	2661.1 ± 63.3	2409.1 ± 215.5

REFERENCES

1. Wang, J., Jia, S. T., & Jia, S. (2016). New Insights into the Regulation of Heterochromatin. Trends in genetics : TIG, 32(5), 284–294. https://doi.org/10.1016/j.tig.2016.02.005

2. Trojer, P., & Reinberg, D. (2007). Facultative heterochromatin: is there a distinctive molecular signature?. Molecular cell, 28(1), 1–13. https://doi.org/10.1016/j.molcel.2007.09.01

3. Grewal, S. I., & Jia, S. (2007). Heterochromatin revisited. Nature reviews. Genetics, 8(1), 35–46. https://doi.org/10.1038/nrg2008

4. Muller, H. J., & Altenburg, E. (1930). The Frequency of Translocations Produced by X-Rays in Drosophila. Genetics, 15(4), 283–311. https://doi.org/10.1093/genetics/15.4.283

5. Allshire, R. C., & Madhani, H. D. (2018). Ten principles of heterochromatin formation and function. Nature reviews. Molecular cell biology, 19(4), 229–244. https://doi.org/10.1038/nrm.2017.119

6. Nielsen, S. J., Schneider, R., Bauer, U. M., Bannister, A. J., Morrison, A., O'Carroll, D., Firestein, R., Cleary, M., Jenuwein, T., Herrera, R. E., & Kouzarides, T. (2001). Rb targets histone H3 methylation and HP1 to promoters. Nature, 412(6846), 561–565. https://doi.org/10.1038/35087620

7. Indovina, P., Marcelli, E., Casini, N., Rizzo, V., & Giordano, A. (2013). Emerging roles of RB family: new defense mechanisms against tumor progression. Journal of cellular physiology, 228(3), 525–535. https://doi.org/10.1002/jcp.24170

8. Isaac, C. E., Francis, S. M., Martens, A. L., Julian, L. M., Seifried, L. A., Erdmann, N., Binné, U. K., Harrington, L., Sicinski, P., Bérubé, N. G., Dyson, N. J., & Dick, F. A. (2006). The retinoblastoma protein regulates pericentric heterochromatin. Molecular and cellular biology, 26(9), 3659–3671. https://doi.org/10.1128/MCB.26.9.3659-3671.2006

9. Zhu, Q., Pao, G. M., Huynh, A. M., Suh, H., Tonnu, N., Nederlof, P. M., Gage, F. H., & Verma, I. M. (2011). BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. Nature, 477(7363), 179–184. https://doi.org/10.1038/nature10371

10. Liu, J., Ali, M., & Zhou, Q. (2020). Establishment and evolution of heterochromatin. Annals of the New York Academy of Sciences, 1476(1), 59–77. https://doi.org/10.1111/nyas.14303

11. Wang, J., Lawry, S. T., Cohen, A. L., & Jia, S. (2014). Chromosome boundary elements and regulation of heterochromatin spreading. Cellular and molecular life sciences : CMLS, 71(24), 4841–4852. https://doi.org/10.1007/s00018-014-1725-x

12. Volpe, T. A., Kidner, C., Hall, I. M., Teng, G., Grewal, S. I., & Martienssen, R. A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by

RNAi. Science (New York, N.Y.), 297(5588), 1833–1837. https://doi.org/10.1126/science.1074973

13. Jia, S., Noma, K., & Grewal, S. I. (2004). RNAi-independent heterochromatin nucleation by the stress-activated ATF/CREB family proteins. Science (New York, N.Y.), 304(5679), 1971–1976. https://doi.org/10.1126/science.1099035

14. Bulut-Karslioglu, A., Perrera, V., Scaranaro, M., de la Rosa-Velazquez, I. A., van de Nobelen, S., Shukeir, N., Popow, J., Gerle, B., Opravil, S., Pagani, M., Meidhof, S., Brabletz, T., Manke, T., Lachner, M., & Jenuwein, T. (2012). A transcription factor-based mechanism for mouse heterochromatin formation. Nature structural & molecular biology, 19(10), 1023–1030. https://doi.org/10.1038/nsmb.2382

15. Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., & Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature, 410(6824), 120–124. https://doi.org/10.1038/35065138

16. Nakayama , J., Rice, J. C., Strahl, B. D., Allis, C. D., & Grewal, S. I. (2001). Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. Science (New York, N.Y.), 292(5514), 110–113. https://doi.org/10.1126/science.1060118

