RSC Chemical Biology

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Check for updates

Cite this: *RSC Chem. Biol.*, 2022, 3, 739

Navigating complex peptide structures using macrocycle conformational maps†

Timothy J. McTiernan, Diego B. Diaz, George J. Saunders, Fiona Sprang and Andrei K. Yudin *

Received 18th January 2022,
Accepted 14th April 2022

DOI: 10.1039/d2cb00016d

rsc.li/rsc-chembio

Identification of turn motifs that are stabilized by intramolecular hydrogen bonds can be useful in describing the conformation of peptide systems. However, this approach is somewhat insufficient for cyclic peptides because peptide regions that are not positioned within a hydrogen bond can be left with no description. Furthermore, non-regular secondary structures and other rarely-observed conformations can be left without detailed evaluation. Herein, we describe "higher-order" ϕ/ψ plots termed macrocycle conformational maps (MCMs) as a tool for evaluating and comparing the conformations of a series of structurally related macrocyclic peptides.

Open Access Article. Published on 19 April 2022. Downloaded on 7/28/2024 10:21:48 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Introduction

Macrocyclic peptides are prominent structures among natural products as well as synthetic bioactive molecules.[1–4] Due to the number of rotatable bonds in their backbone, macrocyclic peptides can access a broad range of feasible conformations that are close in free energy.[5,6] Relatively minor modifications to the structure of such systems can result in drastic conformational changes.[7–9] The network of interacting structural elements in macrocyclic peptides gives rise to conformational organisation that is challenging to predict and characterise.[10–12]

In broad terms, macrocyclic peptide structures are distinguished for their "hub-and-rotor" architecture.[13–15] In addition to fully rotatable bonds (rotors), the peptide backbone contains hub sub-structures that possess limited rotational opportunities. Backbone amide bonds generally adopt the *trans*-configuration ($\omega = 180^\circ$) and are considered hubs due to their restricted rotational freedom and planarity. Rotors in the backbone are described by the phi (ϕ) and psi (ψ) dihedral angles.[16–18] In synthetically modified peptide systems, unnatural structural elements can also be judged based on their rotational properties (Fig. 1a). The global minimum energy conformation of a macrocyclic peptide structure arises from the dynamic interplay of hubs and rotors in the backbone, as these elements engage in coupled bond rotations.[7,19] Coupling in bond rotations describes the phenomenon wherein the rotation of one bond in a molecule influences the rotation of neighbouring bonds, propagating conformational effects throughout the molecule as a function of the degree of rotational coupling.[20–22]

The Ramachandran plot is an analytical tool describing the geometry of the backbone rotors ϕ and ψ in proteins and peptides (Fig. 1b).[23] The ϕ/ψ plot has been conventionally used for protein structure validation, and has recently found utility in the evaluation of both protein geometry and conformational trends among amino acids.[24–29] Recently, Jayaram and coworkers demonstrated that the three-dimensional conformation of a protein or peptide can be reconstructed from an alphanumeric string representing both the ϕ/ψ coordinates and the structural identity of each amino acid in a sequence.[30] In this context, "higher-order" ϕ/ψ plots representing sequential residues can be created to map the global conformation of peptide structures and to probe the local conformational effects of neighbouring residues. An illustrative example of such multi-residue ϕ/ψ plots can be found in the analysis of distinct β-turn types. Numerous β-turn types have been described. These motifs share the property of $(i \rightarrow i + 2)$ hydrogen bonding, but differ in the distribution of ϕ/ψ angles, which annotate the amino acid residues constrained within the turn and define the turn's type.[31] Two-residue ϕ/ψ plots representing distinct β-turn types map the unique regions of conformational space occupied by each of these constructs (Fig. 1c).

We sought to apply the logic of ϕ/ψ conformational mapping to develop a system for classifying macrocyclic peptide structures. This approach became particularly meaningful after we outlined work on conformational "dark space" in macrocyclic peptides. As part of this investigation, we were able to stabilize certain conformations in solution that were difficult to describe in conventional terms. In general, dark space refers to metastable peptide conformations not frequently observed in nature. A specific example of dark space can be found when

Davenport Research Laboratories, Department of Chemistry, University of Toronto, Toronto, ON M5S 3H6, Canada. E-mail: andrei.yudin@utoronto.ca

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d2cb00016d

© 2022 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 19 April 2022. Downloaded on 7/28/2024 10:21:48 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

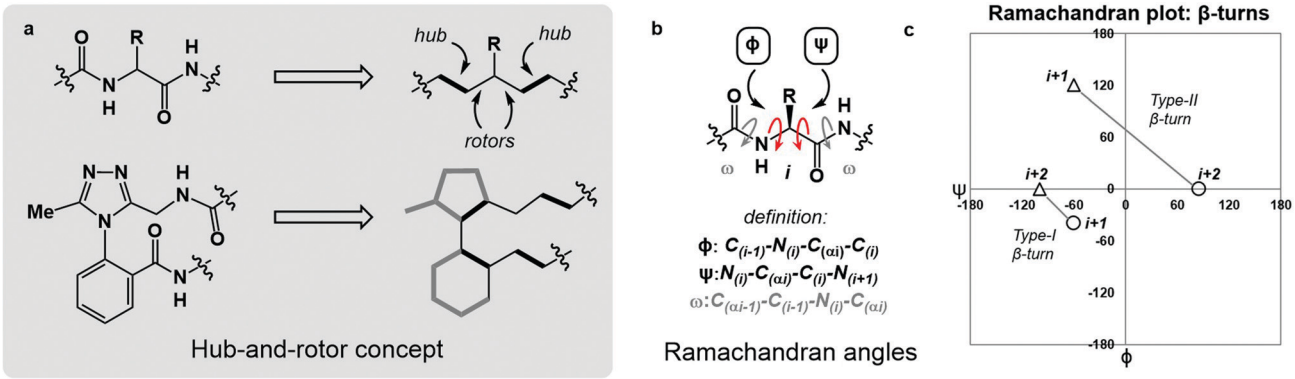


Fig. 1 An overview of angular constraints in peptidic and non-peptidic backbones. (a) The hub-and-rotor concept; (b) definition of Ramachandran angles ϕ, ψ and ω; (c) plotting ϕ and ψ for i + 1 and i + 2 residues of both Type I and II β-turns illustrates their respective conformational maps.

backbone amino acids exhibiting rare left-handed helical conformations are over-represented because of conformational restriction in a macrocycle.[32] In order to observe dark space conformations on a relevant timescale, structural constraints must be integrated into macrocyclic systems. Accordingly, we developed a synthetic motif termed the dominant rotor. This motif bears an internal atropisomer exhibiting a large kinetic barrier to rotation ($\Delta G^{\ddagger} > 27$ kcal mol^{-1}). Upon incorporation into a macrocycle, the dominant rotor gives rise to two isolable atropisomeric species that are different in free energy. The difference in free energy between atropisomers largely reflects the backbone's degree of strain resulting from the conformational restriction imparted by the dominant rotor. Overall, the dominant rotor allows isolation of metastable ''dark space'' macrocycles that can exhibit unconventional conformational behaviour.

