

# Kinetic Models of Cyclosporin A in Polar and Apolar Environments Reveal Multiple Congruent Conformational States

Jagna Witek,<sup>†</sup> Bettina G. Keller,<sup>‡</sup> Markus Blatter,<sup>§</sup> Axel Meissner,<sup>§</sup> Trixie Wagner,<sup>§</sup> and Sereina Riniker\*,<sup>†</sup>

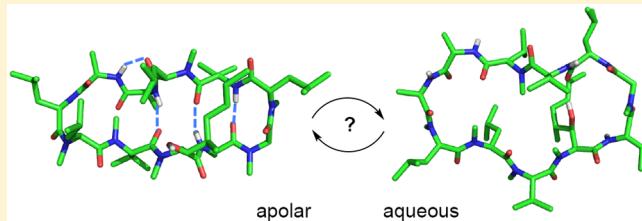
<sup>†</sup>Laboratory of Physical Chemistry, ETH Zürich, Vladimir-Prelog-Weg 2, 8093 Zürich, Switzerland

<sup>‡</sup>Department of Biology, Chemistry, Pharmacy, Freie Universität Berlin, Takustrasse 3, 14195 Berlin, Germany

<sup>§</sup>Novartis Institutes for BioMedical Research, Novartis Pharma AG, Novartis Campus, 4002 Basel, Switzerland

## Supporting Information

**ABSTRACT:** The membrane permeability of cyclic peptides and peptidomimetics, which are generally larger and more complex than typical drug molecules, is likely strongly influenced by the conformational behavior of these compounds in polar and apolar environments. The size and complexity of peptides often limit their bioavailability, but there are known examples of peptide natural products such as cyclosporin A (CsA) that can cross cell membranes by passive diffusion. CsA is an undecapeptide with seven methylated backbone amides. Its crystal structure shows a “closed” twisted  $\beta$ -pleated sheet conformation with four intramolecular hydrogen bonds that is also observed in NMR measurements of CsA in chloroform. When binding to its target cyclophilin, on the other hand, CsA adopts an “open” conformation without intramolecular hydrogen bonds. In this study, we attempted to sample the complete conformational space of CsA in chloroform and in water by molecular dynamics simulations in order to better understand its conformational behavior in these two environments and to rationalize the good membrane permeability of CsA observed experimentally. From 10  $\mu$ s molecular dynamics simulations in each solvent, Markov state models were constructed to characterize the metastable conformational states. The model in chloroform is compared to nuclear Overhauser effect NMR spectroscopy data reported in this study and taken from the literature. The conformational landscapes in the two solvents show significant overlap but also clearly distinct features.



## INTRODUCTION

Cyclic peptides and peptidomimetics have recently gained renewed interest as potential inhibitors of important therapeutic targets such as class-B G-protein coupled receptors and protein–protein interfaces that are difficult to target using small organic molecules.<sup>1–7</sup> However, bioavailability is often a problem for peptide and protein drugs.<sup>8</sup> Structural features such as cyclization and selective methylation of backbone amine residues can facilitate the passive membrane permeability of such compounds.<sup>9–12</sup> An early example of a cyclic peptide therapeutic is cyclosporin A (CsA), used as a potent immunosuppressive drug to prevent graft rejection.<sup>13–15</sup> CsA indirectly suppresses the inflammatory response of the body by binding to cyclophilin, which inhibits calcineurin and leads to suppression of the response from T-cells and mast cells.<sup>16,17</sup> CsA has an unusual property: it is bioavailable despite its relatively high molecular weight and lipophilicity, which violate most of the conventional physicochemical guidelines for permeability and bioavailability (e.g., Lipinski’s rule of five<sup>18</sup>). It has been hypothesized that the reason for the unusually high permeability of CsA is a conformational change from an “open” conformation in water (where the backbone amides form H-bonds with the solvent) to a “closed” conformation in the membrane interior (where intramolecular H-bonds are formed).<sup>10</sup> Thus, the polar groups of the peptide become

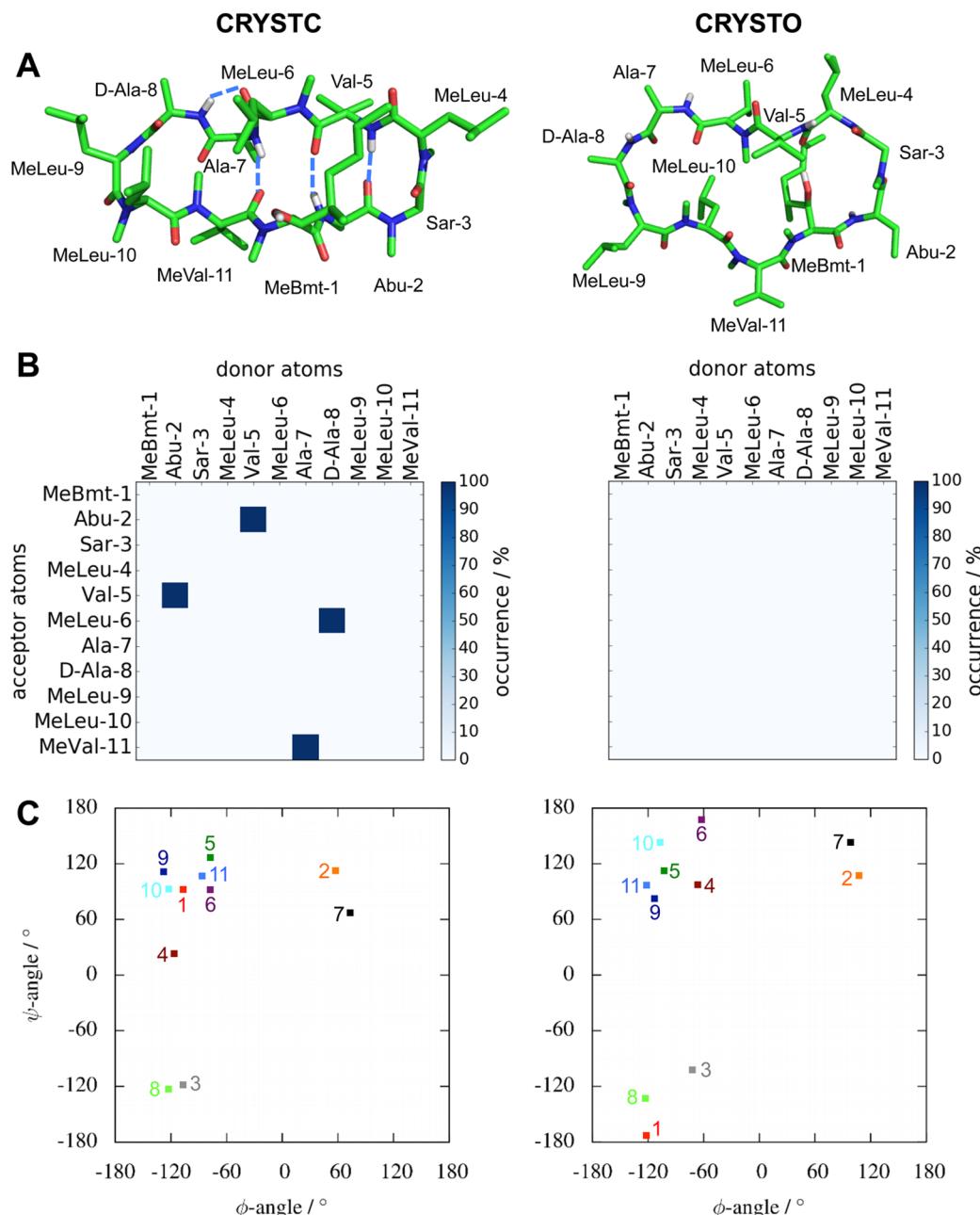
shielded from the apolar environment. The hypothesis suggests that in practice the membrane permeability of small peptides and peptidomimetics may crucially depend on their ability to adopt such apolar conformations.

