

Cation- π interactions in CREBBP bromodomain inhibition: an electrostatic model for small-molecule binding affinity and selectivity

Wilian A. Cortopassi[‡], Kiran Kumar[‡], and Robert S. Paton^{*}

Chemistry Research Laboratory, University of Oxford, 12 Mansfield Road, Oxford OX1 3TA, UK

Abstract: CREBBP bromodomains, epigenetic “reader” proteins that recognize acetylated histone lysine residues, are a current target for cancer therapy. We show that experimental CREBBP binding affinities of small-molecules with aromatic or heteroaromatic functional groups are strongly influenced by a cation- π interaction with a positively charged arginine residue. For a series of fifteen 5-isoxazolylbenzimidazole derivatives, the strength of this non-covalent interaction is directly related to improvements in binding to CREBBP. The aromatic substituents’ inductive and resonance effects are not obviously correlated with observed structure and affinity relationships. In contrast, a coulombic electrostatic model can quantitatively predict the interaction strength. We have assessed different Molecular Mechanics (MM) and Quantum Mechanics (QM) descriptions of the protein-ligand interaction. Quantitative models for binding affinity were generated from: (1) Poisson Boltzmann Surface Area (MM-PBSA) and Generalized Born Surface Area (MM-GBSA) scoring functions that incorporated the entire ligand and (2) QM-complexation energies and (3) Electrostatic Potential Surface values (ESPs) that analyzed the varying aromatic group. A linear relationship between QM-computed ESP values is established for the cation- π interaction strength, and gives the best correlation ($R^2 = 0.84$) with experimental binding affinities. This model also ranks ligand affinity most accurately ($r_s = 0.91$) from the models tested. Consideration of the electrostatic potential in response to the local effects of substituents in addition to that of the aromatic ring is necessary to understand and describe the interaction with the cationic guanidinium ion. This leads to an improved understanding and the ability to quantitatively predict the magnitude of non-covalent interactions in the CREBBP active site.

Introduction

Chromatin remodeling and histone modifications by epigenetic proteins regulate several biological processes, including DNA replication and repair.¹ These enzymes interact with histone lysine and arginine residues and can be classified as either histone epigenetic writers, erasers, or readers due to their ability to add, remove, or recognize (e.g. methyl, acetyl) functional groups.² Targeting enzymes that mediate histone modifications has been established as a promising strategy for the treatment of diseases related to cell function, such as cancer, male infertility and adult obesity.³ While the majority of small molecules that have been clinically approved target proteins involved in the writing and erasing regulator processes, inhibitors of epigenetic readers, such as bromodomains, are still in early stages of clinical trials.²

The bromodomain motif is composed of a left handed bundle of four α -helices (αZ , αA , αB , and αC) linked by two loops (ZA and BC) whose function is the recognition (i.e. *reading*) of acetylated histones. Their binding site is largely hydrophobic and characterized by a conserved asparagine/tyrosine residue that forms a hydrogen bond with the carbonyl of the acetylated lysine. Bromodomain-containing protein (BCP) inhibitors have been studied as promising targets in the treatment of cancer and inflammation, mostly by targeting the bromodomain extra C-terminal domain (BET) family.⁴ The design of small-molecule inhibitors for bromodomains outside of the BET family (e.g. the cAMP response element-binding protein (CREB) binding protein (CREBBP) bromodomain) with high selectivity for a single target is particularly challenging,⁵ due to high sequence similarity among bromodomains.⁶ Recently, Conway and co-workers synthesized small molecule inhibitors with nanomolar affinities for non-BET family BCPs.⁷ The most promising compound, a dihydroquinoxalinone derivative, (*R*)-2, (**Figure 1**) successfully inhibits the binding of CREBBP bromodomain to chromatin in U2OS cells. (*R*)-2 was also selective for CREBBP over seven BCPs that varied phylogenetically. The creation of an “induced fit” pocket was observed in the co-crystal structure, driven by the formation of a stabilizing non-covalent interaction between the ligand’s aromatic group and an arginine residue (Arg1173) that is maintained during 100 ns of

MD simulation.⁷ Due to the high basicity of the guanidine side-chain, arginine residues are predominantly charged in proteins at (and above) physiological pH – this interaction therefore formally constitutes a cation– π interaction between a positively charged amino acid side-chain and an electron-rich π -system. As we show in this work, the binding strength is affected by changes in the aromatic electrostatic potential, justifying the characterization as a cation– π , rather than a π - π (i.e. dispersion-dominated) interaction.

Bromodomains have high sequence homology; however, the Arg1173 residue is specific to CREBBP and therefore provides a target for selective small-molecule inhibitors. This conserved arginine differs structurally from BET bromodomains, such as BRD4, which contain an aspartic acid in the same spatial location on α B (**Figure 1**), thus providing a basis for selectivity. Nevado and Caflisch⁸ have exploited this difference in polarity by incorporating a negatively-charged benzoate group, designing selective inhibitors with low micromolar affinity for CREBBP. Favorable electrostatic interactions with Arg1173 provide selectivity over other bromodomains. Further optimization to nanomolar binding affinity was accomplished by improving the polar interactions with this arginine residue.⁹ Fedorov and co-workers¹⁰ found that a benzoaxazepine-based inhibitor, I-CBP112 (**Figure 1C**, PDB ID: 4NR6), interacts with the cationic Arg1173 residue in CREBBP; in contrast, no such interaction was observed between this compound and BET proteins. Brennan and co-workers¹¹ have shown that 5-isoxazolylbenzimidazole derivatives establish H-bonding interactions with the isoxazole functional group (the acetylated lysine mimic), and that the attached aromatic (3-chloro-4-methoxyphenyl) group interacts closely with Arg1173 (**Figure 1A**, PDB ID: 4NR7). Modification of the aromatic substituents or using a heteroaromatic group led to a significant impact on measured CREBBP affinity. Relative binding affinities were obtained from thermal melting experiments with differential scanning fluorimetry. Ligands with a greater binding affinity form more strongly stabilized complexes with the protein, requiring a higher temperature than the uncomplexed protein (ΔT_m) to unfold. Direct correlation between thermal shift and binding assays have been demonstrated previously,¹¹ and 5-

isoxazolybenzimidazole derivatives with greater values of ΔT_m (higher than 8 °C) have lower IC_{50} (half maximal inhibitory concentration) values against CREBBP.

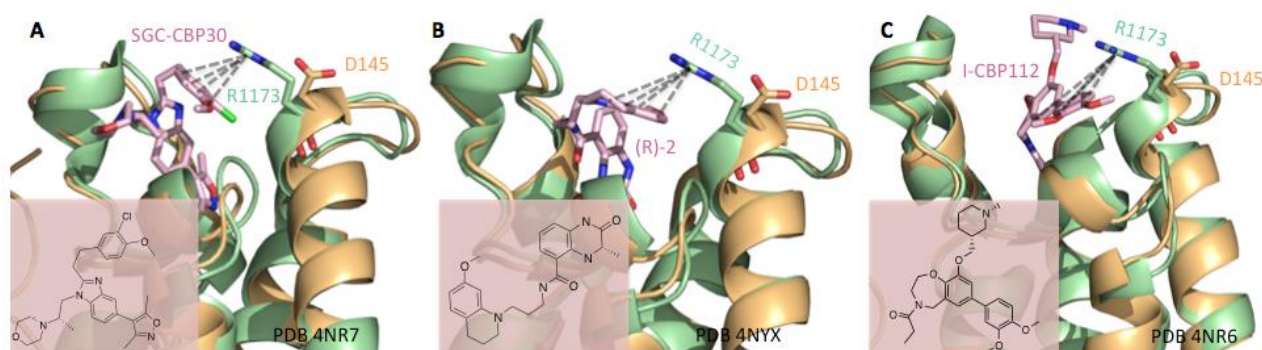


Figure 1. Close contacts between ligands and a conserved arginine residue are a driving force in binding selectivity for CREBBP bromodomains (green) over BRD4 (yellow, PDB ID: 3UVW) for a diverse set of inhibitors including A) 5-isoxazolybenzimidazole derivative, SGC-CBP30 (**L11**); B) dihydroquinoxalinone derivative, (*R*)-2; and C) benzoaxazepine-based inhibitor, I-CBP112.

In this paper, we investigate the structure-affinity relationship of CREBBP binding for a series of structurally related small-molecules (**Chart 1**) for which experimental data is available. These fifteen ligands share a common 5-isoxazolybenzimidazole core, while the attached aromatic/heteroaromatic group is varied. Thus far, attempts to rationalize this data has precluded simple explanations based on expected inductive/resonance effects upon the π -electron density: although the incorporation of electron donating groups (EDGs, such as OMe) increased binding strength, electron withdrawing groups (EWGs, such as halides) also led to further improvements. We have examined the ability of different computational techniques to predict the relative binding affinities: in particular, we focus on quantifying the role of electrostatic effects. The strong correlation of this coulombic term with observed affinities confirms that the aromatic-arginine interaction is polar in origin (i.e. cation- π rather than dispersive). Previous experimental¹² and computational^{12a, 12d, 13} studies have shown analysis of thermal shift data is a valuable approach for identifying the most promising protein inhibitors, when exploring a data set of small molecules with similar physicochemical properties.¹⁴ Different MM and QM-based approaches included: (1) QM-

complexation energies using Density Functional Theory (DFT); (2) Electrostatic Potential Surface (ESP) values of substituted aromatic groups computed with DFT; and (3) Binding energies from MM-PBSA and MM-GBSA scoring functions.