Macrocyclic peptide systems can engage in intramolecular reactions that impact their conformation. The net conformation of a macrocyclic peptide represents an interplay of all stabilizing and destabilizing interactions in the system.[33] The so-called composite barriers are associated with processes that induce coupled bond rotations. For instance, intramolecular chemical transformations in macrocycles such as the Boulton–Katritzky rearrangement can be used as tuneable barriers that impact the system's conformation.[13,34] Atropisomerisation of dominant rotor-containing macrocycles can lead to another type of composite barrier when the high-energy rotation of the dominant rotor gets coupled to distant rotors in the system, giving rise to conformational differentiation.

The lack of robust methodology for describing macrocyclic peptide conformation is at the heart of the challenge in evaluating these constructs. Classically, ''turn types'' have dominated the discourse on macrocyclic peptide conformation. While classical turn types adequately describe some local regions of proteins and linear peptides, the cyclic nature of macrocyclic systems can render this approach inadequate. Exclusive evaluation of macrocyclic peptides according to their internal turn structures inevitably leaves regions of the ring that do not engage in classically defined turns without analysis. Furthermore, this approach ultimately leaves macrocycles that engage in non-regular secondary structures (NORS) without adequate options for conformational description.[35–37] A comprehensive approach to evaluating macrocyclic peptide conformation must be established to meet these challenges.

In this paper, we define multi-residue graphical representations of macrocyclic peptide conformation as ''macrocycle conformation maps'' (MCMs). The sequential pattern of datapoints on the MCM describes the global conformation of a given macrocycle, and the coordinates of each individual datapoint describe the local conformation at each constituent amino acid. Furthermore, the conformational relationship between two macrocycles of identical ring size and related rotor structure can be described by their differences in MCM coordinates (ΔMCM). The difference in ϕ/ψ for each residue in a pair of macrocycles gives rise to a pattern of ΔMCM ''vectors'' corresponding to the dihedral angular differences between conformations. A vector oriented parallel to the ϕ-axis suggests that the conformational difference between species at this residue occurs primarily in the ϕ-dimension. Similarly, a vector oriented parallel to the ψ-axis represents a conformational difference occurring primarily in the ψ-dimension. The magnitude of a ΔMCM vector represents the magnitude of angular difference between analogous macrocycles but does not represent a mechanistic pathway of conformational interconversion in cases where such interconversion is possible. The use of MCMs and ΔMCMs allows for graphical evaluation of macrocycle conformation and identification of regions of conformational change among related macrocycles. Using MCM methodology we have evaluated a range of macrocyclic peptide conformations, including those that do not possess a hydrogen bond defining a classical turn structure.

Results and discussion

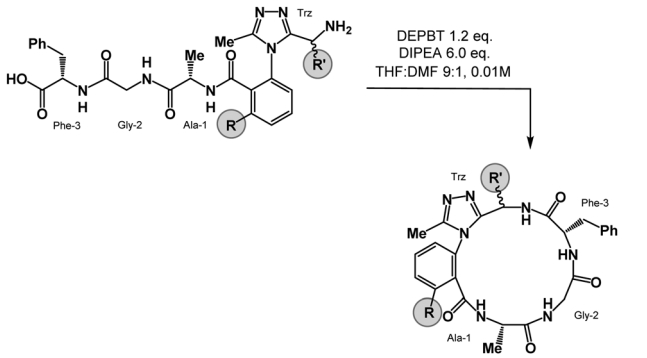
Synthesis of peptide macrocycles

We prepared a series of 16-membered dominant rotor-containing macrocyclic peptides. Each of these species bears the Ala-Gly-Phe (AGF) tripeptide motif constrained by a dominant rotor. Our scope sought to evaluate the divergent

© 2022 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 19 April 2022. Downloaded on 7/28/2024 10:21:48 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Table 1 Synthesis of dominant rotor-containing macrocycles. For full details, see the ESI



Macrocycle	R	R′	R_a yield (%)	S_a yield (%)
1	H	H	26	24
2	Me	H	20	28
3	H	(*R*)-Me	—	31
4	H	(*S*)-Me	11[a]	—

[a] ***R*$_a$-4**: Three minor species identified but not isolated after macrocyclisation.

conformational properties of this motif upon integrating a single methyl group in a range of positions about the dominant rotor. In medicinal chemistry, the ''magic methyl'' effect describes the introduction of a methyl group that impacts the conformation of a drug candidate, consequently affecting binding affinity and/or selectivity for its ligand.[38] Inspired by this logic, we introduced methyl groups in distinct positions and recorded the differences in global conformation induced by methylation. Macrocycles **1–4** were synthesized using solution phase DEPBT-mediated macrolactamisation of linear peptides obtained from solid-phase peptide synthesis (SPPS) (Table 1). The dominant rotor was introduced as a racemic mixture (**1**,**2**) or mixture of diastereomers (**3**,**4**). In principle, two macrocyclic atropdiastereomers can be formed for each system.

Understanding macrocyclic backbones using MCMs

Each of the macrocycles reported in Table 1 were fully characterized using ^{1}H and ^{13}C NMR spectroscopy. We performed variable temperature (VT) ^{1}H NMR to calculate chemical shift/temperature coefficients for each amide NH in the molecule. A temperature coefficient of less than 4 ppb K^{-1} typically indicates that a proton is involved in hydrogen bonding or is not exposed to solvent.[39] In addition, distance restraints were calculated using correlations in 2D ROESY spectra. These distances were used to generate structures that then underwent 100 ns molecular dynamics (MD) simulations in DMSO.[40,41] From these MD simulations, cluster analysis was performed to provide the ten most populated clusters generated from the simulation, which were checked for distance restraint violations. The dihedral angles of each structure were measured and used for MCM analysis. Full experimental procedures, computational details and experimental data can be found in the ESI†.

Macrocycle **1** was synthesized as a mixture of atropdiastereomers, ***R*$_a$-1** and ***S*$_a$-1**, which were individually isolated. To understand the differences in free energy between the two structures, the equilibrium constant for the interconversion between the two atropdiastereomers was calculated by performing a thermal equilibration experiment followed by subsequent analysis using VT-NMR. The entropic and enthalpic components were extrapolated using a Van't Hoff plot (see the ESI†).[32] This analysis established that ***S*$_a$-1** was lower in free energy than ***R*$_a$-1** by 1.15 kcal mol^{-1} (Fig. 2a). Classical turn structure evaluation concluded that the dark-space ***R*$_a$-1** contained two nested intramolecular hydrogen bonds defining a Type II α_{RU} turn. Similarly, the major conformer of ***S*$_a$-1** contained two intramolecular hydrogen bonds defining a Type II α_{RS} turn. In a study on the classification of 356 α-turns observed in natural proteins, only 28 were found to be Type II α_{RU} turns, representing approximately 8% of the structures in this dataset. Furthermore, 39 α-turns were found to be Type II α_{RS} turns, representing approximately 11% of these structures.[42]

MCM analysis was undertaken for both ***R*$_a$-1** and **S*$_a$*-1** to understand specific structural differences between the two macrocycles. The dihedral angles derived from the ten most populated clusters after 100 ns MD simulations were plotted (Fig. 1b). MCM analysis of ***R*$_a$-1** revealed that the ring adopts a single conformation with relatively tight ϕ/ψ distribution over the course of the MD simulation. Both of Ala-1 and Gly-2 occupy the positive ϕ-space in ***R*$_a$-1**. The positioning of Ala-1 in positive ϕ-space represents a non-ground-state conformation, as this sidechain incurs 1,3-allylic strain with the neighbouring amide. Evaluation of the Ramachandran plot of ***S*$_a$-1** revealed that the ring features more conformational mobility over the course of the MD simulation, adopting two related conformations. Furthermore, Gly-2 is the only residue to occupy positive ϕ-space in this structure. The positioning of all L-amino acids in low-energy negative ϕ-space in ***S*$_a$-1** compared to ***R*$_a$-1** reflects ΔG between these species.

