Discovery of a hidden transient state in all bromodomain families

Lluís Raich*, Katharina Meier*, Judith Günther†, Clara D. Christ‡, Frank Noé§,‖, and Simon Olsson*,†,‡

*Department of Mathematics and Computer Science, Freie Universität Berlin, 14195 Berlin, Germany; †Computational Molecular Design, Pharmaceuticals, R&D, Bayer AG, 42096 Wuppertal, Germany; ‡Computational Molecular Design, Pharmaceuticals, R&D, Bayer AG, 13342 Berlin, Germany; and §Department of Chemistry, Rice University, Houston, TX 77005

Edited by David Baker, University of Washington, Seattle, WA, and approved December 1, 2020 (received for review August 17, 2020)

Bromodomains (BDs) are small protein modules that interact with acetylated marks in histones. These posttranslational modifications are pivotal to regulate gene expression, making BDs promising targets to treat several diseases. While the general structure of BDs is well known, their dynamical features and their interplay with other macromolecules are poorly understood, hampering the rational design of potent and selective inhibitors. Here, we combine extensive molecular dynamics simulations, Markov state modeling, and available structural data to reveal a transiently formed state that is conserved across all BD families. It involves the breaking of two backbone hydrogen bonds that anchor the ZA-loop with the α4 helix, opening a cryptic pocket that partially occludes the one associated to histone binding. By analyzing more than 1,900 experimental structures, we unveil just two adopting the hidden state, explaining why it has been previously unnoticed and providing direct structural evidence for its existence. Our results suggest that this state is an allosteric regulatory switch for BDs, potentially related to a recently unveiled BD-DNA-binding mode.

Proteins are highly dynamic biomolecules, often adopting multiple conformational states that can be critical to their function, regulation, and evolvability (1–3). Transitions between the different conformations of a protein may directly connect to changes in several properties, such as protein stability, interactions with binding partners, or the appearance of “cryptic” pockets that could be druggable but are hidden in native structures (4–7). Characterizing and determining such conformations are therefore essential to understanding the intricate relationship between structure, dynamics, and function of proteins, with a direct impact on drug design and protein engineering (8–10). Unfortunately, most of the conformations that a protein can adopt have low populations and are consequently difficult to characterize with conventional biochemical techniques. Indeed, minor conformational states may be masked under certain experimental conditions, making the detection challenging (11, 12). Nevertheless, several methods have emerged to overcome these limitations, including ambient-temperature X-ray crystallography (13, 14), NMR relaxation dispersion experiments (15, 16), and extensive computer simulations (17–20), all of which have been successful in the discovery of such minor but relevant states.

Human bromodomains (BDs) are important epigenetic regulators, recognizing or “reading” acetylation marks in histone tails (21). These domains are part of larger proteins that may also include enzymatic domains to modify chromatin structure or binding domains that can serve as platforms to recruit and activate transcription factors (22). Their central role make BD-containing proteins pivotal to regulate gene expression, representing crucial targets for the treatment of several diseases (23). Significant efforts have gone into characterizing the structure of BDs, aiming to discern the differences between the eight BD families. These studies revealed that all BDs share a common fold comprised by a four-helix bundle (α2, α3, α4, and αC) with two principal loops that connect them (ZA- and BC-loops), conforming a hydrophobic pocket that is suited for acetyl-lysine binding (Fig. 1) (24). The BC-loop contains a well-conserved hydrogen bond donor or “reader”—usually an asparagine—that is key for recognition. The ZA-loop is the longest loop and the one that displays more variability in terms of sequence and structure, presenting short helices, turns, and even hairpin insertions in BD family VIII (Fig. 1). This loop is anchored to the α4 helix via two conserved backbone hydrogen bonds (h-bonds), forming a cavity that is known as the ZA channel, an important structural feature for optimizing the potency and selectivity of BD inhibitors (23, 25). Characterization of BD flexibility has been, on the other hand, limited to a few studies reporting different crystallographic structures and molecular dynamics (MD) simulations that show subtle motions of the ZA-loop, as well as transitions between rotameric states of residues involved in ligand binding (26–29).