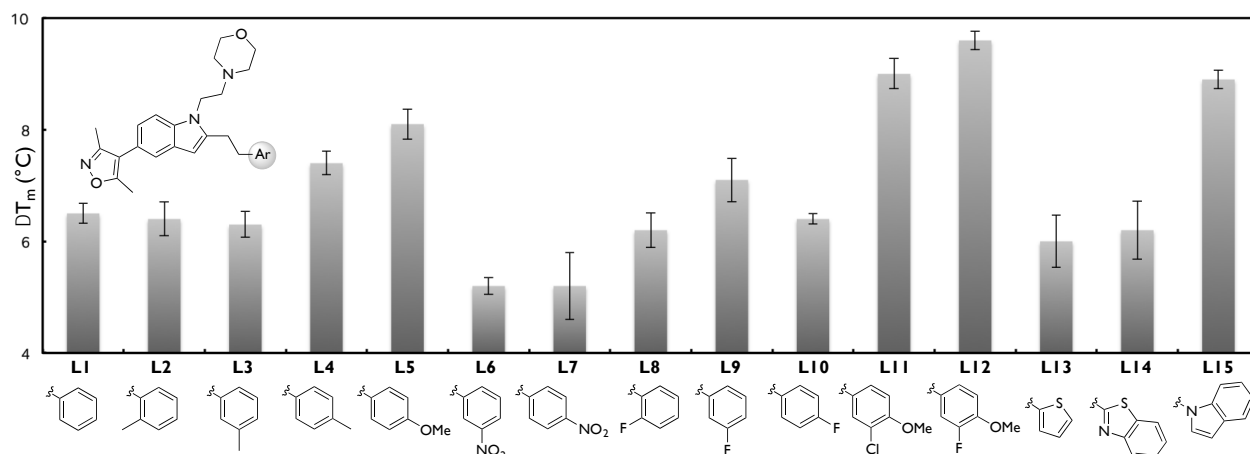


Chart 1. CREBBP bromodomain binding assay for fifteen 5-isoxazolylbenzimidazole derivatives with variable aromatic (Ar) groups; ΔT_m data taken from ref. 11).

Methods

Since the discovery of the cation- π interaction in the 1980s,¹⁵ computational methods have been used to quantify and understand the basis for its magnitude in both gas and condensed phases. We present a combination of quantum and classical computational methods to quantify the arene-arginine interaction present in CREBBP binding. Molecular dynamics (MD) simulations have been previously used for describing the recognition of acetylated histones by bromodomains,¹⁶ as well as for the design of new bromodomain inhibitors.^{7, 17} Wheeler and Houk¹⁸ showed that, despite the neglect of polarization of the aryl π -system by substituents, MM-based approaches using fixed atom-centered point charges are capable of describing electrostatically dominated non-covalent interactions. Coulombic terms from traditional force fields, such as OPLS, correlate well with *ab initio* computations of cation- π interactions in biological environments.¹⁹ Classical descriptions of the cation- π interaction, however, are highly dependent on the point charges used.²⁰

Long-time MD simulations of bromodomains can accurately compute *absolute* binding free energies of small-molecules: Biggin and co-workers demonstrated previously unprecedented accuracy using free energy perturbation (FEP) based on X-ray structures to within 1.0 kcal/mol²¹ of experiment. MD simulations also provide insight into the dynamics of bromodomains with bound substrates and inhibitors.²² However, due to their high computational cost, we employed short timescale (5 ns) MD simulations combined with the MM-PBSA and MM-GBSA end-state free energy calculations.²³ Such approaches correlate with experimental binding values to different proteins ($R^2 > 0.6$),²⁴ and successfully rank a set of inhibitors by binding affinity (Spearman rank-order correlation coefficient, $r_s > 0.6$).^{24b, 24c, 25}

The non-additivity of non-covalent interactions may be problematic for traditional force fields²⁶ and so we turned to electronic structure calculations in the form of (dispersion-corrected) DFT calculations. Previously both QM and QM/MM techniques have been used to quantify protein-ligand binding energies;²⁷ however, due to their high computational cost, there are still limitations in their application in the analysis of large data sets or to flexible protein-ligand complexes due to sampling limitations. Thus the use of truncated model systems is commonplace in QM-studies of enzymatic binding and catalysis, including only the interactions that require QM-based approaches for an accurate description. Dougherty²⁸ demonstrated a strong correlation between interaction energies computed at the Hartree-Fock (HF)/6-31G(d,p) level of theory for model Na⁺-tryptophan complexes and the half maximal effective concentration (EC₅₀) values for acetylcholine at the nicotinic receptor measured *in vivo*. In another study, Dougherty used the methylammonium ion: benzene complex to mimic the interaction between lysine and phenylalanine in protein structures and calculated binding energies at the HF/6-31+G(d) level of theory, since this method agrees well (<1 kcal/mol error) with experimental estimates of polar interactions in solution.²⁹ Cation- π interactions are dominated by through-space electrostatic effects rather than the inductive and mesomeric polarization of aromatic rings by their substituents.³⁰ Houk and Wheeler^{20d} have shown substituent effects upon the molecular ESP, calculated at the M05-2X/6-311+G(2df,2p) level of

theory, result from direct through-space interactions rather than polarization, and that these direct interactions account for the variance in interaction energies between substituted aromatic rings and cations. Direct, through-space electrostatic interactions with the substituents, rather than polarization of the benzene π -system, are responsible for differences in binding strength. However, non-electrostatic effects such as induction still play a role for an accurate description of cation- π interaction energies, especially when comparing structures with very different aromatic cores.³¹ The cation- π interaction is also of growing interest in the development of small-molecule asymmetric catalysts, where QM calculations (typically at the DFT level of theory) have been used to obtain stereochemical models for the observed selectivities.³²

QM-complexation energies. All QM calculations were performed using Gaussian 09.³³ Coordinates for the guanidinium-inhibitor complex were taken from PDB ID: 4NR7. We extracted a cluster model consisting of the ligand aromatic group and the Arg1173 guanidinium group, where the attached CH₂-groups of both were terminated with an H atom (**Figure 2**). In a standard theozyme/cluster-model approach, non-hydrogen guanidinium atoms as well as two carbon atoms from the inhibitor were fixed to maintain orientations within the protein environment. This was necessary to avoid substantial reorganization, which would be sterically prohibited in the full enzyme:ligand complex. The positions of all other atoms were optimized with the meta-generalized gradient approximation (GGA) functional TPSS and valence triple- ζ def2-TZVPP basis set. Boys-Bernadi counterpoise corrections³⁴ were used for quantification of the complexation energies comparing different DFT descriptions (the hybrid B3LYP and the meta-GGA functionals M06-2X and TPSS) and basis sets. The M06-2X functional of Truhlar and Zhao³⁵ has been used recently for accurate calculations of C₆H₅X \cdots Na⁺ binding energies relative to the benchmark CCSD(T) values.³⁶ In all our calculations we used Grimme's D3 density-independent correction for atom-pairwise dispersion with a Becke-Johnson damping function at short range.³⁷ The use of dispersion corrections with the B3LYP functional is encouraged since this fails to describe London Dispersion

Interactions.^{31c} The inclusion of dispersion energies in cation– π interactions have been evaluated by Kim.^{31b} Although such contribution has a negligible effect on Na^+/π and Li^+/π complexes, they are important for accurate energy estimation of complexation between organic cations and π -systems.

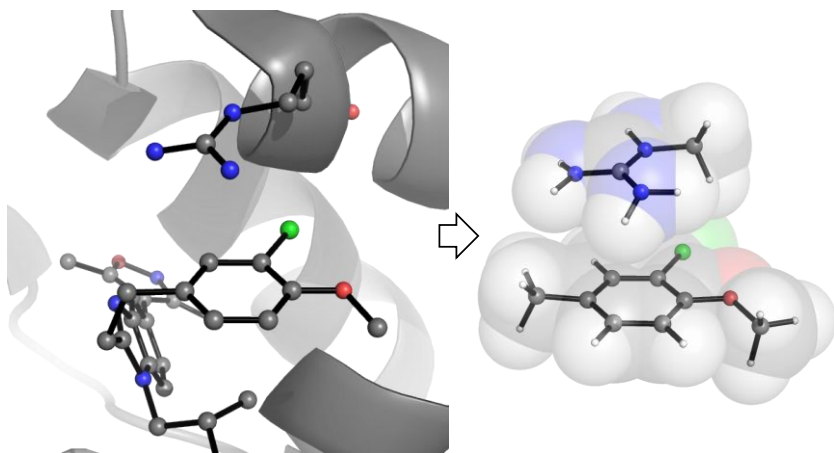


Figure 2. Cluster model of the aromatic:Arg1173 interaction taken from PDB 4NR7 used in QM studies.

Qualitative and quantitative analysis of molecular electrostatic potentials. Geometries of the aromatic and heteroaromatic substituents of ligand **L1-L15** were fully optimized using the B3LYP/6-31G+(d) level of theory. ESPs were obtained from DFT densities at the B3LYP/6-31G(d,p) level of theory (gas and solution-phase); with B3LYP/6-31G++(d,p) and M06-2X/6-31G(d,p). Calculations used the Solvation Model Density (SMD) and SM8 Minnesota implicit solvation models³⁸ of diethyl ether ($\epsilon = 4.24$) to mimic the hydrophobic environment of the binding site of CREBBP, with little effect upon the resulting ESP.³⁹ Although basis set size can have a profound effect upon computed atomic populations, particularly where basis functions approach linear-dependence, the ESP is more robust and we observed no large sensitivities across a range of basis sets including diffuse and polarization functions. ESP maps were generated with *Spartan 14*.⁴⁰ In addition to generating ESP maps onto the molecular isosurface, we evaluated a horizontal slice through the molecular ESP parallel to the plane of each ring, at a distance of 3.5 Å (**Figure 3a** and **3b**) above (or equivalently, below) the ring center. This distance approximates the crystallographic arene – arginine separation: using a distance of 4.0 Å (**Figure S1**) resulted in qualitatively similar

results. This picture more closely represents the potential experienced by the guanidinium cation, which is not in van der Waals contact with the arene. These horizontal slices through the molecular ESP are qualitatively related to the ESP mapped onto the molecular surface (Figure 3c), although are more representative of the potential experienced by interacting molecules at a distance of 3-4 Å away, as is the case for this enzyme:ligand complex.

