Biologically relevant small variations of intra-cellular pH can have significant effect on stability of protein–DNA complexes, including the nucleosome

Alexey V. Onufriev[1,2,3]*

[1]Department of Physics, Virginia Tech, Blacksburg, Blacksburg, VA, United States, [2]Department of Computer Science, Virginia Tech, Blacksburg, Blacksburg, VA, United States, [3]Center from Soft Matter and Biological Physics, Virginia Tech, Blacksburg, VA, United States

Stability of a protein-ligand complex may be sensitive to pH of its environment. Here we explore, computationally, stability of a set of protein-nucleic acid complexes using fundamental thermodynamic linkage relationship. The nucleosome, as well as an essentially random selection of 20 protein complexes with DNA or RNA, are included in the analysis. An increase in intra-cellular/intra-nuclear pH destabilizes most complexes, including the nucleosome. We propose to quantify the effect by $\Delta\Delta G^{0.3}$—the change in the binding free energy due to pH increase of 0.3 units, corresponding to doubling of the H^+ activity; variations of pH of this amplitude can occur in living cells, including in the course of the cell cycle, and in cancer cells relative to normal ones. We suggest, based on relevant experimental findings, a threshold of biological significance of $\frac{1}{2}k_BT$ (~ 0.3 *kcal/mol*) for changes of stability of chromatin-related protein-DNA complexes: a change in the binding affinity above the threshold may have biological consequences. We find that for 70% of the examined complexes, $\Delta\Delta G^{0.3} > \frac{1}{2}k_BT$ (for 10%, $\Delta\Delta G^{0.3}$ is between 3 and 4 k_BT). Thus, small but relevant variations of intra-nuclear pH of 0.3 may have biological consequences for many protein-nucleic acid complexes. The binding affinity between the histone octamer and its DNA, which directly affects the DNA accessibility in the nucleosome, is predicted to be highly sensitive to intra-nuclear pH. A variation of 0.3 units results in $\Delta\Delta G^{0.3} \sim 10k_BT$ (~ 6 *kcal/mol*); for spontaneous unwrapping of 20 bp long entry/exit fragments of the nucleosomal DNA, $\Delta\Delta G^{0.3} = 2.2k_BT$; partial disassembly of the nucleosome into the tetrasome is characterized by $\Delta\Delta G^{0.3} = 5.2k_BT$. The predicted pH -induced modulations of the nucleosome stability are significant enough to suggest that they may have consequences relevant to the biological function of the nucleosome. Accessibility of the nucleosomal DNA is predicted to positively correlate with pH variations during the cell cycle; an increase in intra-cellular pH seen in cancer cells is predicted to lead to a more accessible nucleosomal DNA; a drop in pH associated with apoptosis is predicted to make nucleosomal DNA less accessible. We speculate that processes that depend on accessibility to the DNA in the nucleosomes, such as transcription or DNA replication, might become upregulated due to relatively small, but nevertheless realistic increases of intra-nuclear pH.

1 Introduction

The nucleus of a eukaryotic cell contains its DNA bound to various proteins, surrounded by an atmosphere of diverse counterions (Korolev et al., 2007). This giant protein-DNA complex, referred to as "chromatin (Misteli, 2007)," is central to many of the most important cellular processes such as cell differentiation, DNA replication, repair, transcription, and epigenetic inheritance, *i.e.,* inheritance that is not coded by the DNA sequence (Henikoff, 2008). The "hydrogen atom" of chromatin is the nucleosome (Woodcock, 1973; Kornberg, 1974; Olins and Olins, 1974)—the primary, fundamental level of the DNA compaction (Onufriev and Schiessel, 2019). The nucleosome core particle (Luger et al., 1997), (H2A·H2B)$_2$· (H3·H4)$_2$ · DNA, which we refer to as the nucleosome for simplicity, consists of 147 base pairs of DNA tightly wrapped around a protein core made of two copies of each of the four histone proteins H2A, H2B, H3, H4. Chromatin compaction at the nucleosome level is believed to be the most relevant to gene access, recognition (Misteli, 2007) and regulation (Kornberg and Lorch, 2020). The strength of the protein-DNA association in the nucleosome affects accessibility (Onufriev and Schiessel, 2019) of the nucleosomal DNA to cellular machines that need access to its information content; changes in the DNA accessibility can have direct biological consequences. For example, small increases in nucleosomal DNA accessibility (Onufriev and Schiessel, 2019) can lead to major increases in steady-state transcript levels (Zhu and Thiele, 1996) and promoter activity (Raveh-Sadka et al., 2012). The state of the nucleosome and chromatin, and thus access to its DNA, can be modulated in various ways. One example is reversible structural modifications to the histone proteins (Tessarz and Kouzarides, 2014) such as acetylation, in which normally positively charged Lysine groups become neutralized, generally leading to decreased histone-DNA binding affinity and the more accessible DNA (Fenley et al., 2010; Fenley et al., 2018). Another example is spontaneous and transient unwrapping of the nucleosomal DNA fragments at each end (Gansen et al., 2009; Koopmans et al., 2009; Tims et al., 2011), the corresponding lifetimes of these partially unwrapped states may be just long enough for functionally relevant access to DNA target sites located there (Tims et al., 2011). Another mechanism that can enhance access to the nucleosomal DNA is progressive disassembly of the histone octamer itself (Zlatanova et al., 2009; Andrews and Luger, 2011; Luger et al., 2012; Kato et al., 2017), which leads to the formation of partially assembled nucleosome structures (PANS), each lacking several histone proteins (Rychkov et al., 2017). For example, in the tetrasome, (H3·H4)$_2$ · DNA, about 78 bp of the DNA is accessible (Rychkov et al., 2017). The tetrasome is a key intermediate on the nucleosome assembly/disassembly pathway (Andrews and Luger, 2011), including during DNA replication (Zhang et al., 2020). One of the central open questions is exactly how DNA accessibility is controlled, at the level of the nucleosome? Many such mechanisms have been determined, while probably just as many remain to be uncovered. Could variations in pH be one of them?

A number of other protein-DNA complexes are just as vital to chromatin function, including transcription factors and chromatin remodeling complexes. Their specific, high enough, binding affinity for the DNA is critical for the function of chromatin. In general, modulation of the strength (free energy) of protein-DNA association may be expected to have biological consequences if these free energy changes are comparable to other relevant free energy scales in the system. The main question we address in this work is whether biologically relevant variations in ambient pH can bring about large enough changes in the protein-DNA binding free energy to be of potential relevance to chromatin function.

Many key processes in living cells are sensitive, *via* a diverse set of mechanisms, to the pH of the medium where the process takes place (Parker and Boron, 2013). In particular, macromolecular complexes are key players in a myriad of cellular functions: if properties of such a complex depend on pH, one can reasonably expect its function to be also affected by pH. Within a given cell, depending on the specific organelle, macromolecular complexes can be exposed (Chan et al., 2006; Garcia-Moreno, 2009; Kulichikhin et al., 2009) to a wide range of pH from about 5 to 8.

Multiple experimental and computational studies have demonstrated definitively that stability of protein-protein complexes tend to depend on pH, often over a wide range; the many associated nuances and biological consequences have been investigated and discussed in detail, see, e.g., Ref (Kundrotas and Alexov, 2006) and multiple references therein. In contrast, relatively less is firmly established with respect to pH dependence of protein-nucleic acid complexes, especially for those relevant to chromatin function. For the nucleosome, experiment demonstrated (Libertini and Small, 1984) that the structure becomes looser upon pH increase from about 7 to 8, consistent with the notion that the histone-DNA binding depends on pH. Unfortunately, we aren't aware of relevant measurements or predictions directly reporting *quantitative* data on pH dependence for the nucleosome; without these, it is hard to conclude whether the corresponding pH effects are biologically significant over the small pH variations expected within the nucleus, where the complex is found. For protein - nucleic acid complexes in general, an earlier theoretical work (Record et al., 1978) concluded that free energy of protein-DNA interactions should decrease with pH. We are aware of only a few relevant experimental data points (Senear and Ackers, 1990; Wessel et al., 1992; Peng and Acheson, 1998; Torigoe et al., 2003; Hamimes et al., 2006; Deegan et al., 2010; Moreau et al., 2012)—these show consistently that complex stability decreases with pH, however, most of the conclusions are qualitative, the strength of the reported effect varies substantially from study to study. Reported computational predictions also vary with respect to the significance of pH dependence or, equivalently, net proton uptake, in protein - nucleic acid binding. A fairly strong complex destabilization with increasing pH, consistent with experiment, was predicted for the λ-repressor (Misra et al., 1998). On the other hand, a recent large-scale computational study (Peng and Alexov, 2017) concluded that very little net proton uptake/release occurred upon binding between proteins and nucleic acids between pH = 5 to 8. We are unaware

of similar studies focusing specifically on chromatin-related protein-DNA complexes, or on the nucleosome and its partially assembled and partially unwrapped states. This work is intended to fill the gap, focusing on quantitative predictions suggestive of possible biological relevance.

Note that while intra-cellular and intra-nuclear pH is tightly controlled by the cell (Garcia-Moreno, 2009; Parker and Boron, 2013), variations do occur as part of the normal cell function, as well as in various pathologies such as cancer (Parker and Boron, 2013). Specifically, intra-cellular pH varies along the cell cycle (White et al., 2017), with a transient increase from 7.2 to 7.6. In mutant cells in which this increase in pH is attenuated, S phase is delayed, and G2/M transition is impaired (Putney and Barber, 2003), suggesting that pH variation may play a role in the cell cycle progression. Intra-cellular acidification by as much as 1 pH unit was observed to precede apoptosis, and was proposed as a possible effector of the apoptosis program (Gottlieb et al., 1995). In cancer cells, intra-cellular pH is increased by about 0.4 units compared to normal cells (while extra-cellular pH is decreased) (White et al., 2017). Some of these small changes in pH have been shown to correlate with vital cellular processes, likely involving chromatin components. For example, in higher eukaryotes, increases in the intra-cellular pH have been shown to correlate with cell proliferation and differentiation (Orij et al., 2011). In breast cancel cells, cell cycle progression is proposed to involve cell cycle phase-specific pH regulation (Flinck et al., 2018).

While a computational work such as this one can not firmly establish a causality link from a relevant pH variation to biological function, it can attempt to predict whether or not the associated pH induced changes in the binding affinity of at least some of chromatin-related protein-nucleic acid complexes are large enough to be potentially consequential, or whether these changes are non-existent or too small to consider seriously. If these changes are large enough, by a certain set of criteria that we set out to formulate, then predictions can be put forward, follow-up experiments can be suggested, and intriguing speculations can be made regarding possible effect of pH variations on biological function. This study aims to carry out the program just outlined.

