



The Recognition Specificity of the CHD1 Chromodomain with Modified Histone H3 Peptides

Richard S. L. Stein and Wei Wang*

Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0359, USA

Received 3 July 2010;
received in revised form
16 December 2010;
accepted 21 December 2010
Available online
30 December 2010

Edited by M. Levitt

Keywords:

protein recognition;
histone code;
MM-GBSA;
molecular dynamics;
free-energy calculation

Histone tail peptides comprise the flexible portion of chromatin, the substance which serves as the packaging for the eukaryotic genome. According to the histone code hypothesis, reader protein domains (chromodomains) can recognize modifications of amino acid residues within these peptides, regulating the expression of genes. We have performed simulations on models of chromodomain helicase DNA-binding protein 1 complexed with a variety of histone H3 modifications. Binding free energies for both the overall complexes and the individual residues within the protein and peptides were computed with molecular mechanics-generalized Born surface area. The simulation results agree well with experimental data and identify several chromodomain helicase DNA-binding protein 1 residues that play key roles in the interaction with each of the H3 modifications. We identified one class of protein residues that bind to H3 in all of the complexes (generally interacting hydrophobically), and a second class of residues that bind only to particular H3 modifications (generally interacting electrostatically). Additionally, we found that modifications of H3R2 and H3T3 have a dominant effect on the binding affinity; methylation of H3K4 has little effect on the interaction strength when H3R2 or H3T3 is modified. Our findings with regard to the specificity shown by the latter class of protein residues in their binding affinity to certain modifications of H3 support the histone code hypothesis.

© 2010 Elsevier Ltd. All rights reserved.

Introduction

The DNA that constitutes the eukaryotic genome is compactly packaged in the cell nucleus in the form of a coil around histone proteins to form nucleosomes.¹ Histones form octamer complexes with the N-terminal tails protruding from the nucleosome core. The amino acids on this tail can be modified by way of methylations, acetylations,

phosphorylations, and other functional group additions. These posttranslational modifications may form a “histone code”, which is read by “effector” proteins for the regulation of gene expression.^{1–3} The recognition of effector proteins to specific modified histone tails is critical to recruiting additional proteins for chromatin remodeling and transcriptional regulation.^{4–6} Interruption of these protein/histone interactions may have ramifications in the development of human diseases, such as cancer,^{1,7} but the biological consequence of a specific binding interaction often remains to be determined.

A commonly observed histone modification is the single, double, or triple methylation of lysine residues, for example, at positions 4, 9, and 27 of histone H3. Methylation at each of these sites has been observed to affect genetic function.⁸

*Corresponding author. E-mail address:
wei-wang@ucsd.edu.

Abbreviations used: CHD, chromodomain helicase DNA-binding protein; MD, molecular dynamics; MM-GBSA, molecular mechanics-generalized Born surface area; PDB, Protein Data Bank.

Hydrophobic residues in reader proteins, particularly those containing an aromatic ring (tyrosine, phenylalanine, and tryptophan), have been observed to form attractive interactions with methylated lysine residues of histone tails.^{1,9,10} Chromatin organization modifier domains (chromodomains) have been shown to interact with chromatin by binding to DNA, RNA, and the methylated tails of histones.^{11–14} It is believed, though unproven, that chromodomain/protein binding to histone tails maintains a high-density conformation of the structure of chromatin fiber.^{15,16} Human chromodomain helicase DNA-binding protein 1 (CHD1) is a member of the CHD family of proteins, containing tandem double chromodomains near the N-terminus, an enzymatic helicase domain, and a DNA-binding domain at the C-terminus.¹⁷ Nine distinct chromodomain proteins in the CHD family, labeled CHD1–CHD9, have been observed and isolated in the human genome.¹¹ The CHD family of proteins are believed to operate by modifying chromatin structure,^{11,18} and CHD1 has been shown to be involved in transcription elongation¹⁹ and nucleosome assembly.²⁰ It has been observed to bind strongly to trimethylated H3K4 and, to a lesser extent, to monomethylated H3K4 as well.^{14,21} Experimental studies involving mutations have been performed on the CHD1/H3 complex, showing that mutating a key tryptophan residue in the CHD1 to glutamic acid, as occurs in yeast, reduces the binding interaction strength.¹

As far as we know, this study is the first of its kind to simulate, and calculate the binding energies of, the interactions between histones and an effector protein. We know of one publication describing a protein/histone simulation²² that involved a complex of histone H3 and chromodomain protein HP1. That study focused on basic features of the dynamics trajectory, such as the displacement of the molecules in the complex during the simulation; the binding free energies of the interaction were not reported. Another study performed simulations of histone/chromodomain complexes,²³ including CHD1, but it mainly focused on optimizing force field parameters, while a third paper has reported simulation results of the H3 peptide tail,²⁴ but in the absence of chromodomain protein. Calculating the individual contributions of critical protein and

peptide residues to binding affinity and estimating how these contributions are affected by residue modification should be instrumental in evaluating the histone code hypothesis. Because of the large variety of available crystal structures of CHD1 chromodomains complexed with different modifications of H3, this is an ideal system for the purpose of studying the effect of histone modifications at the molecular level and deciphering the histone code.

Results and Discussion

Calculated binding free energies correlate well with the experimental data

The binding free energies of CHD1 to each of the modified H3 peptides were calculated with molecular mechanics–generalized Born surface area (MM–GBSA). The highly charged nature of the system, in particular, the positively charged R2 and K4 residues of H3 and the negatively charged CHD1 residues occupying the active site,^{14,25} makes the calculation of the binding free energy nontrivial and complex and strongly indicates that an interior dielectric constant ϵ higher than 1.0 will be required to estimate the binding free energy accurately.^{23,26} The results for all three values of ϵ (1.0, 2.0, and 4.0) are given in Table 1. In order to quantitatively compare the binding free energies to the experimental dissociation constants given by Flanagan *et al.*,¹⁴ we graphed $\ln(K_d)$ against those binding free energies for which experimental data were available, on a Pearson plot for all three values of ϵ , as shown in Fig. 1. The coefficients r^2 for the Pearson plots were calculated to be 0.7055, 0.9211, and 0.8651, respectively. Because the binding energies for the interior dielectric constant of 2.0 provided the best fit to the experimental data, this value was used exclusively for all of the in-depth analysis involving residue decomposition. While this method of analysis may not be completely reliable because there are only four points available to obtain the fitting coefficients, the quality of the fit in Fig. 1b makes it reasonable to believe that 2.0 is a promising choice of the interior dielectric constant for this system.

Table 1. CHD1/H3 complex binding free energies

Complex	Binding free energy (kcal/mol) for $\epsilon=1.0$	Binding free energy (kcal/mol) for $\epsilon=2.0$	Binding free energy (kcal/mol) for $\epsilon=4.0$	Experimental dissociation constant (μM) ¹⁴
H3K4me1	-26.48 ± 4.79	-39.38 ± 3.59	-45.84 ± 3.49	17 ± 2
H3K4me3	-25.42 ± 4.31	-40.36 ± 3.45	-47.83 ± 3.41	5.2 ± 0.6
H3(K4me3)(R2me2)	-13.57 ± 5.35	-35.85 ± 4.42	-46.99 ± 4.21	24 ± 4
H3(K4me3)(T3ph)	-9.53 ± 3.53	-30.23 ± 3.04	-40.58 ± 3.61	140 ± 34
H3R2me2	-19.90 ± 2.75	-35.94 ± 2.53	-43.96 ± 2.86	N/A
H3T3ph	-8.30 ± 3.58	-30.00 ± 3.07	-40.85 ± 3.41	N/A

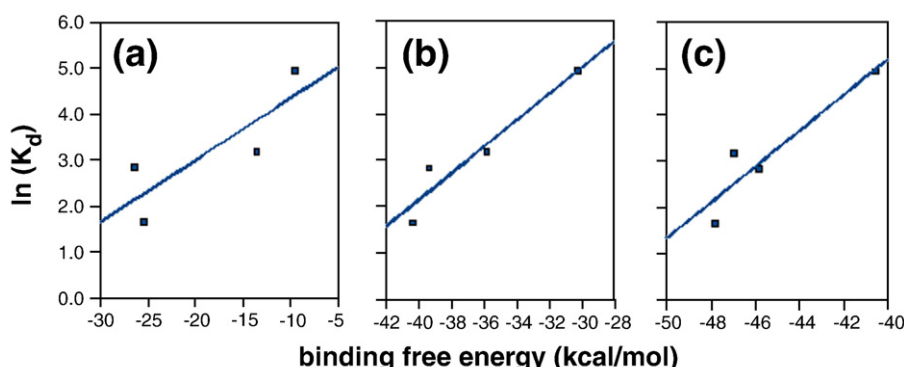


Fig. 1. The experimental dissociation constants given as a function of the binding free energies obtained from the simulations, for internal dielectric constant of 1.0 (a), 2.0 (b), and 4.0 (c). The straight-line fits for the least-squares linear regressions are also shown; the coefficients r^2 are 0.7055 (a), 0.9211 (b), and 0.8651 (c).

