

Molecular Dynamics Simulations of Bromodomains Reveal Binding-Site Flexibility and Multiple Binding Modes of the Natural Ligand Acetyl-Lysine

Dimitrios Spiliotopoulos* and Amedeo Caflisch*[a]

Abstract: Experimental protein structures provide spatial information at the atomic level. A further dimension, time, is supplemented by molecular dynamics. Since the pioneering work on the 58-residue inhibitor of bovine pancreatic trypsin in the group of Martin Karplus in the seventies, molecular dynamics simulations have shown that the intrinsic flexibility of proteins is essential for their function. Here, we review simulation studies of bromodomains. These protein modules are involved in the recognition of acetylated lysine side

chains, a post-translational modification frequently observed in histone tails. The molecular dynamics simulations have unmasked: (i) the large plasticity of the loops lining the acetyl-lysine binding site (coupled to its self-occlusion), and (ii) multiple binding modes of acetyl-lysine. These simulation results suggest that recognition of histone tails by bromodomains is modulated by their intrinsic flexibility, and further corroborate the utility of molecular dynamics in understanding (macro)molecular recognition.

Keywords: atomistic simulations · computational chemistry · drug design · epigenetics · protein–protein interactions

1. Introduction

The number of protein structures solved by X-ray crystallography and/or NMR spectroscopy has grown almost exponentially since the determination of the first protein structure.^[1] Despite their unquestionable usefulness, experimentally determined structures lack most of the information concerning the dynamic behavior of a protein. Richard P. Feynman acutely observed in 1963, “One of the great triumphs in recent times (since 1960), was at last to discover the exact spatial atomic arrangement of certain proteins [...]. One of the sad aspects of this discovery is that we cannot see anything from the pattern; we do not understand why it works the way it does.”^[2] Indeed, protein motions are often essential for protein function.^[3] This is particularly true for protein–ligand interactions, as the current binding models assume that conformational rearrangements (including random thermal motions) occur at both the protein and the ligand level.^[4] Therefore, both structural and thermodynamic properties of the interaction are affected. NMR order parameters and calorimetric measurements can provide insights into protein motion at the residue level and free energy difference upon binding, respectively. However, their limitations hamper the full understanding of the biochemical processes.^[5] Molecular dynamics (MD) simulations can fill this void (“the next problem to be attacked”, in Feynman’s words)^[2] by assisting in the interpretation of experimental data and guiding further experiments.^[3]

Historically, classical MD simulations (i.e., the temporal evolution of a set of atoms according to Newton’s second law of motion) have been performed starting from the 1950s, when liquid systems consisting of relatively few atoms were simulated with hard^[6] and then soft^[7] spheres. An important progress was the simulation of liquid water, the first polar molecule, in the early 1970s.^[8] Chemists questioned the feasibility of such an approach on a protein structure, and biologists its utility.^[9] The first MD simulation of a protein was carried out in the group of Martin Karplus in 1977.^[10] The available computational resources at the time limited the simulation in terms of time (9.2 ps), amount of atoms (the protein under investigation, i.e., bovine pancreatic trypsin inhibitor, had been chosen also because of its small size) and solvent (the protein was simulated in vacuo). Significant methodological developments in terms of algorithms for MD sampling and force field parameterization as well as the ever improving computer hardware have characterized the past forty years.^[11] A wide variety of MD techniques and force fields are con-

[a] D. Spiliotopoulos, A. Caflisch
Department of Biochemistry
University of Zurich
Winterthurerstrasse 190
CH-8057 Zurich (Switzerland)
Fax: (+41) 44-635-68-62
e-mail: d.spiliotopoulos@bioc.uzh.ch
caflisch@bioc.uzh.ch

stantly being developed and improved,^[12] and the first microsecond-^[13] and millisecond-long^[14] MD simulations were reported within slightly more than a decade.

Experimental data are still necessary to validate the computer simulations.^[9] In 1975, B. Honig, A. Warshel and M. Karplus wrote that, “theoretical chemists tend to use the word ‘prediction’ rather loosely to refer to any calculation that agrees with experiment, even when the latter was done before the former.”^[15] Agreement with available experimental data unquestionably provides an *a posteriori* validation of the simulation model and justifies the in-depth analysis of the MD trajectories which reach temporal and spatial resolutions not accessible to experimental approaches. That being said, it must be stressed that there is a continuous increase in the number of *in silico* investigations that formulate predictions in the true meaning of the word. These computational results provide insights that are used to guide experimental efforts. The predictive ability of atomistic simulations has

played key roles in a wide variety of applications, most notably drug discovery^[16] and protein engineering.^[17] This highlights the prominent role of biomolecular modeling and simulations not only in basic but also in application-oriented research. Indeed, simulation techniques based on classical mechanics (ranging in complexity from simple energy minimization up to sophisticated enhanced sampling MD techniques) are used more and more by research groups consisting mainly of experimentalists. Confirming the adult state of the field and its widespread use, some authors have highlighted the necessity to be well aware of the caveats of biomolecular modeling and simulation tools.^[11,18]

In the following sections, we will describe the advances in the understanding of the dynamic properties of bromodomains, a class of protein modules involved in epigenetic processes that has gained increasing importance in the last years. The combinatorial arrangement of the histone post-translational modifications (PTMs) was suggested to contribute to the regulation of DNA-templated processes, such as gene expression and DNA replication.^[19] This hypothesis, termed “histone code”, is now widely accepted. Among the histone PTMs, the acetylation of the ϵ -amino group of lysine residues (first described 50 years ago)^[20] plays a key role in the epigenetic regulation of gene expression. The acetyl group is covalently added and removed by histone acetyltransferases (HATs) and histone deacetylases (HDAC), respectively.^[21] Upon acetylation the positive charge of the lysine side chain is abolished, which directly impairs the histone electrostatic interactions with DNA. It is thus not surprising that histone lysine acetylation has been related (although not exclusively) to open or active chromatin regions.^[22] This histone PTM also selectively recruits a set of acetyl-lysine binding protein modules^[21] which stabilize the binding of further machineries to mediate the indirect effects of the modification.^[23] Bromodomains constitute the major group of acetyl-lysine interacting modules.