Although these insights have helped in the general understanding of BDs, the poor comprehension of their conformational ensembles, their biological relevance within BD-containing proteins, and their precise role in disease are hampering the advance of BD inhibitors into the clinic (30). A recent study illustrates this challenge through a detailed characterization of the impact of

Significance

Proteins dynamically shape shift, adopting different three-dimensional structures, called conformations. The conformations fulfill different functions and expose different binding surfaces to drug molecules aiming to inhibit the protein. While structural biology experiments can capture the most frequently occurring conformations, those that are more rarely visited are very difficult to capture but may be functionally important. Using extensive computer simulations, we discover an undocumented conformation present in all families of bromodomains—protein domains crucial for epigenetic regulation and associated with several diseases. We also show that available experimental data confirms the existence of this previously overlooked state. The surprising ubiquity of this new conformation makes it a likely player in biological function and opens up new drug development strategies.


Competing interest statement: K.M, J.G., and C.D.C. are employees and/or shareholders of Bayer AG.

This article is a PNAS Direct Submission. Published under the PNAS license.

1To whom correspondence may be addressed. Email: frank.noefu-berlin.de or simonols@chalmers.se.

2Present address: Department of Computer Science and Engineering, Chalmers University of Technology, 41258 Gothenburg, Sweden.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2017427118/-/DCSupplemental.

Published January 19, 2021.

https://doi.org/10.1073/pnas.2017427118
missense variation on BD function (31). The authors show that while specific variants did not alter the global fold of BDs in crystal structures, spectroscopic measurements in solution suggest changes in dynamic properties that could account for the loss of protein stability and inhibitor affinity. Furthermore, another recent study demonstrated that BDs can directly bind to DNA through a basic patch on their surface and that, in fact, DNA-binding drives the association of BDs with chromatin (32). These macroscopic interactions could affect the structure and flexibility of BDs, modulating their function and properties, which opens questions on the relevance and precise role of BDs inside their biological context.

MD simulations have been pivotal to complement experiments in understanding protein function, focusing on their dynamics and flexibility. MD simulations, in principle, grant direct access to proteins at the full spatial and temporal resolution. However, extensive simulations may be needed to sample the rare transitions between long-lived protein conformations. The combination of high-throughput MD simulations with Markov state models (MSMs) provides a powerful framework to recover long-time dynamics from multiple short-time simulations, following a “divide-and-conquer” philosophy (33). These techniques have enabled the quantitative study of critical biological processes, such as protein folding or protein–protein association (34, 35), protein plasticity and ligand-binding kinetics (36), as well as to characterize cryptic pockets in proteins and find common states across protein families (9, 37), making them very useful to evaluate the structural flexibility of BDs.

Here, using MD simulations and MSMs, we survey representative BDs from the eight families to characterize their conformational dynamics. Although the analyzed proteins share little sequence identity, we identify two recurring, transient states in which the two backbone hydrogen bonds that anchor the ZA-loop with the $\alpha_A$ helix break, and the acetyl-lysine binding pocket becomes occluded. The population of this state is generally smaller than that of the crystallographic state, suggesting why so far it has escaped experimental observation. Nonetheless, through a detailed analysis of more than 1,900 experimental structures, we reveal two BDs that adopt the hidden state, confirming its existence. We propose that this conformational change represents a mechanism of autoregulation in BDs, potentially related to a
BD-DNA–binding mode, and reveals allosteric pockets that could be leveraged to modulate both histone and DNA-binding contributions to chromatin, offering new starting points to overcome the current challenges facing the design of potent and selective BD inhibitors.