We conclude this section by introducing some of the relevant terminology and concepts. Investigation of pH dependence of complex formation, either experimental or computational, is aided by the so-called linkage relationship: it follows from basic thermodynamics (Tanford, 1970) that whenever a net proton uptake/release occurs upon complex formation between the ligand and the receptor, the corresponding binding free energy, that is complex stability, depends on pH of the environment. The proton uptake/release is directly related to binding-induced changes (shifts) in pK^1 values of titratable (ionizable) groups in proteins and their ligands. Since micro-environment of many titratable groups changes upon complex formation, their pK values can be expected to change in response; the changes can be predicted by a variety of well-established computational methods (Bashford and Karplus, 1990; Yang et al., 1993; Antosiewicz et al., 1994; DelBuono et al., 1994; Demchuk and

Wade, 1996; Sham et al., 1997; Ullmann and Knapp, 1999; Nielson and Vriend, 2001; Georgescu et al., 2002; Mongan et al., 2004; Khandogin and Brooks, 2006; Bas et al., 2008; Kieseritzky and Knapp, 2008; Spassov and Yan, 2008; Song et al., 2009; Huang et al., 2018; Pahari et al., 2018). Indeed, pK shifts in protein-ligand binding, sometimes significant (Aguilar et al., 2010) ($|\Delta pK| > 1$), are well-documented, both experimentally and computationally, see, e.g., these reviews (Onufriev and Alexov, 2013; Petukh et al., 2013); several distinct physical mechanisms are behind these shifts, including electrostatic (Zhang et al., 2011) and "allosteric" (Aguilar et al., 2010). Importantly, pK shifts induced by ligand binding are only necessary, but not sufficient (Onufriev and Alexov, 2013) for net proton release/uptake required for pH dependence of the binding affinity, which makes the question of whether or not a noticeable pH dependence exists a more subtle one than simply determining if meaningful pK shifts occur upon binding.

2 Material and methods

We begin, Section. 2.1, with a description of the over-all strategy for selecting the atomistic structures used in this work, and their initial preparation. In the following subsections, 2.1.1, 2.1.2, 2.1.3, we present detailed descriptions of each set of structures, starting from smaller protein - nucleic acid complexes, followed by the nucleosome and its partially unwrapped and partially assembled states. Next, in Section. 2.2, we describe the computational protocol employed to estimate pK s (titration curves) of the relevant titratable groups, and the pH dependence of the net charge of each complex. The results of these calculations are used to estimate pH dependence of the complex stability, as described in the final Section. 2.3.

2.1 Structure selection and initial preparation

When available, separate, independent experimental structures for the complex and the unligated (apo) protein were used. This protocol, which explicitly accounts for binding-induced structural re-arrangements in the protein, was applied to the 20 protein-nucleic acid complexes described in Section. 2.1.1. When the corresponding unligated structure of the protein was unavailable, an alternative protocol was used: the unligated protein was constructed from the structure of the complex, by manually separating the nucleic acid from the protein in the PDB file. This protocol was applied to all other complexes described in this work. Limitations of the latter approximation, which does not account for binding-induced structural re-arrangements, are well recognized (Alexov et al., 2011); however, alternatives such as constant pH MD (Wallace and Shen, 2009; Swails et al., 2014), which could mitigate the issue, are still too expensive for large complexes such as the nucleosome. Unless otherwise stated, the structures used as input for pK calculations, Section 2.2 below, were not manipulated beyond the steps described above.

2.1.1 20 protein-nucleic acid complexes

We have employed the same set of 20 protein-nucleic acid complexes that were previously used by us for an analysis of pK shifts in protein-ligand binding in Ref. (Aguilar et al., 2010); the

---

1 Here we do not make a distinction between pK_a and $pK_{1/2}$ (mid-point of titration curve), and denote both by pK, see, e.g., Ref (Onufriev et al., 2001). for an in-depth discussion of this issue.
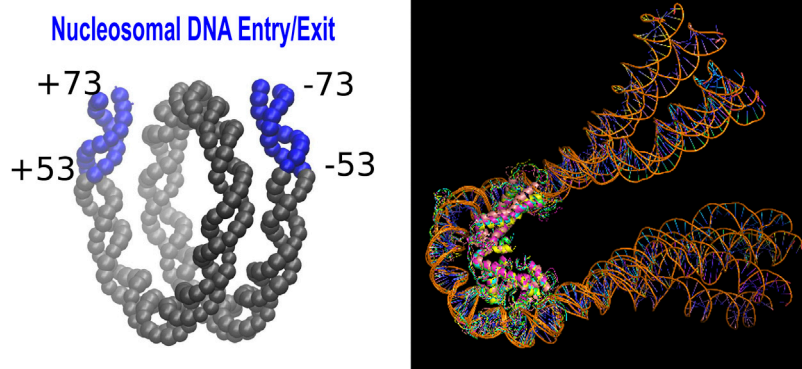
FIGURE 1
Left: A schematic showing the Entry/Exit segments (20 base pairs at each end) of the nucleosomal DNA (Fenley et al., 2018). This region becomes exposed in the partially unwrapped state of the nucleosome. Right: The structural ensemble representing (Rychkov et al., 2017) the tetrasome.

detailed selection criteria, structure preparation procedures can be found *ibid*, a brief description is below. The PDB IDs are given in Supplementary Table S1 of the Supplementary Information.

The structures were selected out of 1932 entries available at the time (Aguilar et al., 2010) in NPIDB database (Spirin et al., 2007) of protein-nucleic acid complexes. Since accuracy of pK calculation and protonation state assignment depend critically on the quality of the input structure, the 20 highest quality structures were chosen that satisfy the following three criteria: 1) no missing residues; 2) 2.5 Å or better resolution; 3) availability of high quality structures for *both* the complex and the unligated protein, as separate PDB entries. These strict selection criteria explain the relatively small size of the set. It also makes the selection essentially random (20 out of near 2000).

The selected complexes represent at least 14 different functional classes; the ligand is DNA in 10 cases, RNA in 9, and in one case the ligand is an RNA/DNA hybrid, see Supplementary Table S1 of the Supplementary Information. The proteins in the complexes vary widely with respect to the number of titratable groups: from 18 to 264.

2.1.2 The lambda repressor complex

The 1.8 Å resolution structure of the λ repressor complex was used, PDB ID: 1LMB. To the best of our knowledge, no high resolution atomistic structure of the corresponding unligated protein is available for this complex, therefore the unligated structure was prepared manually, as described in Section 2.1. The relevant pH -related experimental data points are from Table 3 of Ref. (Senear and Ackers, 1990), corresponding to O_R1; the specific choice is based on the argument presented in Ref. (Misra et al., 1998). The pH dependence of the λ repressor stability is particularly well characterized experimentally, making it a unique reference for a computational study such as this one.

2.1.3 The nucleosome and its partially unwrapped/assembled states
2.1.3.1 The canonical nucleosome

The following PDB structures were used: 5B0Z 1.99 Å (*homo sapiens*), 1AOI, 2.8 Å (*xenopus laevis*), 1KX5 1.9 Å (chimeric). To

the best of our knowledge, no atomistic structures of the "bare" histone core is available, therefore the unligated structures of the nucleosome, as well as of its derivatives described below, were prepared manually, by removing the appropriate DNA fragments from the complex, Section 2.1.

2.1.3.2 Partially unwrapped state of the nucleosome

The partially unwrapped state of the nucleosome, PW(20.20), in which 20 bp on each end "peel off" the histone core was approximated as follows. We removed the corresponding DNA fragments, blue regions in Figure 1 (left), from 1KX5 structure, keeping the DNA in positions from −52 to +52. The computed $\Delta\Delta G^{0.3}$ characterizes the transition $PW(20.20) + 40bp \rightarrow nucleosome$.

2.1.3.3 The tetrasome

No experimental structure of the tetrasome, $(H3 \cdot H4)_2 \cdot DNA$, is available, so we used the one obtained previously by atomistic modeling guided by AFM experiment (Rychkov et al., 2017). To account for non-negligible conformational variability of the tetrasome, we employed models 1,2,3,9, and 10 from Ref (Rychkov et al., 2017).—snapshots from an MD simulation—to represent the tetrasome, Figure 1 (right). These models utilized 1KX5 nucleosome structure (Rychkov et al., 2017) as the basis.

The total charge of the tetrasome complex is the average of the charges obtained for each of the models. The $(H2A \cdot H2B)_2$ dimer structure was taken from 1KX5 nucleosome structure. The transition considered here is $(H3 \cdot H4)_2 \cdot DNA + (H2A \cdot H2B)_2 \rightarrow nucleosome$.

2.2 Calculation of pK values and pH -dependence of net change of the structure

We used H++ server (Gordon et al., 2005; Anandakrishnan et al., 2012), http://biophysics.cs.vt.edu/H++, to prepare the input structures, including assignment of protons. Minimal pre-

processing, detailed in Refs. (Gordon et al., 2005; Anandakrishnan et al., 2012), is performed on the input structures; this step includes assignment of missing heavy atoms within existing residues, and initial assignment of missing protons. No missing residues are added: structures with missing residues in the middle of the chain were not considered. The same H++ server was used to compute the pK (p$K_{1/2}$) values and titration curves of all ionizable residues in the proteins. Version 4.0 was used for individual calculations, except for the 20 protein-nucleic acid complexes for which an in-house (batch) script based on version 3.0 was employed. The calculations were based on the standard continuum electrostatics methodology (Bashford and Karplus, 1990) as implemented in H++ server, proper Boltzmann averages over all protonation states were computed (Beroza et al., 1991; Myers et al., 2006). As is common (Senear and Ackers, 1990; Misra et al., 1998), nucleic acid groups were assumed to have relatively low pK values, and therefore treated as non-titratable at neutral pH of interest; a recent study (Peng and Alexov, 2017) showed that their computed pK values are shifted down upon protein binding, providing further justification for excluding these groups from our calculations that focus on neutral pH. Unless otherwise stated, the protein and its ligand were treated as a low dielectric medium with the default value for the internal dielectric constant $\epsilon_{in} = 10$, while the surrounding solvent was assigned a high dielectric constant $\epsilon_{out} = 80$. The electrostatic screening effects of (monovalent) salt enter *via* the Debye-Huckel screening parameter $\kappa = 0.128\text{Å}^{-1}$ which roughly corresponds to a physiological concentration of [NaCl] = 0.15M. Among the input parameters of these calculations, arguably the largest uncertainty is in the choice of ϵ_{in}, its effect is discussed in Section 3.2. The dependence of pK values in proteins on the solvent salt concentration in the physiological range is known to be relatively weak (Kao et al., 2000), unlikely to bring about changes in the net proton uptake comparable to those due to changes in pH considered here. In addition to the specific protein- (nucleic acid) experimental data points discussed in detail in "Results," the methodology used in this work was tested on a larger set of experimental pK shifts in protein-protein and protein-small molecule complexes (Aguilar et al., 2010).