Bromodomains present four helices (α_Z , α_A , α_B , α_C) folded into a left-handed bundle.^[24] The ZA and BC loops, connecting the former two and latter two helices, respectively, line the opening of the acetyl-lysine binding pocket.^[25] This pocket contains three conserved residues: two tyrosine residues (one in each of the ZA loop and BC loop), whose side chains constitute the hydrophobic pocket, and an asparagine residue, lying immediately downstream of the second conserved tyrosine residue, which secures the acetyl moiety of the acetylated lysine via a hydrogen bond.^[26]

The implication of histone acetylation in the development and progression of human diseases is well established and two HDAC inhibitors are currently used in clinical practice.^[27] Target-based drug design investigations have so far mainly focused on epigenetic enzymes, i.e., “writers” or “erasers”, possibly as a result of the allegedly limited druggability of protein–protein interac-

Amedeo Caflich holds the Chair for Computational Structural Biology at the Department of Biochemistry and has been a member of the Faculty of Science of the University of Zurich since 1996. During 1992–1994 he was a postdoctoral fellow in the research group of Martin Karplus at Harvard University. Previously, he earned a Master in Theoretical Physics and a Ph.D. (1991) at the ETH Zurich. His main research activities are the simulation studies of protein aggregation (using coarse-grained and atomistic models), allosteric effects, and (macro)molecular recognition. Another important activity is computer-aided structure-based drug design with experimental validation in house using biochemical and biophysical assays. A dozen crystal structures of complexes between tyrosine kinases and small-molecule inhibitors (discovered *in silico*) have been solved recently in the Caflich group.



After his M.Sc. in Biology at the Università degli Studi di Milano-Bicocca (2006), Dimitrios Spiliotopoulos worked on neural stem cells in Elena Cattaneo’s lab at Università degli Studi di Milano. He then changed his academic direction to structural biology, was awarded with a three-year fellowship in the Molecular and Cellular Biology PhD program of the Università Vita-Salute San Raffaele in association with the Open University, and investigated the PHD fingers interaction with histone tails using experimental and computational techniques in Giovanna Musco’s group. He is now working in the research group of Amedeo Caflich at the University of Zurich, focusing on the structure and function of bromodomains.



tions^[28] and due to the utilization of enzymatic assays, which can be easily used for small-molecule screening and lead optimization chemistry.^[29] Epigenetic “readers” have recently emerged as legitimate targets for drug discovery. Filippakopoulos and colleagues described a small molecule, JQ1, which acts as a potent and selective inhibitor of the histone binding activity of the bromodomains belonging to the BET family, most notably the BRD4 first bromodomain.^[30] BRD4 is a fused component of a chimeric protein (BRD4-NUT) associated with an aggressive and incurable form of carcinoma. JQ1 was proven to be acetyl-lysine competitive in cell-based assays, which resulted in an anti-proliferative effect in BRD4-dependent cancer cell lines and patient-derived xenograft models.^[30] In the same issue, Nicodeme and co-workers reported I-BET, a pan-BET inhibitor that efficiently suppressed the expression of a subset of lipopolysaccharide-induced genes in LPS-treated bone marrow-derived macrophages by interfering with BET bromodomain recruitment to the corresponding chromatin regions. Of note, prophylactic and therapeutic administration of I-BET in mouse models of sepsis resulted in an increased survival.^[31] Importantly, structural and biophysical investigations provided in both papers the rationale of the ligand binding specificity and these compounds were readily made available for the scientific community. These works provided a proof of concept that a new series of targets for epigenetic drugs, the reader modules, can be used: since protein–protein interactions underpin all epigenetic mechanisms, the target availability is immense. As a matter of fact, numerous works have recently identified selective bromodomain inhibitors.^[32] For this purpose, features such as small structural differences between the proteins to be targeted^[33] and their dynamics^[34] cannot be ignored, making it imperative to perform computational investigations on bromodomains.

Despite their importance, few simulation studies had been performed on bromodomains until recently.^[35] Here we review how recent MD simulations have shed light on the flexibility of bromodomains in their free state and in complex with their physiological ligand.^[36] The simulations add the temporal dimension to the static picture provided by the many crystal and NMR structures of bromodomains. Binding-site flexibility and multiple binding modes emerge thanks to the MD studies.

2. MD Reveals Bromodomain Binding-Site Flexibility and Self-Occlusion

Despite the wealth of available three-dimensional structures of human bromodomains,^[37] little is known about their flexibility and relative plasticity among different families. This hampers the comprehension of the bromodomain–histone tail physiological interactions as differences in selectivity might originate in part from the differ-

ences in intrinsic flexibility. Moreover, a complete understanding of the different dynamics of the individual domains would have crucial relevance for the design of drugs selectively targeting specific bromodomains. These considerations prompted the computational analysis of the binding-site flexibility of 20 bromodomains (representing seven of the eight families of human bromodomains) by means of MD simulations.^[36b]

At 310 K the folded state of all simulated bromodomains is stable over a 1 μ s time scale^[36b] and even along 10 μ s runs (Huang and Caflisch, unpublished results). The sequence profile of the root mean square fluctuations (RMSF), which reports on the flexibility on the nanosecond time scale, indicates that the fluctuations are largest for the ZA loop, BC loop, and the termini in all simulations of bromodomains, as shown for the CREBBP bromodomain in Figure 1B,D.

In striking contrast, the crystallographic B-factors suggest a relatively uniform flexibility along the bromodomain backbone, with a slightly larger disorder only at the N- and C-termini (Figure 1A,C). This comparison shows that the degree of flexibility of the ZA loop and BC loop with respect to the rest of the domain is strongly underestimated by the experimental data, mainly due to spurious contacts involving the loop residues and neighboring molecules in the crystal arrangement.^[36b] Since these loops line the acetyl-lysine binding site, their fluctuations are very likely to affect the biological function of the bromodomain, i.e., the specific recognition of the acetylated lysine and binding of histone tails.

The bromodomain acetyl-lysine binding pocket contains three highly evolutionarily conserved residues (Figure 1C,D). An Asn is present in the BC loop of most bromodomains: co-crystal structures showed that its side chain amide establishes a hydrogen bond with the oxygen atom of the acetyl group of the acetyl-lysine.^[38] Two Tyr residues, lying in the ZA loop and BC loop, partake in the formation of the hydrophobic pocket constituting the binding site.^[26] The MD simulations showed that the motion of the residues adjacent to the conserved Asn influences directly the accessibility of the acetyl-lysine binding site. In particular, the trajectories of some of the 20 simulated bromodomains reveal a significant flexibility of the side chain of the highly conserved Tyr in the BC loop (corresponding to a Phe in BAZ2B and SMARCA4).^[36b] This Tyr (or Phe) residue is upstream in the sequence to the Asn that is involved in the binding to the acetylated lysine. Its aromatic side chain rotates from an orientation towards the center of the binding site to an outward orientation in which it points towards the solvent (Figure 2, top). Interestingly, this conformation is present in only a few apo bromodomain crystal and solution structures, i.e., the bromodomains of BRDT(1) and KIAA1240 (PDB codes 2RFJ and 2DKW, respectively).^[36b] In the MD simulations of BAZ2B and CREBBP, the swapping out of the conserved Tyr in the BC loop results in the re-