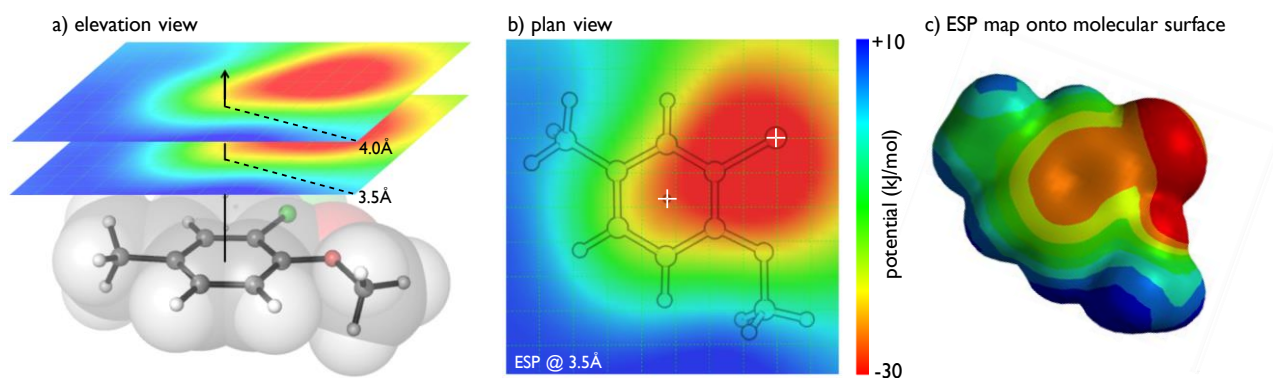


Figure 3. 2D ESP horizontal slice spanning -30 – 10 kJ/mol generated at 3.5 Å and 4.5 Å above the aromatic group of **L11** in a) elevation view and b) plan view; c) ESP map onto the molecular surface of the aromatic group of **L11** spanning -80 – 80 kJ/mol.

As seen in **Figure 3** the more negative regions of the ESP map are not found above the aromatic ring of **L11**, but rather outside, towards the chloro- and methoxy-substituents. Regions closer to the chlorine rather than the oxygen atom have lower (i.e. more negative) values of ESP. Initially, we attempted to correlate ligand binding affinities using a single value of the molecular ESP taken above the aromatic ring. This approach has been successfully applied to monoatomic cations.^{20d} However, such attempts were unsuccessful: due to its size we reasoned that the guanidinium cation, unlike simple cations, experiences an extended region of the arene's electrostatic potential. This led us to consider multivariate models using the electrostatic potential at more than one position. In the spirit of *Occam's razor* we found that a linear model using ESP values at two points gave a good correlation with experimental binding data for ligands **L1-L15**. This model uses the ESP values above the aromatic ring center, and also above the substituent with

the most negative local region of ESP (**Figure 3b**). Linear regression (see **SI** for details) with these two “molecular descriptors” against experimental thermal shift data was used to obtain the coefficients of **Eq. 1**

$$\Delta\Delta T_m = a \cdot ESP_{cation} + b \cdot ESP_{substituent} \quad \text{Eq. 1}$$

The statistical performance was not significantly enhanced by using ESP values at other positions (**Table S2**). With more variables (i.e. ESP values) a greater fit is possible, although to prevent overfitting we used only two independent variables. A third, constant term can be added to the linear model to predict absolute, rather than relative, values of ΔT_m .

Molecular Dynamics simulations. The AMBER12 package⁴¹ was used to simulate all fifteen protein:ligand complexes. Generation of the small-molecule force field with Amberff12SB was performed with Antechamber.⁴² Ligand partial charges were determined with the RESP ESP charge Derive (RED) Server⁴³ using the RESP-A1B methodology. In this approach B3LYP/cc-pVTZ molecular ESPs are computed involving the Connolly surface⁴⁴ algorithm with the integral equation formalism polarizable continuum solvent model (IEFPCM; $\epsilon=4$, ether as solvent).⁴⁵ To ensure planarity for cases in which the ligand contained methoxy or nitro substituents the force constant for dihedral angles was set to 3.6 kcal/mol. CREBBP complexes were created using PDB ID: 4NR7 as a template. Glutamate and aspartate residues were assigned as negatively charged, and lysine and arginine as positively charged. All crystallographic waters were preserved and *tLeap*⁴¹ was used to add hydrogens to the protein. To solvate and neutralize the system we used a 10 Å octahedral box using TIP3P⁴⁶ composed of ~7,100 water molecules and two Na⁺ ions.

Preceding the MD production, 1000 steps of steepest descent minimization were performed while constraining non-heavy complex atoms, followed by a second minimization of 1000 steps for the entire system. This was followed by 1000 ps of equilibration and 5 ns of production. These were carried out using the NPT ensemble and Langevin thermostat at a constant temperature of 310 K.

The SHAKE algorithm was used to constrain all bonds containing hydrogen atoms and allowed the use of a time step of 2 fs. The cut-off of 12 Å was used for long-range interactions. Visualization of the simulation required the use of Visual Molecular Dynamics (VMD) 1.9.2.⁴⁷

Stability of the systems was analyzed by calculating average root-mean-square-deviation (RMSD) values of C_α atoms relative to crystal structures. Frequency calculations of cation- π interactions were performed by determining the distance between the center of mass of the aromatic ring and the Arg1173 guanidinium ion over all frames. Instances in which the distance was shorter than 6 Å were defined as a cation- π event and recorded over 5 ns. This distance based cutoff was first established by Dougherty, showing that 99% of cation- π structures extracted from the PDB satisfying geometric and energetic-based criteria lie inside this value.¹⁹

Classical free energy and electrostatics calculations. The *MMPBSA.py* script developed by Roitberg and co-workers²³ and distributed with AmberTools12⁴¹ was used to derive binding free energies with both the MM-GBSA and MM-PBSA continuum solvation methods from ensembles derived from MD simulation. These techniques have been used to characterize protein-protein interactions and estimation of binding energies, with increasing popularity since its development.⁴⁸ The application of MM-PBSA and MM-GBSA to MD-derived ensembles has been carried out for bromodomain receptors with a diverse set of inhibitors and was similarly successful in calculating binding energies ($R^2 > 0.6$) as in previous studies.^{24a, 49} Energy values were averaged over 100 frames, corresponding to snapshots taken at 50 ps intervals for 5 ns of simulation time. A salt concentration of 0.15 M was used to parallel physiological conditions. The interior (ϵ_{int}) and exterior (ϵ_{ext}) dielectric constants were set to 1 and 80, respectively; alternative values were examined with inferior results (see SI). Default atomic radii were read from the original parameter files generated to run the MD production. Coulombic electrostatic interactions between the aromatic groups of ligands **L1-L15** and arginine guanidinium were calculated at intervals of 10 ps over the

total simulation time using NAMDenergy Plugin 1.4.⁵⁰ The average and standard deviation energy values were computed for each system and correlated against experimental ΔT_m values.

Discussion

Cation- π interactions are important for binding and stable during long time-scale MD simulation. The CREBBP bound X-ray crystal structure of 5-isoxazolylbenzimidazole ligand **L11** shows a close contact between the chloromethoxyphenyl ring and Arg1173 in the hydrophobic binding pocket (**Figure 1A**). We used MD simulation to investigate the temporal stability of this interaction for ligands **L1** and **L11** by performing 100 ns MD simulations in solution using the Amber force field (Figure 4). Starting from the X-ray structure of **L11** in CREBBP (4NR7), we observed intermolecular distances between the chloromethoxyphenyl group and the guanidinium shorter than 6 Å, the distance cutoff used for cation- π interactions,¹⁹ during 98% of the simulation (**Figure 4C**). This illustrates the stability of this interaction and its relevance to recognition of this inhibitor in the protein-binding site. The starting point for simulations with **L1** was taken from the X-ray crystal structure (4NR5) in which the phenyl group of the ligands phenyl group is oriented away from the arginine guanidinium group (**Figure 4A**). Nevertheless, spontaneous formation of a cation- π interaction (as judged by an intermolecular separation closer than 6 Å) occurred after 2 ns of simulation time. This interaction was maintained during 74% of the simulation (**Figure 4C**). These results are in accordance with previous MD simulations showing the temporal stability of the cation- π interaction formed by dihydroquinoxalinone derivative (*R*)-2 in CREBBP (**1**).⁷ Stabilization of both complexes was seen with an average RMSD of 1.58 Å (**L11**) and 1.79 Å (**L1**) with respect to starting structure (**Figure 4B**).

These results indicate the propensity for cation- π interaction formation in both cases, involving one of the strongest binders experimentally, **L11**, but also for **L1**, where close contact was observed for the majority of the MD simulation despite greater flexibility. We therefore built ligand-CREBBP complexes for **L1-L15** based on the 4NR7 crystal structure. We used different

MM and QM-based approaches to quantify these interactions and correlate computed interaction energies (or free energies) with experimental binding affinities.