Atomic level recognition of the histone code

Encouraged by the satisfactory correlation between the calculated and experimental binding affinities, we have investigated the molecular basis of the recognition.

CHD1 interactions with modifications of H3

The binding free energy for each complex was also decomposed to determine the contribution from each of the individual CHD1 and histone H3

residues. Table 2 shows the binding energy contributions for the most important residues for each model.

The residue decomposition data shown in Table 3 identify H3R2, H3K4, Trp67[A] (Trp67 in chain A), and Trp64[A] as the most important contributors to the binding interaction between CHD1 and histone H3. Prior experimental work has mentioned Trp67 and Trp64,^{1,14} in particular, the former, in chain A of CHD1 to be critical in forming a complex with histone H3, on account of cation/ π interactions⁹ between the CHD1 tryptophan residues and the

Table 2. CHD1/H3 binding free-energy residue decomposition

Residue	Contributions in CHD1/H3K4me1	Contributions in CHD1/H3K4me3	Contributions in CHD1/H3 (K4me3)(R2me2)	Contributions in CHD1/H3 (K4me3)(T3ph)
H3A1	-0.18±1.97	-0.32±0.71	+3.39±1.36	+0.47±0.94
H3R2	-5.90±0.81	-7.13±1.10	-6.85±1.11	-5.53±0.97
H3T3	-2.55±0.51	-2.17±0.49	-2.85±0.76	+2.04±0.56
H3K4	-6.65±0.89	-9.28±0.97	-8.21±1.09	-8.00±0.99
H3Q5	-0.48±0.27	+0.46±0.29	+0.48±0.27	+0.40±0.42
Trp67[A]	-6.13±0.66	-6.51±0.73	-6.64±0.78	-5.98±1.02
Trp64[A]	-1.82±0.35	-2.44±0.43	-2.64±0.57	-2.60±0.35
His71[A]	-2.86±0.42	-2.48±0.51	-2.01±0.55	-1.44±0.53
Glu166[A]	-2.79±0.77	-0.78±0.13	-0.34±0.30	-0.48±0.20
Asp167[A]	-1.30±0.28	-1.83±0.46	-2.91±0.88	-1.12±0.41
Ala32[A]	-1.99±0.38	-1.56±0.35	-1.32±0.54	-1.59±0.43
Glu50[C]	-0.97±0.57	-1.22±0.54	-0.70±0.37	-0.20±0.10
Asn52[C]	-0.01±0.04	+0.02±0.01	-1.32±0.37	-0.22±0.40
Tyr37[A]	-0.86±0.37	-1.00±0.45	-0.86±0.51	-0.62±0.32
Asp150[A]	-0.40±0.09	-1.48±0.49	-1.85±0.78	-0.94±0.63
Glu14[A]	-1.46±0.61	-0.94±0.16	-0.76±0.46	-1.24±0.66
Thr73[A]	-0.84±0.22	-1.13±0.21	-0.33±0.14	-0.84±0.33
Leu170[A]	-1.16±0.45	-0.62±0.23	-0.65±0.24	-0.98±0.31
Thr33[A]	-0.75±0.30	-0.15±0.08	-0.11±0.06	-0.38±0.26
Trp165[A]	-1.07±0.51	-0.49±0.18	-0.64±0.39	-0.75±0.30
Lys63[A]	+0.61±0.05	+0.72±0.09	+0.60±0.06	+0.45±0.09
Lys65[A]	+0.37±0.04	+0.51±0.11	+0.35±0.05	+0.22±0.11
Lys53[C]	+0.27±0.02	+0.68±0.67	+0.39±0.11	+0.20±0.13
Lys96[B]	+0.44±0.04	+0.41±0.04	+0.45±0.05	+0.21±0.03
Lys97[B]	+0.45±0.05	+0.50±0.09	+0.42±0.03	+0.28±0.04

Table 3. CHD1/H3 interaction by residue

CHD1 residue	H3 residue(s) of complex interaction
Trp67, chain A	K4me3, R2
Trp64, chain A	K4me3
His71, chain A	T3 (at H3 backbone)
Glu166, chain A	A1, T3
Asp167, chain A	A1, T3
Ala32, chain A	K4 (at chain base), Q5
Glu50, chain C	R2
Asn52, chain C	R2me2
Tyr37, chain A	K4 (at chain base), Q5
Asp150, chain A	A1
Glu14, chain A	K4, K4me1, K4me3, R2, R2me2
Thr73, chain A	K4me3
Leu170, chain A	T3
Thr33, chain A	Q5, T6 (present in H3K4me1 model only)
Trp165, chain A	A1
Lys63, chain A (unfavorable interaction)	K4
Lys65, chain A (unfavorable interaction)	R2
Lys96, chain B (unfavorable interaction)	R2
Lys97, chain B (unfavorable interaction)	R2
Lys53, chain C (unfavorable interaction)	R2

H3K4me3 cation. The residue decomposition data above confirm these observations and also identify several other important CHD1 residues, such as His71[A], Glu166[A], Asp167[A], Ala32[A], and Glu50[C]. A number of lysine residues in CHD1 whose contributions to the H3 binding are overall unfavorable were also found, the most eminent of these being Lys63[A], Lys65[A], Lys96[B], Lys97[B], and Lys53[C]. However, these unfavorable interactions are smaller in magnitude than the favorable interactions contributed by most other CHD1 residues in Table 4. The strength of the binding contributions for all CHD1 residues in Table 2 appearing in chains A and C are shown in Fig. 2, in histogram form with the bars color coded by H3 modification. The histogram confirms, at a glance, the importance of Trp67[A] to the binding affinity, as well as the contributions of Trp64[A], His71[A], Ala32[A], and Asp167[A] to the interaction strength for all H3 modifications. Across the four modifications of H3 featured in the crystal structures, each of these residues contribute, on average, a favorable term to the binding affinity of more than 1.6 kcal/mol. All other CHD1 residues contribute less than 1.2 kcal/mol on average, and they are therefore less critical to the general H3 binding affinity.

A structural analysis of the general CHD1/H3 complex was performed with the use of the molecular graphical visualization program VMD.²⁷ Table 3 shows the most important CHD1 residues in the interaction with H3 and, by virtue of proximity,

Table 4. Residue effects of demethylation in prediction models (in kcal/mol)

Residue	$\Delta(\Delta G_{\text{bind}})$ of demethylating K4 in H3(K4me3)(R2me2)	$\Delta(\Delta G_{\text{bind}})$ of demethylating K4 in H3(K4me3)(T3ph)
H3A1	-2.78	-0.44
H3R2	+0.67	+0.70
H3T3	-0.51	-0.84
H3K4	+0.40	+0.90
H3Q5	-0.01	+0.41
Trp67[A]	-0.25	+0.32
Trp64[A]	+0.92	+0.79
His71[A]	-0.41	-1.00
Glu166[A]	+0.30	+0.31
Asp167[A]	+1.04	-0.24
Ala32[A]	-0.14	-0.03
Glu50[C]	+0.33	-0.02
Asn52[C]	+1.11	-0.33
Tyr37[A]	-0.43	-0.99
Asp150[A]	+0.72	-0.50
Glu14[A]	-0.17	+0.74
Thr73[A]	-0.37	-0.03
Leu170[A]	-0.56	+0.24
Thr33[A]	+0.03	+0.25
Trp165[A]	+0.60	+0.36
Lys63[A]	-0.11	-0.24
Lys65[A]	-0.08	-0.05
Lys53[C]	+0.04	-0.16
Lys96[B]	-0.12	-0.04
Lys97[B]	-0.12	-0.10

which H3 residues each one is observed to interact with. Figure 3 shows a visual representation of the interaction between CHD1 and H3K4me3, both in a general view and with a focus on the most important interacting residues in the protein. The residue Trp67[A], identified as critical in Refs. 1, 14, and 23, serves as a wall separating H3R2 and H3K4 and is in fact a key feature of the shape of the entire CHD1 binding site. Other residues that shape the protein binding pocket are Trp64[A] (which interacts with the H3K4 side chain); His71[A], Thr73[A], Ala32[A], and Tyr37[A] (which interact with either H3Q5 or the H3 backbone); and Leu170[A], Trp165[A], and even Glu14[A] (which interact with the N-terminal residues of H3). The negatively charged residues that influence the specificity of CHD1 binding to a particular H3 modification tend to interact with H3R2 or with the positively charged H3A1 terminal amino group.

In order to best visualize the electrostatic interactions between CHD1 and the histone tail, we created charge distribution maps for the four models corresponding to H3K4me1, H3K4me3, H3(K4me3)(R2me2), and H3(K4me3)(T3ph). The electrostatic potential maps are shown in Fig. 4; histone residues are shown as ball-and-stick models, which the CHD1 molecule is rendered by SURF.²⁸ The ramifications of this map will be discussed in detail in CHD1 interactions specific to H3K4me1 through CHD1 interactions specific to H3(K4me3)(T3ph).