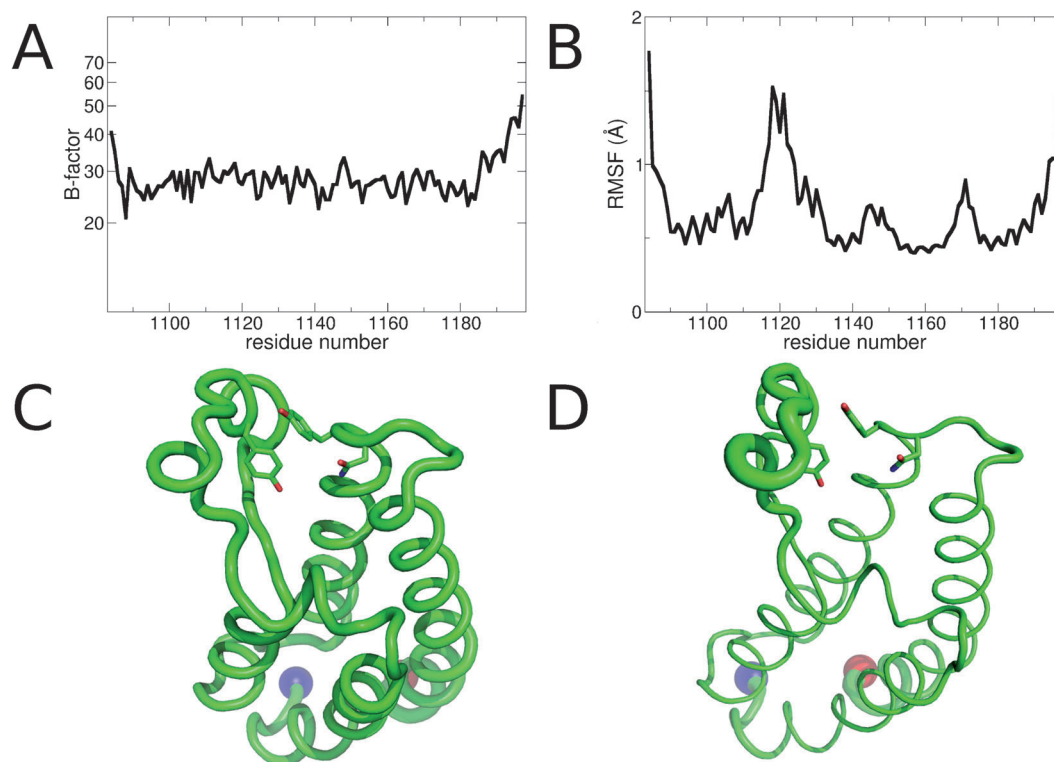


Figure 1. MD simulations show pronounced flexibility of the ZA loop lining the acetyl-lysine binding site of bromodomains. (A) Sequence profile of the measured crystallographic B-factors (PDB code 3DWY) and (B) RMSF values along MD runs at 310 K of the CREBBP bromodomain. The RMSF values were averaged over 5 ns long simulation intervals. (C,D) Tube-like representation of the CREBBP bromodomain. The evolutionarily conserved Tyr and Asn residues are shown as sticks, while the N- and C-termini are indicated by blue and red spheres, respectively. The thickness of the tube is related to the B-factors (C) and the RMSF values (D).

arrangement of the side chains of other residues forming the binding site, which ultimately leads to the partial occlusion of the binding pocket (Figure 2, bottom).

The outward reorientation is often accompanied by large fluctuations of the ZA loop, with the establishment of a hydrogen bond between the side chains of the two conserved residues, the BC loop Asn and the ZA loop Tyr.^[36b] In this conformation, the ZA loop Tyr side chain mimics the acetylated lysine, by occupying its position and being anchored with a hydrogen bond to the side chain amide of the conserved Asn. The self-occluding rearrangement of this residue might constitute a fine autoregulatory mechanism for acetyl-lysine binding. This slight alteration of the binding site is comparable to other well-studied cases of intra-steric regulation involving more dramatic conformational rearrangements, such as the aspartic protease^[39] and protein tyrosine kinases.^[40] The subtlety of this switch could be correlated to the necessity of bromodomains to be readily recruited to and released from chromatin, similar to other epigenetic readers such as PHD fingers,^[41] and would account in part for the relatively low affinity of the bromodomains towards histone tails. Additionally, this rearrangement might also facilitate the release of the ligand. A similar example of a molecular

shovel mechanism has been found in the *Trypanosoma cruzi* trans-sialidase, whose Trp312 side chain displaces the donor aglycone from the active site.^[42]

It is not unprecedented that MD simulations have revealed altered accessibilities in the binding sites of apo proteins.^[43] The partial occlusion of the binding site observed in the MD simulations of the BAZ2B and CREBBP bromodomains is related to a displacement of the ZA loop, which has remarkable differences in length and sequence in the 61 human bromodomains.^[37] This prompts further investigations to assess to what extent this self-occluded metastable state is present in other bromodomains. Analogous alterations in the binding-site accessibility could be present in other epigenetic readers. For example, it has been shown that the binding to histone tails of the PHD finger of the human and *Drosophila* Pygo proteins is enhanced by an alteration of the shape and/or the stability of the binding pocket. In this case, the rearrangement results from the simultaneous interaction with another protein, which binds to a distinct surface of the PHD finger.^[44] On a more application-oriented note, the knowledge of alternative conformation(s) of the bromodomain binding pocket will play a key role in the design of selective inhibitors.