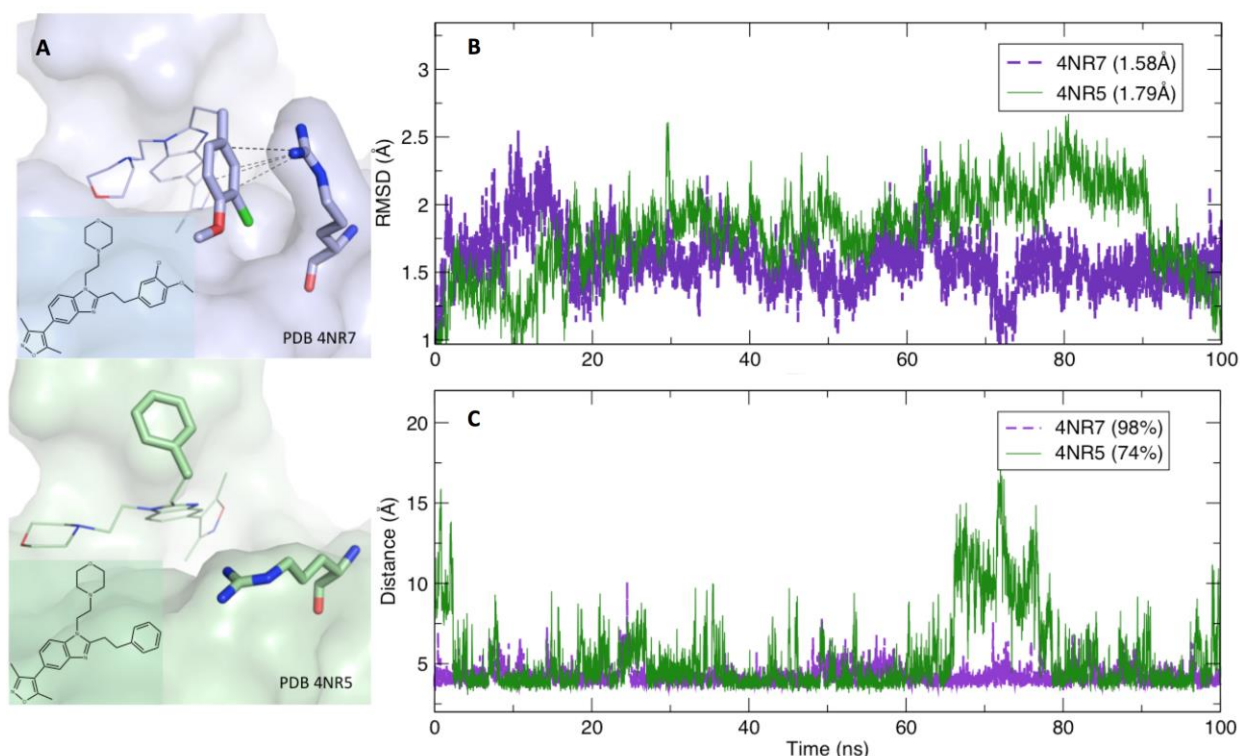


Figure 4. (A) The X-ray crystal structure of the CREBBP-**L11** complex shows a clear cation- π interaction; that of the CREBBP-**L1** complex does not; (B) following 100 ns MD simulation the two complexes are stable, as shown by average RMSD values; (C) both ligands form active site cation- π interactions for a majority of simulation time, using a distance cut-off of 6 Å⁵¹ between guanidinium and aromatic group.

Use of QM energetics and electronic-descriptors for quantification of guanidinium-benzene interactions. Previous work has shown the ability of quantum chemistry to quantify the interaction strength between functionalized aromatic rings and (mainly Na⁺) cations.²⁸⁻²⁹ Houk and Wheeler^{20d} and Dougherty⁵² have observed a strong correlation between QM-computed ESP values above substituted benzenes and the interaction energies of the cation- π complexes. We have evaluated the ability of QM computed arginine:arene binding energies to directly describe the experimental trend in binding affinity, as well the correlation with ESP values above the aromatic ring.

Coordinates of guanidinium and chloro-methoxyphenyl ring of **L11** in CREBBP (PDB ID: 4NR7) were used as a starting point for our QM calculations as shown in **Figure 2**. The interaction energy was computed from the difference between the counterpoise-corrected energy³⁴ of the complex and those of the separated guanidinium and arene. Different DFT functionals (TPSS, B3LYP-D3, M06-2X) and basis sets (triple- and quadruple- ζ) were tested, but had only a marginal effect upon correlation between computed binding energies and experimental affinities (**Table S1**). M06-2X/cc-pVQZ binding energetics showed the best correlation with experiment, being able to rank the ligands according to their binding affinity with $r_s = 0.79$ and $R^2 = 0.64$ (**Chart 2**). The strongest binding ligands with the greatest experimental ΔT_m values (**L5**, **L11**, **L12** and **L15**) have the most favorable DFT complexation energies (**Table 1**), ranging from -12.6 to -15.2 kcal/mol. For the indole group (**L15**) this interaction is stronger than for other complexes, a result of a very negative ESP, as previously reported by Dougherty⁵¹ for the tryptophan side-chain, and expected given the conjugated 10 π -electron system. In accordance with experimental data, strongly-interacting ligands **L5**, **L11** and **L12** have EDGs such as methoxy groups, whereas **L6** and **L7**, with an EWG (nitro-group), have the least favorable interaction energies. However, inductively EWGs such as halogens, as in **L11** (Cl) and **L12** (F), result in a further increase in experimental binding affinity which is reproduced computationally. For example, a more favorable QM binding energy was obtained for **L11** (-13.4 kcal/mol) compared to **L5** (-12.6 kcal/mol). These results suggest that the guanidinium:arene interaction contributes to the experimental affinity and that its magnitude can be used to rank the series of ligands **L1-L15**. The interaction between aromatic ligands and Arg1173 is now apparent from a number of X-ray structures, such as in the recent fragment-based discovery of a selective and cell-active benzodiazepinone CREBBP inhibitor disclosed by Genentech.⁵³

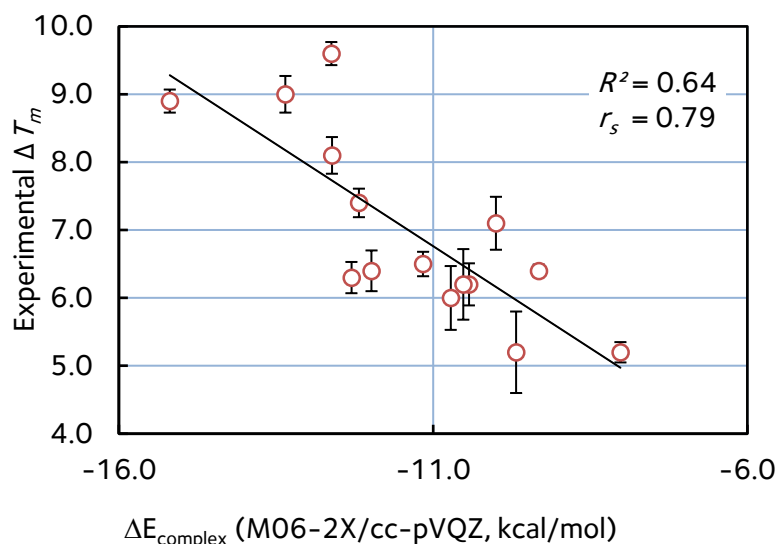


Chart 2. Correlation between experimental ΔT_m and counterpoise-corrected M06-2X/cc-pVQZ complexation energies; experimental error bars shown.

Analysis of the B3LYP/6-31G(d,p) ESP mapped on to molecular isosurfaces (**Figure 5**) confirms that the indole group has the most negative potential, which would be expected to interact most strongly with a guanidinium ion. The QM-complexation energy is indeed most favorable. As pointed out by Wheeler, it should be noted that while a negative ESP value often results from an electron-rich region, the ESP depends not just on the local electron density but upon the entire molecule.¹⁸ The least active compounds contain EWG nitro-substituents: the ESP is least negative above the ring for these ligands; however, it is also the case that a very negative local region of the ESP surface is present directly above the nitro-substituents. EDG methoxy-substitution leads to an increase in the negative potential both above the ring and the substituent itself. This is important, as molecular cations, such as the guanidinium ion extend outside the region of the aromatic ring and interact with both regions of space. Fluorine atoms cause a reduction in ESP above the aromatic ring, but also lead to a corresponding increase in ESP values above these substituents: **L12** is the strongest binder experimentally, which suggests that the cation can experience this negative region of the ESP.^{20d} Finally, heteroaromatic thiazole (**L13**, $\Delta T_m = 6.0$ °C) and benzothiazole (**L14**, $\Delta T_m = 6.2$ °C) have a less negative ESP than an unsubstituted phenyl ring ($\Delta T_m = 6.5$ °C), which result in

slightly lower values of binding affinities. Structures with more negative regions of ESP across their entire surface are among the stronger binders experimentally, and according to QM calculations (**Figure 5**). The most potent ligands have either a combination of a negative potential above the π -system and also one substituent (**L5**, **L11**, **L12**) or simply above an extended (i.e. bicyclic) π -system (**L15**).