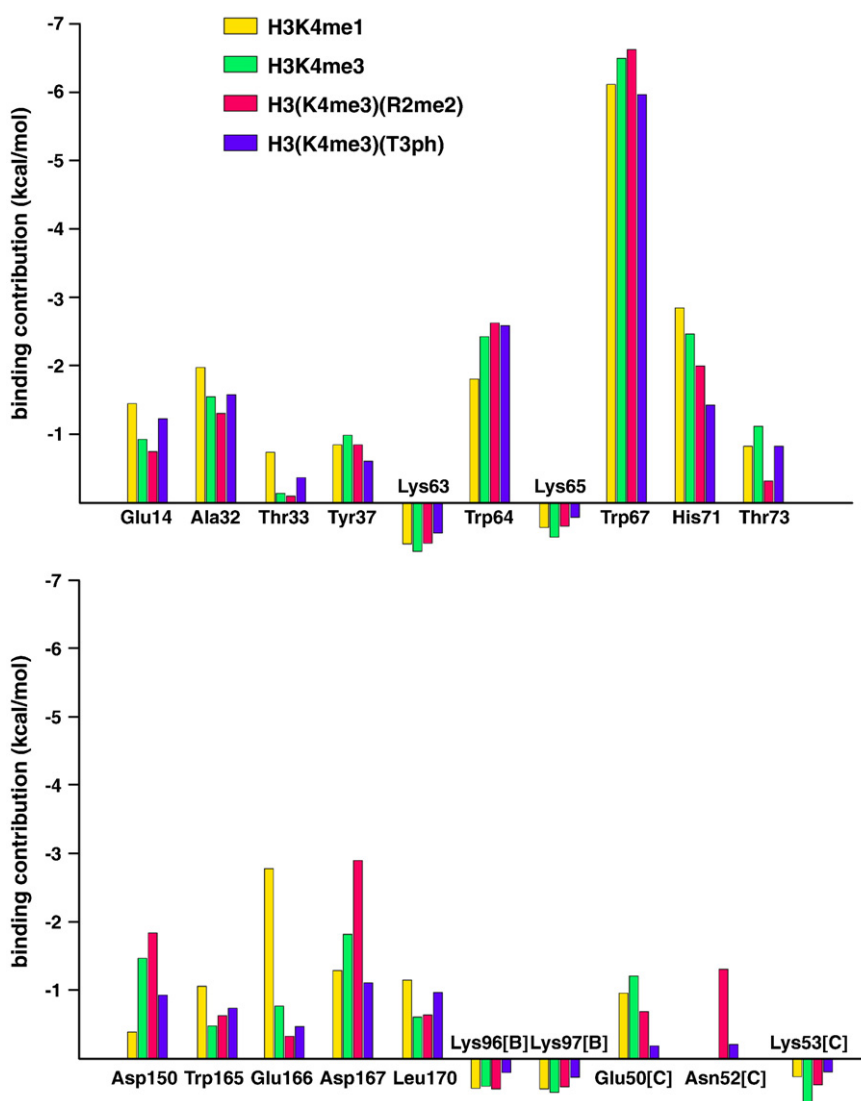


Fig. 2. Binding affinity of CHD1 to various H3 modifications by residue. Higher bars represent a stronger favorable binding interaction; bars pointing below the horizontal axis indicate an unfavorable contribution. Binding contributions for H3K4me1 are shown in yellow; for H3K4me3, green; for H3(K4me3)(R2me2), red; and for H3(K4me3)(T3ph), blue.

CHD1 interactions specific to H3K4me1

The residue Glu166[A] seems to have a strong favorable contribution only in the case of H3K4me1 (Fig. 2). Figure S1 shows the interatomic distance between H3A1 and Glu166[A] over the production run trajectories of both H3K4me1 and H3K4me3. For the purposes of calculating distances, we defined the location of H3A1 as the center of mass of the four heavy atoms, while that for Glu166[A] was defined as the center of mass of the three atoms (O–C–O) making up the side-chain carboxyl group. The average interatomic distance for each run is 7.2 Å for H3K4me3 but only 4.1 Å for H3K4me1, strongly suggesting that trimethylation of H3K4 disrupts the close contact between H3A1 and

Glu166[A], weakening the electrostatic interaction (−1.91 kcal/mol for H3K4me1 and actually positive or unfavorable for other modifications; Fig. 2 and Tables S2–S5). In addition, H3K4me1 is the only modification of H3 for which Thr33[A] contributes appreciably (−0.75 kcal/mol, compared with smaller magnitudes than −0.40 kcal/mol for other modifications; Tables S2–S5). Table S2 shows that the relatively large involvement of Thr33[A] in the attraction to H3K4me1 is due to a van der Waals interaction (−0.65 out of the −0.75 kcal/mol), in spite of the polar nature of H3T6 and Thr33[A]. We think that this is an artifact created by the presence of H3T6 in the crystal structure for CHD1 complexed with H3K4me1; H3T6 was not resolved in the experiments, which gave rise to the other three

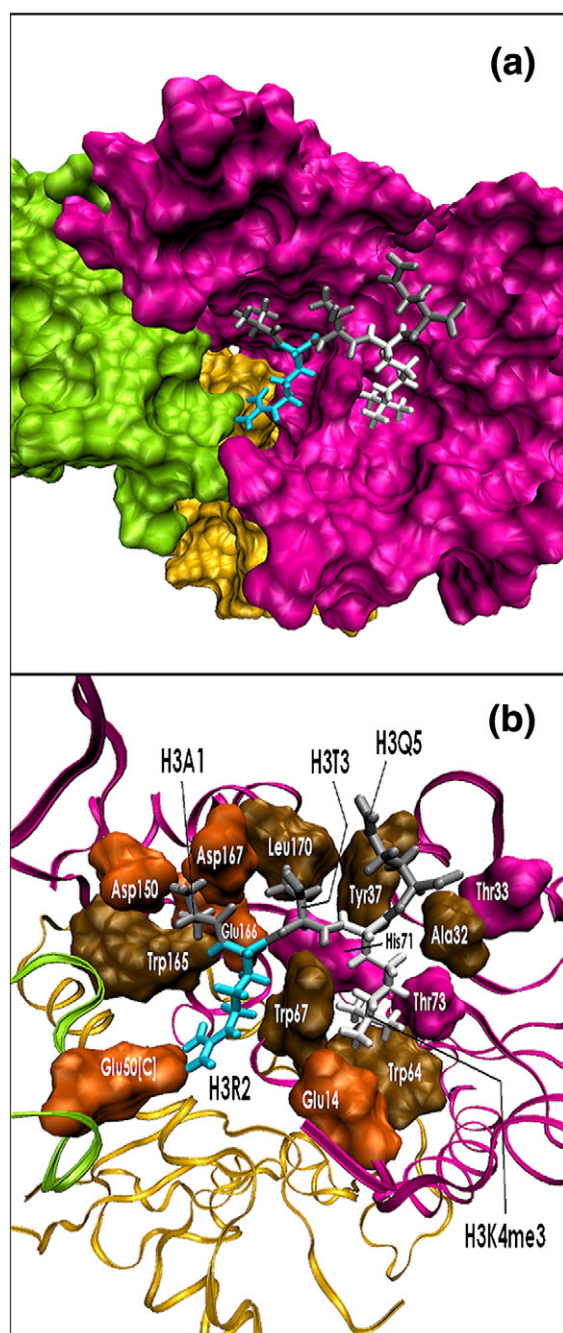


Fig. 3. (a) Surface rendering of the interface between CHD1 and histone H3K4me3. CHD1 chain A is shown in magenta; chain B, in yellow; and chain C, in light green. Histone H3 is drawn by licorice model. H3A1, H3T3, and H3Q5 are colored gray; H3R2, light blue; and H3K4me3, white. (b) Close-up of active site with CHD1 drawn in ribbon representation. All CHD1 residues in Table 2, except for the lysines and Asn52[C], are shown by molecular surface rendering. Hydrophobic residues are shown in brown; negatively charged residues, in orange; and polar neutral residues, in magenta. All other colorings correspond to those used in Fig. 3a. Images are generated by VMD.

crystal structures. Because of the proximity of H3T6 to Thr33[A] in the CHD1 model, we believe that if H3T6 had been resolved in the other crystal structures, the contribution of Thr33[A] in those cases would be much more comparable to that for H3K4me1. To a lesser extent, this may be responsible for Ala32[A] having a stronger interaction with H3K4me1 (-1.99 kcal/mol) than any of the other modifications (weaker than -1.60 kcal/mol).

Other CHD1 residues that have their largest binding contribution in the case of H3K4me1 include Glu14[A], His71[A], Trp165[A], and Leu170[A]. It is likely that those residues are designed to “recognize” the modification H3K4me1 in the conformation of CHD1, which is best suited to bind to that modification. In Fig. 4a, it is evident that the carbonyl backbone carbon of H3A1 has a strong positive charge and is attracted to Glu166[A]. Meanwhile, the positively charged H3R2 appears in a similar attraction to Glu14[A]. The large contributions of these two CHD1 residues to the binding interaction, which are only seen in the complex with H3K4me1, can be attributed to these effects (Fig. 2 and Tables S2–S5). His71[A] and Leu170[A] interact with H3K4me1 through van der Waals attractions (-2.76 and -1.12 kcal/mol, respectively), as expected, while Glu14[A] contributes via an electrostatic attraction of -1.59 kcal/mol (see Table S2). Surprisingly, the contribution of Trp165[A] to the interaction with H3K4me1 is purely electrostatic in nature (-1.25 kcal/mol). Figure 3b suggests that this is likely because it is the backbone of Trp165[A], rather than the hydrophobic side chain, that is in contact with the peptide.