***R*$_a$-1** and ***S*$_a$-1** are atropdiastereomers and can thus be conformationally compared against one another. ΔMCM plots overlay the lowest-violation cluster from each MD simulation on the same plot and establish a dihedral angular comparison of analogous macrocycles when ring size and amino acid substituents are held constant. In cases where multiple conformations were identified over the course of an MD simulation, dashed ovals joined the multiple species, and a vector was drawn only to the most populated species. ΔMCM vectors were drawn between the corresponding residues to represent the conformational changes undergone by each residue of ***R*$_a$-1** and ***S*$_a$-1** (Fig. 2b, bottom). The angular changes undergone by both Gly-2 and Phe-3 are relatively minor and are reflected in the vectors of lesser magnitude representing both residues. In comparison, the Ala-1 vector is significant in magnitude, traversing the $\phi = 0$ axis and representing the localisation of much of the conformational change between the atropisomers. This vector indicates that between ***R*$_a$-1** and ***S*$_a$-1**, the change in AGF conformation is largely attributable to the ϕ-axis of Ala-1. The dihedral angular change characterized by the Ala-1 vector naturally reflects the change in α-turn type from Type II α_{RU} in ***R*$_a$-1** to Type II α_{RS} in ***S*$_a$-1**. MCMs and ΔMCMs provide a view

Open Access Article. Published on 19 April 2022. Downloaded on 7/28/2024 10:21:48 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

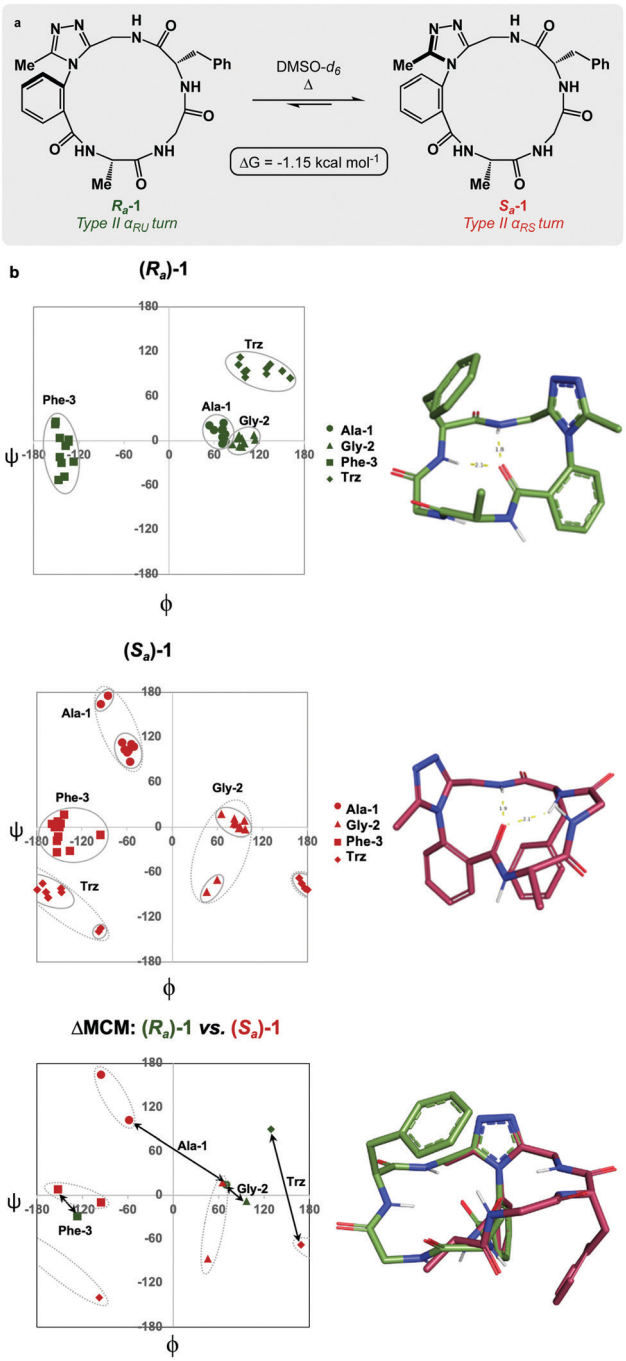


Fig. 2 MCM analysis reveal conformational differences between R_a-1 and S_a-1. (a) 16-membered dominant rotor cyclic peptides R_a- and S_a-1; (b) MCMs and 3D representations of R_a- and S_a-1; ΔMCM representing conformational differences between R_a-1 and S_a-1. MCMs display single-residue conformational ensembles bounded by solid ovals; major and minor clusters of S_a-1 are jointly enclosed by dashed ovals.

of the peptide's conformation that more fully encompasses both its angular structure, conformational plasticity, and relationship to analogues over the course of their respective MD simulations.

We sought to interrogate the effects of substitution to the dominant rotor through a "methyl scan", wherein the location and/or chirality of methyl substituents on the dominant rotor are varied. We hypothesised that this approach might allow us to observe the ensemble of conformations available to an oligopeptide constrained in a macrocycle, and to finely tune the conformation based on this information. Toward this goal, we designed macrocyclic peptides **2–4**, varying the location and/or chirality of methyl groups on the dominant rotor. To understand the downstream effects of methyl placement on the dominant rotor, we prepared macrocycle **2** bearing a methyl group on the *meta* position of the aryl ring. The atropdiastereomers R_a-**2** and S_a-**2** were identified and individually isolated. VT NMR experiments to calculate K_{eq} and ΔG for the equilibrium between the two species established that S_a-**2** was lower in free energy than R_a-**2** by 0.55 kcal mol^{-1} (Fig. 3a). Using classical turn evaluation, we concluded that the "dark-space" R_a-**2** adopts a Type IV β-turn. This type of β-turn corresponds to a backbone conformation that is not conventionally defined.[43] Interestingly, S_a-**2** is not possible to classify using conventional terminology because it does not contain an intramolecular hydrogen bond and is thus designated as a non-regular secondary structure (NORS). As the individual AGF conformation of both R_a-**2** and S_a-**2** cannot be defined by classical terminology, these "dark-space" conformations require description through their relationship of sequential dihedral angles.