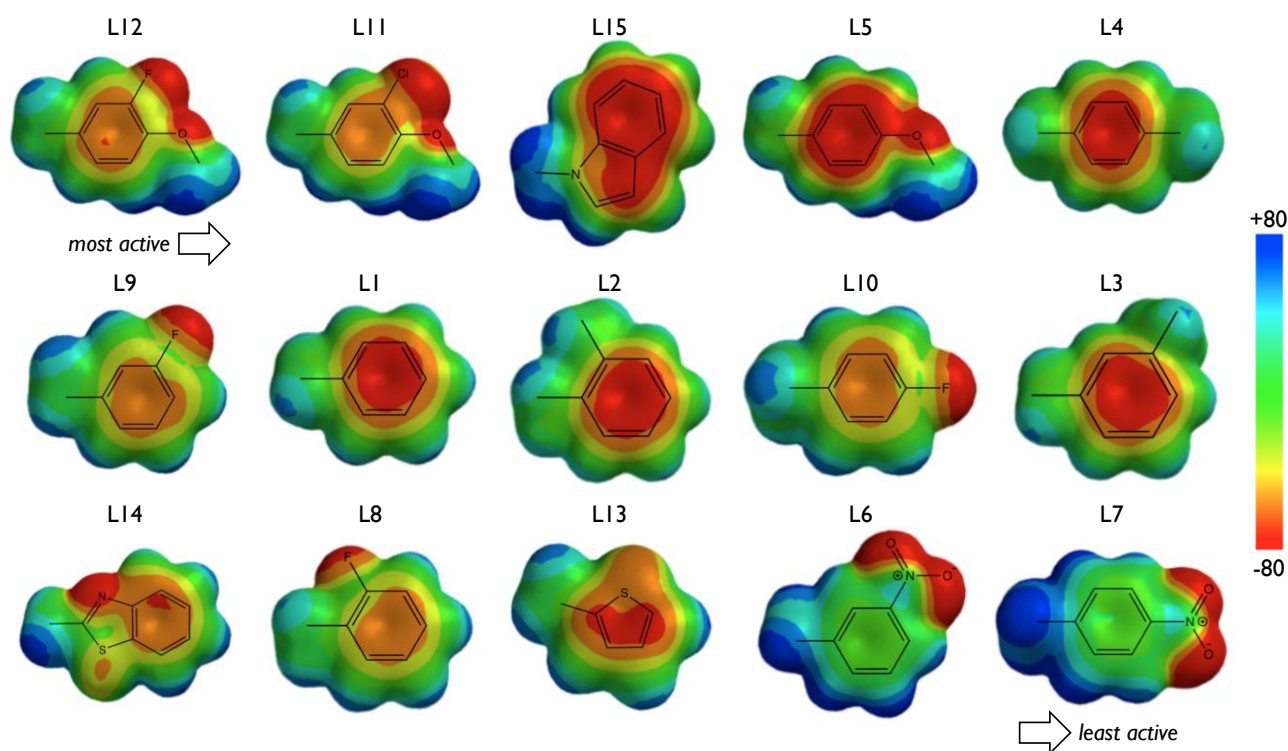


Figure 5. B3LYP/6-31G(d,p) molecular ESP maps (isosurface value 0.002) for aromatic/heteroaromatic groups of **L1-L15**, ordered according to experimental binding affinity for CREBBP.

We next looked to develop a quantitative model based on ESP values for the electrostatic interaction for each ligand with arginine. Values above the ring center do not enable efficient ranking of binding affinity ($r_s = 0.47$, **Table 1**). However, when considering two different points, one above the aromatic ring and another one above the substituent with the most negative potential, we built a linear regression model in the form of **Eq. 1** that was able to rank the inhibitors according to ΔT_m values with good accuracy ($r_s = 0.91$, $R^2 = 0.84$) with no clear outliers (RMSE = 0.52 °C) (**Chart 3**). The linear regression coefficients are -307.7 (ESP values directly above the π -system)

and -238.7 (directly above the substituent).⁵⁴ The fact that both coefficients have the same sign is physically reasonable and indicates the additive nature of the electrostatic contributions in determining complex stability: their ratio (56:44) shows that there is also a nearly equal contribution from both terms. As discussed above, the ESP above the ring is a molecular property, influenced by through-space contributions from substituents (and vice-versa) such that it is not possible to assign the two terms to wholly separate contributions from aromatic ring and substituents. Both functionalities influence the magnitude of the ESP at each point considered.

The use of different DFT descriptions (meta-GGA and hybrid functionals), solvation models and basis sets (**Table S2**), led to small changes in the statistical performance of the bivariate electrostatic model. Relatively cheap B3LYP/6-31G(d,p) ESP calculations can be used without loss of accuracy. Recently, Wheeler⁵⁵ showed that the molecular dipole moment correlates well with theoretical binding affinities between small heterocycles, although a poor correlation was observed for the stacking of fused rings. In our analysis, the dipole moment showed low correlation (**Table S3**, $R^2 < 0.15$) with experimental binding affinities for both small (**L1-L13**) and fused ring systems (**L14-L15**). Nevertheless, our results in this biological setting support the underlying principle outlined by Wheeler: cation- π interactions are not significantly affected by π -resonance effects, but direct electrostatic interactions involving substituents are important.⁵⁶ The recent development of selective CREBBP/P300 inhibitors derived from a different scaffold, but which nonetheless display a cation- π interaction involving a 1,2-dimethoxyphenyl group,⁵⁷ suggests that this model can be applied to understand SAR data for other classes of small-molecule inhibitors.

Ligand	$\Delta E_{\text{complex}}^{\text{a}}$ (kcal/mol)	ESP _{ring} ^b (Hartree)	ESP _{subst.} (Hartree)	Model ^c ΔT_m (°C)	Error ^d (°C)
L1	-11.16	-0.0122	0	6.3	-0.2
L2	-11.98	-0.0128	-0.0014	6.9	+0.5
L3	-12.30	-0.0126	-0.0011	6.7	+0.4
L4	-12.18	-0.0127	-0.0020	7.0	-0.4
L5	-12.61	-0.0134	-0.0074	8.5	+0.4
L6	-9.68	0.0083	-0.0104	4.8	-0.4

L7	-8.02	0.0142	-0.0117	4.9	-0.3
L8	-10.43	-0.0086	-0.0070	6.9	+0.7
L9	-10.00	-0.0082	-0.0086	7.2	+0.1
L10	-9.32	-0.0080	-0.0089	7.2	+0.8
L11	-13.35	-0.0094	-0.0158	9.3	+0.3
L12	-12.62	-0.0099	-0.0124	8.6	-1.0
L13	-10.72	-0.0118	N/A ^e	6.2	+0.2
L14	-10.52	-0.0113	N/A ^e	6.1	-0.1
L15	-15.19	-0.0175	N/A ^e	8.0	-0.9
$r_s = 0.79$		$r_s = 0.47$	$r_s = 0.91$		

Table 1. Correlation of experimental affinity data against QM computations. ^aM06-2X/cc-pVQZ complexation energies; ^bB3LYP/6-31G* ESP values 3.5 Å above the ring-center and substituent with the most negative value. For ligands **L14** and **L15**, the ring with the most negative potential above it was considered; ^cPredicted thermal shift data obtained after bivariate linear regression (**Eq. 1**) using ESP values; ^dModel error relative to experimental ΔT_m values; ^eFor comparison, these ligands were considered as unsubstituted conjugated π -systems.

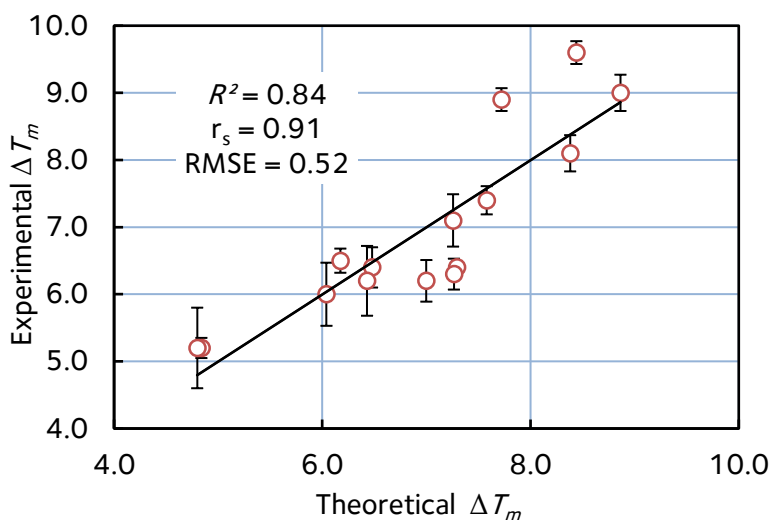


Chart 3. Correlation between experimental and predicted ΔT_m using a parameterized electrostatic model (**Eq. 1**).

Molecular Dynamics simulations for calculation of binding free energies in CREBBP. Since our analysis up to this point had been based on static structures, we used classical MD to investigate the temporal flexibility of the arene:arginine interactions. Force field descriptions of weak non-covalent interactions are often competitive with more demanding levels of theory due to explicit parameterization,⁵⁸ although electrostatic effects are sensitive to the atomic charges used. Following

5 ns simulation of the 15 CREBBP complexes in water, binding free energies between arginine and aromatic group were analyzed (**Figure 6**). Pleasingly, the strongest binding ligands in experiment have more negative binding free energies (LHS, Figure 6), which tends to less negative values for the weaker binding ligands (RHS, Figure 6). The quantitative ability to successfully rank ligands according to ΔT_m data did not vary in regards to the implicit solvation model used (PB versus GB). MM-PBSA performed the best with $r_s = 0.59$ followed by MM-GBSA with $r_s = 0.53$. Ligands **L5**, **L11**, **L12** and **L15** are ranked between the top six strongest binders according to the MM-PBSA approach, in line with the thermal shift data. Changing the internal dielectric constant (ϵ_{int}) from 1 to 4 in the MM-PBSA methodology resulted in a decrease of the correlation to $r_s = 0.38$ (**Table S4**). Although the use of a $\epsilon_{int} = 4$ has been suggested by other authors⁵⁹ for a better correlation with experimental data, this means weakening the electrostatics role in the binding, which would reduce the strength of the interaction with the positively charged arginine of CREBBP. We also analyzed the electrostatic contribution to the interaction energy and frequency of observed cation- π interactions. Coulombic electrostatic energies were able to successfully rank the ligands ($r_s = 0.72$, **Figure 6**). This is consistent with our QM studies showing that these interactions are electrostatically-dominated. In addition to the through-space electrostatic contributions associated with ring substituents identified in QM and ESP studies, these simulations highlight additional close-contacts of ring-substituents with the arginine (alkyl) side-chain. All the analyzed ligands showed the existence of a close-contact - a cation- π interaction based on electrostatic interaction energies - for more than half of the simulation. **L6**, **L9**, **L11** and **L12** had occupancy values close to 1. The formation of this interaction for almost 100% of the simulated time does not necessarily mean strongest guanidinium-benzene interactions. All these compounds have heteroatoms in the meta-position, favoring $-\text{CH}\cdots\text{X}$ interactions with the alkyl chain of the arginine residue. **L7**, containing the inductively and mesomerically electron-withdrawing nitro group in the *para*-position, showed the lowest occupancy (0.65), as expected from a weaker interaction with the guanidinium. This is consistent with our QM computations of ESP and interaction energies and

shows that weakening the cation– π interaction may also result in different structural conformations of the protein in solution.