CHD1 interactions specific to H3K4me3

As expected, the contributions for Trp64[A] and Trp67[A] of CHD1 are greater for H3K4me3 than for H3K4me1 because of the strengthened van der Waals interaction (0.45 kcal/mol stronger for Trp64). Trimethylating H3K4 also diminishes the interaction with Glu14[A] by 0.52 kcal/mol and increases it by 0.25 kcal/mol for Glu50[C], which is observed in Figs. 3 and 4b to be partaking in a strong electrostatic attraction.

These binding free-energy differences could be explained by proposing that modifying H3K4 from a monomethylated to a trimethylated state causes a conformational change in CHD1, which drives H3R2 away from Glu14[A] and into close proximity with Glu50[C]. In order to test this hypothesis, we measured the interatomic distance between H3R2 and each of Glu14[A] and Glu50[C] over the production run trajectories for both H3K4me3 and H3K4me1. The location of H3R2 was defined as the center of mass of the three atoms (N–C–N) at the tip of the positively charged side chain, while those of Glu14[A] and Glu50[C] were defined the same way

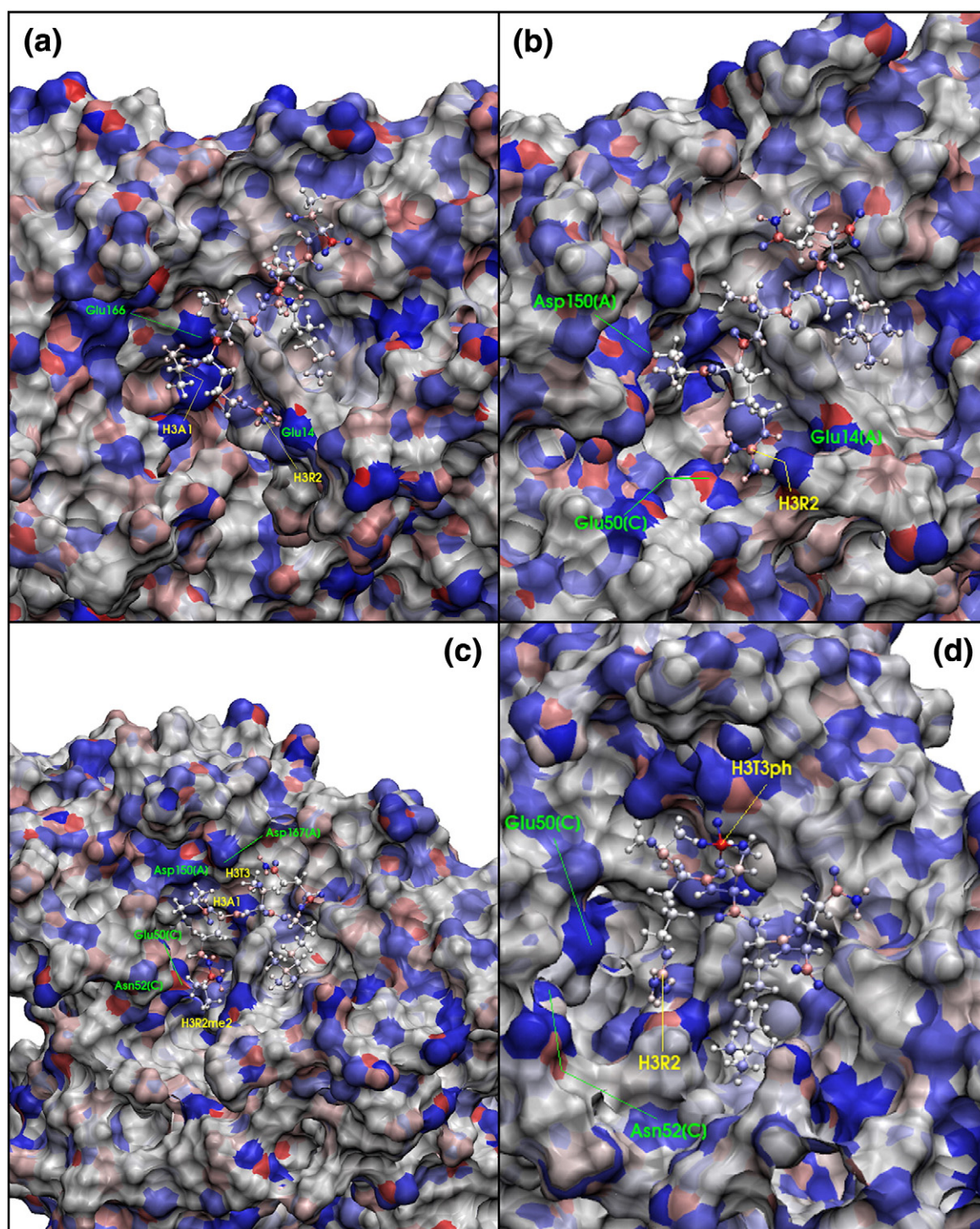


Fig. 4. Electrostatic charge representation of CHD1 complexed with: (a) H3K4me1; (b) H3K4me3; (c) H3(K4me3)(R2me2); and (d) H3(K4me3)(T3ph). Histone residues are drawn as a ball-and-stick model, and CHD1 residues are rendered by SURF. Positively charged atoms are shown in red, while negatively charged ones are shown in blue. Images are generated by VMD.

as for Glu166[A] (see [CHD1 interactions specific to H3K4me1](#)). The interatomic distances for H3R2 and the two glutamic residues are shown in [Figure S2](#).

Over the trajectories, the average interatomic distance between H3R2 and Glu14[A] was 3.3 Å for H3K4me1 but 6.8 Å for H3K4me3; between H3R2

and Glu50[C], it was 9.6 Å for H3K4me1 but only 3.6 Å for H3K4me3. According to Table S3, a large spike in the electrostatic attraction accounts for the increased involvement of Glu50[C] (−1.69 kcal/mol, compared with only −0.52 kcal/mol for H3K4me1). The van der Waals term for the contribution of Glu50[C] in interacting with H3K4me3 is actually unfavorable (+0.47 kcal/mol), as is the same term for the contribution of Glu14[A] in interacting with H3K4me1 (+0.13 kcal/mol). A visualization of the structure (Fig. 5) shows that H3R2 is attracted just as exclusively to Glu14[A] in the H3K4me1 model as it is to Glu50[C] in the case of H3K4me3. Figures 4a 4b and 5 also suggest that trimethylation of H3K4 causes the subsection of CHD1 containing Trp67[A], Trp64[A], and Glu14[A] to roll away from the N-terminus of H3, separating H3R2 and Glu14[A]. Together, these observations constitute overwhelming evidence that the difference in methylation level of H3K4 induces a conformational change in CHD1, and different protein residues recognize each of the modifications.

In Fig. 4b, the histone residue H3K4me3 can be seen in its hydrophobic pocket while H3R2 can be seen in its expected electrostatic interaction with Glu50[C]. Residues that interact with the H3 backbone, such as Tyr37[A] and Thr73[A], also interact more strongly with H3K4me3 than with any

of the other H3 modifications. Asp150[A] exhibits its most favorable interactions to the highly methylated H3 forms, H3K4me3 and H3(K4me3)(R2me2). As shown by Fig. 4b and c, these modifications place H3A1 at a suitable distance for a strong electrostatic interaction between its positively charged N-terminal amino group and Asp150[A]. Table S3 confirms that the increased binding contribution from Asp150[A] is electrostatic in nature (−1.36 out of −1.48 kcal/mol). Finally, the unfavorable contributions of CHD1 lysine residues such as Lys63[A], Lys65[A], and Lys53[C] increase slightly in strength with the trimethylation of H3K4, though not nearly as much compared with the increase in favorable binding affinity for the other residues. On average, the electrostatic terms for the contributions of Lys63[A], Lys65[A], and Lys53[C] are more unfavorable by +0.45 kcal/mol for H3K4me3 than for H3K4me1 (+0.98 kcal/mol in the case of Lys53[C]). This trend provides evidence for H3K4me3 “pulling” hydrophobic residues Trp64[A] and Trp67[A] closer via the van der Waals attraction, and as a consequence, nearby residues on the CHD1 backbone, Lys63[A] and Lys65[A], are also “pulled” closer to H3K4me3, magnifying the strength of the electrostatic repulsion. Note that this increase in unfavorable interaction strength for H3K4me3 is not observed for the CHD1 residues Lys96[B] and Lys97[B], which are