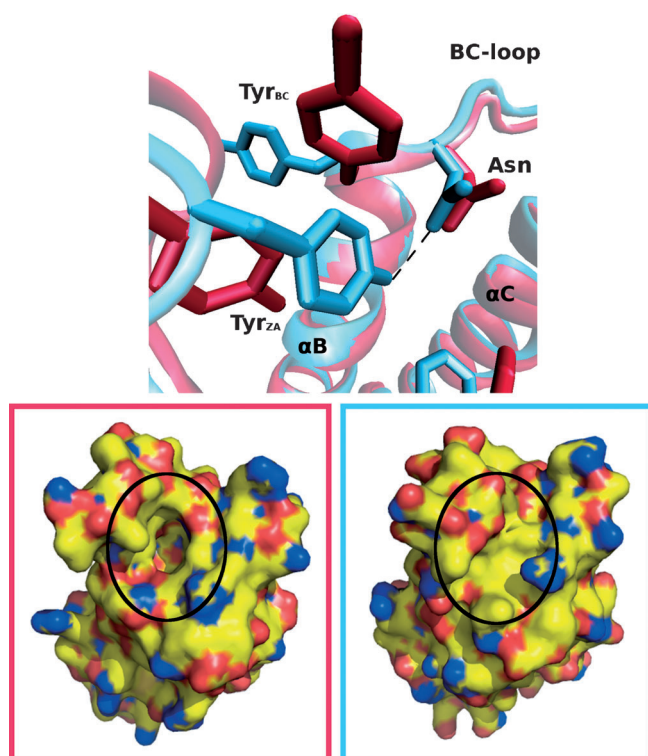


Figure 2. (Top) Two representative snapshots extracted from the MD simulations of CREBBP show the swapping out of the side chain of the Tyr residue preceding the conserved Asn in the BC loop. The dashed line indicates the hydrogen bond between the Asn and the Tyr in the ZA loop. (Bottom) Surface representation of the crystal structure of CREBBP (PDB code 3P1C; left) and of a representative MD snapshot with a partially occluded, i.e., shallower, binding site (right). The colors of the boxes correspond to those in the two structures in the top panel while the color coding of the surface reflects atomic elements (carbon, nitrogen, and oxygen atoms in yellow, blue, and red, respectively). The location of the acetyl-lysine binding site is indicated with a black oval. Figure and legend adapted from reference [36b].

3. MD Reveals Multiple Binding Modes of Acetyl-Lysine

The reversible binding of acetyl-lysine to the second bromodomain of the human TAF1 (TAF1(2); PDB code 3UV4) has been studied by unbiased MD simulations.^[36a] This study reports on the first simulations of the binding of a post-translationally modified residue to an epigenetic reader module. Twenty-four independent MD runs were started after placing the acetylated lysine in a random position at a distance of more than 20 Å from the bromodomain. Spontaneous binding of the ligand to the acetyl-lysine pocket was observed in 16 of the 24 trajectories on a 500 ns time scale; in some cases the entrance of the acetylated lysine into the binding pocket was followed by an unbinding event.

Upon spontaneous binding to the TAF1(2) bromodomain in the simulations, the acetyl-lysine adopted the

binding mode observed in the co-crystal structures, named the “N-binding mode”^[36a] (Figure 3, left). The MD trajectories revealed that this conformation reversibly converted into a more buried binding pose, which has

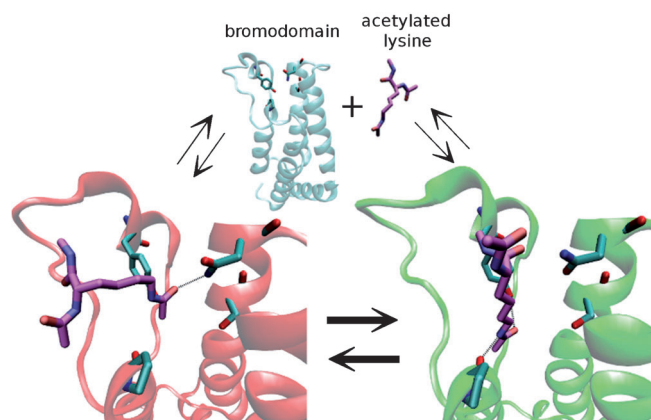


Figure 3. The two major binding modes of acetyl-lysine in the CREBBP bromodomain as observed in MD simulations of spontaneous binding. The N-binding mode (left) corresponds to the crystal structure, while the P-binding mode (right) emerges from the MD simulations and has not yet been reported in experimental studies. The thickness of the arrows reflects the number of direct transitions between states in the MD runs, i.e., the interconversions between N- and P-binding modes are much more frequent than the binding or unbinding process.^[36a] Figure adapted from reference [36a].

been called the “P-binding mode”. In this pose, the oxygen atom of the backbone carbonyl of the proline of the so-called WPF shelf (a Trp-Pro-Phe motif in the ZA loop) acts as acceptor for a hydrogen bond with the N_{ϵ} atom of the acetylated lysine (Figure 3, right). Interestingly, this previously unknown conformation was slightly more populated than the N-binding mode.^[36a] The cut-based free energy profile method^[45] isolated a third metastable conformation, the “P/N-intermediate”. The mean first-passage time indicated that the interconversion between the P- and N-binding modes is relatively fast (20 ns) with respect to the (un)binding events (220–320 ns).^[36a]

Alternative binding modes are an essential feature of protein–protein interactions.^[46] NMR spectroscopy^[47] and X-ray crystallography^[48] approaches have been described to capture transiently populated conformations. Nonetheless, experimental techniques suffer from only reporting a few properties, which are relatively poorly resolved averages over time and space.^[49] Conversely, the insights provided by computational techniques are mainly limited by sampling errors, i.e., atomistic MD simulations might not reach the time scales required for certain fundamental processes, such as folding of large proteins. Another limitation (which plays a key role in *in silico* screening) is the limited accuracy in binding affinity calculations.^[50]

Nonetheless, atomistic simulations can provide insights that are elusive to experimental approaches. Modeling and simulations have unveiled multiple binding modes in other systems, in some cases being confirmed by experimental data.^[51] In this context, it is noteworthy that after the discovery by means of unbiased MD simulations of the acetyl-lysine “P-binding mode” in the TAF1(2) bromodomain,^[36a] two independent studies have reported inhibitors targeting different bromodomains that exhibited a similar binding pose with a hydrogen bond to the carbonyl oxygen of the Pro in the WPF shelf.^[32j,l]

4. Other MD Studies of Binding

Historically, the vast majority of computational studies addressing molecular recognition have focused on the investigation of the end points of the binding process. This means that only the free molecular species and the complex are analyzed, either in their experimentally determined conformation or in their docked lowest energy pose. This two-state approach is inadequately oversimplified.^[52] The intrinsic plasticity of proteins is essential in many cases for their function,^[53] including enzyme-catalyzed reactions^[54] and molecular recognition.^[55] Upon binding to a ligand, proteins often experience a wide range of movements^[56] and alterations (either increases or decreases) of backbone and side chain flexibility.^[57] As a result, the free and bound conformations of a target protein can be different,^[58] and small changes in the protein and solvent geometries can dramatically affect the affinity with a ligand.^[51,57] Moreover, the entropic terms and the changes in solvent structure, which play a crucial role in many binding processes, are elusive to the static structures.^[51] These observations ultimately indicate that any investigation focusing on a single structure is possibly incomplete, and thus error-prone.