NMR spectroscopy can provide valuable information about the structural dynamics of peptides and proteins. One of the physical phenomena observed by NMR spectroscopy is the nuclear Overhauser effect (NOE). Its interpretation provides distances (or upper distance bounds) between pairs of protons, giving insight into the spatial arrangement of residues within a molecule. Since the NOE data are obtained in solution, the upper bounds represent distances averaged over an ensemble of conformations interconverting on the NMR time scale (i.e., interconversion rates <100  $\mu$ s). The structural dynamics of CsA has been extensively investigated by NMR spectroscopy and X-ray diffraction as well as molecular dynamics (MD) simulations.<sup>19–22</sup> The crystal structure of CsA shows a conformation characterized by a twisted  $\beta$ -pleated sheet (type II'  $\beta$ -turn) with a cis peptide bond between MeLeu-9 and MeLeu-10, as shown schematically in the left panel of Figure 1A.<sup>23</sup> A similar backbone conformation has been derived from the interpretation of NMR data in apolar solvents such as

Received: May 2, 2016

Published: July 7, 2016





**Figure 1.** Crystal structures of CsA: (left) DEKSAN<sup>23</sup> (source: Cambridge Structure Database (CSD)), denoted as CRYSTC = closed crystal conformation; (right) 2Z6W<sup>44</sup> (source: Protein Data Bank (PDB), cocrystallized with cyclophilin), denoted as CRYSTO = open crystal conformation. (A) Schematic representations of the crystal structures. H-bonds are marked with dashed light-blue lines. Figures were generated with PyMOL.<sup>45</sup> (B) Backbone–backbone H-bonds. (C) Backbone dihedral angles  $\phi$  and  $\psi$  of the 11 residues in the crystal structures.

chloroform and benzene.<sup>23</sup> NOE measurements indicate that CsA adopts predominately this closed conformation in chloroform.<sup>23–25</sup> Interestingly, similar structural features have also been observed in permeable hexa-Ala peptides with one D-amino acid.<sup>26</sup> In addition to the major closed conformation, the presence of a minor conformation (approximately 6%) was observed in NOE experiments of CsA in chloroform.<sup>20</sup> This finding hints at a more complex conformational landscape of CsA in an apolar environment. Only the structure of the major conformation has been resolved.<sup>20</sup> The conformation of CsA in polar solvents has not been studied as thoroughly because of the poor solubility of CsA in these solvents. However, it is suggested that in a polar environment CsA adopts several distinct conformations that interconvert slowly on the NMR time scale.<sup>19,23,25,27,28</sup>

The preferred conformation of CsA in water is thought to resemble the cyclophilin-bound conformation observed in the NMR solution structure<sup>29–31</sup> or the crystal structure of the complex.<sup>27</sup> This is an “open” conformation with all trans peptide bonds, where the backbone amide groups can potentially form H-bonds with the solvent. Kofron et al.<sup>32</sup> have shown that a conformation with all trans peptide bonds is necessary for efficient binding in the active site of cyclophilin.

The aim of this study was to investigate the conformational landscapes of CsA in apolar and aqueous environments to gain insights into the conformational changes associated with membrane permeation. As NMR measurements can provide only time- and molecule-averaged data and previous MD

simulations<sup>20</sup> have been too short to observe conformational changes, we attempted to extensively sample the conformational space by simulating CsA in water and chloroform. Chloroform with its low dielectric constant was used to mimic the membrane interior. For both solvents, we conducted a total of 10  $\mu$ s of MD simulations that were seeded from 1.1  $\mu$ s of standard and enhanced-sampling MD simulations. From the MD data, Markov state models (MSMs)<sup>33–36</sup> were constructed to identify the essential features of the conformational dynamics of CsA in the two different environments. In recent years, MSMs have become a powerful tool for the construction of kinetic models from MD simulations. For example, they have been used to elucidate protein folding,<sup>37</sup> ligand binding processes,<sup>38</sup> and the dynamics of intrinsically disordered proteins,<sup>39</sup> for the identification of allosteric networks,<sup>40,41</sup> and for the interpretation of time-resolved experimental data.<sup>42,43</sup> In this study, MSMs are used to cluster conformations that interconvert on short time scales (kinetic clustering<sup>33–36</sup>) and to calculate relaxation time scales and equilibrium populations for these clusters. The results are then compared with the new NMR data for CsA in chloroform reported in this study as well as previous data from the literature.

## METHODS

**NMR Measurements.** Sample preparation, data acquisition and processing, and resonance assignments are described in the Supporting Information.

**List of Upper Distances.** Initial peak picking and NOE assignments were performed using the ATNOSCANDID package.<sup>46,47</sup> Peak lists of the first cycle were used as an input for the program CYANA 3.97.<sup>48</sup> The peak list was analyzed in SPARKY, and interstrand NOEs were assigned manually. In presence of dihedral restraints based on homo and hetero  $^3J$  scalar couplings using the Karplus equation, the noeassign protocol of CYANA was used to assign unassigned peaks and calibrate the NOE signals with an average NOE distance of 0.42 nm. The resulting upper distance list was cleaned by applying a cutoff for the support factor of 0.3 and by reviewing the peak lists and inspecting the NOESY spectra. The NOE upper distance bounds are listed in Table S1 in the Supporting Information, and the amide temperature coefficients and the  $^3J$  scalar couplings are compiled in Tables S2 and S3, respectively.

**Preparation of NOE Upper Bounds for Comparison with MD Simulations.** Pseudoatom distance corrections as given by Wüthrich and co-workers<sup>49</sup> were added. The proton–proton distances in the simulation were calculated using the procedure described in ref 50. Processing and analysis of NOE data were performed using the prep\_noe, noe, and post\_noe programs from the GROMOS package of analysis programs.<sup>51</sup>

**Molecular Dynamics Simulations.** The workflow used in this study is composed of three stages: (i) initial equilibrium and enhanced sampling to obtain diverse seed conformations, (ii) exhaustive sampling by performing MD simulations starting from the seed conformations, and (iii) construction of the MSMs.

All of the simulations were performed using the GROMOS software package<sup>52</sup> and the GROMOS 54A7 force field.<sup>50</sup> Simulations were carried out under NPT conditions with periodic boundaries. Newton's equations of motion were integrated using the leapfrog scheme<sup>53</sup> with a time step of 0.002 ps. The simple-point-charge (SPC) water model<sup>54</sup> and the GROMOS chloroform model<sup>55</sup> were used. To keep the

temperature close to its reference value, weak coupling to a temperature bath with a relaxation time of 0.1 ps was applied.<sup>56</sup> The peptide and the solvent were coupled to separate temperature baths. The pressure was maintained close to 1.013 bar (1 atm) by weak coupling to a pressure bath with a relaxation time of 0.5 ps and using the experimental isothermal compressibilities of water and chloroform, respectively. A twin-range cutoff scheme with cutoffs of 0.8 and 1.4 nm was used for the nonbonded interactions. A reaction field force<sup>57</sup> with the relative dielectric permittivity of the corresponding solvent (61<sup>58</sup> for water or 4.8<sup>59</sup> for chloroform) was applied to treat electrostatic interactions beyond the long-range cutoff. Bond lengths in the peptide and in the solvent were constrained to ideal values using the SHAKE algorithm<sup>60</sup> with a tolerance of 10<sup>-4</sup> nm. The coordinates and energies were written out every 5 ps for analysis.

The topology of CsA is provided in the Supporting Information.

**Seed Conformations.** In order to generate seed conformations, both standard and enhanced-sampling MD simulations were carried out at two temperatures, 300 and 400 K. Starting conformations for the MD simulations were taken from the available crystal structures of CsA: (i) the crystal structure of CsA (ID: DEKSAN,<sup>23</sup> source: CSD), which represents the closed conformation predominant in chloroform (left panels in Figure 1), and (ii) the crystal structure of CsA cocrystallized with cyclophilin (ID: 2Z6W,<sup>44</sup> source: PDB) which represents the open conformation assumed to be predominant in water (right panels in Figure 1). In the following they will be called the CRYSTC and CRYSTO conformations, respectively. CRYSTC was solvated in 369 chloroform molecules or 1603 water molecules (box length of 3.698 nm), and CRYSTO was solvated in 302 chloroform or 1327 water molecules (box length of 3.470 nm).