2.3 Calculation of the pH dependence of the complex stability, ΔG(pH)

Once the titration curve of the protein was obtained, and the charge of the structure computed (`charge.dat` in H++), a general thermodynamic linkage relationship (Tanford, 1970) was be used to estimate the pH -dependent correction to the binding energy:

$$\Delta G(\text{pH}) = \Delta G(\text{pH}_{ref}) + 2.3k_BT \int_{\text{pH}_{ref}}^{\text{pH}} \Delta Q(\text{pH})d\text{pH} \qquad (1)$$

Where ΔQ(pH) is the difference between the net charge of the complex and the corresponding free protein and the ligand at the given pH, $\Delta Q(\text{pH}) = Q_{complex}(\text{pH}) - Q_{unbound}(\text{pH}) = Q_{complex}(\text{pH}) - Q_{protein}(\text{pH}) - Q_{ligand}(\text{pH})$. The reference pH$_{ref}$ can be chosen arbitrarily, and is specified in the text if relevant. We use $k_BT = 0.59\ kcal/mol$ where appropriate.

Another convenient form of the linkage relationship in Eq. 1 is

$$\frac{\partial \Delta G}{\partial \text{pH}} = 2.3k_BT\Delta Q(\text{pH}), \qquad (2)$$

$\frac{\partial \Delta G}{\partial \text{pH}} > 0$ means that the complex is destabilized with increasing pH, which obviously requires that the corresponding $\Delta Q > 0$.

3 Results and discussion

The main objective of this work is to determine whether biologically relevant variations of pH of the environment may cause potentially biologically relevant changes in thermodynamic stability of protein - (nucleic acid) complexes relevant to chromatin function. Our focus is only on ΔG(pH), and its possible biological significance, for chromatin -specific complexes, many other aspects of protein-nucleic acid binding (Luscombe et al., 2001; Lejeune et al., 2005; Rohs et al., 2009; Peng and Alexov, 2017) aren't explored here. Due to the relative paucity of chromatin specific complexes such as chromatin remodeling and transcription factors among protein - nucleic complex that pass our selection criteria for suitability for pK calculations, see Methods, our strategy is as follows. To be able to make statistically significant statements we broaden the pool of complexes to include not only chromatin specific complexes, but several other classes as well. Our justification is that the underlying physics of pH sensitivity of protein-ligand complexes stability is universal (Kundrotas and Alexov, 2006; Peng and Alexov, 2017), and thus our general conclusions made for a diverse set of complexes can serve as a reasonable approximation to the specific class of interest. We recognize that the nucleosome is a special case due to its unique and unusual structure, this complex is investigated separately.

For small variations ΔpH, that is to the first order, it follows from Eq. 2 that the corresponding change of the complex stability:

$$\Delta\Delta G \approx 2.3k_BT\Delta Q(\text{pH})\Delta\text{pH} \qquad (3)$$

As seen from Eq. 3, the complex stability change is directly proportional to the pH variation itself, ΔpH; it also depends on the pH of the environment *via* $\Delta Q(\text{pH}) = Q_{complex}(\text{pH}) - Q_{unbound}(\text{pH})$. To utilize Eq. 3 for suggesting whether each calculated ΔQ(pH) may have biological consequences, we need three additional components: 1) the relevant pH of the environment; 2) ΔpH of biologically relevance; 3) a threshold of biological significance for the corresponding $\Delta\Delta G$. Below is our reasoning for each of the three components.

1) One can easily argue that for chromatin components of eukaryotes, including the nucleosome and transcription factors, the most relevant pH is that inside the nucleus where the genetic material resides. Thus, unless otherwise stated, we assume pH = 7.4, which is likely close, to within a few decimal points, to mean intra-nuclear pH in higher organisms (Casey et al., 2009). Note that a more precise value of intra-nuclear pH is unimportant for us here because contribution to $\Delta\Delta G$ due to variation of ΔQ(pH) as a function of pH would appear only in the second order in ΔpH, and thus can be neglected in Eq. 3 for small enough ΔpH. We also assume that intra-nuclear pH ≈ intra-cellular pH, based on the common argument (Casey et al., 2009) that nuclear pores are large

enough to allow virtually unimpeded passage of protons between the two compartments.

2) A more difficult issue is picking the specific ΔpH value most relevant to chromatin components, which should characterize typical intra-nuclear pH changes seen in living cells. We propose that the relevant ΔpH should be consistent with, and reflect the general magnitude of: natural cell-to-cell variations (Orij et al., 2012) of pH, its variation in response to external stimuli (Santos et al., 2016), shifts in intra-cellular pH seen in cancer cells (White et al., 2017) relative to normal ones, as well as temporal variations of intra-nuclear pH, e.g., related to the cell cycle progression (da Veiga Moreira et al., 2015). Based on the specific values of pH changes reported in these examples, we conservatively set the relevant ΔpH = 0.3, corresponding to a factor of 2 ($log_{10}2 \approx 0.3$) change in the proton activity of the solution. Note that we have deliberately excluded from the list the relatively large change in pH seen in apoptosis (Gottlieb et al., 1995), as we aim for the relevant ΔpH to be conservative and maximally inclusive. Once the relevant ΔpH is chosen, the corresponding complex stability change from Eq. 3 is $\Delta\Delta G_{pH}^{0.3} = 2.3 k_B T \Delta Q (pH)^* 0.3$. Since our main interest is pH = 7.4, we will simplify the notation by dropping the lower index: $\Delta\Delta G^{0.3} = \Delta\Delta G_{pH=7.4}^{0.3}$.

3) Finally, we have to decide on the threshold of biological significance for ΔΔG[0.3] in the context of chromatin function. One possible criterion (Fenley et al., 2018) is that, to be of potential biological significance, ΔΔG[0.3] should be comparable with the free energy change characterizing some of the functionally relevant transitions in the given protein-nucleic acid complex. To proceed, we focus on chromatin associated protein-nucleic acid complexes. We need to decide on the smallest ΔΔG[0.3] that can typically lead to meaningful biological consequences at the cellular level. One could make the usual physicist's argument that ΔΔG[0.3] values smaller than the thermal noise $\sim k_B T$ (~ 0.6 kcal/mol) should be inconsequential. However appealing the "universal" $1 k_B T$ threshold may appear, we argue that a more biologically motivated significance threshold for ΔΔG[0.3] should be lower: $\frac{1}{2} k_B T$, corresponding to $\sqrt{e} \approx 1.6$ fold change in the dissociation constant. The significance of the dissociation constant change of this magnitude can be argued for based on the fact that about 1.5 change in the transcription factor residence time can lead to a $\sim$ 3-fold change in the transcription rate (mRNA molecules of the reporter gene) (Popp et al., 2021), which is significant. A similar argument is presented below in the context of the nucleosomal DNA accessibility. In what follows we assume the $\frac{1}{2} k_B T$ threshold of biological significance for predicted changes in the free energy of binding. To conclude this important part, we present a different argument for why the threshold for biological significance of ΔΔG[0.3] may be fairly low. Let's compare the proposed threshold to experimental binding affinities of chromatin-related complexes. A recent analysis of 83 protein-DNA complexes (Peng et al., 2021) revealed a broad distribution of experimentally measured complex binding affinities was observed, with the most likely value at $\sim 15 k_B T$; the same study showed that the stability of *some* of them is less than 10 $k_B T$. Importantly, the experimental binding affinities referred to above are reported at the standard conditions, which implies 1M concentration. In reality, concentration of transcription factors *in vivo* can be (Milo and Phillips, 2015) as low as nM to μM, which would reduce their relevant binding
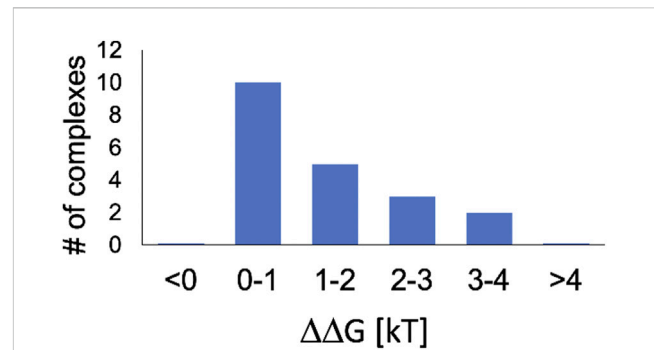

FIGURE 2
Most protein - nucleic acid complexes in the pH environment relevant to the nucleus are destabilized by increasing pH. Shown in the computed distribution of the change in complex stability, ΔΔG[0.3], due to pH increase by 0.3 units. Within the $[0, 1]k_B T$ interval, ΔΔG[0.3] is greater than the threshold of biological significance, $\frac{1}{2} k_B T$, for 4 out of 10 complexes.

affinities by 12 to 8 $k_B T$ from their standard state values, meaning that ΔΔG[0.3] of even a fraction $k_B T$ may affect the corresponding biological function of the protein-nucleic acid complex.

3.1 Twenty protein-nucleic acid complexes

Our main result for twenty, essentially random (see Methods for important details), protein-nucleic acid complexes is summarized in Figure 2. For biologically relevant variation of pH of 0.3 units, the corresponding change in the complex stability, ΔΔG[0.3], varies widely, but for all of them ΔΔG[0.3] ≥ 0. That is at pH = 7.4 relevant to typical intra-cellular conditions, all of the complexes are either unaffected or destabilized by increasing pH, meaning that the affinity between the protein and the nucleic acid decreases at higher pH; conversely, a drop in pH leads to stabilization of the complex. Moreover, we predict $\frac{\partial \Delta G}{\partial pH} \geq 0$ for all of the twenty complexes in the range from pH = 5 to 8, with $\frac{\partial \Delta G}{\partial pH} \approx 0$ for a minority of the complexes over only a portion of the range (results not shown). An illustrative example of this behavior for a transcription factor is discussed below, Figure 3.

Accepting the $\frac{1}{2} k_B T$ biological significance threshold suggested above, we find that 30% of the complexes fall below it. On the other hand, complex stability changes larger than $\frac{1}{2} k_B T$ occur for 70%, and changes larger than $k_B T$ for 50% of the complexes; for 10% of them the effect is between 3 and 4 $k_B T$. We argue that some of these effects must have biological consequences.