Results

All BD Families Share a Conserved and Hidden Metastable State. We started our study focusing on BRD4(1), a thoroughly characterized BD of family II. We ran 64 independent, 1-μs MD simulations in explicit solvent to build an MSM using the PyEMMA software (see SI Appendix, Computational details and methodology for details) (38). Analysis of the resulting model reveals a minor metastable state apart from the one that resembles the crystallographic structure, which involves the displacement of the ZA-loop from the α₅ helix, opening a space beneath it that increases the solvent accessible surface area of the well-conserved aspartate (Asp106, see Fig. 2). Interestingly, the conformational change disrupts the ZA channel, a structural feature that is relevant for inhibitor selectivity. In terms of interactions, the opening involves the breaking of the two conserved backbone h-bonds, whose interactions are partially compensated by the h-bonds that Gin84 establishes with Asp106, acting as a “latch.” The free energy profile along the slowest independent change timescales, and structural features (Fig. 3 and SI Appendix, Tables S4 and Figs. S10 and S11), there are subtle differences between systems that are worth noting. For instance, CECR2 is the only BD that retains a substantial percentage of h-bond 1 in the open state, and thus, the opening process is better described by the breaking of h-bond 2. BRD1 is another interesting system in this regard, since the two key h-bonds are only transiently formed in the closed state. This happens as BRD1 switches between the crystallographic state and an alternative state in which the ZA-loop is slightly displaced toward the α₅ helix, breaking the h-bonds (SI Appendix, Fig. S12). The loss of these key interactions is compensated by backbone amides of the ZA-loop, establishing h-bonds with the conserved aspartate to stabilize the closed state, resembling the interactions occurring in TRIM28 and in most BDs of family VIII (SI Appendix, Figs. S7A and S94). Interestingly, this alternative state is also present in CECR2, PHIP(2), and BAZ2B, albeit much less populated. Finally, we note that the majority of the BDs analyzed here display sporadic opening events in which either one or the two h-bonds break. These openings occur within a microsecond timescale and are generally not metastable (SI Appendix, Figs. S13 and S14). This observation suggests a duality between the classical mechanisms of “fast” induced fit and “slow” conformational selection (41), both of which may be interesting for drug design.

A Cryptic and Allosteric Pocket Opens in the Hidden State. An intriguing aspect of the conformational change is that it always disrupts the ZA channel, reducing the accessibility to the acetyllysine pocket (SI Appendix, Fig. S15). This is relevant since the ZA channel is a key structural element for the design of inhibitors, conferring them potency and selectivity (46–48). Therefore, finding pockets whose targeting could stabilize the open state would be suitable for controlling BD function. To that end, we have used the MDpocket software to analyze a thousand representative structures of each metastable state for every BD (49). The results show the presence of binding pockets in the open state, just beneath the ZA-loop (Fig. 3). We classify these cryptic pockets in two groups depending on their opening region. For instance, in CECR2, BRD4(1), BRD1, and SMARCA2, the pockets mostly open in the region where the ZA-loop was lining, next to the well-conserved aspartate, while in BAZ2B and ZMYND11, the pockets open more in between the ZA-loop and the α9 and α₅ helices. We attribute this difference to a lysine residue found in the upper part of the ZA-loop (Fig. 1B, orange star) that reorients to interact with the exposed aspartate (SI Appendix, Figs. S6 and S8). We note that BRD1 also has an alternative conformational state, not included in the model, in which the pocket is displaced toward the α9 and α₅ helices by a lysine residue (SI Appendix, Fig. S16). Interestingly, the presence of this lysine in SMARCA2 does not have the same effect, which may be due to the electrostatic shielding of two charged residues that are just at the beginning of the α₅ helix, flanking the conserved aspartate (SI Appendix, Fig. S9).

The binding sites of PHIP(2) are less defined, both in the open and closed states, given the large fluctuations of the ZA-loop,
which partially obstructs the acetyl-lysine pocket. However, relying on the opening motion and the lack of a pocket in the central part, we can include this BD in the first group together with CECR2, BRD4(1), BRD1, and SMARCA2. Interestingly, the partial obstruction of the acetyl-lysine pocket is in agreement with the fact that PHIP(2) has been difficult to target, and thus, only a recent study reports a few hits that weakly bind the outermost region of the pocket (50). These results highlight the importance of considering structural ensembles instead of single structures to determine the druggability of a pocket (51).

TRIM28 is an exception in terms of pockets given the unusual conformation of the ZA-loop. In the closed state, for instance, it increases the solvent exposure of the acetyl-lysine pocket such that it is not anymore defined, and in the open state it fluctuates so much that any clear pocket appears in the region of interest.

We assessed the druggability of the cryptic pockets in the open state compared to the acetyl-lysine pocket in the closed state (SI Appendix, Fig. S17), using a similar protocol as in the work of Porter et al. (20). We observe the distribution of scores spanning the entire range of values, with the cryptic pockets generally showing a lower druggability than the acetyl-lysine one, possibly due to their solvent exposition.

Finally, we have superposed the closed state of our models with crystal structures that have drugs bound in the acetyl-lysine pocket, finding an excellent agreement between the pose of all drugs and the computed isosurfaces (SI Appendix, Figs. S2–S9), including a deeply buried drug in SMARCA2, for which a set of conserved water molecules of the pocket are known to be displaceable (48, 52).

