Well-Tempered Metadynamics Simulations Predict the Structural and Dynamic Properties of a Chiral 24-Atom Macrocycle in Solution

Riccardo Capelli, Alexander J. Menke, Hongjun Pan, Benjamin G. Janesko, Eric E. Simanek,* and Giovanni M. Pavan*

ABSTRACT: Inspired by therapeutic potential, the molecular engineering of macrocycles is garnering increased interest. Exercising control with design, however, is challenging due to the dynamic behavior that these molecules must demonstrate in order to be bioactive. Herein, the value of metadynamics simulations is demonstrated: the free-energy surfaces calculated reveal folded and flattened accessible conformations of a 24-atom macrocycle separated by barriers of ∼6 kT under experimentally relevant conditions. Simulations reveal that the dominant conformer is folded—an observation consistent with a solid-state structure determined by X-ray crystallography and a network of Roses established by 1H NMR. Simulations suggest that the macrocycle exists as a rapidly interconverting pair of enantiomeric, folded structures. Experimentally, 1H NMR shows a single species at room temperature. However, at lower temperature, the interconversion rate between these enantiomers becomes markedly slower, resulting in the decoalescence of enantiotopic methylene protons into diastereotopic, distinguishable resonances due to the persistence of conformational chirality. The emergence of conformational chirality provides critical experimental support for the simulations, revealing the dynamic nature of the scaffold—a trait deemed critical for oral bioactivity.

INTRODUCTION

Current interest in macrocycles derives in great part from their potential application in medicine.1−5 Unlike small-molecule drugs that traditionally bind to well-defined active or allosteric sites, macrocycles can also access an alternate therapeutic paradigm by targeting protein−protein interfaces which are often disordered.6 Recognition of conformationally fluid regions on proteins can occur if a macrocycle adopts an extended conformation to display the requisite array of functional groups necessary for productive binding. However, this strategy presents a challenge. That is, an array of polar interactions can preclude transport across membranes and/or oral bioavailability.7,8 Accordingly, a dynamic “chameleon-like” character is desired wherein the macrocycle can adopt extended conformations that facilitate molecular recognition and compact structures that enhance lipophilicity to promote cellular access.9,10

Macrocycle G−G represents the simplest, most flexible macrocyclic scaffold of a panoply of potential compounds which can be accessed by functionalization of any of the 12 sites on this scaffold including the choices of (1) the auxiliary group (here, dimethylamine), (2) N-alkylation state of the hydrazine (here, H), (3) length of the acetal (here, three carbons), choice of the amino acid (here, G), and its N-alkylation state (here, H). Indeed, the opportunity for structural diversity on this scaffold that sits—based on its size and composition—on the border of molecules that are expected to follow Lipinski’s rules of 5 (Ro5) and those that are beyond them (BRo5) makes it a compelling choice for us to study.8−10 The glycine macrocycle, G−G, derives from acid-catalyzed dimerization of a simple monomer comprising a central triazine and auxiliary group (dimethylamine), a protected hydrazine group, and an acetal pendant on the glycine linker. It is available rapidly in three synthetic steps requiring two chromatographic purifications.

Predicting the conformational flexibility of a macrocycle prior to synthesis can greatly impact molecular design.11,12 Molecular simulations are a useful tool to this end. While classical molecular dynamics (MD) simulations allow the exploration of configurations that differ by ∼2 kT in free energy, enhanced sampling approaches such as metadynamics are ideally suited for this task13,14 by preventing entrapment into local free energy minima. Metadynamics allows for an exhaustive exploration of the conformational landscape of complex macromolecules.15 Here, we use well-tempered metadynamics (WT-MetaD)16 simulations to characterize the conformational landscape of the 24-atom macrocycle referred
to as G−G (Chart 1) in the explicit solvent and under experimentally relevant conditions.