As a specific example of our reasoning, consider the methionine repressor MetJ (PDB IDs 1MJM, 1MJK) from our set of twenty protein-nucleic acid complexes. This is a transcription factor, which represses the transcription of genes involved in methionine biosynthesis by binding to certain DNA sequences (Garvie and Phillips, 2000). In the cell, sensitivity of MetJ for the DNA is enhanced by binding of a specific co-repressor, resulting in a 10-fold increase in the binding constant of MetJ to the DNA (Saint-Girons et al., 1986). The
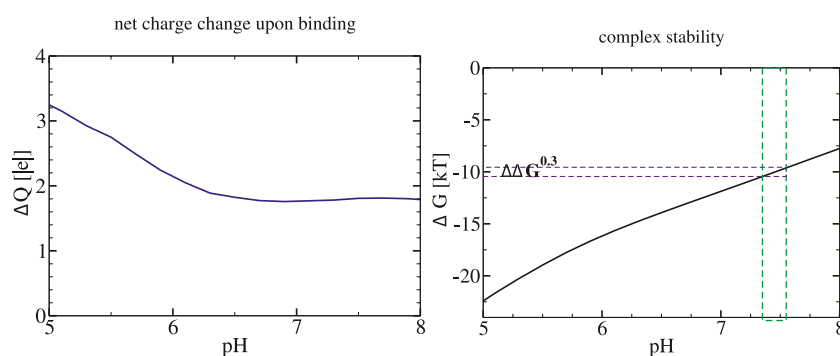
FIGURE 3
Stability of Methionine Repressor (MetJ) transcription factor complex with its DNA is predicted to decrease with pH of the environment. Left: pH dependence of the net charge change (proton uptake) upon complex formation. Right: Stability of the protein-DNA complex. The ΔG(pH) is estimated *via* Eq. 1 with ΔG(pH$_{ref}$)=9.8$k_B T$ (−5.8 kcal/mol (Hyre and Spicer, 1995)) at pH$_{ref}$ =7.4. The vertical green bar around pH$_{ref}$ indicates the proposed magnitude of biologically relevant variation of intra-nuclear pH; the corresponding $\Delta\Delta G^{0.3}$is also indicated.

computed proton uptake upon MetJ complex formation ΔQ(pH = 7.4) = 1.9, yielding $\Delta\Delta G^{0.3}$ = 1.3$k_B T$ (0.77 kcal/mol). Therefore, a pH increase of 0.3, which reduces the intrinsic binding constant of MetJ to the DNA by a factor of 3.6, can lower the beneficial effect of the co-repressor binding significantly.

Relatively few experimental data points on pH dependence of protein-nucleic acid binding free energy are available, but every single one that we have identified so far corroborates our over-all conclusion that protein-nucleic acid complexes are destabilized with increasing pH. The experimental data points are briefly discussed below; direct comparison with the predicted $\Delta\Delta G^{0.3}$ is made where possible. In the list below, quantitative findings are discussed first, followed by more qualitative statements.

1. For the λ-repressor, the experimental ΔG(pH) curve (Senear and Ackers, 1990), shows progressive complex destabilization, $\frac{\partial \Delta G}{\partial \text{pH}} > 0$, over the entire "biologically relevant" range from pH = 5.0 to 8.0. Our prediction for the λ-repressor is the same, $\frac{\partial \Delta G}{\partial \text{pH}} > 0$ over the same range of pH. Quantitatively, we predict $\Delta\Delta G^{0.3}$ = 1.1 kT for the λ-repressor, in reasonable agreement with the experimental (Senear and Ackers, 1990; Misra et al., 1998) $\Delta\Delta G^{0.3}$ = 0.77 ± 0.2kT, inferred from the experimental ΔQ(pH = 7.5) = 1 ± 0.2.

2. It was shown (Deegan et al., 2010) that the binding affinity of a specific transcription factor (estrogen nuclear receptor) decreases steadily with pH, the sharpest decrease occurring at around pH = 7.4; the over-all drop in the binding free energy is about 1.4$k_B T$ over the range from pH = 7 to 8. From the reported experimental ΔG(pH) curve we estimate $\Delta\Delta G^{0.3} \approx 0.5 k_B T$, which means that a biologically relevant variation in intra-cellular pH can change the receptor occupancy of the DNA binding site by a factor of ~ 1.6. By our own criterion, this effect is borderline significant for biological consequences; this complex was proposed to act as a pH sensor (Deegan et al., 2010). We did not perform pK and $\Delta\Delta G^{0.3}$ calculations for this complex because the available X-ray structure did not satisfy a single selection criterion used in this work, see "Methods."

The five experimental data points below all confirm our general conclusion that protein-nucleic acid complex stability decreases with increasing pH.

- The binding activity of Sp1 protein (member of zinc finger family) to the CG box DNA linearly increased, 7-fold, as the pH was lowered from pH = 8.0 to 6.0. (Torigoe et al., 2003).
- A steady and significant decrease of the protein-DNA complex stability with increasing pH from 5 to 8 was observed experimentally for the complex of DNA with a GFP-bound protein (Moreau et al., 2012).
- Binding efficiency of vertebrate protein RDM1 for a 52-mer oligonucleotide gradually decreased as the pH was increased from 6.4, reaching near zero at pH = 7.0 (Hamimes et al., 2006).
- T-antigen (a specific DNA-binding protein from simian virus) binds with high affinity to unspecific DNA at pH = 6, with the affinity decreasing substantially at pH = 7.5 (Wessel et al., 1992).
- DNA-binding activity of polyomavirus large T-antigen decreases steadily from pH = 6.0 to 8.0 (Peng and Acheson, 1998).

From the linkage relationship, Eq. 3, the destabilization of protein-nucleic acid complexes with increasing pH occurs when the complex formation is followed by a net proton uptake, $\Delta Q = Q(complex) - Q(unbound) > 0$. The physical origin of the net uptake is the attraction between the oppositely charged protons and the nucleic acid ligand, which favors protonated states of titratable groups of the bound protein at the given pH, or, equivalently, an increase in pK of these groups: pK (complex) > pK (unbound). Thus, for example, a negatively charged GLU or ASP is likely to pick up a proton from the solution and become neutral upon complex formation, while a LYS is more likely to be found positively charged in the complex then in the free protein. Essentially the same qualitative conclusion was made previously (Peng and Alexov, 2017), based on more than 300 protein-nucleic acid complexes. For the twenty complexes examined here, pK (complex) > pK

(unbound) for the overwhelming majority, 95%, of the titratable groups, whether acidic or basic. However, in apparent contrast to Ref. (Peng and Alexov, 2017), we find a fairly substantial number of titratable groups from at least 2 to as many as 30, in each complex (an average of ∼ 12% of the total number of titratable groups per complex), for which the computed pK shifts are appreciable, pK (*complex*) – pK (*unbound*) > 1. We suggest[2] that it is these larger pK shifts that explain the non-negligible $\Delta Q(pH = 7.4) > 0$ seen for most of the complexes explored in this work, leading to potentially biologically relevant effect on the complex stability predicted for some of the complexes. We note several methodological differences with Ref. (Peng and Alexov, 2017), which concluded that $\Delta Q(pH = 7.4) \approx 0$ for many complexes. In particular, we used separately determined X-ray structures for the complex and the free (unligated) protein, where available. This computational procedure accounts for binding-induced structural re-arrangements (Ellis et al., 2016) in the protein, see "Methods" and Ref. (Aguilar et al., 2010), where it was demonstrated that substantial pK shifts were more pronounced when these structural re-arrangements were included in the computational protocol.

3.2 The nucleosome

Our first set of results in this section, Table 1, characterizes the predicted pH dependence of the nucleosome complex formation, that is the transition *DNA + histone octamer → nucleosome*.

Consistent with what we have seen for other protein-nucleic complexes, the nucleosome is destabilized by increasing pH. Conversely, a drop in intra-nuclear pH is predicted to have a stabilizing effect. Notably, the magnitude of the effect for ΔpH = 0.3, $\Delta\Delta G^{0.3} \sim 10k_BT$ (6 kcal/mol), is considerably larger than that seen in all other protein-nucleic acid complexes examined so far, which probably reflects the key role of electrostatic interactions in the structure and function[13, 77 (Kunze and Netz, 2000; Beard and Schlick, 2001; Kunze and Netz, 2002; Schiessel, 2003),] of this complex, as well as the sheer length of its DNA—147 bp. The relatively large effect appears robust to the structure employed for the computation: the variation of the predicted $\Delta\Delta G^{0.3}$ between the different structures is within $2k_BT$, or about 20%, for the default computational protocol, Table 1. Equally importantly, the estimate is robust to the key parameter[47, 48, 50] of the computational protocol—the internal dielectric constant. In fact, we can take the average difference between the $\Delta\Delta G^{0.3}$ in the 2nd and 3rd rows of Table 1 to arrive at a useful "error range" (Fenley et al., 2018) of ∼ 40 % on the computed $\Delta\Delta G^{0.3}$ for the nucleosome. In what follows we use $\Delta\Delta G^{0.3} \sim 10k_BT$, which corresponds to the bottom of that range.

Experimentally (Libertini and Small, 1984), the nucleosome core particle becomes looser as the pH of the environment is increased from 7 to 8, which can be seen as a qualitative confirmation of our

TABLE 1 The effect of intra-nuclear pH variation of 0.3 units on the nucleosome stability can be substantial. Shown are the computed values of the reduction of the binding affinity between the DNA and the histone octamer due to pH increase by 0.3 units. The change is the binding affinity is expressed as $\Delta\Delta G^{0.3}$ in units of k_BT. The predicted relatively strong effect is robust to the nucleosome structure (species) as well as the key parameter of the computation—the internal dielectric constant ϵ_{in}. The *default* refers to ϵ_{in} =10 commonly used, while ϵ_{in} =4 is the lower bound on ϵ_{in} in proteins (Simonson, 2013).

Methodology	Organism		
	chimera	human	*xenopus*
default	10.2	9.2	8.6
ϵ_{in} = 4	16.4	13.0	11.1

result above. To the best of our knowledge, the free energy (stability) of the nucleosome complex as a function of pH has not been measured quantitatively, possibly because of the associated experimental difficulties (Thåström et al., 2004). We argue that the effect on the nucleosome stability of intra-cellular pH variations should be biologically relevant. First, the value of $\Delta\Delta G^{0.3} \sim 10k_BT$ for the nucleosome is not negligible compared even to the high stability of the whole complex, which is about $64k_BT$ (38 kcal/mol) at *in-vivo* conditions (Fenley et al., 2010). An arguably more informative comparison of the $\Delta\Delta G^{0.3}$ is with changes of the binding affinity between the DNA and the histone octamer that occur due to various biologically relevant alterations of the nucleosome structure, which ultimately alter accessibility of its DNA. Note that a $\frac{1}{2}k_BT$ change in the protein-DNA binding affinity is enough to bring about a factor of 1.6 change in the probability of the DNA to be in the bound state. A change in the genomic DNA accessibility as small as a factor of 1.5, can have direct functional consequences such as transcription modulation, see, e.g., a discussion in Ref. (Fenley et al., 2018). In particular, a number of charge-altering post-translational modifications (Tessarz and Kouzarides, 2014), or PTMs, directly modulate the strength of association between the histone octamer and nucleosomal DNA (Manohar et al., 2009; Materese et al., 2009; Andrews et al., 2010; Bowman and Poirier, 2015; Brehove et al., 2015). For example, acetylation of H3K56 in the globular core of the histone octamer destabilizes the nucleosome by $\Delta\Delta G = 3.4k_BT$ (Andrews et al., 2010); this PTM, which increases DNA accessibility, was shown to increase transcription rates (Tessarz and Kouzarides, 2014). Thus, one can imagine a situation in which a drop in intra-nuclear pH of 0.3 units counterbalances the destabilizing effect of H3K56 thus reducing or completely reversing its effect on nucleosomal DNA accessibility and its functional consequences. In fact, most of lysine acetylations in the globular histone core, some of which are functionally relevant, are characterized (Fenley et al., 2018) by a decrease in DNA binding affinity of less than the nucleosomal $\Delta\Delta G^{0.3}$, which means that their effect on the genomic DNA accessibility can potentially be strongly modulated by biologically relevant intra-nuclear pH variations.