Protein flexibility can be inferred from experimental data, such as crystallographic B-factors^[59] and NMR spin relaxation data.^[60] The former include also static disorder in the crystal, however, while the latter report on an ensemble of molecules. Atomistic biomolecular simulations provide a detailed, single-molecule description of protein motion^[61] as a full *continuum* of conformations^[56] from the fs to μ s time scales,^[3,49d] making them a powerful tool for investigating protein function. MD simulations have revealed that intrinsic protein flexibility governs (un)binding of the natural ligands,^[62] substrates/products,^[63] or drugs.^[64] In particular, simulations have provided valuable insights for a number of small protein-interacting modules in their interaction with their physiological ligands or small molecules. These protein domains are involved in a wide variety of biological processes, such as epigenetics,^[65] signal transduction,^[66] and apoptosis.^[67] A computational clarification might even settle unclear experimental results. For example, the discovery of multiple metastable

states of the ligand in silico has also shed light onto discrepancies in the experimentally determined binding modes, as reported for the *Staphylococcus aureus* Sortase A interaction with the LPATG sorting signal.^[68] Atomistic simulations can also provide information when the experimental data are missing, as shown by Smith and co-workers: three MD-based strategies were devised to identify the binding modes of palmitate to the barley lipid transfer protein where the NOE restraints were not sufficient to clearly define the details of the ligand binding site.^[69]

Recent studies indicate that interactions occurring away from the active site^[70] or the expected binding site^[71] can affect the protein specificity and/or activity. Enhanced sampling techniques have been used to explore the ligand diffusion processes, including steered MD,^[72] random acceleration MD,^[73] metadynamics,^[74] Monte Carlo-based approaches^[75] and conformational sampling methods inspired by robot motion planning algorithms.^[76] Unbiased MD simulations have generated thermodynamic (dissociation constant) and kinetic (on and off rates) data for complex formation. Moreover, they have successfully illustrated the binding pathways for small molecules and drugs to protein targets, either using massive computation^[77] or performing multiple MD runs with a high (i.e., mM) concentration of ligand in the simulation box.^[49c] Interestingly, two features highlighted by the computational studies on bromodomains are common to other systems: (i) the presence of metastable states corresponding to alternative binding poses, and (ii) the crucial importance of solvent molecules in molecular recognition.^[78]

5. Summary and Outlook

Small molecules specifically inhibiting the acetyl-lysine binding activity of bromodomains have been identified and characterized since 2005.^[79] The usefulness of such chemical probes has been extensively described in the past.^[32a,80] As noted by several authors,^[32k,80a] the dynamic properties of bromodomains should not be ignored to efficiently target their acetyl-lysine binding activity. Nonetheless, information on the flexibility of bromodomains and the dynamicity of their interaction with the ligands was limited. The recent computational investigations have revealed key features of these important protein modules, i.e., (i) the flexibility of the loops and its potential impact on histone tail recognition,^[36b] and (ii) the multiple binding modes of the acetylated lysine.^[36a] We wish to stress that these important findings relate with preceding and subsequent experimental observations. The high flexibility of the ZA and BC loops (highlighted for the CREBBP bromodomain by NMR data and preliminary MD simulations^[80a] but underestimated by the crystallographic B-factors), has been quantified recently.^[36b] A novel acetyl-lysine binding mode, described for the first time in silico,

has been supported by similar interactions in recently described bromodomain ligands.^[32j,i] These findings lead one to speculate that the protein plasticity of epigenetic readers might constitute an additional layer of regulation in the dynamic processes mediating epigenetic processes. This could also apply for epigenetic erasers and writers, as hypothesized for the SET and postSET domains of NSD1.^[81] These questions will be addressed by a combination of carefully selected and expertly applied computational, biochemical and functional approaches to be undertaken in the future.

Modeling protein flexibility is currently a major challenge in the theoretical investigations of protein–ligand interactions.^[82] Legitimate optimism for further advances, recently fueled by the enthusiasm kindled by the 2013 Nobel Prize in Chemistry,^[83] accompanies the efforts towards this goal and invigorates the advances of computational techniques aimed to describe long-time-scale processes. Clearly, unbiased MD simulations are not yet suitable for high-throughput virtual screenings, due to their excessive computational cost.^[84] A major challenge for drug design and development will be the investigation of binding and unbinding kinetics, as these might play a significant role in drug efficacy and safety. Significant advances are constantly granted by technological progresses in hardware (as the use of graphics processing units^[84] or *ad hoc* computer architectures^[14]) and software. It will nonetheless be crucial to further improve enhanced sampling techniques, such as those mentioned earlier in this review, accelerated MD,^[85] the free energy guided sampling (FEGS),^[86] and multi-scale approaches.^[83] Both basic research and drug design will benefit as, once mastered, this expertise will allow researchers to identify drug-targetable protein metastable states, including allosteric sites invisible in experimental structures.^[77] Future advances in both experimental and computational structural biology aimed to increase their interplay (e.g., in reference [87]) will constitute a fundamental step for their mutual benefit in the understanding of biomolecular recognition.

Acknowledgements

We are grateful to Lisa Gartenmann, Tim Knehans and Nicolas Blöchliger for critical reading of the manuscript. We also thank all the members of the Caffisch group for interesting discussions. This work was supported financially by a grant of the Swiss SystemsX.ch initiative to D. S.