17. Dialynas, G. K., Vitalini, M. W., & Wallrath, L. L. (2008). Linking Heterochromatin Protein 1 (HP1) to cancer progression. Mutation research, 647(1-2), 13–20. https://doi.org/10.1016/j.mrfmmm.2008.09.007

18. Chang, C., Liu, J., He, W., Qu, M., Huang, X., Deng, Y., Shen, L., Zhao, X., Guo, H., Jiang, J., Fu, X. Y., Huang, R., Zhang, D., & Yan, J. (2018). A regulatory circuit HP1γ/miR-451a/c-Myc promotes prostate cancer progression. Oncogene, 37(4), 415–426. https://doi.org/10.1038/onc.2017.332

19. Chang, S. C., Lai, Y. C., Chen, Y. C., Wang, N. K., Wang, W. S., & Lai, J. I. (2018). CBX3/heterochromatin protein 1 gamma is significantly upregulated in patients with non-small cell lung cancer. Asia-Pacific journal of clinical oncology, 14(5), e283–e288. https://doi.org/10.1111/ajco.12820

20. Chen, L. Y., Cheng, C. S., Qu, C., Wang, P., Chen, H., Meng, Z. Q., & Chen, Z. (2018). CBX3 promotes proliferation and regulates glycolysis via suppressing FBP1 in pancreatic cancer. Biochemical and biophysical research communications, 500(3), 691–697. https://doi.org/10.1016/j.bbrc.2018.04.137

21. Liu, M., Huang, F., Zhang, D., Ju, J., Wu, X. B., Wang, Y., Wang, Y., Wu, Y., Nie, M., Li, Z., Ma, C., Chen, X., Zhou, J. Y., Tan, R., Yang, B. L., Zen, K., Zhang, C. Y., Chen, Y. G., & Zhao, Q. (2015). Heterochromatin protein HP1γ promotes colorectal cancer progression and is regulated by miR-30a. Cancer research, 75(21), 4593–4604. https://doi.org/10.1158/0008-5472.CAN-14-3735

22. Ma, C., Nie, X. G., Wang, Y. L., Liu, X. H., Liang, X., Zhou, Q. L., & Wu, D. P. (2019). CBX3 predicts an unfavorable prognosis and promotes tumorigenesis in osteosarcoma. Molecular medicine reports, 19(5), 4205–4212. https://doi.org/10.3892/mmr.2019.10104

23. Saksouk, N., Hajdari, S., Perez, Y., Pratlong, M., Barrachina, C., Graber, C., Grégoire, D., Zavoriti, A., Sarrazin, A., Pirot, N., Noël, J. Y., Khellaf, L., Fabbrizio, E., Julien, E., & Cammas, F. M. (2020). The mouse HP1 proteins are essential for preventing liver tumorigenesis. Oncogene, 39(13), 2676–2691. https://doi.org/10.1038/s41388-020-1177-8

24. Nishibuchi, G., & Nakayama, J. (2014). Biochemical and structural properties of heterochromatin protein 1: understanding its role in chromatin assembly. Journal of biochemistry, 156(1), 11–20. https://doi.org/10.1093/jb/mvu032

25. Canzio, D., Larson, A., & Narlikar, G. J. (2014). Mechanisms of functional promiscuity by HP1 proteins. Trends in cell biology, 24(6), 377–386. https://doi.org/10.1016/j.tcb.2014.01.002

26. Mendez, D. L., Kim, D., Chruszcz, M., Stephens, G. E., Minor, W., Khorasanizadeh, S., & Elgin, S. C. (2011). The HP1a disordered C terminus and chromo shadow domain cooperate to select target peptide partners. Chembiochem : a European journal of chemical biology, 12(7), 1084–1096. https://doi.org/10.1002/cbic.201000598

27. Brower-Toland, B., Findley, S. D., Jiang, L., Liu, L., Yin, H., Dus, M., Zhou, P., Elgin, S. C., & Lin, H. (2007). Drosophila PIWI associates with chromatin and interacts directly with HP1a. Genes & development, 21(18), 2300–2311. https://doi.org/10.1101/gad.1564307

28. Cheutin, T., McNairn, A. J., Jenuwein, T., Gilbert, D. M., Singh, P. B., & Misteli, T. (2003). Maintenance of stable heterochromatin domains by dynamic HP1 binding. Science (New York, N.Y.), 299(5607), 721–725. https://doi.org/10.1126/science.1078572