The simulation setups of the standard simulations are summarized in Table 1. All of the simulations were 100 ns in

**Table 1. Simulation Setups for the Standard MD Simulations in Water ( $H_2O$ ) or Chloroform ( $CHCl_3$ ); All of the Simulations Were 100 ns in Length**

starting conformation	solvent	temperature [K]
CRYSTO	$H_2O$	300
CRYSTO	$CHCl_3$	300
CRYSTO	$CHCl_3$	400
CRYSTC	$CHCl_3$	300
CRYSTC	$H_2O$	300
CRYSTC	$H_2O$	400

length, preceded by 100 ps of equilibration. The conformations from the 100 ns runs were clustered on the basis of the backbone-atom-positional root-mean-square deviations (RMSDs) relative to the starting structure (using a cutoff of 0.06 nm), and the centroids were collected in a conformation pool (one for each solvent).

Additionally, to enhance sampling, a variety of simulations with unphysical/ altered potential energy terms (summarized in Table 2) were performed at 400 K. These simulations were 50 or 100 ns in length. The trajectories from the enhanced-sampling simulations were reweighted to physical conditions. Each configuration was weighted by the likelihood to be in the unbiased ensemble,

**Table 2. Simulation Setups for the Enhanced-Sampling MD Simulations in Water ( $\text{H}_2\text{O}$ ) or Chloroform ( $\text{CHCl}_3$ ) at 400 K**

starting conformation	solvent	time [ns]	modification
CRYSTC	$\text{H}_2\text{O}$	100	no backbone dihedral angle terms
CRYSTO	$\text{CHCl}_3$	100	no backbone dihedral angle terms
CRYSTC	$\text{H}_2\text{O}$	50	repulsive potential on atoms forming H-bonds
CRYSTO	$\text{CHCl}_3$	50	repulsive potential on atoms forming H-bonds
CRYSTC	$\text{H}_2\text{O}$	50	no charges on atoms forming H-bonds
CRYSTO	$\text{CHCl}_3$	50	no charges on atoms forming H-bonds
CRYSTC	$\text{H}_2\text{O}$	50	reduced charges ( $q = 0.1e$ ) on atoms forming H-bonds
CRYSTO	$\text{CHCl}_3$	50	reduced charges ( $q = 0.1e$ ) on atoms forming H-bonds

$$w_i = \frac{\exp[-(V_{\text{phys}} - V_{\text{unphys}})/k_B T]}{\sum_{j=0}^N \exp[-(V_{\text{phys}} - V_{\text{unphys}})/k_B T]} \quad (1)$$

where  $k_B$  is the Boltzmann constant,  $T$  is the absolute temperature,  $V_{\text{phys}}$  the potential energy under physical conditions, and  $V_{\text{unphys}}$  the potential energy under unphysical conditions. Conformations with weights greater than or equal to  $10^{-9}$  were considered low-energy and added to the respective conformation pool. This cutoff was chosen such that at least 10% of the conformations of a particular simulation were selected.

For each solvent, a backbone-atom-positional RMSD matrix was calculated for the conformations in the pool. On the basis of the RMSD matrix, the 100 most diverse conformations were selected using the LazyPicker program in the RDKit toolkit.<sup>61</sup> This procedure resulted in 100 seed conformations for each solvent.

**Exhaustive Sampling.** From each of the seed conformations obtained in the previous step, an MD simulation with a length of 100 ns was performed at 300 K, resulting in a total of 10  $\mu\text{s}$  of simulation data for each solvent. The simulation trajectories were used further to construct MSMs of CsA.

**Markov State Models.** For the convenience of the reader, we summarize the major features of the MSM theory, which has been discussed in detail elsewhere.<sup>62–64</sup> In MSMs, the conformational space of the molecule,  $\mathbb{R}^d$ , where  $d$  is the number of internal degrees of freedom, is discretized into  $n$  nonoverlapping microstates  $s_i$ :

$$s_i \cap s_j = \emptyset, i \neq j \quad \cup_{i=1}^n s_i = \mathbb{R}^d \quad (2)$$

The trajectory in the continuous conformational space,  $z(t) \in \mathbb{R}^d$ , is projected onto these microstates to obtain a microstate trajectory  $s(t) = i \Leftrightarrow z(t) \in s_i$ . The dynamics in this state space is then modeled as a Markov process, that is, one assumes that the probability of finding the system in state  $s_j$  at time  $t + \tau$  only depends on its state  $s_i$  at time  $t$  and not on any of the previously visited states:

$$t_{ij}(\tau) = \mathbb{P}(z(t + \tau) \in s_j | z(t) \in s_i) \quad (3)$$

The parameter  $\tau$  is called the lag time. The transition probabilities can be estimated from the microstate trajectories by counting the number of transitions  $c_{ij}(\tau)$  from state  $s_i$  to state  $s_j$  within a time window of length  $\tau$  and normalizing by the number of transitions that originate in  $s_i$ :

$$\hat{t}_{ij}(\tau) = \frac{c_{ij}(\tau)}{\sum_{j=1}^n c_{ij}(\tau)} \quad (4)$$

where the hat denotes an estimate. Arranging the transition probabilities into a  $n \times n$  matrix yields the transition matrix  $\mathbf{T}(\tau)$ , whose eigenvalues  $\lambda_i(\tau)$  and eigenvectors  $\mathbf{r}_i$  contain a wealth of information on the conformational dynamics of the system:<sup>36</sup>

$$\mathbf{T}(\tau)\mathbf{r}_i = \lambda_i(\tau)\mathbf{r}_i \quad (5)$$

If the dynamics of the system is indeed Markovian in the chosen state space (the projection typically causes a deviation from full Markovian behavior), the lag-time-dependent eigenvalues decay exponentially. Thus, the time scale of this decay, the so-called implied time scale (ITS), should be lag-time-independent:<sup>34</sup>

$$t_i^* = \frac{-\tau}{\ln(|\lambda_i(\tau)|)} = \text{constant} \quad (6)$$

This relation can be used to check the approximation quality of the MSM.

In this study, the conformational space of CsA was discretized by clustering on the basis of the backbone torsional RMSD, and the index numbers of the cluster centers were used as microstate identifiers. To this end, the trajectories of the first microsecond of simulation (i.e., the first 10 ns of each individual simulation) in water and chloroform were concatenated, and every 20th frame was taken to calculate a backbone torsional RMSD matrix using the `rmsdm` program from the GROMOS package of analysis programs.<sup>51</sup> Subsequently, the conformations were clustered on the basis of this RMSD matrix (cutoff = 20) using the GROMOS `cluster` program. In order to keep the number of microstates reasonable, only clusters containing more than 20 members were considered as microstates. Finally, every conformation in the 10  $\mu\text{s}$  trajectory data set was assigned to the closest cluster  $c_i$  on the basis of the RMSD to the centroid (Voronoi partitioning), such that  $s(t) = \operatorname{argmin}_i d(z(t); c_i)$ , where  $d(z(t), c_i)$  is the distance between the conformation at time  $t$  and the center of cluster  $i$ .

The MSMs were constructed using the EMMA software, version 1.4.<sup>65</sup> Transition matrices were calculated for a series of lag times: 0.1, 0.5, 0.7, 1.0, 1.5, 2.0, 2.5, 3.0, and 5.0 ns. The subsequent analysis was performed on the transition matrix for the lag time, for which the implied time scales were found to be sufficiently constant.