References

- J. C. Kendrew, G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff, D. C. Phillips, *Nature* **1958**, *181*, 662–664.
- R. P. Feynman, R. B. Leighton, M. Sands, “The Feynman Lectures on Physics”, Vol. I, Chapter 3, as accessed on the website: http://www.feynmanlectures.caltech.edu/I_03.html, last accessed on 03/19/2014.
- M. Karplus, J. Kuriyan, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6679–6685.
- J. D. Durrant, J. A. McCammon, *BMC Biol.* **2011**, *9*, 71.
- a) J. Wereszczynski, J. A. McCammon, *Q. Rev. Biophys.* **2012**, *45*, 1–25; b) T. Simonson, G. Archontis, M. Karplus, *Acc. Chem. Res.* **2002**, *35*, 430–437.
- B. J. Alder, T. E. Wainwright, *J. Chem. Phys.* **1957**, *27*, 1208–1209.
- A. Rahman, *Phys. Rev.* **1964**, *136*, A405–A411.
- a) F. H. Stillinger, A. Rahman, *J. Chem. Phys.* **1974**, *60*, 1545; b) A. Rahman, F. H. Stillinger, *J. Chem. Phys.* **1971**, *55*, 3336–3359.
- M. Karplus, *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 1–47.
- J. A. McCammon, B. R. Gelin, M. Karplus, *Nature* **1977**, *267*, 585–590.
- T. Schlick, R. Collepardo-Guevara, L. A. Halvorsen, S. Jung, X. Xiao, *Q. Rev. Biophys.* **2011**, *44*, 191–228.
- a) S. A. Adcock, J. A. McCammon, *Chem. Rev.* **2006**, *106*, 1589–1615; b) A. D. Mackerell, Jr., *J. Comput. Chem.* **2004**, *25*, 1584–1604.
- Y. Duan, P. A. Kollman, *Science* **1998**, *282*, 740–744.
- D. E. Shaw, R. O. Dror, J. K. Salmon, J. P. Grossman, K. M. Mackenzie, J. A. Bank, C. Young, M. M. Deneroff, B. Batson, K. J. Bowers, E. Chow, M. P. Eastwood, D. J. Ierardi, J. L. Klepeis, J. S. Kuskin, R. H. Larson, K. Lindorff-Larsen, P. Maragakis, M. A. Moraes, S. Piana, Y. Shan, B. Towles, *Proceedings of the Conference on High Performance Computing, Networking, Storage and Analysis (SC09)*, Portland, Oregon, USA, November 14–20, **2009**.
- B. Honig, A. Warshel, M. Karplus, *Acc. Chem. Res.* **1975**, *8*, 92–100.
- W. L. Jorgensen, *Science* **2004**, *303*, 1813–1818.
- a) P. Alfano, G. Varadamsetty, C. Ewald, F. Parmeggiani, R. Pellarin, O. Zerbe, A. Pluckthun, A. Caffisch, *Protein Sci.* **2012**, *21*, 1298–1314; b) F. Parmeggiani, R. Pellarin, A. P. Larsen, G. Varadamsetty, M. T. Stumpp, O. Zerbe, A. Caffisch, A. Pluckthun, *J. Mol. Biol.* **2008**, *376*, 1282–1304; c) G. Kiss, N. Çelebi-Ölçüm, R. Moretti, D. Baker, K. N. Houk, *Angew. Chem. Int. Ed.* **2013**, *52*, 5700–5725.
- M. W. van der Kamp, K. E. Shaw, C. J. Woods, A. J. Mulholland, *J. R. Soc., Interface* **2008**, *5* (Suppl. 3), 173–190.
- a) T. Jenuwein, C. D. Allis, *Science* **2001**, *293*, 1074–1080; b) B. D. Strahl, C. D. Allis, *Nature* **2000**, *403*, 41–45.
- a) D. M. P. Phillips, *Biochem. J.* **1963**, *87*, 258–263; b) V. G. Allfrey, R. Faulkner, A. E. Mirsky, *Proc. Natl. Acad. Sci. U.S.A.* **1964**, *51*, 786–794.
- A. J. Bannister, T. Kouzarides, *Cell Res.* **2011**, *21*, 381–395.
- a) F. Winston, C. D. Allis, *Nat. Struct. Biol.* **1999**, *6*, 601–604; b) G. E. Zentner, S. Henikoff, *Nat. Struct. Mol. Biol.* **2013**, *20*, 259–266; c) C. B. Yoo, P. A. Jones, *Nat. Rev. Drug Discovery* **2006**, *5*, 37–50.
- S. D. Taverna, H. Li, A. J. Ruthenburg, C. D. Allis, D. J. Patel, *Nat. Struct. Mol. Biol.* **2007**, *14*, 1025–1040.
- C. Dhalluin, J. E. Carlson, L. Zeng, C. He, A. K. Aggarwal, M. M. Zhou, *Nature* **1999**, *399*, 491–496.
- S. Muller, P. Filippakopoulos, S. Knapp, *Expert Rev. Mol. Med.* **2011**, *13*, e29.
- K. L. Yap, M. M. Zhou, *Crit. Rev. Biochem. Mol. Biol.* **2010**, *45*, 488–505.
- P. A. Marks, W. S. Xu, *J. Cell. Biochem.* **2009**, *107*, 600–608.