Curiously, the MCM of S_a-**2** illustrates the dynamic coupling of rotors through the distribution of adjacent ϕ/ψ coordinates of Ala-1 and Gly-2 (Fig. 3b). Over the course of the MD simulation, this region of the ring was found to be conformationally mobile, with the distribution of ϕ/ψ coordinates extending along the ψ-axis for Ala-1 and along the ϕ-axis for Gly-2. This finding represents the high conformational mobility of the Ala-1 amide plane over the course of the MD simulation. It also demonstrates the fine detail in conformational dynamics that can be elucidated using MCMs. The use of MCMs not only permits classification of the unconventional conformation of S_a-**2**, but also illustrates its regions of high conformational mobility. ΔMCM screening reveals that the angular differences in AGF conformation between R_a-**2** and S_a-**2** are primarily localized at Ala-1 and Gly-2. Ala-1 occupies distinct coordinates in the negative ϕ-space in both structures and Gly-2 undergoes a large angular change, crossing the ϕ axis. Notably, both conformations are distinct from those of R_a-**1** and S_a-**1** despite their structural similarity.

As the dominant rotor's *meta*-methyl constitutes the only structural difference between **1** and **2**, the conformational differences observed can be attributed to its downstream effects. In particular, the methyl group perturbs the ring's rotor structure by clashing with the proximal benzamide motif. This perturbation changes the benzamide rotor's orientation, resulting in downstream conformational effects due to coupling in bond rotations. In general, the most stable dihedral angle for a benzamide motif is $\theta = \pm 30/150°$.[44] Macrocycle **1** adopts dihedral angles that align with the most stable ones: R_a-**1** $\theta = -141°$; S_a-**1** $\theta = 143°$. As a result of structural perturbation by the rotor methyl, R_a-**2** $\theta = -117°$; S_a-**2** $\theta = -76°$. The global conformational differences in **2** in relation to **1** stem from the downstream effects of this benzamide bond rotation.

© 2022 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 19 April 2022. Downloaded on 7/28/2024 10:21:48 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

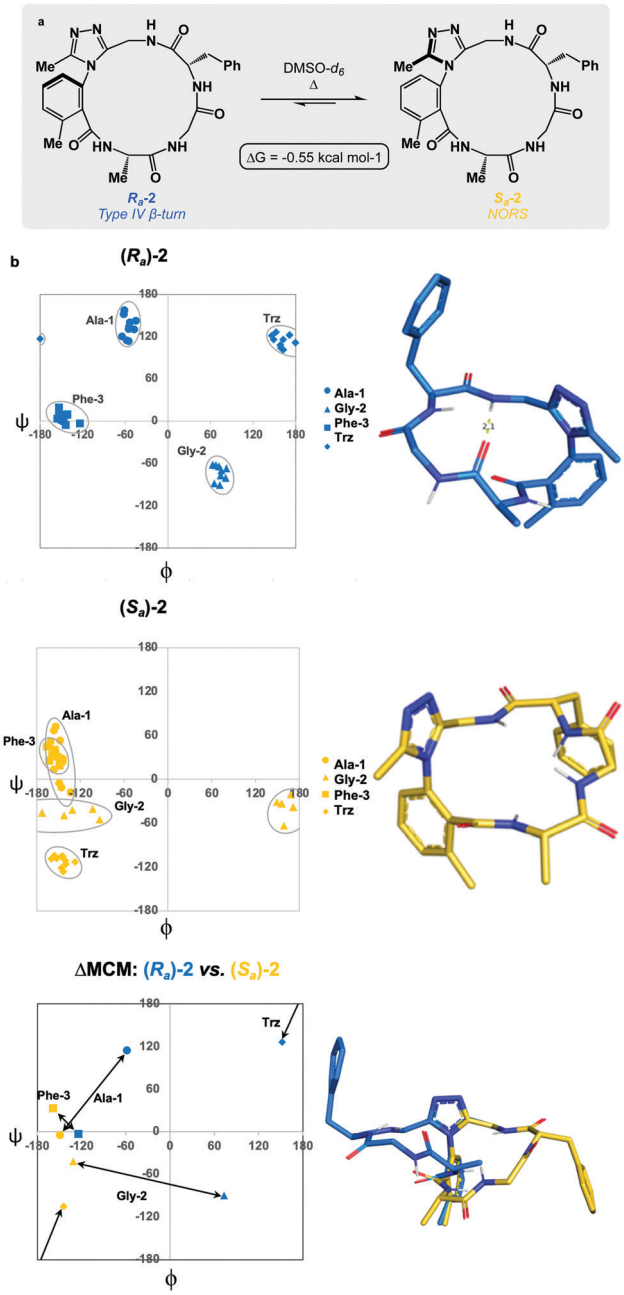


Fig. 3 MCM analysis reveals conformational flexibility of Gly-2 in S_a-2. (a) R_a-2 and S_a-2; (b) MCMs and 3D representations of R_a- and S_a-2; ΔMCM representing conformational differences between R_a-2 and S_a-2. MCMs display single-residue conformational ensembles bounded by solid ovals.