To interpret the MSMs, the microstates were grouped into metastable sets using Perron cluster-cluster analysis plus (PCCA+).<sup>33,62,66</sup> This algorithm uses the first  $m$  eigenvectors of the transition matrix to identify  $m$  metastable sets such that the interconversions between microstates within a metastable set are faster than transitions to other metastable sets. These metastable sets then represent long-lived conformations of the molecule. The PCCA+ algorithm yields fuzzy assignments to metastable sets. That is, each microstate  $i$  is partially assigned to each metastable set  $k$  with a membership value

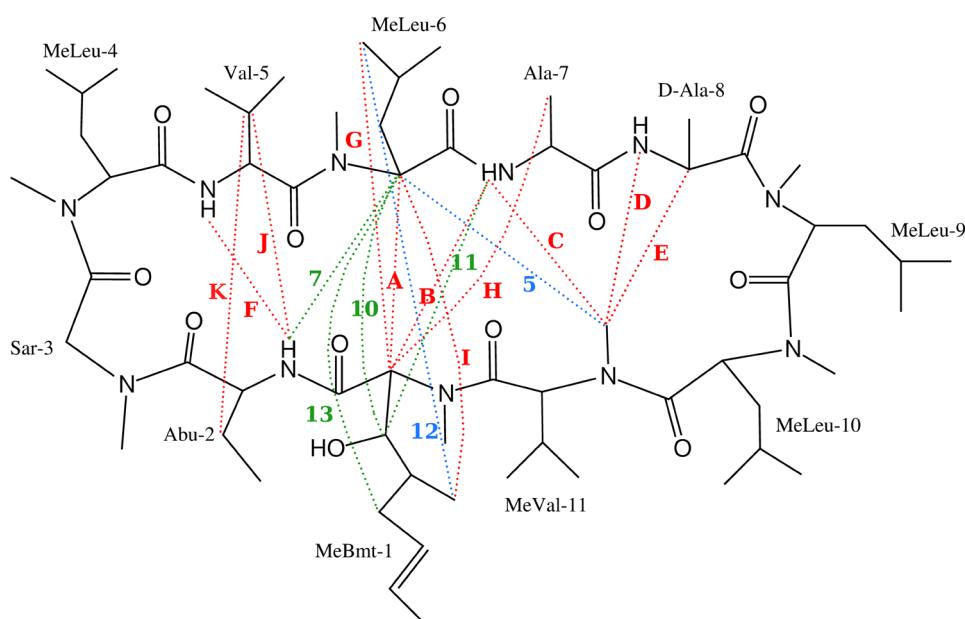
$$w_{ik} \in [0, 1] \quad \forall i, k \quad \sum_{k=1}^m w_{ik} = 1 \quad \forall i \quad (7)$$

with  $i \in [1, \dots, n]$  and  $k \in [1, \dots, m]$ . We converted this fuzzy assignment to a crisp assignment by assigning each microstate

Table 3. Intercycle NOE Upper Distance Bounds of CsA in Chloroform from Two Sets: 1990<sup>20</sup> and 2016 (This Study)<sup>a</sup>

label	1990				2016			
	index	residue 1	residue 2	upper bound [nm]	index	residue 1	residue 2	upper bound [nm]
A	1	1 HA	6 HA	0.269	1	1 HA	6 HA	0.415
B	2	1 HA	7 HN	0.300	2	1 HA	7 HN	0.462
C	3	7 HN	11 HCN	0.394	3	7 HN	11 HCN	0.416
D	4	8 HN	11 HCN	0.400	4*	8 HN	11 HCN	0.435
—					5	6 HA	11 HCN	0.393
E	5	8 HA	11 HCN	0.399	6*	8 HA	11 HCN	0.461
F	6	2 HN	5 HN	0.321	7	2 HN	5 HN	0.465
—	7	2 HN	6 HA	0.357	—			
G	8	1 HA	6 HD	0.460	8*	1 HA	6 HD	0.463
H	9	1 HA	7 HB	0.517	9*	1 HA	7 HB	0.550
—	10	1 HB	6 HA	0.432	—			
—	11	1 HB	7 HN	0.472	—			
I	12	1 HD1	6 HA	0.397	10	1 HD1	6 HA	0.413
—	13	1 HD2	6 HA	0.351	—			
J	14	2 HN	5 HB	0.319	11	2 HN	5 HB	0.479
—	—				12	1 HD1	6 HD	0.499
K	15	2 HB	5 HB	0.298	13	2 HB	5 HB	0.416

<sup>a</sup>Atom labels and indices follow the GROMOS topology provided in the Supporting Information. The NOE peaks marked with asterisks (\*) in the 2016 set were manually picked. The distances are shown graphically in Figure 2.



**Figure 2.** Schematic representation of the intercycle NOE upper distance bounds from the two sets 1990<sup>20</sup> and 2016. NOE distances that are reported in both sets (red) are labeled according to the letter labels in Table 3. NOE distances reported in only one set (green for the 1990 set and blue for the 2016 set) are labeled according to their indexes in Table 3.

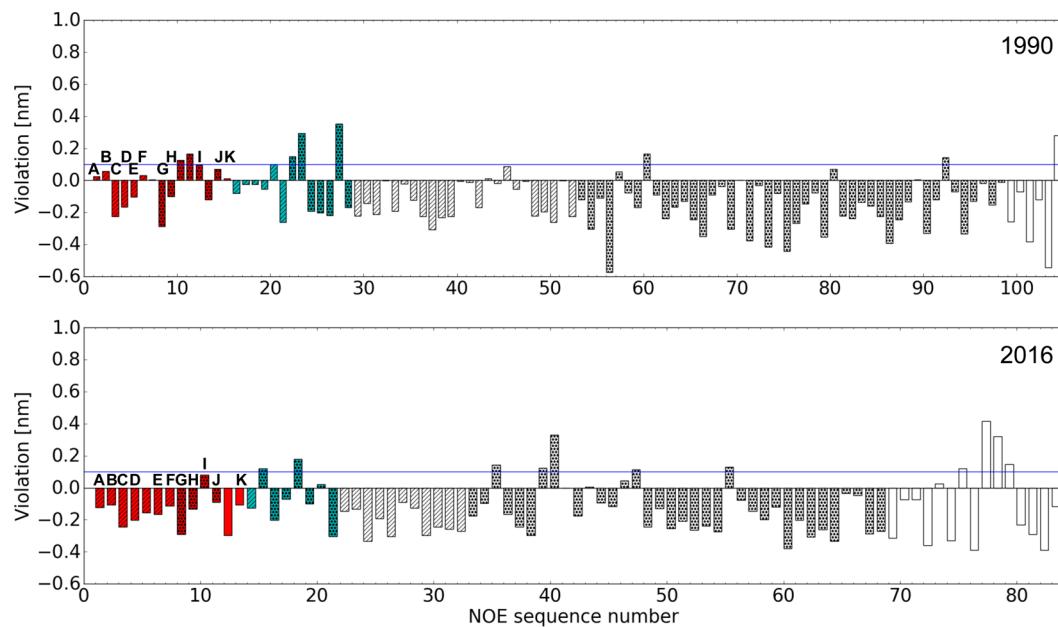
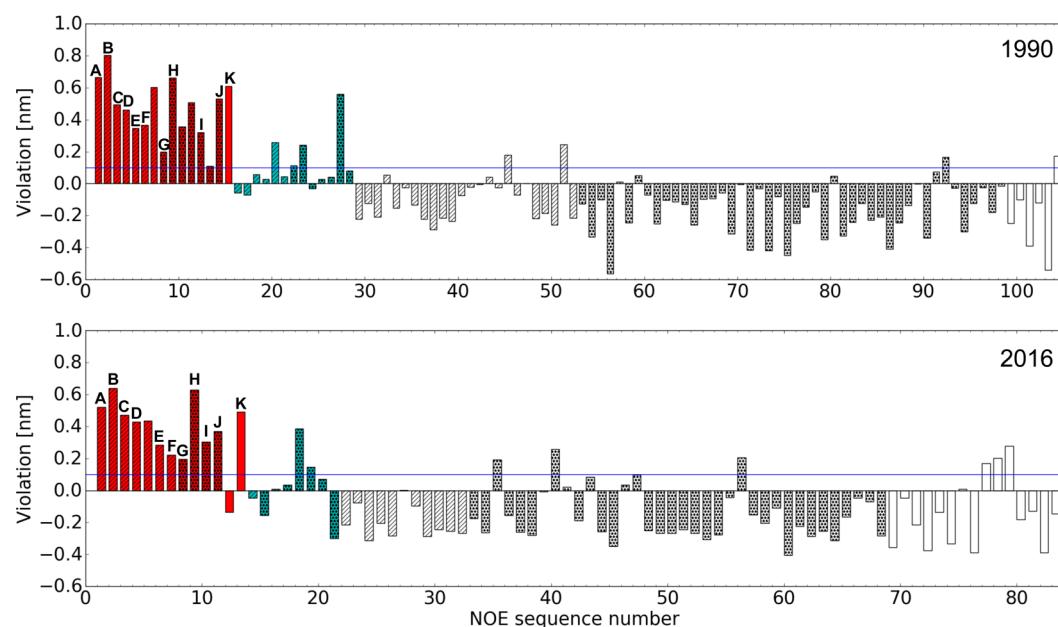
to the metastable set for which it had the highest membership value.