Chart 1. G−G Macrocycle, D1 and D2 WT-MetaD Collective Variables (CVs), and Labels Used in the NMR Spectra. The Solid-State Structure of a Morpholine Derivative Is Shown with the Solvent and Counterions Omitted for Clarity

■ RESULTS AND DISCUSSION

An all-atom G−G macrocycle model was built using Antechamber\(^\text{17}\) with atomic point charges parameterized using the restrained electrostatic potential (RESP) approach\(^\text{18}\) based on HF/(6-31)G\(^*\) calculations performed using Gaussian \(^\text{16}.\)^\(^19\) After an equilibration protocol (see details in the Computational Methods section), a 200 ns-long WT-MetaD simulation was run using D1 and D2 as collective variables (CVs) which discriminate between folded and open conformations. All the simulations were carried out using GROMACS\(^\text{20}\) patched with PLUMED.\(^\text{21,22}\) After reaching convergence, the WT-MetaD results were reweighted using the Tiwary–Parrinello free energy estimator\(^\text{23}\) to obtain free energy surfaces (FESs) (Figure 1). For the CVs, descriptors of the conformational changes in the macrocycle structure, we use the distances indicated as D1 and D2 of G−G in Chart 1 (in red). These two distance parameters, D1 and D2, were found to work well as CVs in the MetaD simulations. Distance D1 samples the folding/unfolding of G−G (recrossing between the folded and extended macrocycle conformations). Distance D2 probes the repulsion/interaction between the protonated triazines (significant for the folding/unfolding of G−G).

During the WT-MetaD simulation, the system is biased to explore all the possible combinations of D1 and D2, where the macrocycle visits different conformations. From well-converged WT-MetaD, we reconstruct complete FESs describing the conformations accessible by the macrocycle under various conditions. Figure 1 (left) shows the FESs, projected on the D1−D2 plane, obtained from the WT-MetaD simulations of the macrocycle in different solvents. All the conformations that belong to every identified minimum can be extracted using a clustering algorithm (based on the all-atom root mean square deviation of all the conformations), yielding a statistically significant centroid of the conformational ensemble. The FESs also provide information on the magnitude of the energetic barriers between the dominant minima. The data obtained by our calculations offer a valuable qualitative insight into the structural dynamics of this macrocycle and into its “chameleon-like” behavior.

WT-MetaD reveals a FES with relatively shallow minima. That is, in water, the regions indicated A/A′ and B are only 6
kT (<4 kcal/mol) lower than the highest-energy conformations. Although the height of barriers estimated via WT-MetaD should be considered qualitatively, the obtained barrier heights are consistent with expectation (e.g., for what pertains to the interconversion of these isomers via rotation around σ-bonds of the aliphatic domain between the hydrazone and amino acid). Similar “shallow minima” are essentially present in all the FESs that are calculated herein, leading us to conclude that G−G exists as a highly dynamic molecule under the simulated conditions. At room temperature, we would expect to see G−G exists as a highly dynamic molecule under the simulated conditions. At room temperature, we would expect to see 3H NMR reflecting resonances corresponding to the average of these conformations with rapid interconversion between A/A’ and B. Indeed, we see only one set of resonances in all of these solvents.

Closer inspection of the FESs, however, suggests that the solvent may play some role in conformation preference. For example, the surfaces that describe the conformational landscape in water and methanol are quite similar. At low values of D1 and D2 (lower left corner of the FESs), a pair of enantiomeric, folded structures, A and A’, emerge that show extensive π−π overlap. At higher values of D1 and D2 (upper right in the FESs), an extended structure, B, is seen. Slight differences in the shape and centers of these minima are revealed by representative structures, which show that the regions of π−π overlap can easily shift and twist, generating a cluster of conformations around A/A’ which are relatively similar in free energy.24 The shapes and centers of the minima for structures associated with B are more conserved, which reflects its open, extended conformation. In order to facilitate the comparison of the behavior of G−G in water with that in acetonitrile and dimethyl sulfoxide (DMSO) solvents, we shift from examination of the FESs to the probability distributions that can be calculated from them.

From the FESs, it is also possible to obtain information on the relative probabilities \( [P(r)] \) for the populations of all sampled conformers based on their relative free energy levels (Figure 1: right) as in eq 1

\[
P(r) \propto e^{-\frac{G(r)}{kT}}
\]

The probability maps included in Figure 1 (right) lead us to the conclusion that a folded conformation (A/A’) is populated and present in all solvents, meaning that G−G folding is possible in all cases. Methanol and acetonitrile show population densities that are similar in dimension/shape. DMSO shows two density peaks corresponding to A/A’. Water shows a larger density peak which suggests the additive sum of the previous three leading to the conclusion that in water, the folded state is more dynamic/diverse. In methanol, we observe that the peaks of the folded and unfolded conformers are very similar, which suggests that in this solvent, G−G is better solvated, and it has equal probability to exist in the folded or unfolded state. These observations must be tempered with recognition that barriers between these structures on the surface are low and the experimental observation that a single set of resonances is seen in the 1H NMR spectra of these species (a signature of rapidly interconverting conformers). Even in methanol, where a second, extended conformation (B) is populated, only one set of resonances is observed by NMR, which suggests that the folded and unfolded conformations interconvert.