- [28] Q. Lu, A. M. Quinn, M. P. Patel, S. F. Semus, A. P. Graves, D. Bandyopadhyay, A. J. Pope, S. H. Thrall, *J. Biomol. Screening* **2012**, *17*, 555–571.
- [29] G. Drewes, *Drug Discovery Today: Ther. Strategies* **2012**, *9*, e121–e127.
- [30] P. Filippakopoulos, J. Qi, S. Picaud, Y. Shen, W. B. Smith, O. Fedorov, E. M. Morse, T. Keates, T. T. Hickman, I. Felletar, M. Philpott, S. Munro, M. R. McKeown, Y. Wang, A. L. Christie, N. West, M. J. Cameron, B. Schwartz, T. D. Heightman, N. La Thangue, C. A. French, O. Wiest, A. L. Kung, S. Knapp, J. E. Bradner, *Nature* **2010**, *468*, 1067–1073.
- [31] E. Nicodeme, K. L. Jeffrey, U. Schaefer, S. Beinke, S. Dewell, C. W. Chung, R. Chandwani, I. Marazzi, P. Wilson, H. Coste, J. White, J. Kirilovsky, C. M. Rice, J. M. Lora, R. K. Prinjha, K. Lee, A. Tarakhovskiy, *Nature* **2010**, *468*, 1119–1123.
- [32] a) J. C. Borah, S. Mujtaba, I. Karakikes, L. Zeng, M. Muller, J. Patel, N. Moshkina, K. Morohashi, W. Zhang, G. Gerona-Navarro, R. J. Hajjar, M. M. Zhou, *Chem. Biol.* **2011**, *18*, 531–541; b) C. W. Chung, H. Coste, J. H. White, O. Mirguet, J. Wilde, R. L. Gosmini, C. Delves, S. M. Magny, R. Woodward, S. A. Hughes, E. V. Boursier, H. Flynn, A. M. Bouillot, P. Bamborough, J. M. Brusq, F. J. Gellibert, E. J. Jones, A. M. Riou, P. Homes, S. L. Martin, I. J. Uings, J. Toun, C. A. Clement, A. B. Boullay, R. L. Grimley, F. M. Blandel, R. K. Prinjha, K. Lee, J. Kirilovsky, E. Nicodeme, *J. Med. Chem.* **2011**, *54*, 3827–3838; c) M. A. Dawson, R. K. Prinjha, A. Dittmann, G. Giotopoulos, M. Bantscheff, W. I. Chan, S. C. Robson, C. W. Chung, C. Hopf, M. M. Savitski, C. Huthmacher, E. Gudgin, D. Lugo, S. Beinke, T. D. Chapman, E. J. Roberts, P. E. Soden, K. R. Auger, O. Mirguet, K. Doehner, R. Delwel, A. K. Burnett, P. Jeffrey, G. Drewes, K. Lee, B. J. Huntly, T. Kouzarides, *Nature* **2011**, *478*, 529–533; d) D. S. Hewings, M. Wang, M. Philpott, O. Fedorov, S. Uttarkar, P. Filippakopoulos, S. Picaud, C. Vuppasetty, B. Marsden, S. Knapp, S. J. Conway, T. D. Heightman, *J. Med. Chem.* **2011**, *54*, 6761–6770; e) P. Bamborough, H. Diallo, J. D. Goodacre, L. Gordon, A. Lewis, J. T. Seal, D. M. Wilson, M. D. Woodrow, C. W. Chung, *J. Med. Chem.* **2012**, *55*, 587–596; f) P. V. Fish, P. Filippakopoulos, G. Bish, P. E. Brennan, M. E. Bunnage, A. S. Cook, O. Fedorov, B. S. Gerstenberger, H. Jones, S. Knapp, B. Marsden, K. Nocka, D. R. Owen, M. Philpott, S. Picaud, M. J. Primiano, M. J. Ralph, N. Sciammetta, J. D. Trzupke, *J. Med. Chem.* **2012**, *55*, 9831–9837; g) G. Zhang, A. N. Plotnikov, E. Rusinova, T. Shen, K. Morohashi, J. Joshua, L. Zeng, S. Mujtaba, M. Ohlmeyer, M. M. Zhou, *J. Med. Chem.* **2013**, *56*, 9251–9264; h) D. Hay, O. Fedorov, P. Filippakopoulos, S. Martin, M. Philpott, S. Picaud, D. S. Hewings, S. Uttakar, T. D. Heightman, S. J. Conway, S. Knapp, P. E. Brennan, *MedChemComm* **2013**, *4*, 140–144; i) O. Fedorov, H. Lingard, C. Wells, O. P. Monteiro, S. Picaud, T. Keates, C. Yapp, M. Philpott, S. J. Martin, I. Felletar, B. Marsden, P. Filippakopoulos, S. Muller-Knapp, S. Knapp, P. E. Brennan, *J. Med. Chem.* **2014**, *57*, 462–476; j) F. M. Ferguson, O. Fedorov, A. Chaikuad, M. Philpott, J. R. C. Muniz, I. Felletar, F. von Delft, T. Heightman, S. Knapp, C. Abell, A. Ciulli, *J. Med. Chem.* **2013**, *56*, 10183–10187; k) S. Picaud, C. Wells, I. Felletar, D. Brotherton, S. Martin, P. Savitski, B. Diez-Dacal, M. Philpott, C. Bountra, H. Lingard, O. Fedorov, S. Muller, P. E. Brennan, S. Knapp, P. Filippakopoulos, *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 19754–19759; l) X. Lucas, D. Wohlwend, M. Hugle, K. Schmidt-kunz, S. Gerhardt, R. Schule, M. Jung, O. Einsle, S. Gunther, *Angew. Chem. Int. Ed.* **2013**, *52*, 14055–14059.
- [33] P. Chi, C. D. Allis, G. G. Wang, *Nat. Rev. Cancer* **2010**, *10*, 457–469.
- [34] C. Santiago, K. Nguyen, M. Schapira, *J. Comput.-Aided Mol. Des.* **2011**, *25*, 1171–1178.
- [35] a) V. C. Quy, S. Pantano, G. Rossetti, M. Giacca, P. Carloni, *Biology* **2012**, *1*, 277–296; b) F. Pizzitutti, A. Giansanti, P. Ballario, P. Ornaghi, P. Torreri, G. Ciccotti, P. Filetici, *J. Mol. Recognit.* **2006**, *19*, 1–9; c) K. D. Eichenbaum, Y. Rodriguez, M. Mezei, R. Osman, *Proteins* **2010**, *78*, 447–456.
- [36] a) A. Magno, S. Steiner, A. Caffisch, *J. Chem. Theory Comput.* **2013**, *9*, 4225–4232; b) S. Steiner, A. Magno, D. Huang, A. Caffisch, *FEBS Lett.* **2013**, *587*, 2158–2163.
- [37] P. Filippakopoulos, S. Picaud, M. Mangos, T. Keates, J. P. Lambert, D. Barsyte-Lovejoy, I. Felletar, R. Volkmer, S. Muller, T. Pawson, A. C. Gingras, C. H. Arrowsmith, S. Knapp, *Cell* **2012**, *149*, 214–231.
- [38] R. Sanchez, M. M. Zhou, *Curr. Opin. Drug Discovery Dev.* **2009**, *12*, 659–665.
- [39] A. Y. Lee, S. V. Gulnik, J. W. Erickson, *Nat. Struct. Biol.* **1998**, *5*, 866–871.
- [40] S. R. Hubbard, M. Mohammadi, J. Schlessinger, *J. Biol. Chem.* **1998**, *273*, 11987–11990.
- [41] C. A. Musselman, T. G. Kutateladze, *Nucleic Acids Res.* **2011**, *39*, 9061–9071.
- [42] F. L. Mitchell, S. M. Miles, J. Neres, E. V. Bichenkova, R. A. Bryce, *Biophys. J.* **2010**, *98*, L38–L40.
- [43] L. Fang, Y. Pan, J. L. Muzyka, C. G. Zhan, *J. Phys. Chem. B* **2011**, *115*, 8797–8805.
- [44] M. Fiedler, M. J. Sanchez-Barrena, M. Nekrasov, J. Mieszczyzanek, V. Rybin, J. Muller, P. Evans, M. Bienz, *Mol. Cell* **2008**, *30*, 507–518.
- [45] S. V. Krivov, S. Muff, A. Caffisch, M. Karplus, *J. Phys. Chem. B* **2008**, *112*, 8701–8714.
- [46] P. J. Kundrotas, I. A. Vakser, *Protein Sci.* **2013**, *22*, 1141–1145.
- [47] a) A. J. Baldwin, L. E. Kay, *Nat. Chem. Biol.* **2009**, *5*, 808–814; b) A. Sekhar, L. E. Kay, *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 12867–12874.
- [48] J. S. Fraser, H. van den Bedem, A. J. Samelson, P. T. Lang, J. M. Holton, N. Echols, T. Alber, *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 16247–16252.
- [49] a) W. F. van Gunsteren, D. Bakowies, R. Baron, I. Chandrasekhar, M. Christen, X. Daura, P. Gee, D. P. Geerke, A. Glattli, P. H. Hunenberger, M. A. Kastenholtz, C. Oostenbrink, M. Schenk, D. Trzesniak, N. F. van der Vegt, H. B. Yu, *Angew. Chem. Int. Ed.* **2006**, *45*, 4064–4092; b) W. F. van Gunsteren, J. Dolenc, A. E. Mark, *Curr. Opin. Struct. Biol.* **2008**, *18*, 149–153; c) D. Huang, A. Caffisch, *ChemMedChem* **2011**, *6*, 1578–1580; d) R. O. Dror, R. M. Dirks, J. P. Grossman, H. Xu, D. E. Shaw, *Annu. Rev. Biophys.* **2012**, *41*, 429–452.
- [50] M. K. Gilson, H. X. Zhou, *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 21–42.
- [51] D. L. Mobley, K. A. Dill, *Structure* **2009**, *17*, 489–498.
- [52] R. Baron, J. A. McCammon, *Annu. Rev. Phys. Chem.* **2013**, *64*, 151–175.
- [53] K. Henzler-Wildman, D. Kern, *Nature* **2007**, *450*, 964–972.
- [54] J. D. McGeagh, K. E. Ranaghan, A. J. Mulholland, *Biochim. Biophys. Acta* **2011**, *1814*, 1077–1092.
- [55] E. Marco, F. Gago, *ChemMedChem* **2007**, *2*, 1388–1401.
- [56] J. D. Durrant, J. A. McCammon, *Curr. Opin. Pharmacol.* **2010**, *10*, 770–774.