Each metastable set was analyzed and characterized in terms of its pattern of backbone H-bonds, the Ramachandran plots of the backbone dihedral angles, potential energies, and the backbone-atom-positional RMSDs relative to the crystal structures using the GROMOS programs<sup>51</sup> hbond, tser, ene\_ana, and rmsd, respectively. The criterion for a H-bond to be considered present was a hydrogen-acceptor distance less than 0.25 nm and an angle larger than 135°.<sup>24</sup> For selected metastable sets as well as combinations of them, the violations of the NOE upper distance bounds were calculated as discussed below.

Reactive transitions between selected metastable sets were analyzed by transition path theory<sup>67</sup> as implemented in the EMMA software. The transitions were characterized by the forward committor ( $q+$ ) and fluxes between the metastable sets.

## RESULTS AND DISCUSSION

**NOE Upper Distance Bounds.** Apart from the NOE data reported in this study, two other NOE sets from the literature were considered. In the following the three sets will be denoted by their years of publication: 1985,<sup>23</sup> 1990,<sup>20</sup> and 2016. As the distance bounds from 1985 are rather coarse (i.e., they were simply divided into the three categories weak, medium, and

**A CRYSTC****B CRYSTO**

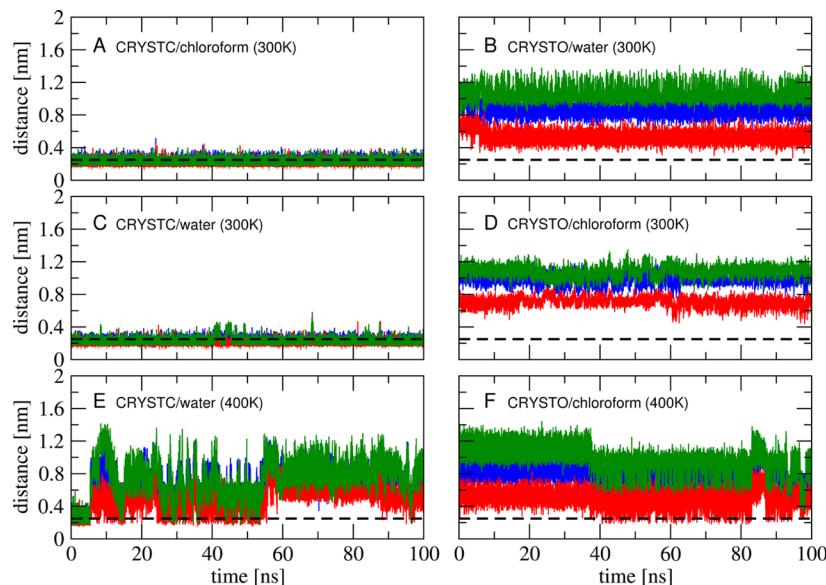
**Figure 3.** Violations of the experimental NOE upper distance bounds as a function of the NOE sequence number for the crystal structures of CsA. (A) Conformation CRYSTC (ID: DEKSAN,<sup>23</sup> source: CSD). (B) Conformation CRYSTO from the CsA–cyclophilin complex (ID: 2Z6W,<sup>44</sup> source: PDB). The top and bottom graphs in each panel correspond to violations from the 1990 and 2016 data sets, respectively. Hydrogen pairs that were reported in both sets are marked with capital letters according to the labels in the first column of Table 3. The NOE distances are grouped into “intercycle” (red), “intermediate” (cyan), and NOE distances from neighboring or the same residues (white), and they are labeled on the basis of the nature of the protons: backbone–backbone (stripes), backbone–side chain (dots), and side chain–side chain (plain).

strong) and the data from 1990 represent a follow-up by the same authors, we will focus on the latter data set. The NOE upper bounds from 1985 are listed in Table S4.

The “intercycle” NOE upper distance bounds of the two sets 1990<sup>20</sup> and 2016 are listed in Table 3 and shown graphically in Figure 2. An NOE distance bound is classified as “intercycle” when the atoms are separated by at least three residues, while “intermediate” NOEs are two residues apart. The main focus of the analysis was put on the intercycle NOEs, as these contain information about the backbone conformation of CsA. NOE

distances between neighboring residues provide information about their relative spatial arrangement with respect to the ring but not the global structure of the peptide. The full list of distances for each set are given in Tables S1 and S5.

The sets of NOE upper distance bounds reported in this study and in 1990<sup>20</sup> were both recorded in chloroform. However, because of differences in the calibration of the NOE intensities and the temperature at which the NMR experiments were performed, differences between the two sets were observed (see Table 3 and Figure 2). NOE bounds no. 4, 5,



**Figure 4.** Time series of distances between backbone atoms in 100 ns simulations. Distances between Abu-2 N–H and Val-5 C=O are shown in blue, distances between Val-5 N–H and Abu-2 C=O in red, and distances between Ala-7 N–H and MeVal-11 C=O in green. The interatomic distance cutoff for H-bonds is shown by the horizontal dotted black lines. (A) Simulation in chloroform started from CRYSTC at 300 K. (B) Simulation in water started from CRYSTO at 300 K. (C) Simulation in water started from CRYSTC at 300 K. (D) Simulation in chloroform started from CRYSTO at 300 K. (E) Simulation in water started from CRYSTC at 400 K. (F) Simulation in chloroform started from CRYSTO at 400 K.

8, and 9 reported in the 1990 set (4, 6, 8, and 9 in the 2016 set) were not automatically detected in our NOE spectra, and thus, the peaks were manually picked. NOEs no. 7 and 13 in the 1990 set could not be determined in our spectra because of artifacts or overlap with other signals. NOEs no. 10 and 11 were not visible in our spectra, likely because of the higher temperature at which the spectra were recorded (the NOE distances from 1990 were measured at 250 K in order to obtain stronger intensities by shifting to slower motion). Differences between the 1990 set and the 2016 set can also be seen in the upper distance bounds. For example, the upper distance bound of NOE no. 1 was reported in 1990 as 0.269 nm versus 0.415 nm determined in this study. This is likely due to the modern approach that was employed here to calibrate the NOE intensities, which is described in ref 68. This is a more conservative method that takes into account the flexibility of a molecule as well as the errors in NOE intensities. As this reflects better what is obtained from MD simulations, in the following the focus of the analysis was put on the 2016 set. As reported previously,<sup>20</sup> the presence of a minor conformation was observed in our spectra (at about 6%), but it could not be resolved.

The amide temperature coefficients (Table S2) indicate that the amide moieties of Abu-2, Val-5, and D-Ala-8 are involved in strong intramolecular H-bonds, whereas the intramolecular H-bond formed by the N–H of Ala-7 is weak.

**Comparison of NOE Upper Distance Bounds and Crystal Structures.** The NOE upper distance bounds were compared with the crystal conformations CRYSTC (ID: DEKSAN,<sup>23</sup> source: CSD) and CRYSTO (ID: 2Z6W,<sup>44</sup> source: PDB). The NOE violations of the two structures are shown in Figure 3. Only violations above 0.1 nm are considered to be significant. CRYSTC fulfills all of the intercycle NOE distances of the 2016 set within this limit but violates two intermediate NOE distances (NOE no. 15 is between the side chain of MeBmt-1 and the backbone of Sar-3, and NOE no. 18 is between the backbone of D-Ala-8 and the side chain of MeLeu-

6) as well as a few intraresidue NOE distances (Figure 3A). The number of violations of the 1990 set by CRYSTC is higher, as these upper bounds are generally more tight. Of the four  $^3J(H_N-H_{C\alpha})$  coupling constants, only two are fulfilled within 1 Hz by CRYSTC (Figure S1 in the Supporting Information), indicating that a more diverse conformational ensemble is present in chloroform.

CRYSTO violates all of the intercycle NOE distances of both sets, together with many intermediate NOEs (Figure 3B). No experimental NOE data for CsA in water are available because CsA is not soluble enough.

**Initial MD Simulations.** Standard MD simulations with a length of 100 ns were performed in chloroform and water using two different starting conformations: CRYSTC (ID: DEKSAN,<sup>23</sup> source: CSD) and CRYSTO (ID: 2Z6W,<sup>44</sup> source: PDB). The time series of the distances between backbone atoms that form intercycle H-bonds in the CRYSTC and CRYSTO conformations are shown for the different simulations in Figure 4. The corresponding time series of the H-bonds and H-bonding patterns are presented in Figures S2–S7.