Table 1 shows the estimated A/A’ and B populations calculated in four solvents. The designation “other” refers to all other sampled conformations obtained by summing the remaining FES points. The areas associated with A/A’ and B were calculated by including all conformations with \( \Delta G < 3 \) kT from the minima.

The predictions from WT-MetaD simulations, both in terms of the structure and dynamic behavior of G−G, are supported by the experiments. A folded structure for a derivative of G−G is observed using single-crystal X-ray diffraction.

Consistent with the prediction of a compact structure, the solid-state structure is found to be symmetric and shows extensive π−π stacking of the triazine ring upon the hydrazone. Both 1D and 2D NMR spectroscopy experiments further support the evidence from the WT-MetaD simulations. At room temperature, the NMR spectra of G−G in DMSO-d6, CD3OD/H2O, CD3CN, and D2O/H2O show a single set of resonances corresponding to equivalent environments for each half of the symmetric dimer.

Consistent with prediction, the 1H NMR spectra show similar chemical shifts across the solvents surveyed. The spectrum in DMSO-d6 (Figure 2, top) is well resolved except

**Table 1. Populations (in %) of Conformers as Predicted by WT-MetaD Simulations in Various Explicit Solvents**

| isomer | water | DMSO | MeCN | MeOH |
|--------|-------|------|------|------|
| A/A’   | 80    | 52   | 65   | 43   |
| B      | 5     | 11   | 2    | 20   |
| other  | 15    | 37   | 33   | 37   |

for the broad resonance of C, a feature we attribute as indicative of the locus for dynamic motion for interconversion of A ⇔ B ⇔ A’. Critical rOes include H’ to α-NH (site of protonation), NHN−CH− (trans-hydrazone), and C−NH to both α and H2CNCH3 (amide orientation).

The most interesting experimental support for the dynamic interconversion of A ⇔ B ⇔ A’ derives from low-temperature
NMR obtained in CD$_3$CN (Figure 2). As the temperature is decreased, in the fingerprint region of the spectra from 7 to 13 ppm, the number of resonances remains constant. Rapidly exchanging, downfield NH protons sharpen, and other resonances—that is, the hydrazone methine proton (NHN=CH−) and amide NH—broaden as expected due to viscosity. Upfield, however, the α and C methylene protons of G−G decoalesce.

This markedly different behavior is not consistent with the emergence of a new conformation per se because we would expect the chemical environment of all resonances to be affected. Instead, this behavior reflects the emergence of diastereomers resulting from the slowing of the interconversion process A ⇔ B ⇔ A’. That is, methylenes α, B, and C are enantiotopic. They are interconverted by a transition between enantiomeric conformations of the macrocycle at higher temperatures. However, at −35 °C, the folded conformations persist longer than the NMR timescale, and we observe the emergence of conformational chirality—reviewed recently by De Riccardis.26 Accordingly, these enantiotopic methylenes are rendered diastereotopic and are resolved in the spectra. At −20 °C in CD$_3$OD, the B methylene becomes diastereotopic as well.

CONCLUSIONS

In this work, we show how WT-MetaD simulations can provide precious qualitative insights, useful to predict and characterize the key structural and dynamical features of the G−G macrocycle in various solvents, providing evidence of the “chameleon-like” dynamic behavior of such a macrocycle under experimentally relevant conditions. As increased attention turns to controlling macrocycle conformation,27,28 identifying accessible conformations and characterizing the barriers of interconversion between them offer compelling motivation for its broadening application.29,30 Its use to survey the effects of substitution on different sites on this class of macrocycles is ongoing.