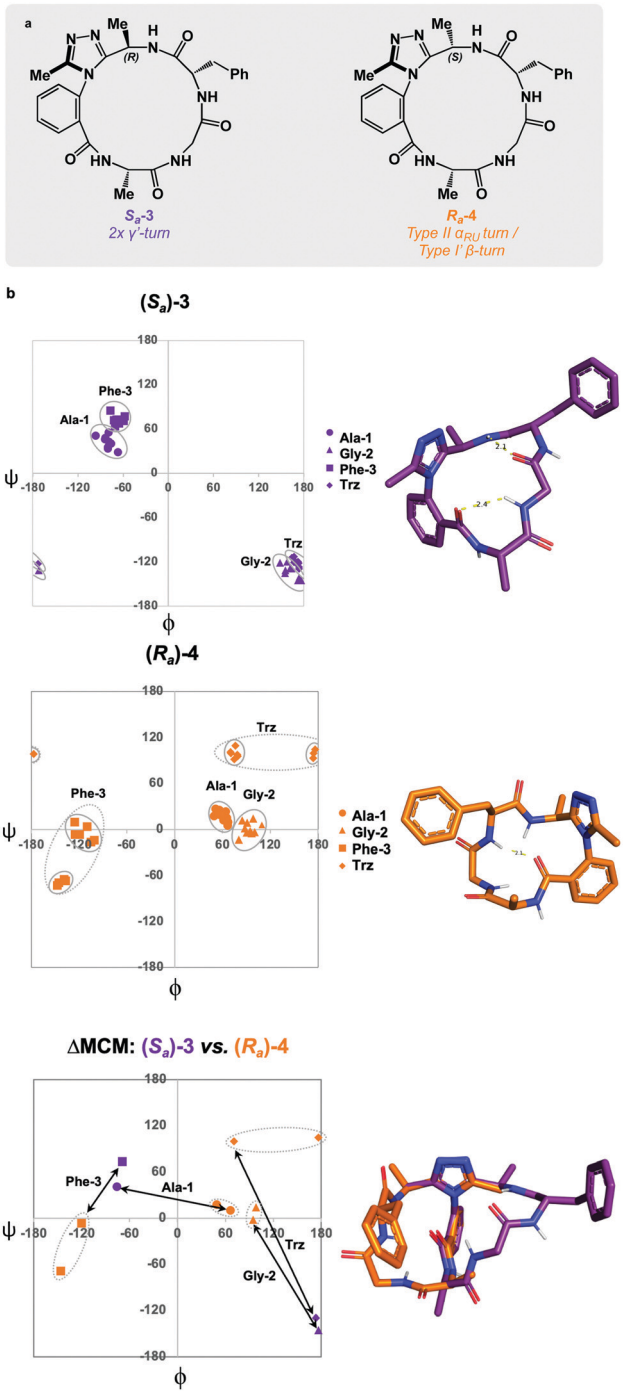


Fig. 4 Comparison of S_a-3 and R_a-4 using MCM methodology. (a) S_a-3 and R_a-4; (b) MCMs and 3D representation of S_a-3 and R_a-4; ΔMCM representing conformational differences between S_a-3 and R_a-4. MCMs display single-residue conformational ensembles bounded by solid ovals. Major and minor clusters of R_a-4 are jointly enclosed by dashed ovals.

To complete the methyl scan of the dominant rotor, both (*R*)- and (*S*)-methyl groups were individually introduced in the vicinity of the triazole fragment (Fig. 4a). We hypothesized that a chiral sidechain in this position may sterically clash with the dominant rotor in one well, but not the other. In this way, the R_a:S_a product distribution could be controlled through the introduction of a chiral centre on the dominant rotor, wherein the chirality may influence the distribution of atropisomers. We first tested this theory by incorporating a (*R*)-methyl group on the dominant rotor to form a 16-membered AGF macrocycle. Excitingly, a single species was observed upon macrocyclization, and S_a-3 was isolated as a single species (Fig. 4a). Conformational evaluation of S_a-3 revealed that it formed a largely rigid conformation bearing two intramolecular hydrogen bonds, classically defining

© 2022 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 19 April 2022. Downloaded on 7/28/2024 10:21:48 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

γ′-turns about each of Ala-1 and Phe-3 (Fig. 4b). In a similar experiment, the point chirality at this position was changed to (*S*), yielding the 16-membered AGF macrocycle **4**. One major species was isolated and assigned $\boldsymbol{R_a}$**-4** (Fig. 4a). Alongside the major species, three minor species were identified by LC-MS analysis but were not analyzed in detail. Conformational evaluation of $\boldsymbol{R_a}$**-4** revealed a set of two related conformations, one bearing a Type II α_{RU} turn and the other bearing a Type I′ β-turn. This effect was intriguing, as the introduction of the (*S*)-methyl group in $\boldsymbol{R_a}$**-4** resulted in a conformational deviation from the methylene-bearing $\boldsymbol{R_a}$**-1**.

ΔMCM screening of $\boldsymbol{S_a}$**-3** and $\boldsymbol{R_a}$**-4** reveals conformational differences at each residue. Ala-1 traverses the ϕ = 0 axis between the two species. Gly-2 occupies positive ϕ-space in both species, yet occupies distinct regions in each. Furthermore, Phe-3 occupies two unique regions of the negative ϕ-space in each conformer. This plot illustrates the potential value in the introduction and screening of chiral substituents on the dominant rotor. We found that the point chirality at this position can bias the atropisomer formed, resulting in divergent conformations of the same peptide sequence. This further displays the conformational tuning and downstream effects that can be achieved by modifying the dominant rotor itself.

We undertook ΔMCM analysis for each of the macrocycles in the methyl scan to highlight conformational differences between these analogues. Since $\boldsymbol{R_a}$**-1** and $\boldsymbol{R_a}$**-2** share the same ring size and similar hub/rotor structure, these species can be appropriately evaluated using this method. An overlay of ϕ/ψ coordinates for both rings reveal a range of conformational differences that arise downstream from the methyl group's perturbation of the benzamide rotor (Fig. 5a). The most dramatic difference between these two species is that Ala-1 occupies a positive ϕ-angle in $\boldsymbol{R_a}$**-1** but a negative ϕ-angle in $\boldsymbol{R_a}$**-2**. The methyl group perturbs the benzamide dihedral angle to turn into the ring, allowing the peptide chain to relax at Ala-1.

A similar assessment can be made to evaluate the conformational differences between $\boldsymbol{S_a}$**-1** and $\boldsymbol{S_a}$**-2** (Fig. 5b). In this example, the dihedral angular differences are again localized primarily at Ala-1 and Gly-2. This conformational effect is similar to that in the R_a-well, as the dominant rotor's aryl methyl group perturbs the benzamide rotor. The bond rotation induces downstream conformational effects that propagate throughout the system but are primarily localized to those residues most proximal to the dominant rotor.

ΔMCM screening of $\boldsymbol{R_a}$**-1** and $\boldsymbol{R_a}$**-4** reveals that the conformations of these systems are rather similar (Fig. 5c). This is not surprising, as these systems bear related rotor structure. A curious component of the behaviour of $\boldsymbol{R_a}$**-4** is its conformational dynamics. $\boldsymbol{R_a}$**-4** can access both a Type II α_{RU} turn and a Type I′ β-turn. These conformations are similar, differing primarily in the identity of the hydrogen bonding NH: hydrogen bonding of Trz-NH leads to a Type II α_{RU} turn and hydrogen bonding of Phe-3-NH leads to a Type I′ β-turn. These conformations are under exchange in the MD simulation, reflecting the component of conformational instability introduced by the (*S*)-methyl of the dominant rotor of $\boldsymbol{R_a}$**-4** in comparison to the methylene of $\boldsymbol{R_a}$**-1**. This effect illustrates that relative conformational dynamics can also be evaluated using the ΔMCM plot.

ΔMCM screening allows the comparison of $\boldsymbol{S_a}$**-1** with $\boldsymbol{S_a}$**-3** (Fig. 5d). These structures differ only in the placement (*R*)-methyl in $\boldsymbol{S_a}$**-3**. This minor structural change gives rise to significant conformational differences. Angular changes are observed in each residue, wherein both Ala-1 and Phe-3 adopt ϕ/ψ regions characteristic of γ′-turns with Gly-2 occupying an extended conformation in $\boldsymbol{S_a}$**-3**. This is compared to the Type II α_{RS} turn adopted by the major conformer of $\boldsymbol{S_a}$**-1**. The comparison of $\boldsymbol{S_a}$**-2** with $\boldsymbol{S_a}$**-3** constitutes a particularly interesting example, as these species have the same dominant rotor geometry but differ in the position of the methyl group (Fig. 5f). While each species adopts a conformation characterized by all-negative ϕ-angles, each conformation presents as stable over the course of the MD simulation and each residue occupies distinct regions that characterize their respective turn structure.