- [57] S. J. Teague, *Nat. Rev. Drug Discovery* **2003**, *2*, 527–541.
- [58] A. Kahraman, R. J. Morris, R. A. Laskowski, J. M. Thornton, *J. Mol. Biol.* **2007**, *368*, 283–301.
- [59] Z. Yuan, T. L. Bailey, R. D. Teasdale, *Proteins* **2005**, *58*, 905–912.
- [60] S. Morin, *Prog. Nucl. Magn. Reson. Spectrosc.* **2011**, *59*, 245–262.
- [61] M. Karplus, J. A. McCammon, *Nat. Struct. Biol.* **2002**, *9*, 646–652.
- [62] a) R. Elber, M. Karplus, *Science* **1987**, *235*, 318–321; b) M. Karplus, G. A. Petsko, *Nature* **1990**, *347*, 631–639.
- [63] S. Mishra, A. Caflisch, *Biochemistry* **2011**, *50*, 9328–9339.
- [64] a) H. Park, M. S. Yeom, S. Lee, *ChemBioChem* **2004**, *5*, 1662–1672; b) A. A. Gorfe, A. Caflisch, *Structure* **2005**, *13*, 1487–1498; c) D. Huang, A. Caflisch, *J. Chem. Theory Comput.* **2012**, *8*, 1786–1794.
- [65] a) M. Ozboyaci, A. Gursoy, B. Erman, O. Keskin, *PLoS One* **2011**, *6*, e14765; b) D. Spiliotopoulos, A. Spitaleri, G. Musco, *PLoS One* **2012**, *7*, e46902.
- [66] a) I. H. Park, C. Li, *J. Mol. Recognit.* **2011**, *24*, 254–265; b) S. Steiner, A. Caflisch, *Proteins: Struct., Funct., Bioinf.* **2012**, *80*, 2562–2572.
- [67] Q. Yan, J. M. McDonald, T. Zhou, Y. Song, *Proteins* **2013**, *81*, 377–385.
- [68] K. Kappel, J. Wereszczynski, R. T. Clubb, J. A. McCammon, *Protein Sci.* **2012**, *21*, 1858–1871.
- [69] L. J. Smith, W. F. Gunsteren, J. R. Allison, *Protein Sci.* **2013**, *22*, 56–64.
- [70] a) D. Devaurs, L. Bouard, M. Vaisset, C. Zanon, I. Al-Bluwi, R. Iehl, T. Simeon, J. Cortes, *Nucleic Acids Res.* **2013**, *41*, W297–W302; b) M. Novinec, M. Korenč, A. Caflisch, R. Ranganathan, B. Lenarčič, A. Baici, *Nat. Commun.* **2014**, *5*, 3287.
- [71] S. Mostarda, D. Gfeller, F. Rao, *PLoS Comput. Biol.* **2012**, *8*, e1002429.
- [72] F. Colizzi, R. Perozzo, L. Scapozza, M. Recanatini, A. Cavalli, *J. Am. Chem. Soc.* **2010**, *132*, 7361–7371.
- [73] a) D. Long, Y. Mu, D. Yang, *PLoS One* **2009**, *4*, e6081; b) S. K. Lüdemann, V. Lounnas, R. C. Wade, *J. Mol. Biol.* **2000**, *303*, 797–811.
- [74] a) F. L. Gervasio, A. Laio, M. Parrinello, *J. Am. Chem. Soc.* **2005**, *127*, 2600–2607; b) F. Pietrucci, F. Marinelli, C. Carloni, A. Laio, *J. Am. Chem. Soc.* **2009**, *131*, 11811–11818.
- [75] K. W. Borrelli, A. Vitalis, R. Alcantara, V. Guallar, *J. Chem. Theory Comput.* **2005**, *1*, 1304–1311.
- [76] J. Cortés, T. Siméon, V. Ruiz de Angulo, D. Guieysse, M. Remaud-Siméon, V. Tran, *Bioinformatics* **2005**, *21*, i116–i125.
- [77] a) R. O. Dror, A. C. Pan, D. H. Arlow, D. W. Borhani, P. Maragakis, Y. Shan, X. Huafeng, D. E. Shaw, *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 13118–13123; b) Y. Shan, E. T. Kim, M. P. Eastwood, R. O. Dror, M. A. Seeliger, D. E. Shaw, *J. Am. Chem. Soc.* **2011**, *133*, 9181–9183.
- [78] D. Huang, E. Rossini, S. Steiner, A. Caflisch, *ChemMedChem* **2014**, *9*, 573–579.
- [79] L. Zeng, J. Li, M. Muller, S. Yan, S. Mujtaba, C. Pan, Z. Wang, M. M. Zhou, *J. Am. Chem. Soc.* **2005**, *127*, 2376–2377.
- [80] a) Sachchidanand, L. Resnick-Silverman, S. Yan, S. Mujtaba, W. Liu, L. Zeng, J. J. Manfredi, M.-M. Zhou, *Chem. Biol.* **2006**, *13*, 81–90; b) C. W. Chung, J. Witherington, *J. Biomol. Screening* **2011**, *16*, 1170–1185.
- [81] M. Morishita, E. di Luccio, *Biochem. Biophys. Res. Commun.* **2011**, *412*, 214–219.
- [82] P. Tuffery, P. Derreumaux, *J. R. Soc., Interface* **2012**, *9*, 20–33.
- [83] “The Nobel Prize in Chemistry 2013”, as accessed on the website: http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2013/, last accessed on 01/14/2014.
- [84] I. Buch, T. Giorgino, G. De Fabritiis, *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 10184–10189.
- [85] D. Hamelberg, J. Mongan, J. A. McCammon, *J. Chem. Phys.* **2004**, *120*, 11919–11929.
- [86] T. Zhou, A. Caflisch, *J. Chem. Theory Comput.* **2012**, *8*, 2134–2140.
- [87] A. Vitalis, A. Caflisch, *Structure* **2014**, *22*, 156–167.

Received: January 14, 2014
Accepted: February 17, 2014
Published online: ■■■■■, 0000

REVIEWS

D. Spiliotopoulos, A. Caflisch**

■■ – ■■

Molecular Dynamics Simulations of Bromodomains Reveal Binding-Site Flexibility and Multiple Binding Modes of the Natural Ligand Acetyl-Lysine