COMPUTATIONAL METHODS

Parametrization, Minimization, and Equilibration. The atomistic G−G macrocycle model was parametrized using Antechamber17 according to the Generalized Amber Force Field (GAFF)31 with atomic point charges parameterized using the RESP approach18 based on HF/(6-31)G* calculations performed using Gaussian.16,19 The G−G macrocycle model was inserted in a periodic simulation box that was filled with different types of explicit solvents. In detail, we obtained atomistic models for the G−G macrocycle solvated in explicit water (TIP3P),32 DMSO,33 acetonitrile,34 or methanol.35 All simulations were run under periodic boundary conditions. The systems were initially minimized using the steepest descent algorithm for 2000 steps and then with a conjugate gradient algorithm for further 2000 steps. After minimization, each model system underwent preliminary MD simulations to allow the systems to reach a temperature of 298 K (under NVT conditions for 5 ns) and a pressure of 1 bar (under NPT conditions). In the NPT MD phase, we used the Berendsen barostat36 for 1 ns, imposing a pressure of 1 bar (coupling time of 1 ps), followed by the last NPT equilibration run using the Parrinello—Rahman barostat37 set at 1 bar (coupling time of 1 ps) for 1 ns. All simulations ran in this work used a velocity-rescale thermostat38 (at 298 K, a coupling time of 0.2 ps) and a simulation time step of 2 fs, allowed by the use of LINCS39 constraints for all the bonds that involve hydrogen atoms. All simulations were performed with GROMACS 2021.2,40 patched with PLUMED 2.7.21,22

Well-Tempered Metadynamics Simulations. For all the four solvents considered herein, we performed 200 ns-long46 WT-MetaD simulations using the same MD parameters as those used in the last equilibration step (see the previous section). During this simulation time, all systems reached convergence. As descriptors of the G−G macrocycle conformations (and conformational changes), we used two CVs, namely, distances D1 and D2 in Chart 1 in main text. For both variables, we set a sigma of 0.025 nm, a Gaussian height of 1.2 kJ/mol, and a bias factor of 25. During the WT-MetaD runs, we deposited a Gaussian on the landscape every 1 ps. After the WT-MetaD simulations, a reweighting procedure was performed using the Twary—Parrinello estimator.23 For the analyses of the FESs (and population probabilities), we considered the last 100 ns of our simulations.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.2c03536.

Detailed experimental procedures, spectra, and additional computational calculations (PDF)

AUTHOR INFORMATION

Corresponding Authors

Eric E. Simanek — Department of Chemistry & Biochemistry, Texas Christian University, Fort Worth, Texas 76129, United States; orcid.org/0000-0002-3195-4523; Email: e.simanek@tcu.edu

Giovanni M. Pavan — Department of Applied Science and Technology, Politecnico di Torino, 10129 Torino, Italy; Department of Innovative Technologies, University of Applied Sciences and Arts of Southern Switzerland, 6962 Lugano-Viganello, Switzerland; orcid.org/0000-0002-3473-8471; Email: giovanni.pavan@polito.it

Authors

Riccardo Capelli — Department of Applied Science and Technology, Politecnico di Torino, 10129 Torino, Italy; orcid.org/0000-0001-9522-3132

Alexander J. Menke — Department of Chemistry & Biochemistry, Texas Christian University, Fort Worth, Texas 76129, United States

Hongjun Pan — Department of Chemistry, University of North Texas, Denton, Texas 76129, United States

Benjamin G. Janesko — Department of Chemistry & Biochemistry, Texas Christian University, Fort Worth, Texas 76129, United States; orcid.org/0000-0002-2572-5273

Complete contact information is available at: https://pubs.acs.org/doi/10.1021/acsomega.2c03536

Notes

The authors declare no competing financial interest. Complete input, simulation, and data analysis files are available at the Zenodo repository https://doi.org/10.5281/zenodo.6967170.
G.M.P. acknowledges the funding received from the European Research Council (ERC) under European Union’s Horizon 2020 research and innovation programme (grant agreement no. 818776—DYNAPOL). E.E.S. thanks the Robert A. Welch Foundation (P-0008) and the NIH (NIGMS R15 GM135900) for support. B.G.J. thanks the TCU High-Performance Computing Center used for portions of the computations.