3.2.1 Partial DNA unwrapping in the entry/exit region

Access to nucleosomal DNA can also be facilitated by spontaneous unwrapping of its end fragments (Tims et al., 2011),

---

2   For a single titratable site that experiences a small pK shift |ΔpK| ≪ 1, the maximum proton uptake $\Delta Q \approx 0.57\Delta$p$K$, at pH = p$K$. Therefore, if all of the pK shifts in the protein are small, no appreciable net proton uptake can be expected. Conversely, if at least some ΔpK >1 at pH ≈ pK, these sites will experience net proton uptake $\Delta Q \sim 1$, potentially leading to noticeable pH dependence of the complex stability.

which can "peel off" transiently (Armeev et al., 2021). Here we have calculated $\Delta\Delta G^{0.3}$ that characterizes pH dependence of transient unwrapping 20 bp fragments at each end, see Fig. 1 (left) in "Methods." The transient opening of this so-called "entry/exit" region (Fenley et al., 2018) can facilitate "invasion" of the nucleosome by key factors involved in initiating processes such as transcription (Tims et al., 2011). Our estimate is $\Delta\Delta G^{0.3} = 2.2 k_B T$ for the entry/exit region sensitivity to pH. This magnitude of the $\Delta\Delta G^{0.3}$ is significantly higher than the proposed general threshold of biological significance, $\frac{1}{2} k_B T$, corresponding to 1.6-fold change in the nucleosomal DNA accessibility (probability to be in the open state) (Fenley et al., 2018). Let's explore more specific comparisons. The predicted $\Delta\Delta G^{0.3}$ for the entry/exit region corresponds to a 9-fold change in the nucleosomal DNA accessibility, which is much larger than the threshold of 1.5-fold change of biological significance proposed earlier (Fenley et al., 2018) for the entry/exit DNA accessibility change (Fenley et al., 2018). In particular, acetylation of H4K31, known to be associated with an order of magnitude increase in steady-state transcript levels (Zhu and Thiele, 1996), and promoter activity (Raveh-Sadka et al., 2012), leads to an estimated 1.5- fold change in the entry/exit DNA accessibility (Fenley et al., 2018). Based on the above, we make a specific prediction for most lysine acetylations that affect accessibility of the nucleosomal DNA in the entry/exit region: their effect can be significantly modulated by biologically relevant pH variations. These variations may negate, or significantly enhance, the effect of a specific acetylation on the DNA accessibility.

To conclude this subsection, let's compare the free energy cost of unwrapping the entry/exit region to the corresponding $\Delta\Delta G^{0.3}$. The cost of unwrapping a ~ 10 bp long DNA fragment at each end is ~ $1.7 k_B T$ (Koopmans et al., 2009; Wei et al., 2015), increasing roughly proportionally to the fragment length (Blossey and Schiessel, 2011; Culkin et al., 2017). Thus, the cost of unwrapping two 20 bp fragments is ~ $6.8 k_B T$; our predicted pH effect, $\Delta\Delta G^{0.3} = 2.2 k_B T$, is 1/3 of the total unwrapping cost, which is significant, further supporting the claim of the significance of the biologically relevant pH variations on the DNA accessibility in the entry/exit region of the nucleosome.

3.2.2 Partially assembled states of the nucleosome

Yet another mechanism that facilitates access to the nucleosomal DNA is progressive disassembly of the histone octamer itself (Andrews and Luger, 2011; Luger et al., 2012), leading to the formation of partially assembled nucleosome structures, PANS (Rychkov et al., 2017). Among these, the tetrasome, $(H3 \cdot H4)_2 \cdot DNA$, Figure 1 (right), is believed to be highly relevant. Experimentally reported (Andrews et al., 2010) stability of the nucleosome with respect to this specific disassembly pathway, that is the free energy cost of removing H2A and H2B histones from the nucleosome to form the tetrasome, is ~ $17 k_B T$. Assuming (Fenley et al., 2018) ~ 300 μM nucleosome concentration *in-vivo*, this value is reduced to about ~ $9 k_B T$, see also the discussion above. Our estimate of the pH dependence of the $(H3 \cdot H4)_2 \cdot DNA + (H2A \cdot H2B)_2 \rightarrow$ *nucleosome* reaction, $\Delta\Delta G^{0.3} = 5.2 k_B T$, has comparable magnitude, implying that the nucleosome assembly/disassembly process can be modulated by biologically relevant small variations of intra-nuclear pH.

4 Conclusion

In this paper we have explored, computationally, the extent to which thermodynamic stability of protein-nucleic acid complexes may be affected by biologically relevant, small variations of ambient pH around its average value. Since our main interest is in protein-nucleic acid complexes relevant to chromatin function, the focus is on intra-nuclear/intra-cellular pH. We have argued, based on examples of pH variations reported for various biological processes, that a reasonable scale for the "biologically relevant" variation of intra-nuclear pH should be 0.3 units (around a mean value of pH ≈ 7.4). In some processes, e.g., *in apoptosis* (Gottlieb et al., 1995), experimentally reported relevant variation of pH were noticeably larger than 0.3, this is why we believe that our estimate of the significance threshold of 0.3 is conservative. The metric we have introduced to characterize the response of a protein-ligand complex to the change of ambient pH by 0.3 is the corresponding change, $\Delta\Delta G^{0.3}$, of the complex stability, or, equivalently, the change in the free energy of the protein-ligand binding, computed at pH = 7.4. In this work, estimation of $\Delta\Delta G^{0.3}$ for each complex utilized a set of standard computational tools and the fundamental thermodynamic linkage relationship that connects, exactly, the $\Delta\Delta G^{0.3}$ with the predicted net proton update (net charge change) upon complex formation.

We have computed $\Delta\Delta G^{0.3}$ for a set of twenty carefully selected protein-DNA and protein-RNA complexes that cover a range of biological functions. We have found that, in most cases, increasing pH leads to a drop in the binding affinity between the protein and nucleic acid, and *vice versa*. An intuitive, albeit very qualitative, explanation for this general trend is that the binding of a negatively charged nucleic acid can lead to some of the titratable groups of the receptor protein acquiring a proton due to its favorable charge-charge interaction with the oppositely charged ligand. The resulting net gain in the attractive free energy of the complex is higher when protons are more readily available in the ambient solution, that is at lower pH. Only for a small minority of the complexes examined here, it so happens that there is no net proton uptake upon nucleic acid binding around pH ~ 7, and hence the complex stability does not depend on pH in this range. Further analysis has revealed a perhaps unexpected result: the effect of the relatively small pH variations can lead to non-negligible effect $\Delta\Delta G^{0.3} > 1/2 k_B T$ (0.3 kcal/mol) on stability of about 70% of the diverse protein-nucleic acid complexes tested. And for at least 10% of the complexes the effect can be as large as 3 -4 $k_B T$. We argue for possible biological relevance of the variations of complex stability of $1/2 k_B T$ or above, in the context relevant to main chromatin functions such as transcription and DNA replication. Note that a change in the protein-ligand binding affinity of just $1/2 k_B T$ leads to a 1.6-fold change of the ligand occupancy, which is not insignificant. For a more quantitative argument, we have compared the $\Delta\Delta G^{0.3}$ with several known free energy scales characterizing some of the relevant transitions in the complexes of interest or the biological processes that depend on it. As an example, we have shown that binding of a specific transcription factor can be noticeably affected by a pH variation of 0.3 units. Comparison with available experimental data points, albeit not very many, corroborate our general conclusion regarding destabilization of protein-nucleic acid complexes with increasing pH.

We have applied the same computational approach and reasoning to the fundamental unit of DNA compaction in eukaryotes: the nucleosome. The biological function of the nucleosome is diverse, including packaging and protecting of the genomic DNA, as well as gene regulation (Kornberg and Lorch, 2020). Multiple structural transitions in the nucleosome system can change accessibility of its DNA to cellular machines, e.g., those involved in transcription or DNA replication. Increases in nucleosomal DNA accessibility as small as 1.5-fold can have significant biological consequences, e.g., up to 10-fold increase in transcription rate and promoter activity (Zhu and Thiele, 1996; Raveh-Sadka et al., 2012). These consequences of increased DNA accessibility are not sequence-specific, i.e., the effects are due to the increase of the DNA accessibility itself (Raveh-Sadka et al., 2012). Here we have explored several nucleosome core particle structures, considering the effect of pH on the stability of the entire complex between the histone octamer core and its DNA. We have also explored other possible structural transitions in the nucleosome that make its DNA more accessible to cellular machinery. In particular, we have considered the formation of a partially unwrapped state in which 20 base pairs "peel off" transiently at each end of the nucleosomal DNA, as well as a partially assembled state of the nucleosome (the tetrasome), where (H2A·H2B)$_2$ dimer dissociates, exposing about 78 base pairs of the DNA.

We have compared the computed $\Delta\Delta G^{0.3}$ values with some of the biologically relevant energy scales in the above systems/transitions, and argued that $\Delta\Delta G^{0.3}$ is large enough to have biological significance. According to our predictions, the nucleosome complex is destabilized by about $10k_BT$ (6 kcal/mol) by a 0.3 unit increase in the ambient pH. The result is robust to a key parameter of the computational protocol and the specific atomistic structure employed for the computation. We have argued that, compared to relevant energy scales in the nucleosome, $\Delta\Delta G^{0.3} = 10k_BT$ is large enough to be biologically relevant. One such relevant energy scale is the change of the nucleosome stability due to charge-altering post-translation modifications (Fenley et al., 2018), many of which have been shown to lead to biologically relevant effects *in vivo*, e.g., increased transcription rates. Likewise, we predict that one of the key nucleosome disassembly pathways that exposes its DNA, *via* the tetrasome complex of four core histones with the DNA, is also sensitive to small pH variations, with $\Delta\Delta G^{0.3} = 5.2k_BT$ (3 kcal/mol). For spontaneous, transient unwrapping of 20 base pair long end fragments of the nucleosomal DNA, $\Delta\Delta G^{0.3} = 2.2k_BT$. We have argued that this magnitude of sensitivity to ambient pH may also be relevant biologically.