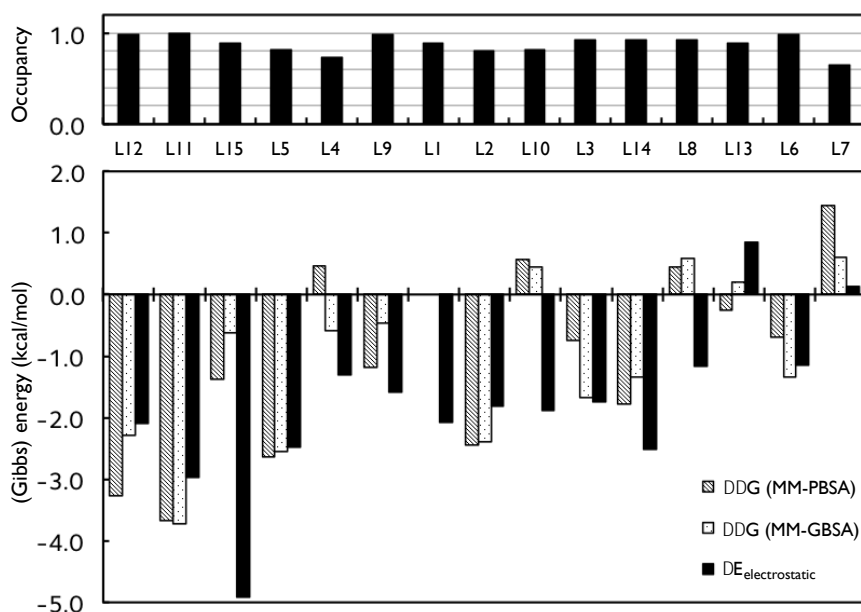


Figure 6. Binding free energies and arginine electrostatics calculated for ligands **L1-L15** using MM-PBSA and MM-GBSA analysis over 5 ns of MD simulation. The ligands are ordered from strongest to weakest experimental affinity.

Conclusion

Elucidation of inhibitory mechanisms for epigenetic reader proteins is important in the design of anti-cancer therapeutics. Computational models have been able to account for small-molecule structure-affinity relationships in binding to the CREBBP bromodomain, which is challenging for chemical intuition alone. Using QM and MM approaches, we have identified electrostatically-dominated cation– π interactions between ligand aromatic groups and an arginine residue in the active site of CREBBP. QM approaches were able to successfully rank the 5-isoxazolylbenzimidazole derivatives in line with their experimental affinities ($r_s = 0.79$). MD-based scoring functions (MM-PBSA and MM-GBSA) performed reasonably, with MM-PBSA giving a slightly better correlation than MM-GBSA ($r_s = 0.59$ and $r_s = 0.53$). For both MM and QM methodologies, analysis focused on through-space electrostatic effects gave the best quantitative

description, showing that measured binding affinities is strongly correlated with the cation- π interaction strength.

An electrostatic model to predict binding affinity outperforms QM and MM approaches. This model gives a good correlation with binding affinity data ($R^2 = 0.84$) and was able to successfully rank the ligands ($r_s = 0.91$) with higher accuracy than alternative methods. Due to the size of the guanidinium cation, it is necessary to consider a ligand's electrostatic potential away from the ring center in addition to the above the ring itself. The resulting additive model, although neglecting the effects of the π -polarization, provides insight into the nature of CREBBP inhibition and demonstrates the utility of QM-computed molecular descriptors in the analysis of biological non-covalent interactions. The relatively modest cost of DFT-computed electrostatic potential maps enables predictions of ligand affinity in cation- π dependent binding sites to be performed quickly and routinely, and presents the possibility of e.g. scaffold-hopping based on similar ESP values instead of steric descriptors.

ASSOCIATED CONTENT

Supporting Information.

Cartesian coordinates of DFT-optimized guanidinium:arene complexes and absolute energies; functional and basis set comparison of counterpoise-corrected binding energetics; ESP values and statistical model fitting protocol; RESP charges of 5-isoxazolylbenzimidazole derivatives used in MD simulations; raw data for MM-PBSA and MM-GBSA binding energies of 5-isoxazolylbenzimidazole derivatives.

AUTHOR INFORMATION

Corresponding Author

* robert.paton@chem.ox.ac.uk

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. ‡These authors contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

W. A. C. is supported by a Science Without Borders (CAPES) scholarship. K. K. is supported by a World Bank Education Grant. We thank NVIDIA for the donation of Tesla GPUs through the academic partnership program and acknowledge the use of the EPSRC UK National Service for Computational Chemistry Software (CHEM870), the University of Oxford Advanced Research Computing (ARC) facility (<http://dx.doi.org/10.5281/zenodo.22558>) and the dirac cluster (EP/L015722/1) in carrying out this work. We thank Michael Brand for constructive comments.

REFERENCES

1. Cortopassi, W. A.; Kumar, K.; Duarte, F.; Pimentel, A. S.; Paton, R. S., Mechanisms of histone lysine-modifying enzymes: A computational perspective on the role of the protein environment. *J. Mol. Graph. Mod.* **2016**, *67*, 69-84.
2. Brand, M.; Measures, A. M.; Wilson, B. G.; Cortopassi, W. A.; Alexander, R.; Hoss, M.; Hewings, D. S.; Rooney, T. P.; Paton, R. S.; Conway, S. J., Small molecule inhibitors of bromodomain-acetyl-lysine interactions. *ACS Chem. Biol.* **2015**, *10* (1), 22-39.
3. (a) Uemura, M.; Yamamoto, H.; Takemasa, I.; Mimori, K.; Hemmi, H.; Mizushima, T.; Ikeda, M.; Sekimoto, M.; Matsuura, N.; Doki, Y.; Mori, M., Jumonji domain containing 1A is a novel prognostic marker for colorectal cancer: in vivo identification from hypoxic tumor cells. *Clin Cancer Res.* **2010**, *16* (18), 4636-4646; (b) Okada, Y.; Scott, G.; Ray, M. K.; Mishina, Y.; Zhang, Y., Histone demethylase JHDM2A is critical for Tnp1 and Prm1 transcription and spermatogenesis. *Nature* **2007**, *450* (7166), 119-123.
4. Filippakopoulos, P.; Knapp, S., Targeting bromodomains: epigenetic readers of lysine acetylation. *Nat. Rev. Drug. Discov.* **2014**, *13* (5), 337-356.
5. (a) Pan, C.; Mezei, M.; Mujtaba, S.; Muller, M.; Zeng, L.; Li, J.; Wang, Z.; Zhou, M. M., Structure-guided optimization of small molecules inhibiting human immunodeficiency virus 1 Tat association with the human coactivator p300/CREB binding protein-associated factor. *J. Med. Chem.* **2007**, *50* (10), 2285-2288; (b) Borah, J. C.; Mujtaba, S.; Karakikes, I.; Zeng, L.; Muller, M.; Patel, J.; Moshkina, N.; Morohashi, K.; Zhang, W.; Gerona-Navarro, G.; Hajjar, R. J.; Zhou, M. M., A small molecule binding to the coactivator CREB-binding protein blocks apoptosis in cardiomyocytes. *Chem. Biol.* **2011**, *18* (4), 531-541; (c) Gerona-Navarro, G.; Yoel, R.; Mujtaba, S.; Frasca, A.; Patel, J.; Zeng, L.; Plotnikov, A. N.; Osman, R.; Zhou, M. M., Rational design of cyclic peptide modulators of the transcriptional coactivator CBP: a new class of p53 inhibitors. *J. Am. Chem. Soc.* **2011**, *133* (7), 2040-2043; (d) Galdeano, C.; Ciulli, A., Selectivity on-target of bromodomain chemical probes by structure-guided medicinal chemistry and chemical biology. *Future Med. Chem.* **2016**, <http://dx.doi.org/10.4155/fmc-2016-0059> [Epub ahead of print]
6. Hewings, D. S.; Fedorov, O.; Filippakopoulos, P.; Martin, S.; Picaud, S.; Tumber, A.; Wells, C.; Olcina, M. M.; Freeman, K.; Gill, A.; Ritchie, A. J.; Sheppard, D. W.; Russell, A. J.; Hammond, E. M.; Knapp, S.; Brennan, P. E.; Conway, S. J., Optimization of 3,5-dimethylisoxazole derivatives as potent bromodomain ligands. *J. Med. Chem.* **2013**, *56* (8), 3217-3227.
7. Rooney, T. P.; Filippakopoulos, P.; Fedorov, O.; Picaud, S.; Cortopassi, W. A.; Hay, D. A.; Martin, S.; Tumber, A.; Rogers, C. M.; Philpott, M.; Wang, M.; Thompson, A. L.; Heightman, T. D.; Pryde, D. C.; Cook, A.; Paton, R. S.; Muller, S.; Knapp, S.; Brennan, P. E.; Conway, S. J., A