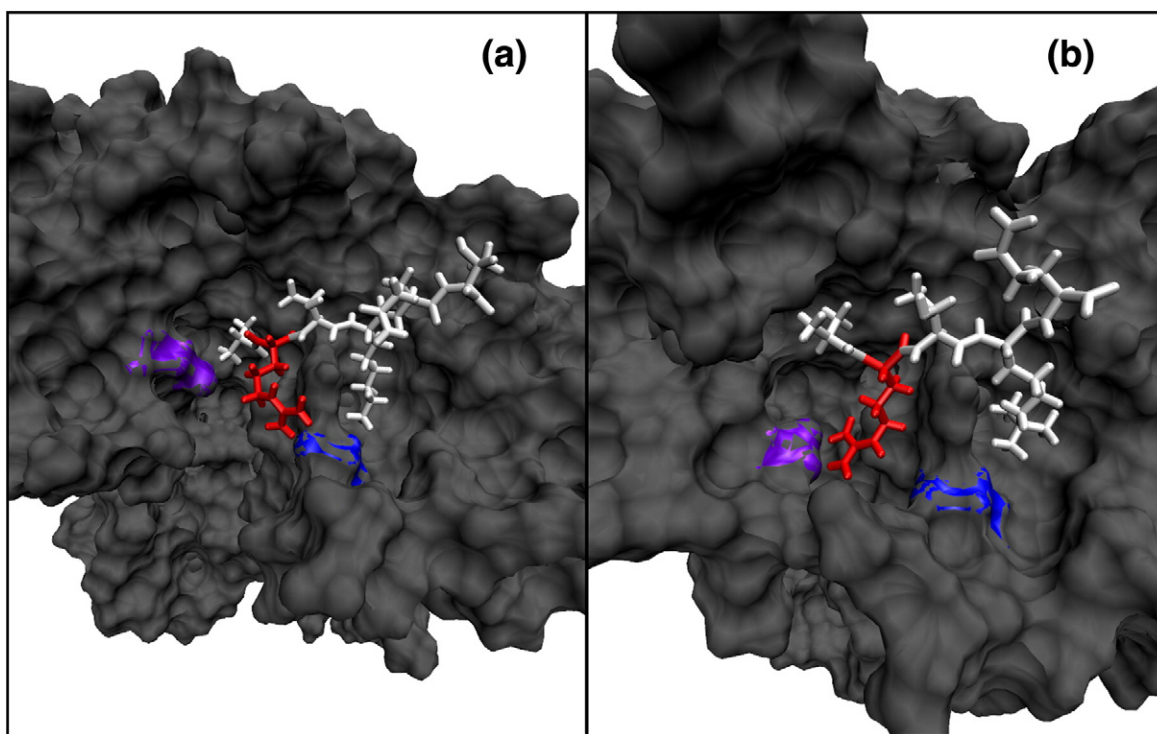


Fig. 5. Surface rendering of the interface between CHD1 and (a) H3K4me1 and between CHD1 and (b) H3K4me3. The CHD1 protein is shown in dark gray, while Glu14[A] (blue) and Glu50[C] (violet) have been highlighted. H3 is drawn by licorice model and is shown in white except for H3R2 (red). Images are generated by VMD.

not near to any residues whose attraction to H3 is magnified when H3K4 is trimethylated, and therefore are not subject to the “pulling” effect that influences Lys63[A], Lys65[A], and Lys53[C].

CHD1 interactions specific to H3(K4me3)(R2me2)

On the other hand, Asp167[A] interacts most favorably with H3(K4me3)(R2me2), apparently because methylation of both H3K4 and H3R2 places H3T3 at the right distance to interact with Asp167[A] via van der Waals forces (−1.74 kcal/mol) and electrostatically (−1.17 kcal/mol). In the complex of CHD1 with H3(K4me3)(R2me2), the residues Asp150[A] and Asp167[A] have large contributions compared to complexes with other histone modifications; it is not immediately clear how this arises on inspection of Fig. 4c. Although it is plausible that the hydroxyl group of H3T3 could hydrogen-bond with Asp167[A], the carbonyl backbone carbon of H3A1 is the only charged histone atom in the close proximity of Asp150[A]. However, the effect of dimethylating H3R2 is exhibited drastically in that Glu50[C] is displaced from the binding site with H3R2me2, and in fact it is Asn52[C] that partakes in the main interaction with H3R2me2, in particular, the C^δ and unmethylated N^H of H3R2me2, which carry large positive and negative charges, respectively. The contribution of Glu50[C] is greatly diminished (to −0.70 kcal/mol) when H3R2 is dimethylated, and visual inspection (see Fig. 4c) in fact shows that H3R2me2 directly disrupts the electrostatic attraction with Glu50[C]. In the case of H3R2me2, Asn52[C] is closer to H3R2 and shows a stronger favorable interaction (−1.32 kcal/mol) with H3R2 than Glu50[C]. The attraction of Asn52[C] to H3R2me2 is hydrophobic in nature (the van der Waals term is −1.62 kcal/mol); the contributions of Trp64[A] and Trp67[A] also increase in the case of H3(K4me3)(R2me2) due to the strengthened van der Waals attractions, −0.45 kcal/mol for Trp67 and −0.33 kcal/mol for Trp64 (see Table S4). However, the interaction strength is hurt for many of the H3-backbone-binding residues, notably Ala32[A], Tyr37[A], His71[A], Thr73[A], and Glu166[A].

CHD1 interactions specific to H3(K4me3)(T3ph)

The phosphorylation of H3T3 generally weakens interactions between H3 and CHD1 residues across the board. The complex of CHD1 with H3(K4me3)(T3ph) is shown in Fig. 4d, where it is clear that no CHD1 residue has a strong interaction with H3 compared to the other modification models, as the phosphorylation of H3T3 leads to the driving of the other H3 side chains from their optimal binding sites. For instance, the binding contribution of His71[A] is weakened (by 1.04 kcal/mol) by phosphorylating H3T3, since this drives the histone peptide out

of a position where H3T3 can bind to His71[A] in its natural manner. In contrast, Glu14[A] interacts most favorably with H3K4me1 and H3(K4me3)(T3ph) because this residue binds through electrostatic interactions with H3K4 and H3R2, which are damped when these histone residues are methylated. In general, the phosphorylation of H3T3 disrupts the binding interaction by preventing the CHD1 protein from assuming a natural conformation for binding to H3 (see Fig. 4d). The electrostatic contribution of Glu50[C] to the binding, at −0.07 kcal/mol, is effectively muted by this H3 modification (see Table S5). It is worth noting that the unfavorable repulsive electrostatic interactions of the CHD1 lysine residues are also weaker for H3(K4me3)(T3ph) than for any other modification of H3.

Simulation procedure reveals atomic-level interactions

The above survey of interactions between the residues of CHD1 protein and H3 peptide indicates that our simulation and data analysis procedures have successfully identified the major atomistic mechanisms by which the chromodomain recognizes the histone code. The protein contains residues that interact strongly with all H3 modifications, such as Trp67[A], Trp64[A], His71[A], and Ala32[A], as well as residues that are designed to interact with only a specific H3 modification, such as Glu166[A], Asp167[A], Asp150[A], and Glu50[C]. Therefore, we have found that hydrophobic and neutral polar CHD1 residues generally contribute to the H3 binding affinity, while negatively charged protein residues are responsible for H3 modification binding specificity. The variance of the contributions of key CHD1 binding residues among the four H3 modifications with crystal structures is shown in Fig. 6. The overall contribution variances are compared with each of the interaction component variances, strongly indicating that residues with a large electrostatic component are responsible for the specificity of CHD1 modification recognition.

The tools of analysis available to us have the capacity to quantitatively determine the magnitude and extent of each residue to the binding interaction, whether it is favorable or unfavorable and whether the interaction arises from electrostatic or van der Waals effects or from a mixture of the two. Furthermore, the dependence of these interactions on the specific histone modification, as well as the contributions of the histone tail residues themselves, is quantifiable with a high degree of accuracy. Visualization software allows for the confirmation of interactions between protein and peptide residues by observing their proximity, as well as the observation of conformational changes in the CHD1 protein that allow distinct residues to recognize specific H3 modifications.