29. Darnell, J. E., Jr, Kerr, I. M., & Stark, G. R. (1994). Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. Science (New York, N.Y.), 264(5164), 1415–1421. https://doi.org/10.1126/science.8197455

30. Zhang, T., Kee, W. H., Seow, K. T., Fung, W., & Cao, X. (2000). The coiled-coil domain of Stat3 is essential for its SH2 domain-mediated receptor binding and subsequent activation induced by epidermal growth factor and interleukin-6. Molecular and cellular biology, 20(19), 7132–7139. https://doi.org/10.1128/MCB.20.19.7132-7139.2000

31. Rawlings, J. S., Rosler, K. M., & Harrison, D. A. (2004). The JAK/STAT signaling pathway. Journal of cell science, 117(Pt 8), 1281–1283. https://doi.org/10.1242/jcs.00963

32. Owen, K. L., Brockwell, N. K., & Parker, B. S. (2019). JAK-STAT Signaling: A Double-Edged Sword of Immune Regulation and Cancer Progression. Cancers, 11(12), 2002. https://doi.org/10.3390/cancers11122002

33. Harrison, D. A., Binari, R., Nahreini, T. S., Gilman, M., & Perrimon, N. (1995). Activation of a Drosophila Janus kinase (JAK) causes hematopoietic neoplasia and developmental defects. The EMBO journal, 14(12), 2857–2865.

34. Luo, H., Hanratty, W. P., & Dearolf, C. R. (1995). An amino acid substitution in the Drosophila hopTum-l Jak kinase causes leukemia-like hematopoietic defects. The EMBO journal, 14(7), 1412–1420.

35. Shi, S., Calhoun, H. C., Xia, F., Li, J., Le, L., & Li, W. X. (2006). JAK signaling globally counteracts heterochromatic gene silencing. Nature genetics, 38(9), 1071–1076. https://doi.org/10.1038/ng1860

36. Shi, S., Larson, K., Guo, D., Lim, S. J., Dutta, P., Yan, S. J., & Li, W. X. (2008). Drosophila STAT is required for directly maintaining HP1 localization and heterochromatin stability. Nature cell biology, 10(4), 489–496. https://doi.org/10.1038/ncb1713

37. The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.

38. Hu, X., Dutta, P., Tsurumi, A., Li, J., Wang, J., Land, H., & Li, W. X. (2013). Unphosphorylated STAT5A stabilizes heterochromatin and suppresses tumor growth. Proceedings of the National Academy of Sciences of the United States of America, 110(25), 10213–10218. https://doi.org/10.1073/pnas.1221243110

39. Keenen, M. M., Brown, D., Brennan, L. D., Renger, R., Khoo, H., Carlson, C. R., Huang, B., Grill, S. W., Narlikar, G. J., & Redding, S. (2021). HP1 proteins compact DNA into mechanically and positionally stable phase separated domains. eLife, 10, e64563. https://doi.org/10.7554/eLife.64563

40. Cooper A. (1999). Thermodynamic analysis of biomolecular interactions. Current opinion in chemical biology, 3(5), 557–563. https://doi.org/10.1016/s1367-5931(99)00008-3

41. Nkansah, E., Shah, R., Collie, G. W., Parkinson, G. N., Palmer, J., Rahman, K. M., Bui, T. T., Drake, A. F., Husby, J., Neidle, S., Zinzalla, G., Thurston, D. E., & Wilderspin, A. F. (2013). Observation of unphosphorylated STAT3 core protein binding to target dsDNA by PEMSA and X-ray crystallography. FEBS letters, 587(7), 833–839. https://doi.org/10.1016/j.febslet.2013.01.065

42. Raz, R., Durbin, J. E., & Levy, D. E. (1994). Acute phase response factor and additional members of the interferon-stimulated gene factor 3 family integrate diverse signals from cytokines, interferons, and growth factors. The Journal of biological chemistry, 269(39), 24391–24395.