Crystallographic Evidence Confirms the Existence of the Open State.

To validate our results, we analyzed all the structures that are present in the bromodomain entry (PF00439) of the Pfam database (53), accounting a total of 1,903 structures (see data repository for the complete list of PDBs). We have computed the distances of the two backbone h-bonds for all structures and projected them onto the free energy landscape of BRD4(1) obtained with the same descriptors. Remarkably, the experimental distribution of h-bonds closely overlaps with the low free energy regions of the landscape, showing a small variance for the first h-bond and a more pronounced variance for the second (see Fig. 4A and SI Appendix, Fig. S18 for the projection onto the other BDs). This is a surprising result since the analyzed structures are of BDs with very different sequences, including organisms that are genetically far from humans, and yet the free energy landscape of a single BD encloses all this variation. Most importantly, this analysis has allowed us to detect outliers in the distribution and find two crystal structures that are stable in the open state (Fig. 4B and C). These structures are the ones of ZMYND11 and PB1(6), BDs that present substantial sequence variation—particularly in the ZA-loop region—compared to the general trend.

A close inspection of ZMYND11 reveals the presence of a proline residue (Pro199) in place of the residue that acts as donor for the second h-bond, impeding its formation. This chemical modification presumably contributes to destabilize the closed state in this BD. It is worth noting that in most crystal structures of ZMYND11 the ZA-loop is not resolved. In a notable exception (PDB 4N4G), authors proved that contacts with another crystallographic unit stabilize this flexible region, making it observable (54). This is consistent with our simulations, as we

Fig. 3. All bromodomain families share a hidden state that opens a cryptic pocket beneath the ZA-loop. Ensemble of structures of the “closed” crystal-like state (cyan) and the “open” state (yellow) for each BD family. MPocket frequency maps are represented by blue isosurfaces at 0.25 (light) and 0.50 (intense) isovalues, highlighting different binding sites. Percentages refer to medians of MSM populations obtained from a bootstrapped distribution (SI Appendix, Table S2). Populations for TRIM28 have not been determined due to its complex conformational landscape and neither have the ones of ZMYND11 since all structures are in the open state.
suggesting that similar interactions with other biological entities possibility to modulate BD flexibility with macromolecular contacts, Importantly, this observation provides direct evidence of the shown in Fig. 3, making it difficult to capture its electron density. 

find the ZA-loop switching between the two states that are shown in Fig. 3, making it difficult to capture its electron density. Importantly, this observation provides direct evidence of the possibility to modulate BD flexibility with macromolecular contacts, suggesting that similar interactions with other biological entities—for example, DNA—could also lead to such conformational changes.

The other crystallographic evidence is a structure of PBI(6), which is also an atypical BD having an unusually short ZA-loop (24). In comparison with SMARCA2, a member of the same family, it shows a very similar opening despite having a low sequence identity (Fig. 4C). The presence of a bulky threonine residue (Thr789) in place of the highly conserved aspartate may be one of the reasons why this BD is not stable in the closed state. We note that there are a few other BDs lacking this aspartate, and yet their crystal structures are stable in the closed state (SI Appendix, Fig. S19). In these BDs, the aspartate is replaced by residues like serine, alanine, or tryptophan, which represent drastic changes in terms of amino acid properties. Nonetheless, these modifications are accompanied by changes in the surrounding residues, leading to complementary interactions. This highlights that epistatic effects can compensate for the lack of the conserved aspartate, adapting local interactions to stabilize the closed state and retain BD function.