The main limitation of the above quantitative analysis stems from the well known limitations of the underlying continuum solvent methodology we have employed for our $\Delta\Delta G^{0.3}$ estimates. Still, even a hypothetical factor of 2 downward error in our $\Delta\Delta G^{0.3}$ estimates should not invalidate our main conclusion that the pH effect on stability of the nucleosome, and on various other transitions that expose its DNA, is not insignificant compared to the relevant energy scales. Also note that our choice of biologically relevant ΔpH is quite conservative, as some biologically relevant processes are characterized by higher pH changes. For example, if we computed apoptosis-specific $\Delta\Delta G^{0.3}$ values, these would be about three times larger than those used to make the above conclusions, meaning that the predicted effect on the protein-nucleic acid

stability would be about three times stronger. In particular, assuming ΔpH = 1.0 in Eq. 3, we obtain ~ $33k_BT$, ~ $17k_BT$, and ~ $7k_BT$ for the corresponding stability change (increase), $\Delta\Delta G^{0.3}$, of the whole nucleosome, its dis-assmbly into the tetrasome, and unwrapping of the entry/exit region, respectively. These are large effects, diretly comparable, or even equal, to the corresponding free energies themselves, which means that the equilibrium is shifted strongly towards the intact, "frozen," nucleosome away from its more "open" states. Accordingly, the drop in pH by 1 unit associated with apoptosis is predicted to lead to a very significant reduction of the nucleosomal DNA accessibility.

To use the above conclusions for making predictions *in-vivo* requires a leap of faith—faith both in simplicity and the power of reductionsim. Making the leap, we can predict that small, naturally occurring variations of intra-cellular pH can have detectable effect on processes that depend on the DNA accessibility in the nucleosome, e.g., transcription and replication. For example, we speculate that an increase in pH during the S-phase of the cell cycle (Putney and Barber, 2003) may upregulate the DNA replications rate; conversely, if this increase is somehow blocked, the cell cycle may elongate or halt altogether. Likewise, we speculate that both transcription and DNA replication should be upregulated in cancer cells because their intra-cellular pH is shifted upward relative to normal cells, meaning that the DNA association in the nucleosomes becomes relatively looser.[3] Perhaps the strongest speculative prediction can be made for the effect of the significant pH drop seen in apoptosis. Assuming that the drop is an effector of the apoptosis program (Gottlieb et al., 1995), we predict significant downregulation of processess that depend on accessibility of the nucleosomal DNA, in particular DNA replication and transcription. The main limitation of this kind of reasoning is that the reality is likely more complex than the one assumed by our simplistic picture. For example, even considering only the results of this work, one can notice that, *e.g.*, transcription rate as a function of increased pH can be affected by two factors, working in the opposite directions: the more accessible DNA in the nucleosome is countered by the diminished effectiveness of transcription factor(s) due to their lower DNA binding affinity at a higher pH. A more complex, quantitative model may be required to make stronger connections between our $\Delta\Delta G^{0.3}$ predictions and what happens *in-vivo*. Going well beyond this work, it would be useful to establish whether the magnitude of biologically relevant intra-cellular pH variations are somehow "universal," that is are not dissimilar between similar processes in higher eukaryotes. In our view, the following fact hints at this possibility: a large variety of different molecular mechanisms are involved in regulating intra-cellular pH, and these mechanisms appear highly redundant (Doyen et al., 2022). Tractable mathematical models (Doyen et al., 2022) can provide initial insights and guide future experiments. We suggest that the proposed "minimally relevant unit of pH variation" of 0.3, or equivalently, 2-fold change in the H^+ activity, may be adopted as a convenient threshold for classification of experimentally determined

---

3 In light of this, one may ask a hypothetical question what would be the effect of "alkaline diet" (sometimes advocated as an anti-cancer remedy), if the relevant pH could actually be altered by diet.

pH changes or differences as biologically relevant, at least in the context of chromatin function.

In summary, our predictions can be divided into two groups. First, we have made predictions that can be verified *in-vitro* by relatively standard biochemical methods such as isothermal titration calorimetry, which can revel the binding free energy. These measurements taken on a reasonably large sample of chromatin remodeling/transcription factors, at several points around pH = 7.4, can verify our main prediction that the pH effect on stability of these complexes is non-negligible, statistically speaking. In our opinion, pH dependence of the DNA accessibility in the nucleosome warrants a special investigation due to its potential importance. For example, pH dependence of spontaneous unwrapping of end fragments of nucleosomal DNA may potentially be tested within the same types of experiments (Gansen et al., 2009; Koopmans et al., 2009; Tims et al., 2011; Chen et al., 2014; Chen et al., 2017) that were used to investigate the original effect. Presumably, the same idea applies to experimental determination of pH dependence of the transitions involved in partial assembly/dis-assembly of the nucleosome (Andrews and Luger, 2011).

The second group of predictions includes the speculative connections to what may happen *in-vivo* as a result of small variations of intra-cellular pH. One may start to verify these connections by "minimal" assays, e.g., those that can test replication rates in nucleosomes arrays *in-vitro*. If these experiments do demonstrate meaningful pH dependence, then fully *in-vivo* experiment can begin to assess the most complex picture.

We conclude by proposing a curious, and potentially useful analogy between the nucleosome and the hydrogen atom. There are several layers to the analogy. First, this unique protein-DNA complex is the simplest fundamental unit of chromatin compaction in eukaryotes, its main building block. Second, by analogy with the electronic energy levels in the hydrogen atom, structural transitions in the nucleosome system are characterized by the appropriate energy scales, which determine which processes are allowed or forbidden. Further, the energy scales and the corresponding transitions can be affected by "external modulators": fields, such as magnetic or electric field in the case of the real hydrogen atom or, as we have seen here, by pH in the case of the nucleosome. The list of the biologically relevant external modulators can be continued, potentially leading to further insights. For example, multiple species of mobile ions, including polyamines, are present in the nucleus; their relative abundances fluctuate, e.g., along the cell cycle. These ions can affect protein-nucleic acid binding *via* a variety of physical mechanisms, such as non-specific Debye screening of charge-charge interactions, and counterion-condensation effects. We suggest that the general approach advocated in this work should be applicable: the effect of each modulator can be quantified by the properly defined $\Delta\Delta G_Y^X$, and compared with the biologically relevant variations of the modulator in question.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://www.rcsb.org/.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Funding

Acknowledgments

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be .construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2023.1067787/full#supplementary-material

References

Aguilar, B., Anandakrishnan, R., Ruscio, J. Z., and Onufriev, A. V. (2010). Statistics and physical origins of pK and ionization state changes upon protein-ligand binding. *Biophys. J.* 98, 872–880. doi:10.1016/j.bpj.2009.11.016

Alexov, E., Mehler, E. L., Baker, N., Baptista, A. M., Huang, Y., Milletti, F., et al. (2011). Progress in the prediction of pKa values in proteins. *Proteins* 79, 3260–3275. doi:10.1002/prot.23189

Anandakrishnan, R., Aguilar, B., and Onufriev, A. V. (2012). H++ 3.0: Automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* 40, 537–541. doi:10.1093/nar/gks375

Andrews, A. J., Chen, X., Zevin, A., Stargell, L. A., and Luger, K. (2010). The histone chaperone Nap1 promotes nucleosome assembly by eliminating nonnucleosomal histone DNA interactions. *Mol. Cell* 37, 834–842. doi:10.1016/j.molcel.2010.01.037

Andrews, A. J., and Luger, K. (2011). Nucleosome structure(s) and stability: Variations on a theme. *Annu. Rev. Biophysics* 40, 99–117. doi:10.1146/annurev-biophys-042910-155329

Antosiewicz, J., McCammon, J. A., and Gilson, M. K. (1994). Prediction of pH-dependent properties of proteins. *J. Mol. Biol.* 238, 415–436. doi:10.1006/jmbi.1994.1301

Armeev, G. A., Kniazeva, A. S., Komarova, G. A., Kirpichnikov, M. P., and Shaytan, A. K. (2021). Histone dynamics mediate dna unwrapping and sliding in nucleosomes. *Nat. Commun.* 12, 2387. doi:10.1038/s41467-021-22636-9

Bas, D. C., Rogers, D. M., and Jensen, J. H. (2008). Very fast prediction and rationalization of pka values for protein-ligand complexes. *Proteins* 73, 765–783. doi:10.1002/prot.22102

Bashford, D., and Karplus, M. (1990). pK$_a$'s of ionizable groups in proteins: atomic detail from a continuum electrostatic model. *Biochemistry* 29, 10219–10225. doi:10.1021/bi00496a010

Beard, D., and Schlick, T. (2001). Modeling salt-mediated electrostatics of macromolecules: The discrete surface charge optimization algorithm and its application to the nucleosome. *Biopolymers* 58, 106–115. doi:10.1002/1097-0282(200101)58:1<106::AID-BIP100>3.0.CO;2-#

Beroza, P., Fredkin, D., Okamura, M., and Feher, G. (1991). Protonation of interacting residues in a protein by a Monte Carlo method: Application to lysozyme and the photosynthetic reaction center of rhodobacter sphaeroides. *Proc. Natl. Acad. Sci. U. S. A.* 88, 5804–5808. doi:10.1073/pnas.88.13.5804

Blossey, R., and Schiessel, H. (2011). The dynamics of the nucleosome: Thermal effects, external forces and ATP. *FEBS J.* 278, 3619–3632. doi:10.1111/j.1742-4658.2011.08283.x

Bowman, G. D., and Poirier, M. G. (2015). Post-Translational modifications of histones that influence nucleosome dynamics. *Chem. Rev.* 115, 2274–2295. doi:10.1021/cr500350x

Brehove, M., Wang, T., North, J., Luo, Y., Dreher, S. J., Shimko, J. C., et al. (2015). Histone core phosphorylation regulates DNA accessibility. *J. Biol. Chem.* 290, 22612–22621. doi:10.1074/jbc.m115.661363

Casey, J., Grinstein, S., and Orlowski, J. (2009). Sensors and regulators of intracellular ph. *Nat. Rev. Mol. Cell Biol.* 11, 50–61. doi:10.1038/nrm2820

Chan, P., Lovric, J., and Warwicker, J. (2006). Subcellular pH and predicted pH-dependent features of proteins. *Proteomics* 6, 3494–3501. doi:10.1002/pmic.200500534

Chen, Y., Tokuda, J., Topping, T., Meisburger, S., Pabit, S., Gloss, L., et al. (2017). Asymmetric unwrapping of nucleosomal dna propagates asymmetric opening and dissociation of the histone core. *Proc. Natl. Acad. Sci.* 114, 334–339. doi:10.1073/pnas.1611118114

Chen, Y., Tokuda, J., Topping, T., Sutton, J., Meisburger, S., Pabit, S., et al. (2014). Revealing transient structures of nucleosomes as DNA unwinds. *Nucleic Acids Res.* 42, 8767–8776. doi:10.1093/nar/gku562