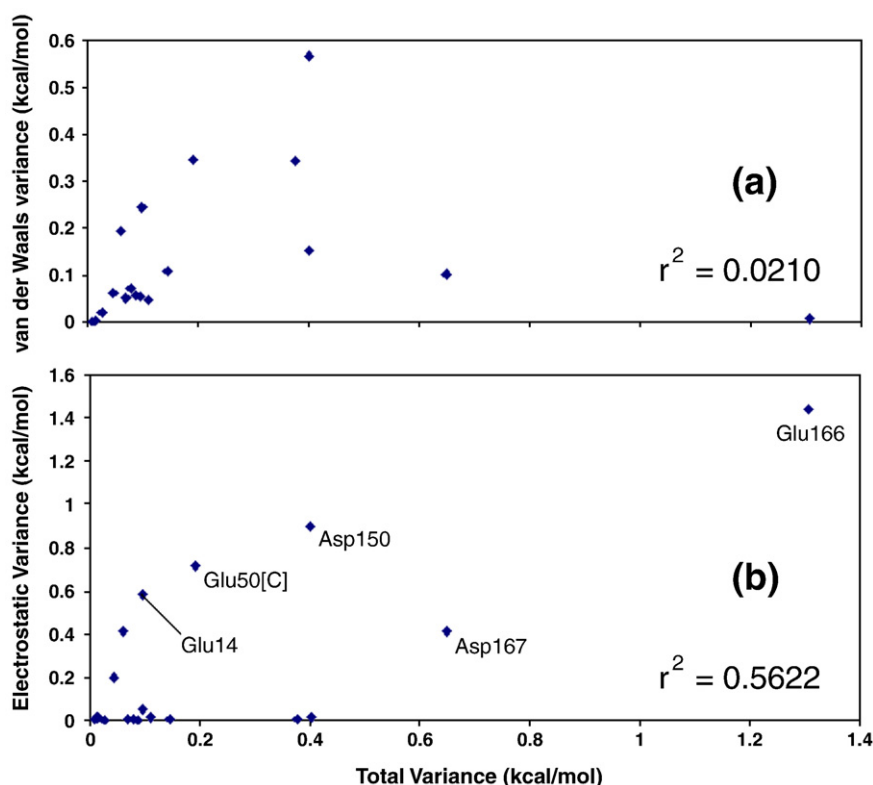


Fig. 6. Comparison of the variance of CHD1 residue overall interaction contributions with the variance of the residue contribution components. The contribution components are van der Waals in (a) and electrostatic and solvation in (b). Each point represents one of the 20 critical residues in CHD1 for binding to the modified H3 tails. The correlation coefficients for each comparison are (a) 0.0210 and (b) 0.5622, indicating that it is the electrostatically interacting residues that are responsible for overall variance in binding to H3 modifications (i.e., specificity).

Relative motion of residues during the trajectory

In order to measure individual residue mobility, we calculated the root-mean-square deviation (RMSD) of all the individual residues for H3 and several selected residues for CHD1 in a manner similar to that described in Ref. 22. This was computed for the CHD1/H3K4me3 complex, using the *ptraj* module in AMBER 9.0, for the production run of the trajectory. The three backbone atoms of each residue (N-C α -C) and all heavy atoms in each side chain were considered. Figure 7 shows the RMSD for six residues in CHD1 critical for binding to H3 (a) and for each of the residues in H3K4me3 (b).

Among the CHD1 residues, the negatively charged residues Asp167[A] and Glu50[C] show the most mobility while the small Ala32[A] is the most stationary of the key binding residues. The high and sharply defined plateaus of the RMSD for Asp167[A] look interesting, as it has been observed to interact with both H3A1 and H3T3 (see Table 4). However, analyzing the interatomic distances between Asp167[A] and both of H3A1 and H3T3

showed nothing extraordinary, and visualization of the trajectory with VMD indicated that the RMSD fluctuation is due only to a rocking motion of the subunit of CHD1 containing Asp167[A]. As Asp167[A] is relatively close to the C-terminus of chain A, it is more free to “rock” than other more interior protein residues. The large residues Trp67[A] and Trp64[A], which are critical for binding to H3, as well as His71[A], have larger RMSD fluctuations than Ala32[A] only because their side chains are larger; in general, they are immobile.

For the H3 residues, the large flexible H3R2 and H3K4me3, which are most responsible for binding to CHD1, and H3Q5 exhibit the most movement, while H3A1, the N-terminal H3 residue, is the least mobile. H3T3 is generally immobile, although its short-interval spikes suggest that it briefly adapts an alternative conformation during the trajectory. The two spikes in the RMSD for H3T3 coincide with points in the trajectory where the RMSD for Asp167[A] rises to a plateau; observation of the trajectory showed that this is merely an effect of H3T3, following along with the rocking motion of Asp167[A].

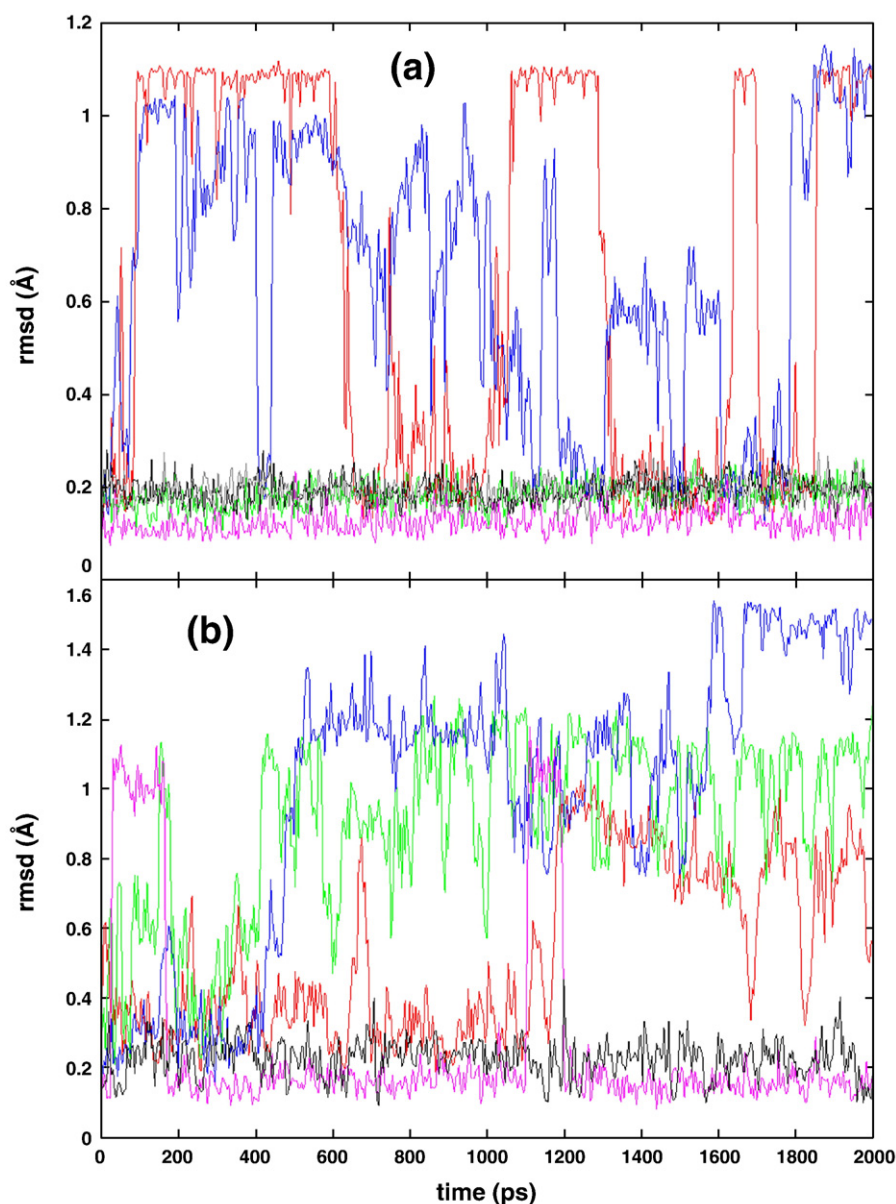


Fig. 7. The RMSD for the backbone atoms of selected individual protein (a) and peptide (b) residues in the CHD1 complex with H3K4me3. Top: Asp167[A] (red), Glu50[C] (blue), Trp64[A] (dark brown), Trp67[A] (silver), His71[A] (green), and Ala32[A] (magenta). Bottom: H3K4me3 (blue), H3Q5 (green), H3R2 (red), H3T3 (magenta), and H3A1 (dark brown).

Prediction of binding energies for new modifications of H3

In addition, simulations were run on two other complexes to investigate the interactions between modifications on K4 and R2/T3: CHD1 with H3R2me2 and CHD1 with H3T3ph. For neither of these systems were either crystal structures or experimental binding affinity data available. The models were constructed by opening the Protein Data Bank (PDB) files for the crystal structures of H3

(K4me3)(R2me2) and H3(K4me3)(T3ph), respectively, and modifying the trimethylated K4 back to standard lysine residues. Then the simulation and data analysis protocol were performed on both models, exactly as done for the four crystal structures as described in the [Materials and Methods](#) section. The calculated binding free energies, given in [Table 1](#), are very close to their respective trimethylated K4 counterparts (for $\epsilon=2.0$). This observation suggests that R2me2 and T3ph are unfavorable for the CHD1 chromodomain and that

they are dominant over methylations on K4. Based on the residue contributions calculated and given in Tables S6 and S7, we hypothesize that the removal of the methyl groups from K4 leads to strengthened hydrophobic contributions from CHD1 residues that interact with the backbone of the peptide, serving to compensate for the favorable van der Waals interactions that arise out of interactions between specific CHD1 residues and trimethylated H3K4. One exception is the electrostatic contribution of Asp150[A], which is more favorable by 0.8 kcal/mol for H3T3ph than for H3(K4me3)(T3ph).