REFERENCES

(1) Kelly, C. N.; Townsend, C. E.; Jain, A. N.; Naylor, M. R.; Pye, C. R.; Schwochott, J.; Lokey, R. S. Geometrically Diverse Lariat Peptide Scaffolds Reveal an Untapped Chemical Space of High Membrane Permeability. J. Am. Chem. Soc. 2021, 143, 795–714.

(2) Vinogradov, A. A.; Yin, Y.; Suga, H. Macrocyclic Peptides as Drug Candidates: Recent Progress and Remaining Challenges. J. Am. Chem. Soc. 2019, 141, 4167–4181.

(3) Reuther, J. F.; Goodrich, A. C.; Escamilla, P. R.; Lu, T. A.; Del Rio, V.; Davies, B. W.; Ansynl, E. V. A Versatile Approach to Noncanonical, Dynamic Covalent Single- and Multi-Loop Peptide Macrocycles for Enhancing Antimicrobial Activity. J. Am. Chem. Soc. 2018, 140, 3768–3774.

(4) Kobori, S.; Huh, S.; Appavoo, S. D.; Yudin, A. K. Two-Dimensional Barriers for Probing Conformational Shifts in Macrocycles. J. Am. Chem. Soc. 2021, 143, 5166–5171.

(5) Usanov, D. L.; Chan, A. I.; Maianti, J. P.; Liu, D. R. Second-Generation DNA- Templated Macrocyclic Libraries for the Discovery of Bioactive Small Molecules. Nat. Chem. 2018, 10, 704–714.

(6) Villar, E. A.; Beglov, D.; Chennamadhavuni, S.; Porco, J. A.; Kozakov, D.; Vajda, S.; Whitty, A. How Proteins Bind Macromolecules. Nat. Chem. Biol. 2014, 10, 723–731.

(7) Dougherty, P. G.; Qian, Z.; Pei, D. Macromolecules as Protein-Protein Interaction Inhibitors. Biochem. J. 2017, 474, 1109–1125.

(8) DeGoey, D. A.; Chen, H. J.; Cox, P. B.; Wendt, M. D. Beyond the Rule of 5: Lessons Learned from AbbVie’s Drugs and Compound Collection. J. Med. Chem. 2018, 61, 2636–2651.

(9) Furukawa, A.; Schwochott, J.; Pye, C. R.; Asano, D.; Edmondson, Q. D.; Turmon, A. C.; Klein, V. G.; Ono, S.; Okada, O.; Lokey, R. S. Drug-Like Properties in Macromolecules above MW 1000: Backbone Rigidity versus Side-Chain Lipophilicity. Angew. Chem., Int. Ed. 2020, 59, 21571–21577.

(10)Rossi Sebastiano, M.; Doak, B. C.; Backlund, M.; Pavan, G. M.; Barducci, A.; Albertazzi, L.; McTiernan, T.; Jwad, R.; Weissberger, D.; Hunter, L.; McConkey, D.; Zollinger, A.; Zollinger, M.; Musco, D.; Kollman, P. A. Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. J. Phys. Chem. B 2015, 119, 736–742.

(11)Sharma, V. R.; Mehmoood, A.; Janesko, B. G.; Simonek, E. Efficient Syntheses of Macrocycles Ranging from 22-28 Atoms through Spontaneous Dimerization to Yield Bis-Hydrzones. RSC Adv. 2020, 10, 3217–3220.

(12) De Riccardis, P.; Hamelryck, T.; Murgia, F.; van Gunsteren, W. F.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. J. Phys. Chem. B 2004, 108, 9473–9789.

(13) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. a.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. Gaussian 16; Gaussian Inc.: Wallingford CT, 2016.

(14) Abraham, M. J.; Murtola, T.; Schulz, R.; Päll, S.; Smith, J. C.; Hess, B.; Lindahl, E. Gromacs: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. SoftwareX 2015, 1-2, 19–25.

(15) Essmann, U.; Perera, L.; Berkowitz, M. S.; Darden, T. A.; and Lee, H. J. Comm. 2002, 1, 2651–2652.

(16) Beckstein, O.; Jorgensen, W. L.; and Noid, D. W. Development and Testing of a General Amber Force Field. J. Phys. Chem. B 2003, 107, 7943–7949.
(37) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* 1981, 52, 7182–7190.

(38) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* 2007, 126, 014101.

(39) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* 1997, 18, 1463–1472.