Culkin, J., de Bruin, L., Tompitak, M., Phillips, R., and Schiessel, H. (2017). The role of DNA sequence in nucleosome breathing. *Eur. Phys. J. E* 40, 106. doi:10.1140/epje/i2017-11596-2

da Veiga Moreira, J., Peres, S., Steyaert, J. M., Bigan, E., Paulevé, L., Nogueira, M. L., et al. (2015). Cell cycle progression is regulated by intertwined redox oscillators. *Theor. Biol. Med. Model* 12, 10. doi:10.1186/s12976-015-0005-2

Deegan, B. J., Seldeen, K. L., McDonald, C. B., Bhat, V., and Farooq, A. (2010). Binding of the eralpha nuclear receptor to dna is coupled to proton uptake. *Biochemistry* 49, 5978–5988. doi:10.1021/bi1004359

DelBuono, G., Figueirido, F., and Levy, R. (1994). Intrinsic pk$_a$s of ionizable residues in proteins: An explicit solvent calculation for lysozyme. *Proteins* 20, 85–97. doi:10.1002/prot.340200109

Demchuk, E., and Wade, R. C. (1996). Improving the continuum dielectric approach to calculating pK$_a$s of ionizable groups in proteins. *J. Phys. Chem.* 100, 17373–17387. doi:10.1021/jp960111d

Doyen, D., Poët, M., Jarretou, G., Pisani, D. F., Tauc, M., Cougnon, M., et al. (2022). Intracellular ph control by membrane transport in mammalian cells. insights into the selective advantages of functional redundancy. *Front. Mol. Biosci.* 9, 825028. doi:10.3389/fmolb.2022.825028

Ellis, C. R., Tsai, C. C., Hou, X., and Shen, J. (2016). Constant pH molecular dynamics reveals pH-modulated binding of two small-molecule bace1 inhibitors. *J. Phys. Chem. Lett.* 7, 944–949. doi:10.1021/acs.jpclett.6b00137

Fenley, A. T., Adams, D. A., and Onufriev, A. V. (2010). Charge state of the globular histone core controls stability of the nucleosome. *Biophys. J.* 99, 1577–1585. PMCID 2931741. doi:10.1016/j.bpj.2010.06.046

Fenley, A. T., Anandakrishnan, R., Kidane, Y. H., and Onufriev, A. V. (2018). Modulation of nucleosomal dna accessibility via charge-altering post-translational modifications in histone core. *Epigenetics Chromatin* 11, 11. doi:10.1186/s13072-018-0181-5

Flinck, M., Kramer, S. H., Schnipper, J., Andersen, A. P., and Pedersen, S. F. (2018). The acid-base transport proteins nhe1 and nbcn1 regulate cell cycle progression in human breast cancer cells. *Cell Cycle* 17, 1056–1067. doi:10.1080/15384101.2018.1464850

Gansen, A., Valeri, A., Hauger, F., Felekyan, S., Kalinin, S., Tóth, K., et al. (2009). Nucleosome disassembly intermediates characterized by single-molecule FRET. *Proc. Natl. Acad. Sci.* 106, 15308–15313. doi:10.1073/pnas.0903005106

Garcia-Moreno, B. (2009). Adaptations of proteins to cellular and subcellular ph. *J. Biol.* 8, 98. doi:10.1186/jbiol199

Garvie, C. W., and Phillips, S. E. (2000). Direct and indirect readout in mutant met repressor-operator complexes. *Structure* 8, 905–914. doi:10.1016/s0969-2126(00)00182-9

Georgescu, R., Alexov, E., and Gunner, M. (2002). Combining conformational flexibility and continuum electrostatics for calculating pK$_a$s in proteins. *Biophy. J.* 83, 1731–1748. doi:10.1016/S0006-3495(02)73940-4

Gordon, J. C., Myers, J. B., Folta, T., Shoja, V., Heath, L. S., and Onufriev, A. (2005). H++: A server for estimating pK$_a$s and adding missing hydrogens to macromolecules. *Nucleic. acids. Res.* 33, 368–371. doi:10.1093/nar/gki464

Gottlieb, R. A., Giesing, H. A., Zhu, J. Y., Engler, R. L., and Babior, B. M. (1995). Cell acidification in apoptosis: Granulocyte colony-stimulating factor delays programmed cell death in neutrophils by up-regulating the vacuolar h(+)-atpase. *Proc. Natl. Acad. Sci. U. S. A.* 92, 5965–5968. doi:10.1073/pnas.92.13.5965

Hamimes, S., Bourgeon, D., Stasiak, A. Z., Stasiak, A., and Van Dyck, E. (2006). Nucleic acid-binding properties of the rrm-containing protein rdm1. *Biochem. Biophysical Res. Commun.* 344, 87–94. doi:10.1016/j.bbrc.2006.03.154

Henikoff, S. (2008). Nucleosome destabilization in the epigenetic regulation of gene expression. *Nat. Rev. Genet.* 9, 15–26. doi:10.1038/nrg2206

Huang, Y., Harris, R., and Shen, J. (2018). Generalized born based continuous constant ph molecular dynamics in amber: Implementation, benchmarking and analysis. *J. Chem. Inf. Model.* 58, 1372–1383. doi:10.1021/acs.jcim.8b00227

Hyre, D. E., and Spicer, L. D. (1995). Thermodynamic evaluation of binding interactions in the methionine repressor system of escherichia coli using isothermal titration calorimetry. *Biochemistry* 34, 3212–3221. doi:10.1021/bi00010a010

Kao, Y. H., Fitch, C. A., Bhattacharya, S., Sarkisian, C. J., Lecomte, J. T. J., and García-Moreno, E. B. (2000). Salt effects on ionization equilibria of histidines in myoglobin. *Biophysical J.* 79, 1637–1654. doi:10.1016/S0006-3495(00)76414-9

Kato, D., Osakabe, A., Arimura, Y., Mizukami, Y., Horikoshi, N., Saikusa, K., et al. (2017). Crystal structure of the overlapping dinucleosome composed of hexasome and octasome. *Science* 356, 205–208. doi:10.1126/science.aak9867

Khandogin, J., and Brooks, C. L. (2006). Toward the accurate first-principles prediction of ionization equilibria in proteins. *Biochemistry* 45, 9363–9373. doi:10.1021/bi060706r

Kieseritzky, G., and Knapp, E. W. (2008). Improved pKa prediction: Combining empirical and semimicroscopic methods. *J. Comput. Chem.* 29, 2575–2581. doi:10.1002/jcc.20999

Koopmans, W. J. A., Buning, R., Schmidt, T., and van Noort, J. (2009). spFRET using alternating excitation and FCS reveals progressive DNA unwrapping in nucleosomes. *Biophysical J.* 97, 195–204. doi:10.1016/j.bpj.2009.04.030

Kornberg, R. (1974). Chromatin structure: A repeating unit of histones and DNA. *Science* 184, 868–871. doi:10.1126/science.184.4139.868

Kornberg, R. D., and Lorch, Y. (2020). Primary role of the nucleosome. *Mol. Cell* 79, 371–375. doi:10.1016/j.molcel.2020.07.020

Korolev, N., Vorontsova, O. V., and Nordenskiöld, L. (2007). Physicochemical analysis of electrostatic foundation for DNA-protein interactions in chromatin transformations. *Prog. biophysics Mol. Biol.* 95, 23–49. doi:10.1016/j.pbiomolbio.2006.11.003

Kulichikhin, K., Greenway, H., Byrne, L., and Colmer, T. (2009). Regulation of intracellular pH during anoxia in rice coleoptiles in acidic and near neutral conditions. *J. Exp. Bot.* 60, 2119–2128. doi:10.1093/jxb/erp090

Kundrotas, P. J., and Alexov, E. (2006). Electrostatic properties of protein-protein complexes. *Biophys. J.* 91, 1724–1736. doi:10.1529/biophysj.106.086025

Kunze, K. K., and Netz, R. R. (2002). Complexes of semiflexible polyelectrolytes and charged spheres as models for salt-modulated nucleosomal structures. *Phys. Rev. E Stat. Nonlin Soft Matter Phys.* 66, 011918. doi:10.1103/PhysRevE.66.011918

Kunze, K. K., and Netz, R. R. (2000). Salt-induced DNA-histone complexation. *Phys. Rev. Lett.* 85, 4389–4392. doi:10.1103/PhysRevLett.85.4389

Lejeune, D., Delsaux, N., Charloteaux, B., Thomas, A., and Brasseur, R. (2005). Protein-nucleic acid recognition: Statistical analysis of atomic interactions and influence of dna structure. *Proteins* 61, 258–271. doi:10.1002/prot.20607

Libertini, L. J., and Small, E. W. (1984). Effects of pH on the stability of chromatin core particles. *Nucleic Acids Res.* 12, 4351–4359. doi:10.1093/nar/12.10.4351

Luger, K., Dechassa, M., and Tremethick, D. (2012). New insights into nucleosome and chromatin structure: An ordered state or a disordered affair? *Nat. Rev. Mol. Cell Bio* 13, 436–447. doi:10.1038/nrm3382

Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251–260. doi:10.1038/38444

Luscombe, N. M., Laskowski, R. A., and Thornton, J. M. (2001). Amino acid-base interactions: A three-dimensional analysis of protein-dna interactions at an atomic level. *Nucleic Acids Res.* 29, 2860–2874. doi:10.1093/nar/29.13.2860

Manohar, M., Mooney, A. M., North, J. A., Nakkula, R. J., Picking, J. W., Edon, A., et al. (2009). Acetylation of histone H3 at the nucleosome dyad alters DNA-histone binding. *J. Biol. Chem.* 284, 23312–23321. doi:10.1074/jbc.m109.003202

Materese, C., Savelyev, A., and Papoian, G. (2009). Counterion atmosphere and hydration patterns near a nucleosome core particle. *J. Am. Chem. Soc.* 131, 15005–15013. doi:10.1021/ja905376q

Milo, R., and Phillips, R. (2015). *Cell biology by the numbers*. 1 edn. New York: Taylor & Francis. doi:10.1201/9780429258770

Misra, V. K., Hecht, J. L., Yang, A. S., and Honig, B. (1998). Electrostatic contributions to the binding free energy of the λci repressor to dna. *Biophys. J.* 75, 2262–2273. doi:10.1016/S0006-3495(98)77671-4

Misteli, T. (2007). Beyond the sequence: Cellular organization of genome function. *Cell* 128, 787–800. doi:10.1016/j.cell.2007.01.028

Mongan, J., Case, D., and McCammon, J. A. (2004). Constant *ph* molecular dynamics in generalized born implicit solvent. *J. Comp. Comp. Chem.* 25, 2038–2048. doi:10.1002/jcc.20139

Moreau, M., Morin, I., Askin, S., Cooper, A., Moreland, N., Vasudevan, S., et al. (2012). Rapid determination of protein stability and ligand binding by differential scanning fluorimetry of gfp-tagged proteins. *RSC Adv.* 2, 11892–11900. doi:10.1039/c2ra22368f