Biologically, the histone modification marks H3R2me2 and H3K4me3 have been demonstrated to be mutually exclusive.²⁹ In our simulation, we found CHD1 to bind to H3R2me2 and H3(R2me2)(K4me3) with virtually equal affinity, as shown in Table 1. The experiment suggests that CHD1 should not favor H3(R2me2)(K4me3) over H3R2me2 in its binding affinity, and our free-energy calculation confirms this. For the case of phosphorylated T3, CHD1 also binds to H3(K4me3)(T3ph) and H3T3ph with nearly equal affinity. These results suggest that modifications of H3 closer to the N-terminus than H3K4 are dominant in determining CHD1 binding affinity, and only when H3R2 and H3T3 are unmodified does the methylation of H3K4 play a significant role.

It is apparent that for low interior dielectric constant ϵ , demethylating H3K4 in H3(K4me3)(R2me2) strengthens the interaction of binding—and yet at the same time, demethylating H3K4 in H3(K4me3)(T3ph) is disruptive to the binding. Table 4 shows the effect on each H3 and CHD1 residue of performing the demethylation for each of the two models. In the case of H3R2me2, demethylating H3K4 weakens the binding most for Trp64[A], Asp150[A], Asp167[A], and Asn52[C], while strengthening it for Thr73[A] and Leu170[A]. For H3T3ph, demethylating H3K4 weakens the binding most drastically for Glu14[A] and Trp64[A], while strengthening it for Tyr37[A], His71[A], and Asp150[A]. Because the models of the CHD1 complexes with H3R2me2 and H3T3ph were not directly derived from the crystal structures, the binding free energies associated with them may not be predicted here with great accuracy. However, these simulations have the capacity to form a basis for future experiments, involving complexation of CHD1 with the modified peptides H3R2me2 and H3T3ph, and measurement of their dissociation constants.

Conclusions

We have performed molecular dynamics (MD) simulations and calculated binding free energies for the human protein CHD1 complexed with six

different histone H3 modifications: H3K4me1, H3K4me3, H3(K4me3)(R2me2), H3(K4me3)(T3ph), H3R2me2, and H3T3ph. As far as we know, this is the first study in which the binding free energy of a protein to a histone tail peptide has been analyzed by decomposing the contributions by residue. With the proper choice of dielectric constant, the binding free energies agree well with the experimental results, which are available for the first four of these complexes. We have estimated the extent to which each individual residue contributes to the binding interaction and the nature of each contribution (electrostatic or van der Waals); these calculations are supported by visual inspection of the model. From our study, the following key points can be taken: firstly, the important residues in CHD1 can generally be divided into two classes—the residues that contribute to the binding affinity to all H3 modifications and the residues that contribute to binding specificity to certain H3 modifications; second, the CHD1 residues that are responsible for specificity tend to be negatively charged; third, the CHD1 residues that bind well to all forms of the H3 peptide tend to be hydrophobic; and finally, demethylating H3K4 in either the H3(K4me3)(R2me2) or the H3(K4me3)(T3ph) complex should have little effect on the strength of the binding interaction with the CHD1 protein. Future experiments designed to obtain dissociation constants for these complexes would test the accuracy of these predictions.

It is to be noted that these simulations are not expected to be a definitive authority on modeling the behavior of the CHD1/H3 complexes. At the time these simulations were conducted, there were a limited number of chromodomain/histone complex crystal structures available in the PDB. As recently as the last year, several additional crystal structures of chromodomain proteins bound to histone tails have become available in the PDB. Therefore, there is much scope for more detailed analysis to refine our understanding of the binding and recognition of histone modifications by the appropriate reader proteins and to provide additional insight into the interaction between chromodomains and histone peptides. Nevertheless, it is encouraging that this methodology can currently produce binding free-energy data that are both in agreement with experiment and consistent with visual inspection. A possible avenue of future experiments might involve performing point mutations of CHD1 residues observed to be critical in this study for binding to H3 and investigating how these mutations affect the binding of the protein to a histone H3 modification. Mutations in the residues that give rise to binding specificity could conceivably compromise the ability of the chromodomain to differentiate between the modifications H3K4me3 and H3K4me1.

Materials and Methods

Modeling procedure

All of the CHD1/histone complex structures were obtained from the PDB.³⁰ For H3K4me1, the PDB entry is 2B2V; for H3K4me3, it is 2B2W; for H3(K4me3)(R2me2), it is 2B2U; and for H3(K4me3)(T3ph), it is 2B2T. The module of *LEaP* in the AMBER 9.0³¹ software package was used to prepare the models for MD simulation. The methylated lysine and arginine and the phosphorylated threonine residues were not recognized by the standard residue library of *LEaP* and were treated as nonstandard residues. The charges for all nonstandard residues were calculated by *Gaussian 03*,³² using the HF/6-31G* basis set, and were subsequently fitted with RESP.³³ The partial charges on all nonstandard residue atoms as calculated by *Gaussian 03* and RESP are listed in Table S1a, b, and c. The nonstandard residues were all handled by *Gaussian 03* individually and capped³⁴ by ACE and NME residues so they could be treated as complete molecules. The FF03 force field³⁵ was used for all interatomic parameters for which it had values. For those parameters not defined within the FF03 force field libraries, typically in the regions of histone modifications, the GAFF force field³⁶ was used to create special “.frmod” files that were loaded into *LEaP* to supply the missing parameters. Each CHD1/histone complex was placed in a rectangular periodic box of TIP3P water molecules.³⁷ The dimensions of the box extended 10 Å past the solute system on all sides. Na⁺ counter ions were added to neutralize the system for each model when necessary.

MD simulation

Each system was minimized in two stages. In the first stage, the water molecules and counterions were relaxed while organic molecules were restrained by a harmonic force of 100 kcal/mol Å². In the second stage, the restraints were released, and the entire system was minimized. Each stage consisted of 400 steps of steepest descent followed by 1600 steps of conjugate gradient minimization. After the minimization stages, the systems were heated from 0 to 300 K over 30 ps. The temperature of 300 K and pressure of 1 atm were maintained by a Berendsen thermostat and barostat with a coupling time³⁸ of 0.2 ps. The SHAKE algorithm was used to restrain all bonds involving hydrogen.³⁹ At the start of the heating portion of the MD, we assigned initial temperatures randomly according to the Maxwell-Boltzmann distribution. The systems were then equilibrated for 200 ps followed by production runs lasting 2 ns.

Trajectory analysis and free-energy calculation

The RMSD for each system during the equilibration and production run phases of the MD was calculated by the *ptraj* module in the AMBER package. As a representative sample, the RMSD for CHD1/H3K4me3 is shown in Fig. S3. The interval for the binding free-energy calculation was chosen for the latter half of each production run, between 1.0 and 2.0 ns, based on the stabilization of the RMSD after about 1000 ps. MM-GBSA^{40–42} was used to

calculate the binding free energies for each CHD1/histone complex using the *mm-pbsa* module in AMBER 9.0, according to the general procedure given in Refs. 43 and 44. The binding free energies, ΔG_{bind} , were calculated from 100 snapshots taken at equal intervals between the 1.0- and 2.0-ns marks of the production run trajectory. ΔG_{bind} was calculated according to:

$$\Delta G_{\text{bind}} = \langle \Delta G_{\text{complex}} \rangle - (\langle \Delta G_{\text{protein}} \rangle + \langle \Delta G_{\text{peptide}} \rangle)$$

Here, $\langle G_{\text{complex}} \rangle$, $\langle G_{\text{protein}} \rangle$, and $\langle G_{\text{peptide}} \rangle$ are the respective individual free energies of the complex, protein, and peptide, and each term is calculated by summing the contributions from electrostatic potential E_{elec} , van der Waals potential E_{vdw} , and solvation free energy G_{solv} . In turn, G_{solv} is the sum of the polar contribution G_{polar} and the nonpolar contribution G_{nonpolar} to the solvation free energy:

$$\begin{aligned} G_{(\text{complex} / \text{protein} / \text{peptide})} &= E_{\text{elec}} + E_{\text{vdw}} + G_{\text{solv}} \\ G_{\text{solv}} &= G_{\text{polar}} + G_{\text{nonpolar}} \end{aligned}$$

G_{polar} was calculated by setting the value of IGB to 2, activating the generalized Born parameters of Onufriev *et al.*⁴⁵ G_{nonpolar} was estimated as 0.0072 times the solvent-accessible surface area, as measured by the LCPO method.⁴⁶ A grid size of 0.5 Å was used to solve the Poisson-Boltzmann equation, and the probe radius⁴⁷ was set to 1.4 Å. Three different values (1.0, 2.0, and 4.0) were used for the interior dielectric constant ϵ (which applies to the region of the protein/histone binding pocket) in separate free-energy calculations, while the external dielectric constant was given throughout as 80.0. Conformational entropy was derived in a separate calculation (see Table S8; more positive values indicate a more favorable binding interaction). Because the standard deviations of each entropy contribution term are comparable in size to the terms, it is arguable that the entropy differences between the models are statistically insignificant. Liu and Duan found considerable entropy effects on free-modified H3 tail peptides.²⁴ However, in our case, the H3 tails are complexed with a large protein, greatly restricting their available conformations and narrowing the entropy differences.