Starting from CRYSTC and using chloroform as the solvent, the closed structure was preserved during the 100 ns simulations (Figure 4A). The intramolecular H-bonds were stable over the course of the simulation, and the MeBmt-1 side chain remained folded over the backbone. Starting from CRYSTO and using water as the solvent, CsA remained in an open form with the backbone amides pointing outward from the ring and forming H-bonds with the water, but as expected, it showed much higher flexibility (Figure 4B). The MeBmt-1 side chain extended into the solvent.

To check whether direct interconversion between the two structures could be observed, simulations in water starting from CRYSTC and in chloroform starting from CRYSTO were carried out. Interestingly, when the simulation was started from CRYSTC with water as the solvent (at 300 K), CsA did not

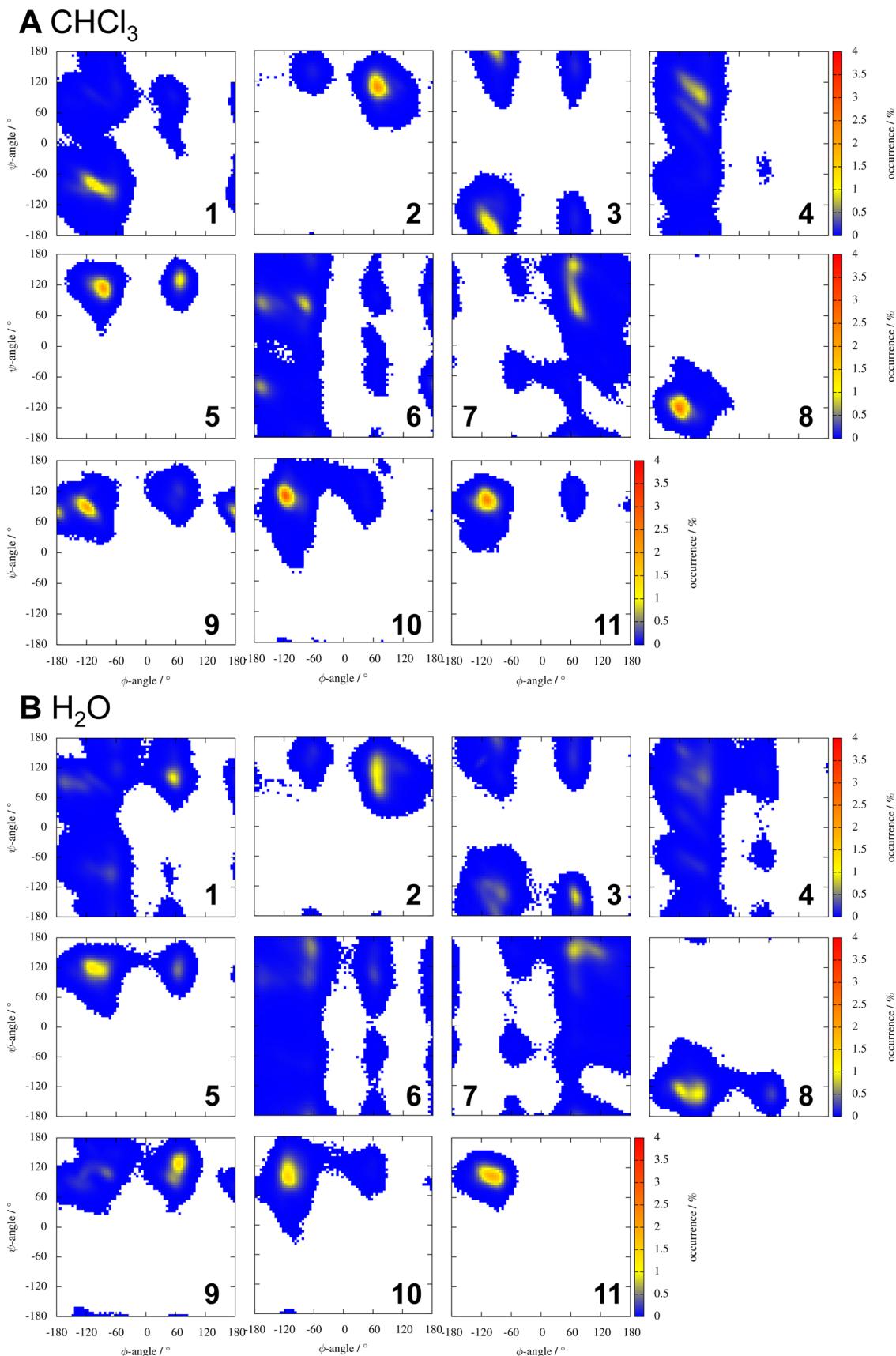
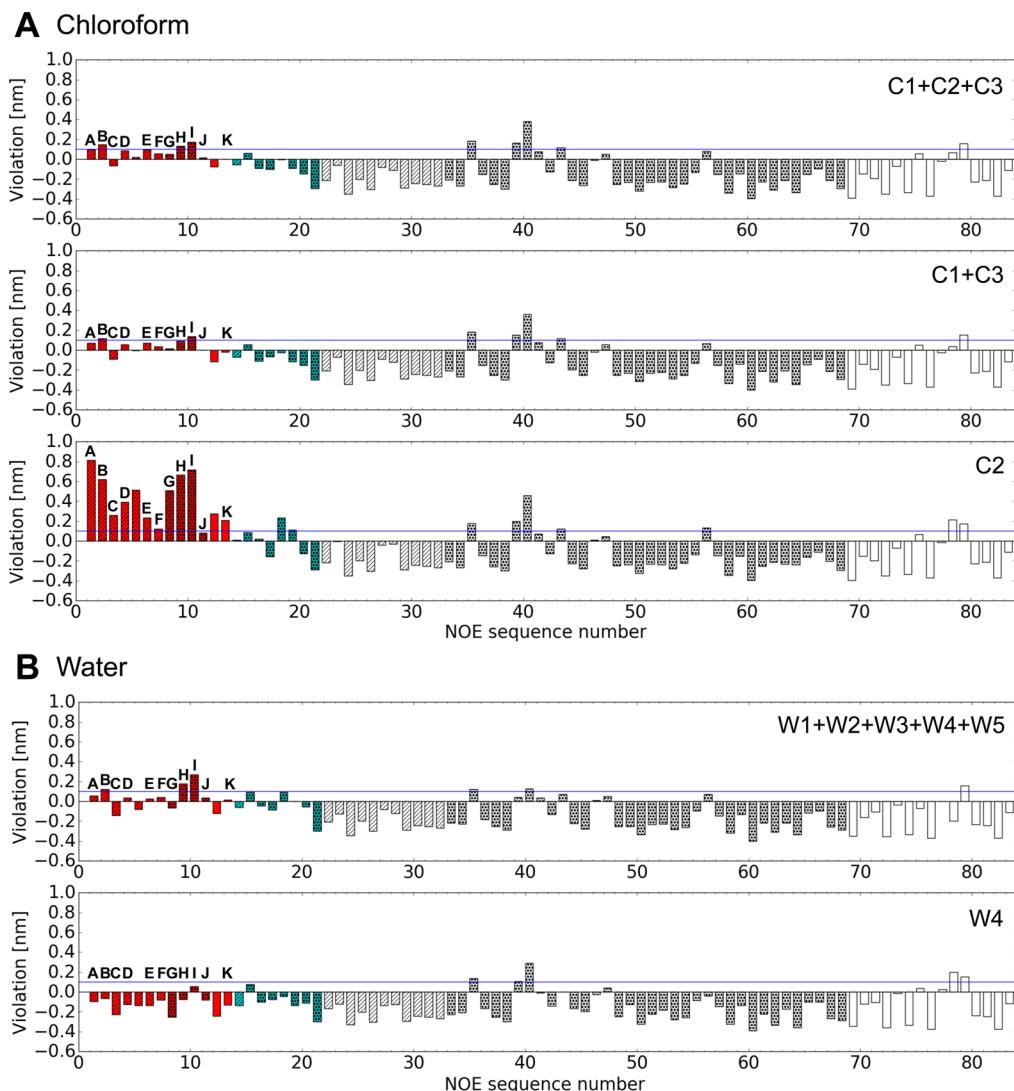


Figure 5. Ramachandran plots of the backbone dihedral angles for 10  $\mu$ s simulations in (A) chloroform and (B) water at 300 K.

open but remained in the closed conformation with all of the intramolecular backbone H-bonds preserved (Figure 4C).