Analysis of non-dominant rotor macrocycles

To further demonstrate the utility of MCM plots, we performed analysis on other macrocyclic structures possessing both a disparate ring size and distinct hub and rotor profile. Specifically, we wanted to see if MCM methodology could be used to understand conformational nuances in biologically relevant molecules, with the downstream possibility to aid in the design of related bioactive molecules. The interaction between the lymphocyte integrin $\alpha_4\beta_7$ and the mucosal address in cell adhesion molecule-1 (MAdCAM-1) have been implicated in several diseases such as Type 1 diabetes, Crohn's disease and chronic inflammatory bowel disease.[45–47] The binding of the integrin to MAdCAM-1 relies on the Leu-Asp-Thr (LDT) tripeptide sequence located on the *N*-terminal Ig domain of MAdCAM-1, which has stimulated the development of inhibitors based on this motif.[48,49] Previously reported macrocycles **5** and **6** each contain the LDT sequence in a unique backbone position, and incorporate a 1,3,4-oxadiazole (Odz).[40] The solution structures of **5** and **6** reported in DMSO (derived from 10 ns MD simulations) each include a Type II β-turn (Fig. 6). Furthermore, the reduced amide and oxadiazole motif stabilizes a β-turn-like conformation with the proximal amide NH. Despite their similar backbone conformations, positional isomers **5** and **6** display significantly different inhibitory activity towards the MAdCAM-1/$\alpha_4\beta_7$ protein-protein interaction (measured by ELISA assays: **5**, IC_{50} = >45 mM; **6**, 1.6 μM).[40]

To visualise the specific differences between the two molecules, individual MCM plots of **5** and **6**, as well as a ΔMCM plot comparing the two structures were prepared from previously reported dihedral angles extracted from the most populated cluster of 10 ns MD simulations (Fig. 6).[40] Here, we show two sets of vectors in the ΔMCM plot: (i) black solid arrows represent the changes in dihedral angle between specific positions around the macrocycle, *e.g.* Gly-2/Leu-2, and (ii) red dashed arrows represent dihedral angular changes between the specific Leu/Asp/Thr residues as a result of movement within the macrocyclic backbone, *e.g.* Thr-5/Thr-4.

© 2022 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 19 April 2022. Downloaded on 7/28/2024 10:21:48 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

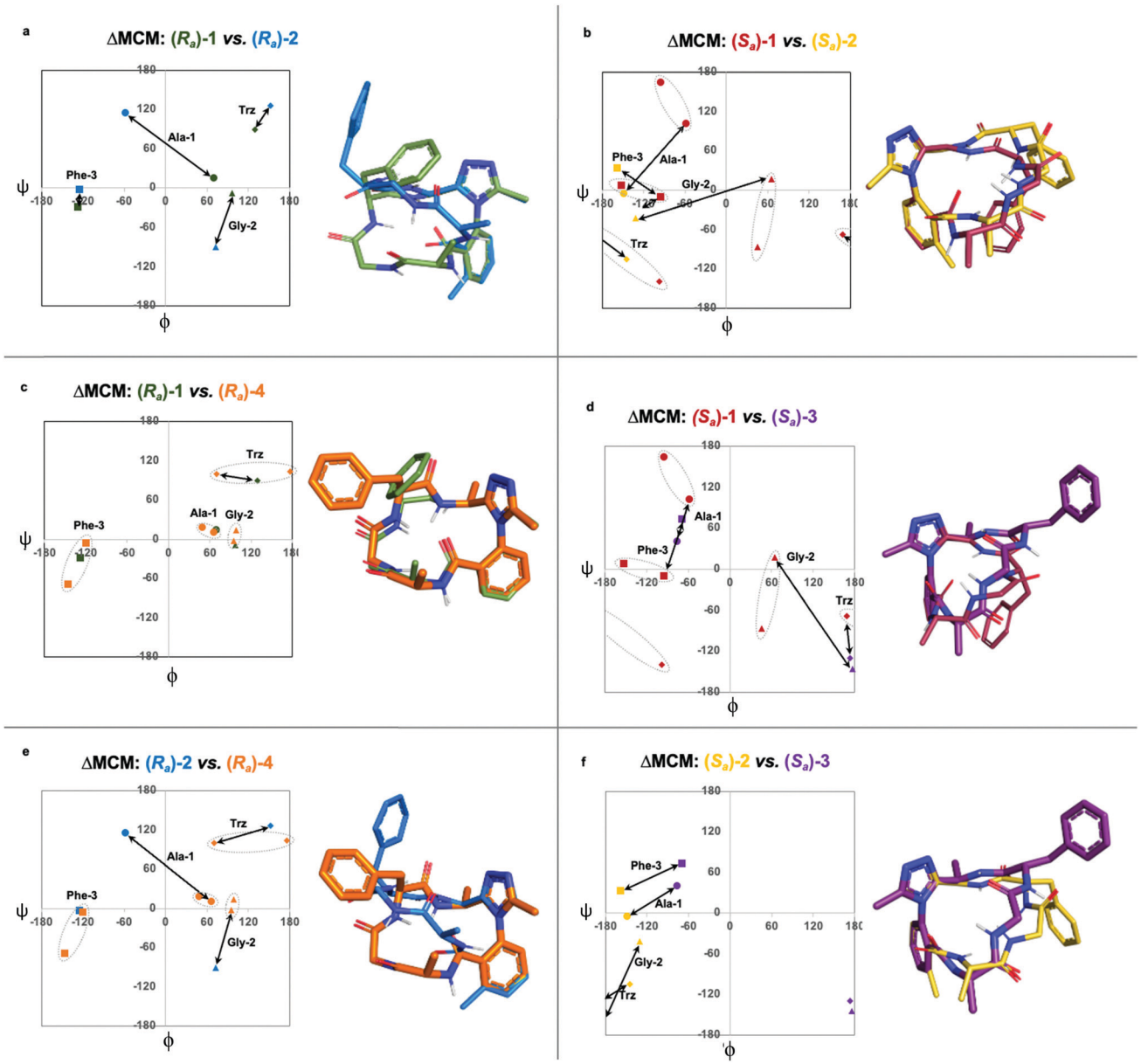


Fig. 5 Further comparison of pairs of dominant rotor containing macrocycles using ΔMCM analysis. ΔMCMs: (a) R_a-**1** *vs.* R_a-**2**; (b) S_a-**1** *vs.* S_a-**2**; (c) R_a-**1** *vs.* R_a-**4**; (d) S_a-**1** *vs.* S_a-**3**; (e) R_a-**2** *vs.* R_a-**4**; (f) S_a-**2** *vs.* S_a-**3**. Dashed ovals represent major and minor conformations of S_a-**1**. Points on the ΔMCMs represent the most populated cluster of each species. Arrows represent dihedral angular changes.

The ΔMCM plot comparing **5** and **6** presents solid black vectors of relatively small magnitude for each residue in the backbone, which highlights the conformational similarities between the two macrocycles. The red dotted vectors displaying movement of the LDT residues within the backbone show that the ϕ/ψ space occupied by the Leu residue does not change dramatically based on its relative position in the turn sequence. In contrast, the Asp and Thr residues occupy distinct regions of conformational space in these two molecules. While it is challenging to identify generalisable conformational characteristics that distinguish high-affinity inhibitors from a single example, we believe that within the context of a broader dataset this type of analysis may prove fruitful in highlighting conformational regions of interest for the design of macrocyclic peptide-based scaffolds. This analysis demonstrates the use of our methodology to visualise conformational relationships in macrocycles with bioactivity, which we anticipate will be of immediate utility to chemists developing other macrocyclic peptide inhibitors.