To investigate the molecular basis of recognition specificity, we conducted component decomposition of the binding free energy. The contributions to the binding free energy were also broken down for individual residues and by component (van der Waals, electrostatic, and solvation) for each of the complexes. Decompositions were performed by invoking the DECOMP module in the MM-GBSA program within AMBER. Each H3 or CHD1 residue was evaluated according to its overall contribution to the binding interaction, as opposed to a pair-wise analysis of every CHD1/H3 residue combination.

Supplementary materials related to this article can be found online at [doi:10.1016/j.jmb.2010.12.030](https://doi.org/10.1016/j.jmb.2010.12.030)

Acknowledgements

This work was partially supported by the National Institutes of Health grant R01GM085188 (to

W.W.). Some of the MD simulations were conducted in a computer cluster of the Center for Theoretical Biological Physics at University of California, San Diego.

References

1. Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D. & Patel, D. J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat. Struct. Mol. Biol.* **14**, 1025–1040.
2. Jenuwein, T. & Allis, C. D. (2001). Translating the histone code. *Science*, **293**, 1074–1080.
3. Allis, C. D., Jenuwein, T., Reinberg, D. & Caparros, M. L. (2006). *Epigenetics*. Cold Spring Harbor Laboratory Press, Woodbury, NY.
4. Berger, S. L. (2007). The complex language of chromatin regulation during transcription. *Nature*, **447**, 407–412.
5. Kouzarides, T. (2007). Chromatin modifications and their function. *Cell*, **128**, 693–705.
6. Chang, B., Chen, Y., Zhao, Y. & Bruick, R. K. (2007). JMJD6 is a histone arginine demethylase. *Science*, **318**, 444–447.
7. Feinberg, A. P. (2007). Phenotypic plasticity and the epigenetics of human disease. *Nature*, **447**, 433–440.
8. Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R. & Young, R. A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. *Cell*, **130**, 77–88.
9. Ma, J. C. & Dougherty, D. A. (1997). The cation- π interaction. *Chem. Rev.* **97**, 1303–1324.
10. Burley, S. K. & Petsko, G. A. (1988). Weakly polar interactions in proteins. *Adv. Protein Chem.* **39**, 125–189.
11. Marfella, C. G. A. & Imbalzano, A. N. (2007). The Chd family of chromatin remodelers. *Mutat. Res.* **618**, 30–40.
12. Bouazoune, K., Mitterweger, A., Langst, G., Imhof, A., Akhtar, A., Becker, P. B. & Brehm, A. (2002). The dMi-2 chromodomains are DNA binding modules important for ATP-dependent nucleosome mobilization. *EMBO J.* **21**, 2430–2440.
13. Akhtar, A., Zink, D. & Becker, P. B. (2000). Chromodomains are protein-RNA interaction modules. *Nature*, **407**, 405–409.
14. Flanagan, J. F., Mi, L. Z., Chruszcz, M., Cymborowski, M., Clines, K. L., Kim, Y. *et al.* (2005). Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature*, **438**, 1181–1185.
15. Nielsen, A. L., Oulad-Abdelghani, M., Ortiz, J. A., Remboutsika, E., Chambon, P. & Losson, R. (2001). Heterochromatin formation in mammalian cells: interaction between histones and HP1 proteins. *Mol. Cell*, **7**, 729–739.
16. Zhao, T., Heyduk, T., Allis, C. D. & Eissenberg, J. C. (2000). Heterochromatin protein 1 binds to nucleosomes and DNA *in vitro*. *J. Biol. Chem.* **275**, 28332–28338.
17. Lusser, A., Urwin, D. L. & Kadonaga, J. T. (2005). Distinct activities of CHD1 and ACF in ATP-dependent chromatin assembly. *Nat. Struct. Mol. Biol.* **12**, 160–166.
18. Woodage, T., Basrai, M. A., Baxevanis, A. D., Hieter, P. & Collins, F. S. (1997). Characterization of the CHD family of proteins. *Proc. Natl Acad. Sci. USA*, **94**, 11472–11477.
19. Sims, R. J., III, Belotserkovskaya, R. & Reinberg, D. (2004). Elongation by RNA polymerase II: the short and long of it. *Genes Dev.* **18**, 2437–2468.
20. Konev, A. Y., Tribus, M., Park, S. Y., Podhraski, V., Lim, C. Y. & Emelyanov, A. V. (2007). CHD1 motor protein is required for deposition of histone variant H3.3 into chromatin *in vivo*. *Science*, **317**, 1087–1090.
21. Sims, R. J., III, Chen, C. F., Santos-Rosa, H., Kouzarides, T., Patel, S. S. & Reinberg, D. (2005). Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. *J. Biol. Chem.* **280**, 41789–41792.
22. Jiang, Y. K., Zou, J. W., Wu, Y. Q., Zhang, N., Yu, Q. S. & Jiang, Y. J. (2008). Molecular dynamics simulation on HP1 protein binding by histone H3 tail methylation and phosphorylation. *Int. J. Quantum Chem.* **109**, 746–755.
23. Grauffel, C., Stote, R. H. & DeJaegere, A. (2010). Force field parameters for the simulation of modified histone tails. *J. Comput. Chem.* **31**, 2434–2451.
24. Liu, H. & Duan, Y. (2008). Effects of posttranslational modifications on the structure and dynamics of histone H3 N-terminal peptide. *Biophys. J.* **94**, 4579–4585.
25. Mellor, J. (2006). It takes a PHD to read the histone code. *Cell*, **126**, 22–24.
26. Wang, W. & Kollman, P. A. (2000). Free energy calculations on dimer stability of the HIV protease using molecular dynamics and a continuum solvent model. *J. Mol. Biol.* **303**, 567–582.
27. Theoretical and Computation Biophysics Group, University of Illinois at Urbana-Champaign, <http://www.ks.uiuc.edu/Research/vmd/>.
28. Varshney, A., Brooks, F. P. & Wright, W. V. (1994). Computing Smooth Molecular Surfaces. *IEEE Comput. Graph. Appl.* **14**, 19–25.
29. Guccione, E., Bassi, C., Casadio, F., Martinato, F., Cesaroni, M., Schuchlautz, H. *et al.* (2007). Methylation of histone H3R2 by PRMT6 and H3K4 by an MLL complex are mutually exclusive. *Nature*, **449**, 933–937.
30. Berman, H. M., Westbrook, J. D., Feng, Z., Gilliland, G. L., Bhat, T. N., Weissig, H. *et al.* (2000). The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242.
31. Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M. *et al.* (2005). The Amber biomolecular simulation programs. *J. Comput. Chem.* **26**, 1668–1688.
32. Gaussian 03, Revision C.02, M. J. Frisch and J. A. Pople *et al.*, Gaussian Inc., <http://www.gaussian.com/>.
33. Bayly, C. I., Cieplak, P., Cornell, W. D. & Kollman, P. A. (1993). A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **97**, 10269–10280.
34. Cieplak, P., Cornell, W. D., Bayly, C. & Kollman, P. A. (1995). Application of the multimolecule and multi-conformational RESP methodology to biopolymers: charge derivation for DNA, RNA, and proteins. *J. Comput. Chem.* **16**, 1357–1377.

35. Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W. *et al.* (2003). A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **24**, 1999–2012.
36. Wang, J. M., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. (2004). Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174.
37. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935.
38. Berendsen, H. J. C., Postma, J. P. M., Vangunsteren, W. F., Dinola, A. & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690.
39. Ryckaert, J., Ciccotti, G. & Berendsen, H. J. C. (1977). Numerical integration of cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comput. Phys.* **23**, 327–341.
40. Qiu, D., Shenkin, P. S., Hollinger, F. P. & Still, W. C. (1997). The GB/SA continuum model for solvation. A fast analytical method for the calculation of approximate Born radii. *J. Phys. Chem. A*, **101**, 3005–3014.
41. Kollman, P. A., Massova, I., Reyes, C. M., Kuhn, B., Huo, S., Chong, L. *et al.* (2000). Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **33**, 889–897.
42. Wang, W., Donini, O., Reyes, C. M. & Kollman, P. A. (2001). Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein–ligand, protein–protein, and protein–nucleic acid noncovalent interactions. *Annu. Rev. Biophys. Biomol. Struct.* **30**, 211–243.
43. Hou, T., McLaughlin, W. A. & Wang, W. (2008). Evaluating the potency of HIV-1 protease drugs to combat resistance. *Proteins*, **71**, 1163–1174.
44. Hou, T., Chen, K., McLaughlin, W. A., Lu, B. & Wang, W. (2006). Computational analysis and prediction of the binding motif and protein interacting partners of the Abl SH3 domain. *PLoS Comput. Biol.* **2**, 0046–0055.
45. Onufriev, A., Bashford, D. & Case, D. A. (2000). Modification of the generalized Born model suitable for macromolecules. *J. Phys. Chem. B*, **104**, 3712–3720.
46. Weiser, J., Shenkin, P. S. & Still, W. C. (1999). Approximate solvent-accessible surface areas from tetrahedrally directed neighbor densities. *Biopolymers*, **50**, 373–380.
47. Pellegrini, E. & Field, M. J. (2002). A generalized-born solvation model for macromolecular hybrid-potential calculations. *J. Phys. Chem. A*, **106**, 1316–1326.