- series of potent CREBBP bromodomain ligands reveals an induced-fit pocket stabilized by a cation- π interaction. *Angew. Chem. Int. Ed.* **2014**, *53* (24), 6126-6130.
8. Xu, M.; Unzue, A.; Dong, J.; Spiliotopoulos, D.; Nevado, C.; Caflisch, A., Discovery of CREBBP Bromodomain Inhibitors by High-Throughput Docking and Hit Optimization Guided by Molecular Dynamics. *J. Med. Chem.* **2016**, *59* (4), 1340-1349.
 9. Unzue, A.; Xu, M.; Dong, J.; Wiedmer, L.; Spiliotopoulos, D.; Caflisch, A.; Nevado, C., Fragment-Based Design of Selective Nanomolar Ligands of the CREBBP Bromodomain. *J. Med. Chem.* **2016**, *59* (4), 1350-1356.
 10. Hammitzsch, A.; Tallant, C.; Fedorov, O.; O'Mahony, A.; Brennan, P. E.; Hay, D. A.; Martinez, F. O.; Al-Mossawi, M. H.; de Wit, J.; Vecellio, M.; Wells, C.; Wordsworth, P.; Muller, S.; Knapp, S.; Bowness, P., CBP30, a selective CBP/p300 bromodomain inhibitor, suppresses human Th17 responses. *Proc. Natl. Acad. Sci. USA* **2015**, *112* (34), 10768-10773.
 11. Hay, D. A.; Fedorov, O.; Martin, S.; Singleton, D. C.; Tallant, C.; Wells, C.; Picaud, S.; Philpott, M.; Monteiro, O. P.; Rogers, C. M.; Conway, S. J.; Rooney, T. P.; Tumber, A.; Yapp, C.; Filippakopoulos, P.; Bunnage, M. E.; Muller, S.; Knapp, S.; Schofield, C. J.; Brennan, P. E., Discovery and optimization of small-molecule ligands for the CBP/p300 bromodomains. *J. Am. Chem. Soc.* **2014**, *136* (26), 9308-9319.
 12. (a) Zolek, T.; Maciejewska, D.; Zabinski, J.; Kazmierczak, P.; Rezler, M., Probing the Relationship between Anti-Pneumocystis carinii Activity and DNA Binding of Bisamidines by Molecular Dynamics Simulations. *Molecules* **2015**, *20* (4), 5942-5964; (b) Shaikh, S. A.; Jayaram, B., A swift all-atom energy-based computational protocol to predict DNA-ligand binding affinity and ΔT_m . *J. Med. Chem.* **2007**, *50* (9), 2240-2244; (c) Martinez, J.; Sanchez, R.; Castellanos, M.; Makarava, N.; Aguzzi, A.; Baskakov, I. V.; Gasset, M., PrP charge structure encodes interdomain interactions. *Sci. Rep.* **2015**, *5*, 13623; (d) Hugle, M.; Lucas, X.; Weitzel, G.; Ostrovskiy, D.; Breit, B.; Gerhardt, S.; Einsle, O.; Gunther, S.; Wohlwend, D., 4-Acyl Pyrrole Derivatives Yield Novel Vectors for Designing Inhibitors of the Acetyl-Lysine Recognition Site of BRD4(1). *J. Med. Chem.* **2016**, *59* (4), 1518-1530.
 13. Helbling, R. E.; Aeschmann, W.; Simona, F.; Stocker, A.; Cascella, M., Engineering Tocopherol Selectivity in α -TTP: A Combined In Vitro/In Silico Study. *PLOS One* **2012**, *7* (11), e49195.
 14. Holdgate, G. A.; Ward, W. H. J., Measurements of binding thermodynamics in drug discovery. *Drug Discov. Today* **2005**, *10* (22), 1543-1550.
 15. (a) Sunner, J.; Nishizawa, K.; Kebarle, P., Ion-Solvent Molecule Interactions in the Gas-Phase - the Potassium-Ion and Benzene. *J. Phys. Chem.* **1981**, *85* (13), 1814-1820; (b) Meotner, M.; Deakyne, C. A., Unconventional Ionic Hydrogen-Bonds .1. $\text{CH}_3\text{-}\Delta^+\text{...X}^-$ - Complexes of

- Quaternary Ions with Normal-Donors and Pi-Donors. *J. Am. Chem. Soc.* **1985**, *107* (2), 469-474; (c) For a review of recent developments see: An, Y.; Wheeler, S. E. Cation- π interactions. *Encyclopedia of Inorganic and Bioinorganic Chemistry*, **2015**, 1–11.
16. (a) Pizzitutti, F.; Giansanti, A.; Ballario, P.; Ornaghi, P.; Torreri, P.; Ciccotti, G.; Filetici, P., The role of loop ZA and Pro371 in the function of yeast Gcn5p bromodomain revealed through molecular dynamics and experiment. *J. Mol. Recognit.* **2006**, *19* (1), 1-9; (b) Magno, A.; Steiner, S.; Caflisch, A., Mechanism and Kinetics of Acetyl-Lysine Binding to Bromodomains. *J. Chem. Theory Comput.* **2013**, *9* (9), 4225-4232.
17. (a) Ferguson, F. M.; Fedorov, O.; Chaikuad, A.; Philpott, M.; Muniz, J. R.; Felletar, I.; von Delft, F.; Heightman, T.; Knapp, S.; Abell, C.; Ciulli, A., Targeting low-druggability bromodomains: fragment based screening and inhibitor design against the BAZ2B bromodomain. *J. Med. Chem.* **2013**, *56* (24), 10183-10187; (b) Lucas, X.; Wohlwend, D.; Hugle, M.; Schmidtkunz, K.; Gerhardt, S.; Schule, R.; Jung, M.; Einsle, O.; Gunther, S., 4-Acyl pyrroles: mimicking acetylated lysines in histone code reading. *Angew. Chem. Int. Ed.* **2013**, *52* (52), 14055-14059.
18. Wheeler, S. E.; Houk, K. N., Through-Space Effects of Substituents Dominate Molecular Electrostatic Potentials of Substituted Arenes. *J. Chem. Theory Comput.* **2009**, *5* (9), 2301-2312.
19. (a) Minoux, H.; Chipot, C., Cation- π Interactions in Proteins: Can Simple Models Provide an Accurate Description? *J. Am. Chem. Soc.* **1999**, *121* (44), 10366-10372; (b) Gallivan, J. P.; Dougherty, D. A., Cation- π interactions in structural biology. *Proc. Natl. Acad. Sci. USA* **1999**, *96* (17), 9459-9464.
20. (a) Dougherty, D. A., Cation- π interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. *Science* **1996**, *271* (5246), 163-168; (b) Fong, T. M.; Cascieri, M. A.; Yu, H.; Bansal, A.; Swain, C.; Strader, C. D., Amino-aromatic interaction between histidine 197 of the neurokinin-1 receptor and CP 96345. *Nature* **1993**, *362* (6418), 350-353; (c) Wintjens, R.; Lievin, J.; Rooman, M.; Buisine, E., Contribution of cation- π interactions to the stability of protein-DNA complexes. *J. Mol. Biol.* **2000**, *302* (2), 395-410; (d) Wheeler, S. E.; Houk, K. N., Substituent effects in cation/ π interactions and electrostatic potentials above the centers of substituted benzenes are due primarily to through-space effects of the substituents. *J. Am. Chem. Soc.* **2009**, *131* (9), 3126-3127.
21. Aldeghi, M.; Heifetz, A.; Bodkin, M. J.; Knapp, S.; Biggin, P. C., Accurate calculation of the absolute free energy of binding for drug molecules. *Chem. Sci.* **2016**, *7* (1), 207-218.
22. Steiner, S.; Magno, A.; Huang, D. Z.; Caflisch, A., Does bromodomain flexibility influence histone recognition? *FEBS Letters* **2013**, *587* (14), 2158-2163.

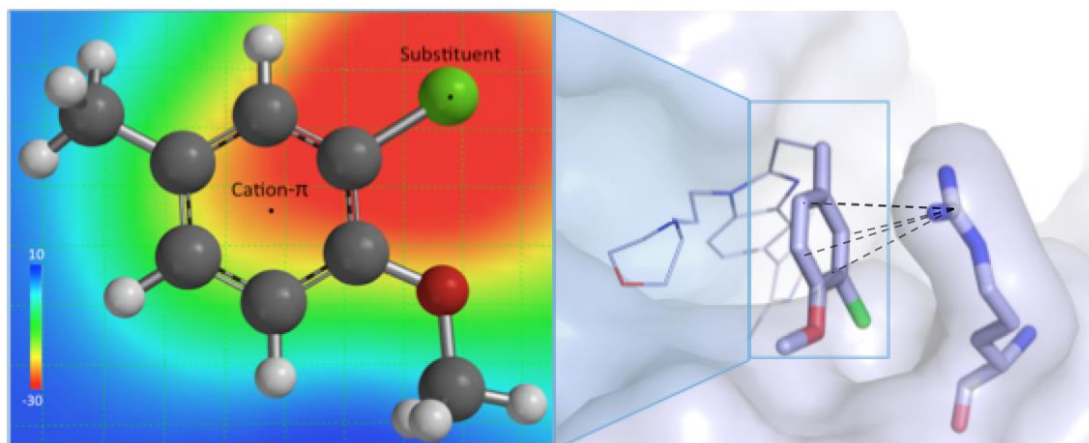
23. Miller, B. R., 3rd; McGee, T. D., Jr.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E., MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* **2012**, 8 (9), 3314-3321.
24. (a) Ran, T.; Zhang, Z. M.; Liu, K. J.; Lu, Y.; Li, H. F.; Xu, J. X.; Xiong, X.; Zhang, Y. M.; Xu, A. Y.; Lu, S.; Liu, H. C.; Lu, T.; Chen, Y. D., Insight into the key interactions of bromodomain inhibitors based on molecular docking, interaction fingerprinting, molecular dynamics and binding free energy calculation. *Mol. Biosyst.* **2015**, 11 (5), 1295-1304; (b) Oehme, D. P.; Brownlee, R. T.; Wilson, D. J., Effect of atomic charge, solvation, entropy, and ligand protonation state on MM-PB(GB)SA binding energies of HIV protease. *J. Comput. Chem.* **2012**, 33 (32), 2566-2580; (c) Wright, D. W.; Hall, B. A.; Kenway, O. A.; Jha, S.; Coveney, P. V., Computing Clinically Relevant Binding Free Energies of HIV-1 Protease Inhibitors. *J. Chem. Theory Comput.* **2014**, 10 (3), 1228-1241.
25. (a) Hou, T. J.; Wang, J. M.; Li, Y. Y.; Wang, W., Assessing the Performance of the Molecular Mechanics/Poisson Boltzmann Surface Area and Molecular Mechanics/Generalized Born Surface Area Methods. II. The Accuracy of Ranking Poses Generated From Docking. *J. Comp. Chem.* **2011**, 32 (5), 866-877; (b) Rapp, C.; Kalyanaraman, C.; Schiffmiller, A.; Schoenbrun, E. L.; Jacobson, M. P., A molecular mechanics approach to modeling protein-ligand interactions: relative binding affinities in congeneric series. *J. Chem. Inf. Model.* **2011**, 51 (9), 2082-2089.
26. (a) Caldwell, J. W.; Kollman, P. A., Cation-Pi Interactions - Nonadditive Effects Are Critical in Their Accurate Representation. *J. Am. Chem. Soc.* **1995**, 117 (14), 4177-4178; (b) Kumpf, R. A.; Dougherty, D. A., A Mechanism for Ion Selectivity in Potassium Channels - Computational Studies of Cation-Pi Interactions. *Science* **1993**, 261 (5129), 1708-1710; (c) Caldwell, J. W.; Kollman, P. A., Cation-pi. Interactions: Nonadditive Effects Are Critical in Their Accurate Representation. *J. Am. Chem. Soc.* **1995**, 117 (14), 4177-4178.
27. (a) Kuang, M.; Zhou, J.; Wang, L.; Liu, Z.; Guo, J.; Wu, R., Binding Kinetics versus Affinities in BRD4 Inhibition. *J. Chem. Inf. Model.* **2015**, 55 (9), 1926-1935; (b) Heifetz, A.; Trani, G.; Aldeghi, M.; MacKinnon, C. H.; McEwan, P. A.; Brookfield, F. A.; Chudyk, E. I.; Bodkin, M.; Pei, Z.; Burch, J. D.; Ortwine, D. F., Fragment Molecular Orbital Method Applied to Lead Optimization of Novel Interleukin-2 Inducible T-Cell Kinase (ITK) Inhibitors. *J. Med. Chem.* **2016**, 9 (9), 4352-4363.
28. Zhong, W.; Gallivan, J. P.; Zhang, Y.; Li, L.; Lester, H. A.; Dougherty, D. A., From ab initio quantum mechanics to molecular neurobiology: a cation-pi binding site in the nicotinic receptor. *Proc. Natl. Acad. Sci. USA* **1998**, 95 (21), 12088-12093.