Increasing the temperature to 400 K resulted in partial opening of the conformation. The distance between the atoms forming



**Figure 6.** Violations of the experimental NOE upper distance bounds from the 2016 set as a function of the NOE sequence number calculated from the  $10\ \mu\text{s}$  of trajectory data. The NOE distances are grouped into “intercycle” (red), “intermediate” (cyan), and NOE distances from neighboring or the same residues (white) and are labeled on the basis of the nature of the protons: backbone–backbone (stripes), backbone–side chain (dots), and side chain–side chain (plain). Hydrogen pairs that were reported in both sets are marked with capital letters according to the labels in the first column of Table 3. (A) MSM in chloroform with the metastable sets C1, C2, and C3. (B) MSM in water with the metastable sets W1, W2, W3, W4, and W5.

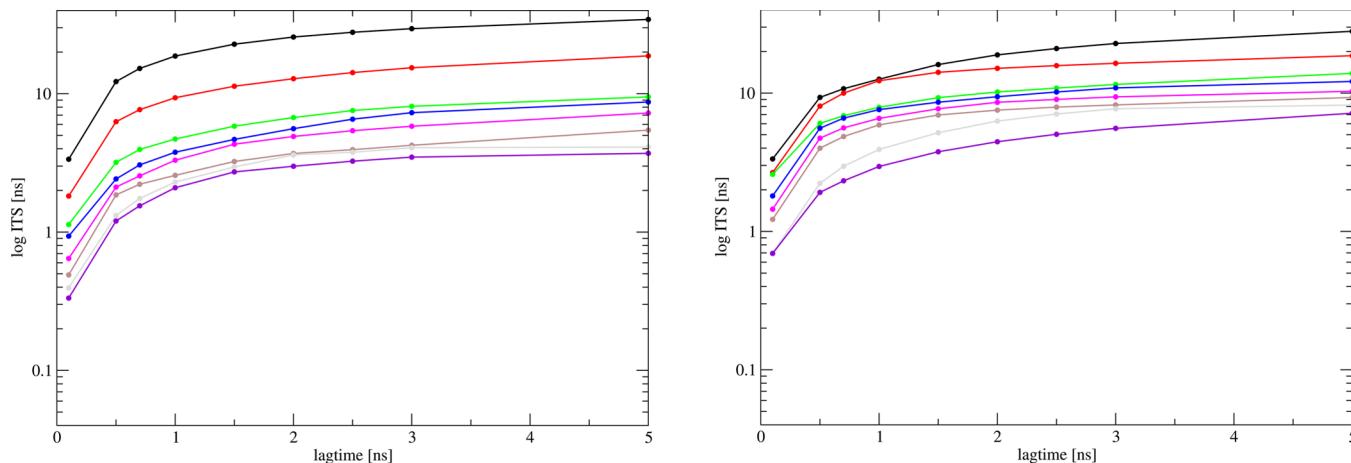
backbone H-bonds in CRYSTC increased significantly (Figure 4E). On the other hand, when the simulation was started from CRYSTO with chloroform as the solvent (at 300 K), the structure was found to undergo hydrophobic collapse to a more closed conformation (Figure 4D) that did not resemble CRYSTC (Figure S8). One major intercycle backbone–backbone H-bond was formed between Val-5 and MeLeu-10, which is not present in CRYSTC (Figure S5). Increasing the temperature to 400 K resulted in enhanced flexibility of CsA, thereby forming other intramolecular H-bonds with shorter lifetimes, but the backbone did not rearrange to the CRYSTC conformation (Figure 4F).

These results indicate that although the sampling time has been increased by 3 orders of magnitude compared with previous MD simulations<sup>20</sup> (100 ns versus 50 ps), it is still insufficient to observe direct interconversions between the open and closed conformations, even at elevated temperatures.

**Attempted Exhaustive Sampling.** By means of the protocol described in Methods, 100 diverse conformations of

CsA in water and chloroform were obtained and used as seed conformations for 100 ns simulations. The Ramachandran plots of the backbone dihedral angles sampled in both solvents are shown in Figure 5. The same plots for the seed conformations are given in Figure S9. Interestingly, the sampled range of the torsional angles was found to be largely similar in chloroform and in water. In general, a slightly wider range was sampled in water, which is to be expected as NMR experiments report a higher flexibility of CsA in polar solvents.<sup>19,25,27,28</sup> Differences are the presence of areas sampled only in water for MeLeu-4 ( $\phi \approx 60^\circ$ ) and D-Ala-8 ( $\phi/\psi \approx 60^\circ/-120^\circ$ ) and areas sampled only in chloroform for MeVal-11 ( $\phi/\psi \approx 60^\circ/120^\circ$ ). For MeBmt-1, Sar-3, MeLeu-6, and MeLeu-9, the position of the main minimum changes in going from water to chloroform. In the case of MeLeu-9, the main minimum is in the  $\beta$ -region ( $\phi/\psi \approx -120^\circ/120^\circ$ ) in chloroform but not in water.

The conformational ensemble generated in the  $10\ \mu\text{s}$  simulations in chloroform was compared with the experimental NOE upper distance bounds from the 2016 set (Figure 6).



**Figure 7.** Implied time scales for interconversion processes (on log scale) described by MSMs in (left) chloroform and (right) water. The linear-scaled graphs are shown in Figure S12.

Overall there is good agreement between the MD ensemble in chloroform and the 2016 NOE set, with three smaller violations of intercycle NOE distances (top panel of Figure 6A). The largest intercycle violation is NOE no. 10, which is between the side chain of MeBmt-1 and the backbone of MeLeu-6. Interestingly, the same NOE distance is also violated to a small degree by CRYSTC. The other two violations are NOEs no. 2 and 9 between the residues MeBmt-1 and Ala-7. It is to be stressed at this point that the NOE data sets represent only the major conformation observed in chloroform (the minor conformation present at about 6% could not be resolved), whereas the MD ensemble contains all of the conformations sampled. The violations of the 1990 NOE set by the MD ensemble were generally larger (Figure S10), indicating that the modern approach for the calibration of NOE intensities better reflects the conformational flexibility observed in the simulations.