Myers, J., Grothaus, G., Narayanan, S., and Onufriev, A. (2006). A simple clustering algorithm can be accurate enough for use in calculations of pKs in macromolecules. *Proteins* 63, 928–938. doi:10.1002/prot.20922

Nielson, J. E., and Vriend, G. (2001). Optimizing the hydrogen-bond network in Poisson-Boltzmann equation-based pK_a calculations. *Proteins* 43, 403–412. doi:10.1002/prot.1053

Olins, A., and Olins, D. (1974). Spheroid chromatin units (ν bodies). *Science* 183, 330–332. doi:10.1126/science.183.4122.330

Onufriev, A., Case, D. A., and Ullmann, G. M. (2001). A novel view of pH titration in biomolecules. *Biochemistry* 40, 3413–3419. doi:10.1021/bi002740q

Onufriev, A. V., and Alexov, E. (2013). Protonation and pK changes in protein-ligand binding. *Q. Rev. biophysics* 46, 181–209. doi:10.1017/S0033583513000024

Onufriev, A. V., and Schiessel, H. The nucleosome: From structure to function through physics. *Curr. Opin. Struct. Biol.* 56 (2019) 119–130. doi:10.1016/j.sbi.2018.11.003*Sequences Topol. Carbohydrates*

Orij, R., Brul, S., and Smits, G. (2011). Intracellular pH is a tightly controlled signal in yeast. *Biochimica Biophysica Acta (BBA) - General Subj.* 1810, 933–944. doi:10.1016/j.bbagen.2011.03.011

Orij, R, Urbanus, M. L., Vizeacoumar, F. J., Giaever, G., Boone, C., Nislow, C., et al. (2012). Genome-wide analysis of intracellular pH reveals quantitative control of cell division rate by pH$_c$ in saccharomyces cerevisiae. *Genome Biol.* 13, R80. doi:10.1186/gb-2012-13-9-r80

Pahari, S., Sun, L., Basu, S., and Alexov, E. (2018). Delphipka: Including salt in the calculations and enabling polar residues to titrate. *Proteins Struct. Funct. Bioinforma.* 86, 1277–1283. doi:10.1002/prot.25608

Parker, M. D., and Boron, W. F. (2013). The divergence, actions, roles, and relatives of sodium-coupled bicarbonate transporters. *Physiol. Rev.* 93, 803–959. doi:10.1152/physrev.00023.2012

Peng, Y., and Alexov, E. (2017). Computational investigation of proton transfer, pka shifts and ph-optimum of protein-dna and protein-rna complexes. *Proteins Struct. Funct. Bioinforma.* 85, 282–295. doi:10.1002/prot.25221

Peng, Y., Li, S., Onufriev, A., Landsman, D., and Panchenko, A. R. (2021). Binding of regulatory proteins to nucleosomes is modulated by dynamic histone tails. *Nat. Commun.* 12, 5280. doi:10.1038/s41467-021-25568-6

Peng, Y. C., and Acheson, N. H. (1998). Polyomavirus large T antigen binds cooperatively to its multiple binding sites in the viral origin of dna replication. *J. Virol.* 72, 7330–7340. doi:10.1128/jvi.72.9.7330-7340.1998

Petukh, M., Stefl, S., and Alexov, E. (2013). The role of protonation states in ligand-receptor recognition and binding. *Curr. Pharm. Des.* 19, 4182–4190. doi:10.2174/1381612811319230004

Popp, A. P., Hettich, J., and Gebhardt, J. (2021). Altering transcription factor binding reveals comprehensive transcriptional kinetics of a basic gene. *Nucleic Acids Res.* 49, 6249–6266. doi:10.1093/nar/gkab443

Putney, L. K., and Barber, D. L. (2003). Na-h exchange-dependent increase in intracellular ph times g2/m entry and transition. *J. Biol. Chem.* 278, 44645–44649. doi:10.1074/jbc.M308099200

Raveh-Sadka, T., Levo, M., Shabi, U., Shany, B., Keren, L., Lotan-Pompan, M., et al. (2012). Manipulating nucleosome disfavoring sequences allows fine-tune regulation of gene expression in yeast. *Nat. Genet.* 44, 743–750. doi:10.1038/ng.2305

Record, M., Anderson, C., and Lohman, T. (1978). Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: The roles of ion association or release, screening, and ion effects on water activity. *Q. Rev. biophysics* 11, 103–178. doi:10.1017/S003358350000202X

Rohs, R., West, S., Sosinsky, A., Liu, P., Mann, R., and Honig, B. (2009). The role of dna shape in protein–dna recognition. *Nature* 461, 1248–1253. doi:10.1038/nature08473

Rychkov, G. N., Ilatovskiy, A. V., Nazarov, I. B., Shvetsov, A. V., Lebedev, D. V., Konev, A. Y., et al. (2017). Partially assembled nucleosome structures at atomic detail. *Biophysical J.* 112, 460–472. doi:10.1016/j.bpj.2016.10.041

Saint-Girons, I., Belfaiza, J., Guillou, Y., Perrin, D., Guiso, N., Bârzu, O., et al. (1986). Interactions of the escherichia coli methionine repressor with the metf operator and its corepressor, s-adenosylmethionine. *J. Biol. Chem.* 261, 10936–10940. doi:10.1016/s0021-9258(18)67477-0

Santos, J. M., Martínez-Zaguilán, R., Facanha, A. R., Hussain, F., and Sennoune, S. R. (2016). Vacuolar h+-atpase in the nuclear membranes regulates nucleo-cytosolic proton gradients. *Am. J. Physiology-Cell Physiology* 311, C547–C558. –C558. doi:10.1152/ajpcell.00019.2016

Schiessel, H. (2003). The physics of chromatin. *J. Phys. Condens Matter* 15, 699–774. doi:10.1088/0953-8984/15/19/203

Senear, D. F., and Ackers, G. K. (1990). Proton-linked contributions to site-specific interactions of λ cI repressor and o$_r$. *Biochemistry* 29, 6568–6577. doi:10.1021/bi00480a004

Sham, Y. Y., Chu, Z. T., and Warshel, A. (1997). Consistent calculations of pK_a's of ionizable residues in proteins: Semi-microscopic and microscopic approaches. *J. Phys. Chem.* 101, 4458–4472. doi:10.1021/jp963412w

Simonson, T. (2013). What is the dielectric constant of a protein when its backbone is fixed? *J. Chem. Theory Comput.* 9, 4603–4608. doi:10.1021/ct400398e

Song, Y., Mao, J., and Gunner, M. R. (2009). MCCE2: Improving protein pKa calculations with extensive side chain rotamer sampling. *J. Comput. Chem.* 30, 2231–2247. doi:10.1002/jcc.21222

Spassov, V. Z., and Yan, L. (2008). A fast and accurate computational approach to protein ionization. *Protein. Sci.* 17, 1955–1970. doi:10.1110/ps.036335.108

Spirin, S., Titov, M., Karyagina, A., and Alexeevski, A. (2007). Npidb: A database of nucleic acids protein interactions. *Bioinformatics* 23, 3247–3248. doi:10.1093/bioinformatics/btm519

Swails, J., York, D., and Roitberg, A. (2014). Constant ph replica exchange molecular dynamics in explicit solvent using discrete protonation states: Implementation, testing, and validation. *J. Chem. Theory Comput.* 10, 1341–1352. doi:10.1021/ct401042b

Tanford, C. (1970). Protein denaturation. c. theoretical models for the mechanism of denaturation. *Adv. Protein Chem.* 24, 1–95.

Tessarz, P., and Kouzarides, T. (2014). Histone core modifications regulating nucleosome structure and dynamics. *Nat. Rev. Mol. Cell Biol.* 15, 703–708. doi:10.1038/nrm3890

Thåström, A., Gottesfeld, J. M., Luger, K., and Widom, J. (2004). Histone-DNA binding free energy cannot be measured in dilution-driven dissociation experiments. *Biochemistry* 43, 736–741. doi:10.1021/bi0302043

Tims, H., Gurunathan, K., Levitus, M., and Widom, J. (2011). Dynamics of nucleosome invasion by DNA binding proteins. *J. Mol. Biol.* 411, 430–448. doi:10.1016/j.jmb.2011.05.044

Torigoe, T., Izumi, H., Yoshida, Y., Ishiguchi, H., Okamoto, T., Itoh, H., et al. (2003). Low pH enhances Sp1 DNA binding activity and interaction with TBP. *Nucleic Acids Res.* 31, 4523–4530. doi:10.1093/nar/gkg487

Ullmann, G. M., and Knapp, E. W. (1999). Electrostatic models for computing protonation and redox equilibria in proteins. *Eur. Biophys. J.* 28, 533–551. doi:10.1007/s002490050236

Wallace, J., and Shen, J. (2009). *Predic*, 466, 455–475. doi:10.1016/s0076-6879(09)66019-5*Ting pKa values with continuous constant pH molecular dynamicsMethods Enzymol*.

Wei, S., Falk, S. J., Black, B. E., and Lee, T. H. H. (2015). A novel hybrid single molecule approach reveals spontaneous DNA motion in the nucleosome. *Nucleic acids Res.* 43, e111. doi:10.1093/nar/gkv549

Wessel, R., Ramsperger, U., Stahl, H., and Knippers, R. (1992). The interaction of sv40 large T antigen with unspecific double-stranded dna: An electron microscopic study. *Virology* 189, 293–303. doi:10.1016/0042-6822(92)90705-T

White, K. A., Grillo-Hill, B. K., and Barber, D. L. (2017). Cancer cell behaviors mediated by dysregulated pH dynamics at a glance. *J. Cell Sci.* 130, 663–669. doi:10.1242/jcs.195297

Woodcock, C. (1973). Ultrastructure of inactive chromatin. *J. Cell. Biol.* 59, A368.

Yang, A. S., Gunner, M. R., Sampogna, R., Sharp, K., and Honig, B. (1993). On the calculation of pK$_a$s in proteins. *Proteins Struct. Funct. Genet.* 15, 252–265. doi:10.1002/prot.340150304

Zhang, W., Feng, J., and Li, Q. (2020). The replisome guides nucleosome assembly during dna replication. *Cell Biosci.* 10, 37. doi:10.1186/s13578-020-00398-z

Zhang, Z., Witham, S., and Alexov, E. (2011). On the role of electrostatics in protein-protein interactions. *Phys. Biol.* 8, 035001. doi:10.1088/1478-3975/8/3/035001

Zhu, Z., and Thiele, D. J. (1996). A specialized nucleosome modulates transcription factor access to a c. glabrata metal responsive promoter. *Cell* 87, 459–470. doi:10.1016/s0092-8674(00)81366-5

Zlatanova, J., Bishop, T. C., Victor, J. M., Jackson, V., and van Holde, K. (2009). The nucleosome family: Dynamic and growing. *Structure* 17, 160–171. doi:10.1016/j.str.2008.12.016