**Chemical Shift Predictions Suggest a Link between the Hidden State and a DNA-Binding Mode.** Understanding the biological relevance of conformational states in a protein is crucial for the success of inhibitor design. However, this is not an easy task given the number of possible interactions that can occur in a cellular context. A recent study has demonstrated that SMARCA2 interacts with DNA through a basic surface patch, whose electrostatic contribution drive BD association to chromatin (32). Reported chemical shifts upon BD-DNA binding show significant perturbations in the region of the ZA-loop and the α helix, adjacent to where we observe the conformational change. This observation prompted us to compute the chemical shift of 1H, 13C, and 15N atoms along the process, using the CamShift empirical predictor (SI Appendix, Figs. S20 and S21) (55). Intriguingly, comparison of the predicted 1H shifts with the experimental values reveal a clear perturbation for Leu1412 (Fig. 5A), which is hydrogen bonded with the conserved Asp1430 (Fig. 5B). This interaction can break both in the closed and the open state, changing the local environment of Leu1412 and lowering its 1H shift. Therefore, it seems plausible that the experimental perturbation also relates to the breaking of the same h-bond, either in the closed or the open state, likely due to direct contacts with DNA. These results suggest that SMARCA2 could undergo a population shift from the closed state to the open as a consequence of the interactions with DNA (Fig. 5C), modulating its flexibility and function in an allosteric manner. This hypothesis opens the door for future experiments to further describe the nature of the open state and its potential relevance to DNA binding.

**Discussion**

In this work, we have used extensive MD simulations and Markov state modeling to characterize the dynamics and flexibility of BDs. Our results reveal a hidden and ubiquitous state in which the ZA-loop displaces from the α helix, disrupting the ZA channel and opening a pocket beneath it. The conservation of this state across families, persisting along evolutionary pressures, of conserved residues that could help in the rationalization of...

---

Raich et al.
Discovery of a hidden transient state in all bromodomain families

PNAS | 5 of 7
https://doi.org/10.1073/pnas.2017427118
and experimental shifts upon its binding to DNA (5′Leu1412 1HN shift for the closed and open states (cyan and yellow, respectively). The numbering of residues is that of Morrison et al. (32). The numbering of residues is that of the open state to occlude the acetyl-lysine pocket and regulate BD function.

Most importantly, our detailed chemical shift analyses of the opening process reveal signals that may be particularly sensitive to the general understanding of BD structure, dynamics, and function, providing a mechanism of potential autoregulation that deepens in BD biology and paves the way toward the design of future inhibitors with higher potency and selectivity.

Materials and Methods

We performed MD simulations at 300 K and 1 bar using OpenMM (63), with the ff14SB (64) force field parameters to describe the solute and the TIP3P (65) water model to describe the solvent. We ran multiple 1-μs trajectories for each bromodomain and subsequently used the PyEMMA (38) software to build MSMs from this data. MDpocket (49) was used to detect small molecule binding sites, and Plumed (66) was used to predict chemical shifts via the CamShift model (55). Bromodomain structures of the Pfam database (PF00439) were downloaded from the PDB and superimposed with Theseus (67). Further methodological details, convergence tests, and analyses are available in the Materials and Methods section of the supplementary information.

Data Availability. All molecular trajectories and notebooks to reproduce the results of this paper are available in Zenodo at https://doi.org/10.5281/zenodo.4343283.

ACKNOWLEDGMENTS. We acknowledge funding from the European Commission (ERC CoG 772230 “ScaleCell”), the Deutsche Forschungsgemeinschaft (CRC1114/C03, NO825/2-2), the MATH+ Berlin Mathematics Research Center (AA1x6, EF1x2), and Bayer AG (LSC project: "QuAnTaStics"). We thank Jacopo Negroni, Joseph Bluck, Michael Edmund Beck, Sikander Hayat, Jeremie Mortier, Holger Steuber, and Nicolas Werbeck for feedback and discussions at different stages during this study. We also thank Emma A. Morrison and Catherine A. Musselman for sharing raw chemical shift data.

Altogether, we believe that our results significantly contribute to the general understanding of BD structure, dynamics, and function, providing a mechanism of potential autoregulation that deepens in BD biology and paves the way toward the design of future inhibitors with higher potency and selectivity.

Materials and Methods

We performed MD simulations at 300 K and 1 bar using OpenMM (63), with the ff14SB (64) force field parameters to describe the solute and the TIP3P (65) water model to describe the solvent. We ran multiple 1-μs trajectories for each bromodomain and subsequently used the PyEMMA (38) software to build MSMs from this data. MDpocket (49) was used to detect small molecule binding sites, and Plumed (66) was used to predict chemical shifts via the CamShift model (55). Bromodomain structures of the Pfam database (PF00439) were downloaded from the PDB and superimposed with Theseus (67). Further methodological details, convergence tests, and analyses are available in the Materials and Methods section of the supplementary information.

Data Availability. All molecular trajectories and notebooks to reproduce the results of this paper are available in Zenodo at https://doi.org/10.5281/zenodo.4343283.