29. Gallivan, J. P.; Dougherty, D. A., A computational study of cation–pi interactions vs salt bridges in aqueous media: Implications for protein engineering. *J. Am. Chem. Soc.* **2000**, *122* (5), 870-874.
30. Dougherty, D. A., The cation–pi interaction. *Acc. Chem. Res.* **2013**, *46* (4), 885-893.
31. (a) Cubero, E.; Luque, F. J.; Orozco, M., Is polarization important in cation–pi interactions? *Proc. Natl. Acad. Sci USA* **1998**, *95* (11), 5976-5980; (b) Tsuzuki, S.; Yoshida, M.; Uchamaru, T.; Mikami, M., The origin of the cation/pi interaction: The significant importance of the induction in Li⁺ and Na⁺ complexes. *J. Phys. Chem. A* **2001**, *105* (4), 769-773; (c) Raju, R. K.; Bloom, J. W.; An, Y.; Wheeler, S. E., Substituent effects on non-covalent interactions with aromatic rings: insights from computational chemistry. *ChemPhysChem* **2011**, *12* (17), 3116-3130.
32. (a) Kennedy, C. R.; Lin, S.; Jacobsen, E. N., The Cation–pi Interaction in Small-Molecule Catalysis. *Angew. Chem. Int. Ed.* **2016**, *55*, 2-31; (b) Johnston, C. P.; Kothari, A.; Sergeieva, T.; Okovytyy, S. I.; Jackson, K. E.; Paton, R. S.; Smith, M. D., Catalytic enantioselective synthesis of indanes by a cation–directed 5-endo-trig cyclization. *Nat. Chem.* **2015**, *7* (2), 171-177.
33. Frisch, M. J.; et al, Gaussian 09, Revision B.01. Wallingford CT, 2009.
34. (a) Boys, S. F.; Bernardi, F., The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* **2002**, *100* (1), 65-73; (b) Simon, S.; Duran, M.; Dannenberg, J. J., How does basis set superposition error change the potential surfaces for hydrogen bonded dimers? *J. Chem. Phys.* **1996**, *105* (24), 11024-11031.
35. Zhao, Y.; Truhlar, D. G., The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120* (1-3), 215-241.
36. Pople, J. A.; Headgordon, M.; Raghavachari, K., Quadratic Configuration-Interaction - a General Technique for Determining Electron Correlation Energies. *J. Chem. Phys.* **1987**, *87* (10), 5968-5975.
37. Grimme, S.; Ehrlich, S.; Goerigk, L., Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32* (7), 1456-1465.
38. a) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G., Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J. Phys. Chem. B* **2009**, *113* (18), 6378-6396. b) Marenich, A. V.; Olson, R. M.; Kelly, C. P.; Cramer, C. J.; Truhlar, D. G., Self-consistent reaction field model for aqueous and nonaqueous solutions based on accurate polarized partial charges. *J. Chem. Theor. Comp.* **2007**, *3* (6), 2011-2033.

39. Mujtaba, S.; Zeng, L.; Zhou, M. M., Structure and acetyl-lysine recognition of the bromodomain. *Oncogene* **2007**, *26* (37), 5521-5527.
40. *Spartan '14*, Wavefunction, Inc. Irvine, CA.
41. Case, D. A. *et al.* AMBER 12. University of California, San Francisco: 2012.
42. Wang, J. M.; Wang, W.; Kollman, P. A.; Case, D. A., Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **2006**, *25* (2), 247-260.
43. Dupradeau, F. Y.; Pigache, A.; Zaffran, T.; Savineau, C.; Lelong, R.; Grivel, N.; Lelong, D.; Rosanski, W.; Cieplak, P., The R.ED. tools: advances in RESP and ESP charge derivation and force field library building. *Phys. Chem. Chem. Phys.* **2010**, *12* (28), 7821-7839.
44. Connolly, M. L., Solvent-Accessible Surfaces of Proteins and Nucleic-Acids. *Science* **1983**, *221* (4612), 709-713.
45. Kendall, R. A.; Dunning, T. H.; Harrison, R. J., Electron-Affinities of the 1st-Row Atoms Revisited - Systematic Basis-Sets and Wave-Functions. *J. Chem. Phys.* **1992**, *96* (9), 6796-6806.
46. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79* (2), 926-935.
47. Humphrey, W.; Dalke, A.; Schulten, K., VMD: visual molecular dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33-38.
48. Genheden, S.; Ryde, U., The MM-PBSA and MM-GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* **2015**, *10* (5), 449-461.
49. Muvva, C.; Singam, E. R.; Raman, S. S.; Subramanian, V., Structure-based virtual screening of novel, high-affinity BRD4 inhibitors. *Mol. Biosyst.* **2014**, *10* (9), 2384-2397.
50. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K., Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781-1802.
51. Dougherty, D. A., Cation- π interactions involving aromatic amino acids. *J. Nutr.* **2007**, *137*, 1504S-1508S.
52. Mecozzi, S.; West, A. P.; Dougherty, D. A., Cation- π interactions in simple aromatics: Electrostatics provide a predictive tool. *J. Am. Chem. Soc.* **1996**, *118* (9), 2307-2308.
53. Taylor, A. M. *et al.* Fragment-Based Discovery of a Selective and Cell-Active Benzodiazepinone CBP/EP300 Bromodomain Inhibitor (CPI-637). *ACS Med. Chem. Lett.* **2016** *7* (5), 531-536
54. ESP values were used in Hartree: the units of the regression coefficients in Eq. 1 are therefore $^{\circ}\text{C}/\text{Hartree}$.

55. An, Y.; Doney, A. C.; Andrade, R. B.; Wheeler, S. E., Stacking Interactions between 9-Methyladenine and Heterocycles Commonly Found in Pharmaceuticals. *J. Chem. Inf. Mod.* **2016**, *56* (5), 906-914.
56. (a) Raju, R. K.; Bloom, J. W. G.; An, Y.; Wheeler, S. E. Substituent Effects on Non-Covalent Interactions with Aromatic Rings: Insights from Computational Chemistry. *ChemPhysChem* **2011**, *12*, 3116-3130; (b) Wheeler, S. E. Understanding Substituent Effects in Noncovalent Interactions Involving Aromatic Rings. *Acc. Chem. Res.*, **2013**, *46* (4), pp 1029–1038.
57. Popp, T. A.; Tallant, C.; Rogers, C.; Fedorov, O.; Brennan, P. E.; Müller, S.; Knapp, S.; Bracher, F. *J. Med. Chem.* **2016** *59*, 8889-8912
- Paton, R. S.; Goodman, J. M. Hydrogen Bonding and Pi-Stacking: How Reliable are Force Fields? A Critical Evaluation of Force Field Descriptions of Non-Bonded Interactions. *J. Chem. Inf. Model.* **2009**, *49*, 944-955.
58. (a) Sun, H. Y.; Li, Y. Y.; Tian, S.; Xu, L.; Hou, T. J., Assessing the performance of MM-PBSA and MM-GBSA methods. 4. Accuracies of MM-PBSA and MM-GBSA methodologies evaluated by various simulation protocols using PDBbind data set. *Phys. Chem. Chem. Phys.* **2014**, *16* (31), 16719-16729; (b) Stein, R. S. L.; Li, N.; He, W.; Komives, E.; Wang, W., Recognition of Methylated Peptides by *Drosophila melanogaster* Polycomb Chromodomain. *J. Proteome Res.* **2013**, *12* (3), 1467-1477.

TOC graphic



Keywords:

Bromodomains - cation- π interactions - CREBBP - electrostatic potentials