43. Catlett-Falcone, R., Dalton, W. S., & Jove, R. (1999). STAT proteins as novel targets for cancer therapy. Signal transducer an activator of transcription. Current opinion in oncology, 11(6), 490–496. https://doi.org/10.1097/00001622-199911000-00010

44. Lin, T. S., Mahajan, S., & Frank, D. A. (2000). STAT signaling in the pathogenesis and treatment of leukemias. Oncogene, 19(21), 2496–2504. https://doi.org/10.1038/sj.onc.1203486

45. Liu, L., McBride, K. M., & Reich, N. C. (2005). STAT3 nuclear import is independent of tyrosine phosphorylation and mediated by importin-alpha3. Proceedings of the National Academy of Sciences of the United States of America, 102(23), 8150–8155. https://doi.org/10.1073/pnas.0501643102

46. Thiru, A., Nietlispach, D., Mott, H. R., Okuwaki, M., Lyon, D., Nielsen, P. R., Hirshberg, M., Verreault, A., Murzina, N. V., & Laue, E. D. (2004). Structural basis of HP1/PXVXL motif peptide interactions and HP1 localisation to heterochromatin. The EMBO journal, 23(3), 489–499. https://doi.org/10.1038/sj.emboj.7600088

47. de Vries, S. J., van Dijk, M., & Bonvin, A. M. (2010). The HADDOCK web server for data-driven biomolecular docking. Nature protocols, 5(5), 883–897. https://doi.org/10.1038/nprot.2010.32

48. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T., Rempfer, C., Bordoli, L., Lepore, R., & Schwede, T. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic acids research, 46(W1), W296–W303. https://doi.org/10.1093/nar/gky427

49. Lomberk, G., Wallrath, L., & Urrutia, R. (2006). The Heterochromatin Protein 1 family. Genome biology, 7(7), 228. https://doi.org/10.1186/gb-2006-7-7-228

50. Strom, A. R., Emelyanov, A. V., Mir, M., Fyodorov, D. V., Darzacq, X., & Karpen, G. H. (2017). Phase separation drives heterochromatin domain formation. Nature, 547(7662), 241–245. https://doi.org/10.1038/nature22989

51. Larson, A. G., Elnatan, D., Keenen, M. M., Trnka, M. J., Johnston, J. B., Burlingame, A. L., Agard, D. A., Redding, S., & Narlikar, G. J. (2017). Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. Nature, 547(7662), 236–240. https://doi.org/10.1038/nature22822

52. Sanulli, S., Trnka, M. J., Dharmarajan, V., Tibble, R. W., Pascal, B. D., Burlingame, A. L., Griffin, P. R., Gross, J. D., & Narlikar, G. J. (2019). HP1 reshapes nucleosome core to promote phase separation of heterochromatin. Nature, 575(7782), 390–394. https://doi.org/10.1038/s41586-019-1669-2

53. Keenen, M. M., Brown, D., Brennan, L. D., Renger, R., Khoo, H., Carlson, C. R., Huang, B., Grill, S. W., Narlikar, G. J., & Redding, S. (2021). HP1 proteins compact DNA into mechanically and positionally stable phase separated domains. eLife, 10, e64563. https://doi.org/10.7554/eLife.64563

54. Dutta, P., Zhang, L., Zhang, H., Peng, Q., Montgrain, P. R., Wang, Y., Song, Y., Li, J., & Li, W. X. (2020). Unphosphorylated STAT3 in heterochromatin formation and tumor

suppression in lung cancer. *BMC cancer*, *20*(1), 145. https://doi.org/10.1186/s12885-020-6649-2

55. Honorato, R. V., Koukos, P. I., Jiménez-García, B., Tsaregorodtsev, A., Verlato, M., Giachetti, A., Rosato, A., & Bonvin, A. (2021). Structural Biology in the Clouds: The WeNMR-EOSC Ecosystem. *Frontiers in molecular biosciences*, *8*, 729513. https://doi.org/10.3389/fmolb.2021.729513

56. van Zundert, G., Rodrigues, J., Trellet, M., Schmitz, C., Kastritis, P. L., Karaca, E., Melquiond, A., van Dijk, M., de Vries, S. J., & Bonvin, A. (2016). The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *Journal of molecular biology*, *428*(4), 720–725. https://doi.org/10.1016/j.jmb.2015.09.014

57. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. Nature, 596(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

58. Vangone A. and Bonvin A.M.J.J. "Contact-based prediction of binding affinity in protein-protein complexes", eLife, 4, e07454 (2015).

59. Xue L., Rodrigues J., Kastritis P., Bonvin A.M.J.J.*, Vangone A.*, "PRODIGY: a web-server for predicting the binding affinity in protein-protein complexes", Bioinformatics, doi:10.1093/bioinformatics/btw514 (2016).