ACKNOWLEDGMENTS. We acknowledge funding from the European Commission (ERC CoG 772230 “ScaleCell”), the Deutsche Forschungsgemeinschaft (CRC1114/C03, NO825/2-2), the MATH+ Berlin Mathematics Research Center (AA1x6, EF1x2), and Bayer AG (LSC project: "QuAnTaStics"). We thank Jacopo Negroni, Joseph Bluck, Michael Edmund Beck, Sikander Hayat, Jeremie Mortier, Holger Steuber, and Nicolas Werbeck for feedback and discussions at different stages during this study. We also thank Emma A. Morrison and Catherine A. Musselman for sharing raw chemical shift data.

Altogether, we believe that our results significantly contribute to the general understanding of BD structure, dynamics, and function, providing a mechanism of potential autoregulation that deepens in BD biology and paves the way toward the design of future inhibitors with higher potency and selectivity.

Materials and Methods

We performed MD simulations at 300 K and 1 bar using OpenMM (63), with the ff14SB (64) force field parameters to describe the solute and the TIP3P (65) water model to describe the solvent. We ran multiple 1-μs trajectories for each bromodomain and subsequently used the PyEMMA (38) software to build MSMs from this data. MDpocket (49) was used to detect small molecule binding sites, and Plumed (66) was used to predict chemical shifts via the CamShift model (55). Bromodomain structures of the Pfam database (PF00439) were downloaded from the PDB and superimposed with Theseus (67). Further methodological details, convergence tests, and analyses are available in the Materials and Methods section of the supplementary information.

Data Availability. All molecular trajectories and notebooks to reproduce the results of this paper are available in Zenodo at https://doi.org/10.5281/zenodo.4343283.

ACKNOWLEDGMENTS. We acknowledge funding from the European Commission (ERC CoG 772230 “ScaleCell”), the Deutsche Forschungsgemeinschaft (CRC1114/C03, NO825/2-2), the MATH+ Berlin Mathematics Research Center (AA1x6, EF1x2), the Alexander von Humboldt Foundation (postdoctoral fellowship to S.O.), and Bayer AG (LSC project: “QuAnTaStics”). We thank Jacopo Negroni, Joseph Bluck, Michael Edmund Beck, Sikander Hayat, Jeremie Mortier, Holger Steuber, and Nicolas Werbeck for feedback and discussions at different stages during this study. We also thank Emma A. Morrison and Catherine A. Musselman for sharing raw chemical shift data.

Altogether, we believe that our results significantly contribute to the general understanding of BD structure, dynamics, and function, providing a mechanism of potential autoregulation that deepens in BD biology and paves the way toward the design of future inhibitors with higher potency and selectivity.

Materials and Methods

We performed MD simulations at 300 K and 1 bar using OpenMM (63), with the ff14SB (64) force field parameters to describe the solute and the TIP3P (65) water model to describe the solvent. We ran multiple 1-μs trajectories for each bromodomain and subsequently used the PyEMMA (38) software to build MSMs from this data. MDpocket (49) was used to detect small molecule binding sites, and Plumed (66) was used to predict chemical shifts via the CamShift model (55). Bromodomain structures of the Pfam database (PF00439) were downloaded from the PDB and superimposed with Theseus (67). Further methodological details, convergence tests, and analyses are available in the Materials and Methods section of the supplementary information.

Data Availability. All molecular trajectories and notebooks to reproduce the results of this paper are available in Zenodo at https://doi.org/10.5281/zenodo.4343283.

ACKNOWLEDGMENTS. We acknowledge funding from the European Commission (ERC CoG 772230 “ScaleCell”), the Deutsche Forschungsgemeinschaft (CRC1114/C03, NO825/2-2), the MATH+ Berlin Mathematics Research Center (AA1x6, EF1x2), the Alexander von Humboldt Foundation (postdoctoral fellowship to S.O.), and Bayer AG (LSC project: “QuAnTaStics”). We thank Jacopo Negroni, Joseph Bluck, Michael Edmund Beck, Sikander Hayat, Jeremie Mortier, Holger Steuber, and Nicolas Werbeck for feedback and discussions at different stages during this study. We also thank Emma A. Morrison and Catherine A. Musselman for sharing raw chemical shift data.