The findings from these ΔMCMs illustrate the breadth of conformations available to short oligopeptide sequences constrained in macrocyclic scaffolds. A current challenge in macrocycle development entails the use and discovery of specific modifications which enable rational and predictable access to specific areas of conformational space. Although this goal has not yet been achieved, we anticipate that future use of the MCM

© 2022 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 19 April 2022. Downloaded on 7/28/2024 10:21:48 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

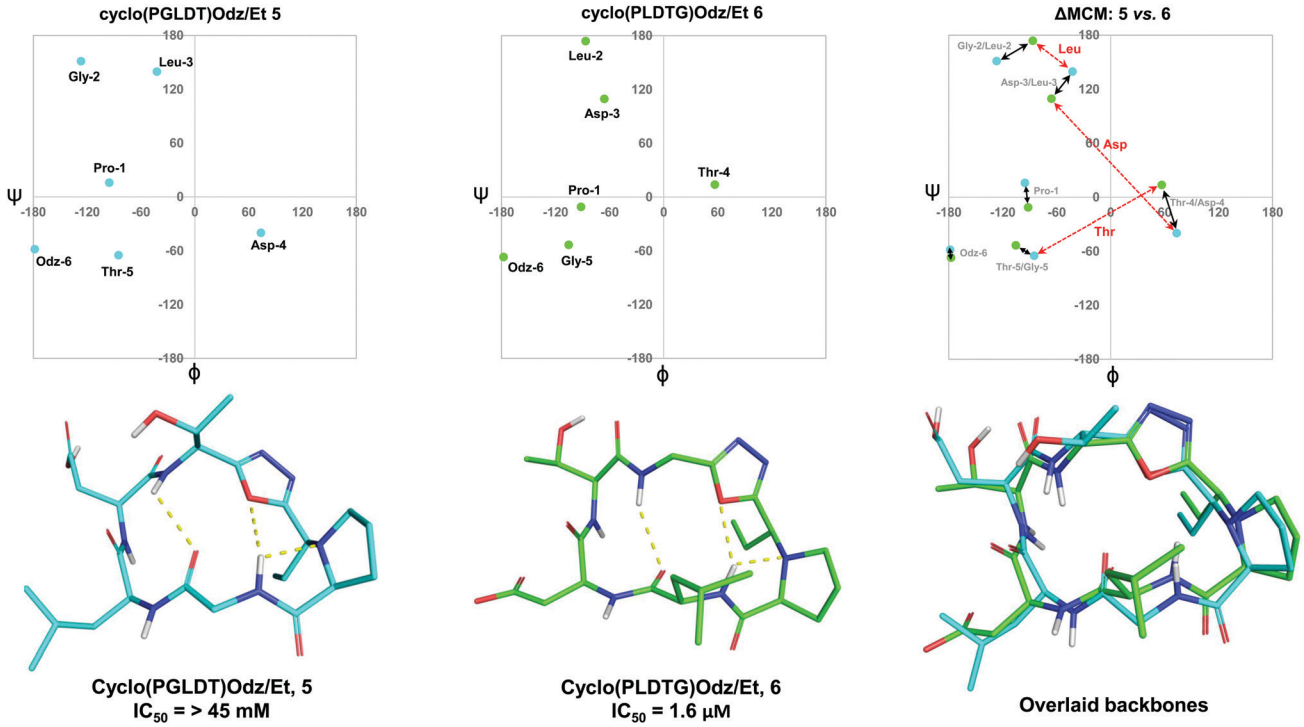


Fig. 6 MCM analysis of the solution structures of cyclo(PGLDT)Odz/Et 5 and cyclo(PLDTG)Odz/Et 6 reveal conformational differences that pertain to biological activity. Points on the maps represent the ϕ/ψ coordinates of the most populated cluster of each species. Black arrows represent dihedral angular changes between residues in same relative position around the macrocycle *e.g.*, Gly-2/Leu-2. Red dashed arrows represent dihedral angular changes between the same residue in different positions around the macrocycle, *e.g.*, Thr-5/Thr-4.

tool will enable better understanding of the conformational nuances caused by synthetic modifications to macrocyclic peptide backbones.

Conclusions

MCM and ΔMCM screening can be useful tools to evaluate conformational ensembles identified for macrocyclic peptides and their structural analogues. During our study, we found that a ''methyl scan'' on dominant rotor fragments in macrocycles can result in significant differences in conformation that are adequately expressed using our method. Additionally, we established that by controlling the chirality of dominant rotor methylation, we can bias the system to favour a specific atropisomer. Furthermore, the dark-space conformations identified in this study were evaluated with MCM analysis, including those with non-regular secondary structures. In this case, MCMs allow for analysis of macrocyclic peptide structures that could not be defined using conventional turn-structure terminology.

ΔMCM screening can be a useful tool for the comparison of macrocyclic peptide conformations and should find utility in conformation-activity studies. This manner of analysing macrocyclic peptides could find applications in tuning the conformation of a series of structurally related systems. Such an approach could result in the design of more potent and selective macrocycle-based medicines as the functional properties of such compounds are intrinsically related to their conformation.

Author contributions

A. K. Y. and T. J. M. conceived the study and designed the experiments; T. J. M., D. B. D., G. J. S., and F. S. carried out macrocycle synthesis, purification and characterisation, T. J. M. and D. B. D. computed the NMR-based structures and performed unrestrained MD simulations, T. J. M. and G. J. S. wrote the paper with help from all the authors.

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgements

We thank NSERC and Genentech for financial support of this work. A. K. Y. has been a Killam Fellow over the course of this project. T. J. M. received funding from the Ontario Graduate Scholarship (OGS) program in support of this work. We thank S. D. Appavoo for providing structural data on LDT-containing macrocycles. We also thank D. C. Burns, J. Sheng, and K. Z. Demmans of the CSICOMP NMR facility for assistance with spectroscopic experiments.

© 2022 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 19 April 2022. Downloaded on 7/28/2024 10:21:48 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Notes and references

1 M. Gurrath, G. Muller, H. Kessler, M. Aumailley and R. Timpl, *Eur. J. Biochem.*, 1992, **210**, 911–921.
2 E. Marsault and M. L. Peterson, *J. Med. Chem.*, 2011, **54**, 1961–2004.
3 A. R. Battersby, C. J. R. Fookes, G. W. J. Matcham and E. McDonald, *Nature*, 1980, **285**, 17–21.
4 L. K. Buckton, M. N. Rahimi and S. R. McAlpine, *Chem. – Eur. J.*, 2021, **27**, 1487–1513.
5 A. S. Kamenik, U. Lessel, J. E. Fuchs, T. Fox and K. R. Liedl, *J. Chem. Inf. Model.*, 2018, **58**, 982–992.
6 K. S. Watts, P. Dalal, A. J. Tebben, D. L. Cheney and J. C. Shelley, *J. Chem. Inf. Model.*, 2014, **54**, 2680–2696.
7 S. D. Appavoo, S. Huh, D. B. Diaz and A. K. Yudin, *Chem. Rev.*, 2019, **119**, 9724–9752.
8 R. Jwad, D. Weissberger and L. Hunter, *Chem. Rev.*, 2020, **120**, 9743–9789.
9 J. Schwochert, Y. Lao, C. R. Pye, M. R. Naylor, P. V. Desai, I. C. Gonzalez Valcarcel, J. A. Barrett, G. Sawada, M.-J. Blanco and R. S. Lokey, *ACS Med. Chem. Lett.*, 2016, **7**, 757–761.
10 C. G. Pappas, P. K. Mandal, B. Liu, B. Kauffmann, X. Miao, D. Komáromy, W. Hoffmann, C. Manz, R. Chang, K. Liu, K. Pagel, I. Huc and S. Otto, *Nat. Chem.*, 2020, **12**, 1180–1186.
11 R. L. Baldwin and G. D. Rose, *Trends Biochem. Sci.*, 1999, **24**, 77–83.
12 H. Kessler, R. Gratias, G. Hessler, M. Gurrath and G. Müller, *Pure Appl. Chem.*, 1996, **68**, 1201–1205.
13 S. Kobori, S. Huh, S. D. Appavoo and A. K. Yudin, *J. Am. Chem. Soc.*, 2021, **143**, 5166–5171.
14 Y. A. Ovchinnikov and V. T. Ivanov, *Tetrahedron*, 1975, **31**, 2177–2209.
15 P. Khalili, C. B. Barnett and K. J. Naidoo, *J. Chem. Phys.*, 2013, **138**, 184110.
16 L. A. LaPlanche and M. T. Rogers, *J. Am. Chem. Soc.*, 1964, **86**, 337–341.
17 M. W. MacArthur and J. M. Thornton, *J. Mol. Biol.*, 1996, **264**, 1180–1195.
18 J. S. Richardson, *Adv. Protein Chem.*, 1981, 167–339.
19 R. W. Hoffmann, *Angew. Chem., Int. Ed. Engl.*, 1992, **31**, 1124–1134.
20 T. Haliloglu, I. Bahar, B. Erman and W. L. Mattice, *Macromolecules*, 1996, **29**, 8942–8947.
21 A. Sharma, S. Sharma, R. P. Tripathi and R. S. Ampapathi, *J. Org. Chem.*, 2012, **77**, 2001–2007.
22 T. J. Sereda, C. T. Mant, F. D. Sönnichsen and R. S. Hodges, *J. Chromatogr. A*, 1994, **676**, 139–153.
23 G. N. Ramachandran and V. Sasisekharan, *Adv. Protein Chem.*, 1968, 283–437.
24 R. W. W. Hooft, C. Sander and G. Vriend, *Bioinformatics*, 1997, **13**, 425–430.
25 D. B. Dahl, Z. Bohannan, Q. Mo, M. Vannucci and J. Tsai, *J. Mol. Biol.*, 2008, **378**, 749–758.
26 D. A. C. Beck, D. O. v. Alonso, D. Inoyama and V. Daggett, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 12259–12264.
27 W.-C. Lo, P.-J. Huang, C.-H. Chang and P.-C. Lyu, *BMC Bioinf.*, 2007, **8**, 307.
28 R. V. Mannige, J. Kundu and S. Whitelam, *PLoS One*, 2016, **11**(8), e0160023.
29 R. V. Mannige, *PeerJ*, 2017, **5**, e3327.
30 D. DasGupta, R. Kaushik and B. Jayaram, *J. Phys. Chem. B*, 2015, **119**, 11136–11145.
31 C. M. Wilmot and J. M. Thornton, *J. Mol. Biol.*, 1988, **203**, 221–232.
32 D. B. Diaz, S. D. Appavoo, A. F. Bogdanchikova, Y. Lebedev, T. J. McTiernan, G. dos Passos Gomes and A. K. Yudin, *Nat. Chem.*, 2021, **13**, 218–225.
33 P. Hosseinzadeh, G. Bhardwaj, V. K. Mulligan, M. D. Shortridge, T. W. Craven, F. Pardo-Avila, S. A. Rettie, D. E. Kim, D.-A. Silva, Y. M. Ibrahim, I. K. Webb, J. R. Cort, J. N. Adkins, G. Varani and D. Baker, *Science*, 2017, **358**, 1461–1466.
34 Z. Zhang, D. C. Burns, J. R. Kumita, O. S. Smart and G. A. Woolley, *Bioconjugate Chem.*, 2003, **14**, 824–829.
35 R. L. Jones, *Nature*, 1978, **272**, 185–187.
36 K. M. Saravanan, Y. Peng and Y. Wei, *J. Biomol. Struct. Dyn.*, 2020, **38**, 268–274.
37 P. M. Cromm, K. Wallraven, A. Glas, D. Bier, A. Fürstner, C. Ottmann and T. N. Grossmann, *ChemBioChem*, 2016, **17**, 1915–1919.
38 H. Schönherr and T. Cernak, *Angew. Chem., Int. Ed.*, 2013, **52**, 12256–12267.
39 H. Kessler, *Angew. Chem., Int. Ed. Engl.*, 1982, **21**, 512–523.
40 S. D. Appavoo, T. Kaji, J. R. Frost, C. C. G. Scully and A. K. Yudin, *J. Am. Chem. Soc.*, 2018, **140**, 8763–8770.
41 J. R. Frost, C. C. G. Scully and A. K. Yudin, *Nat. Chem.*, 2016, **8**, 1105–1111.
42 V. Pavone, G. Gaeta, A. Lombardi, F. Nastri, O. Maglio, C. Isernia and M. Saviano, *Biopolymers*, 1996, **38**, 705–721.
43 A. G. de Brevern, *Sci. Rep.*, 2016, **6**, 33191.
44 K. A. Brameld, B. Kuhn, D. C. Reuter and M. Stahl, *J. Chem. Inf. Model.*, 2008, **48**, 1–24.
45 M. J. Briskin, L. Rott and E. C. Butcher, *J. Immunol.*, 1996, **156**, 719.
46 J. Rivera-Nieves, T. Olson, G. Bamias, A. Bruce, M. Solga, R. F. Knight, S. Hoang, F. Cominelli and K. Ley, *J. Immunol.*, 2005, **174**, 2343–2352.
47 B. Xu, R. E. Cook and S. A. Michie, *J. Autoimmun.*, 2010, **35**, 124–129.
48 J. Boer, D. Gottschling, A. Schuster, M. Semmrich, B. Holzmann and H. Kessler, *J. Med. Chem.*, 2001, **44**, 2586–2592.
49 N. Green, J. Rosebrook, N. Cochran, K. Tan, J.-H. Wang, T. A. Springer and M. J. Briskin, *Cell Adhes. Commun.*, 1999, **7**, 167–181.

© 2022 The Author(s). Published by the Royal Society of Chemistry