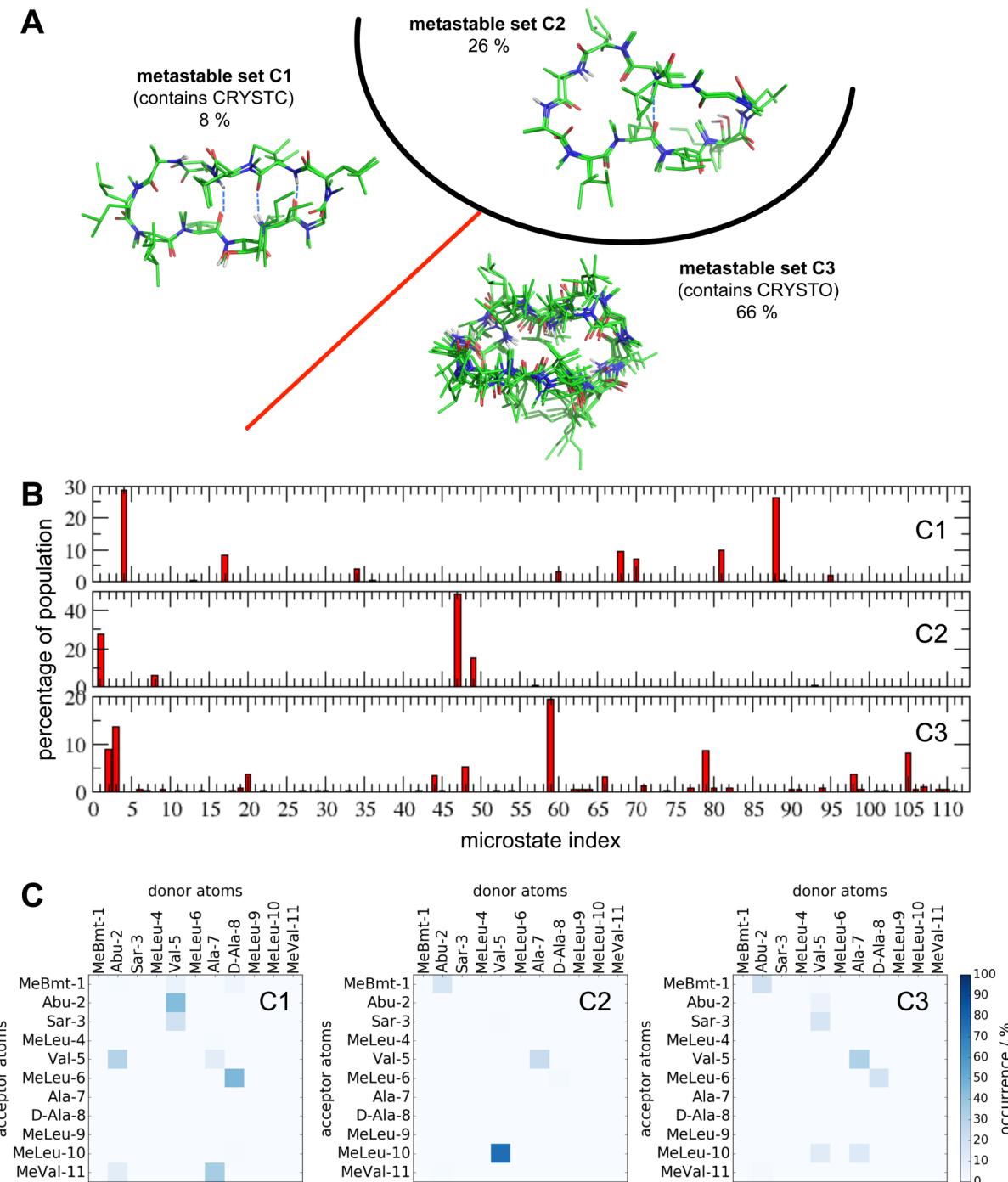
As expected, the violations of the NOEs in chloroform by the MD ensemble in water were found to be larger (top panel in Figure 6B). This indicates that the MD ensembles in water and in chloroform differ despite the similarities in the Ramachandran plots, which has to be the case on the basis of the experimental observations.

**Kinetic Models.** MSMs based on the conformations obtained from 10  $\mu$ s MD simulations in water and chloroform were constructed. As a common discretization was used for the MSMs in the two solvents, the microstates and eigenvectors of the two models can be directly compared. The discretization protocol described in Methods resulted in a total of 112 microstates. On the basis of the transition matrix obtained from the discretized trajectories, the implied time scales (ITSs) of the interconversion processes were calculated (Figure 7). The ITSs reach a plateau at approximately  $\tau = 2.5$  ns, which was thus chosen as the lag time for further analysis. The relative populations of the 112 microstates in the two solvents are shown in Figure S11. While most of the highly populated microstates can be attributed to a particular solvent, there is significant overlap between the conformational spaces of CsA in chloroform and water in the microstates with lower population.

In chloroform, CsA showed two slow processes (>25 ns and >10 ns) that are separated by a spectral gap from a continuum of fast conformational interconversion processes. It should be noted that because of inaccuracies in the force field<sup>69</sup> and

discretization errors,<sup>36,70</sup> the transition rates obtained from MSMs are only rough estimates of the time scales of the interconversion processes occurring in nature. The microstates were thus grouped into three metastable sets termed C1, C2, and C3 (where “C” stands for chloroform) using the PCCA+ method<sup>33,62,66</sup> (Figure 8A). The metastable set C2 was found to be separated from the rest of the conformational space by the highest energy barriers, leading to the lowest interconversion rates to all of the other metastable sets (black line in Figure 7 left). Four microstates (nos. 1, 8, 47, and 49) constitute over 97% of the conformations in C2 (Figure 8B). The structures of the centroids of the two most abundant microstates (nos. 1 and 47) are shown in Figure 8A. C2 is characterized by an intercycle backbone–backbone H-bond between Val-5 and MeLeu-10, which occurs in 76% of the conformations in C2 (Figure 8C) and is not present in CRYSTC. The average backbone-atom-positional RMSDs relative to the crystal structures are relatively high, i.e., 0.342 nm to CRYSTC and 0.209 nm to CRYSTO (Table S6). The metastable set with the next lowest interconversion rate is C1 (red line in Figure 7 left). C1 comprises six microstates with a share of more than 5%, i.e., microstates no. 4, 17, 68, 70, 81, and 88 (Figure 8B). The structures of the two most abundant microstates (nos. 4 and 88) are shown in Figure 8A. Microstate no. 4 corresponds to the CRYSTC conformation. In line with this, C1 is characterized by the four backbone–backbone H-bonds also observed in CRYSTC, which are all present in more than 30% of the conformations in C1 (Figure 8C). Similarly, the average backbone-atom-positional RMSD relative to CRYSTC is 0.122 nm, whereas the average RMSD relative to CRYSTO is 0.225 nm (Table S6). In the metastable set C3, all of the remaining microstates are bundled (Figure 8B), with fast interconversion between them (Figure 7 left). This set was therefore not analyzed in more detail.

In NMR measurements of CsA in chloroform, a major and a minor conformation (or conformational ensembles) were observed. The MSM agrees well with this experimental finding. The metastable set C2 would thereby constitute the minor conformation with a low interconversion rate in terms of the NMR time scale. This hypothesis is supported by the fact that the NOE distance violations of C2 are large, and when only conformations belonging to C1 and C3 were considered, the agreement with the NOE distances of the major conformation improved (Figure 6A). In addition, Kessler et al.<sup>20</sup> noted that



**Figure 8.** Metastable sets in chloroform: C1, C2, and C3. (A) Schematic representation of the MSM. The colors of the dividing lines correspond to the implied time scale plots in the left panel of Figure 7. The structures shown are the centroids of the clusters corresponding to the two most abundant microstates in C1 and C2 and the eight most abundant microstates in C3 (generated using PyMOL<sup>45</sup>). The percentages indicate the relative populations obtained from the MSM. (B) Microstate distributions. (C) Patterns of backbone–backbone H-bonds.

the minor and major conformation likely differ in a cis/trans isomerization about an N-methylated peptide bond. The same conclusion can be drawn from the NMR measurements reported in this study. In the set C1, the peptide bond between MeLeu-9 and MeLeu-10 is in a cis conformation (as in CRYSTC), while in the set C2 all of the peptide bonds are in the trans conformation. The presence of two conformational states C1 and C2 can also be seen in the  $^3J$  coupling constants (Figure S1). The relative populations of the metastable sets are 26% for C2 and 74% for C1 + C3 (Figure 8A), which agree

roughly with the NMR data (i.e., the minor conformation is estimated to be approximately 6% present<sup>20</sup>). The sampling time of 10  $\mu$ s might still be too short to obtain accurate relative populations from the MSM. Another study on cyclic nonamers with sampling times of up to 1 ms found the longest implied time scale to be on the order of a few microseconds,<sup>71</sup> indicating that a sampling time of 10  $\mu$ s is in the appropriate range, yet at the low end. We therefore expect deviations in the relative populations by a few percentage points, but we do not expect the qualitative picture to change. Although the total

sampling time of  $10\ \mu\text{s}$  is much shorter than what is considered NMR slow exchange (approximately 50 ms), we can expect to see the different conformational states because of the use of enhanced sampling techniques for the generation of the diverse seed conformations, which cover the sampled conformational space (see Figure S9). Interestingly, both the average solute–solute potential energy and the average total potential energy of CsA in C2 were found to be comparable to those of C1 (Table 4). This indicates that in chloroform the conformational states

**Table 4.** Average and Median Potential Energies of the Metastable Sets in Chloroform (C1, C2, and C3) and Water (W1, W2, W3, W4, and W5)

metastable set	average potential energy [kJ/mol]					
	solute–solute		solute–solvent		total	
	mean	median	mean	median	mean	median
C1	-245	-247	-528	-526	-773	-773
C2	-242	-243	-542	-542	-785	-786
C3	-228	-230	-541	-540	-770	-770
W1	-218	-219	-603	-600	-822	-821
W2	-190	-191	-634	-634	-825	-825
W3	-178	-178	-676	-677	-855	-855
W4	-237	-241	-561	-554	-799	-798
W5	-173	-173	-688	-688	-861	-861

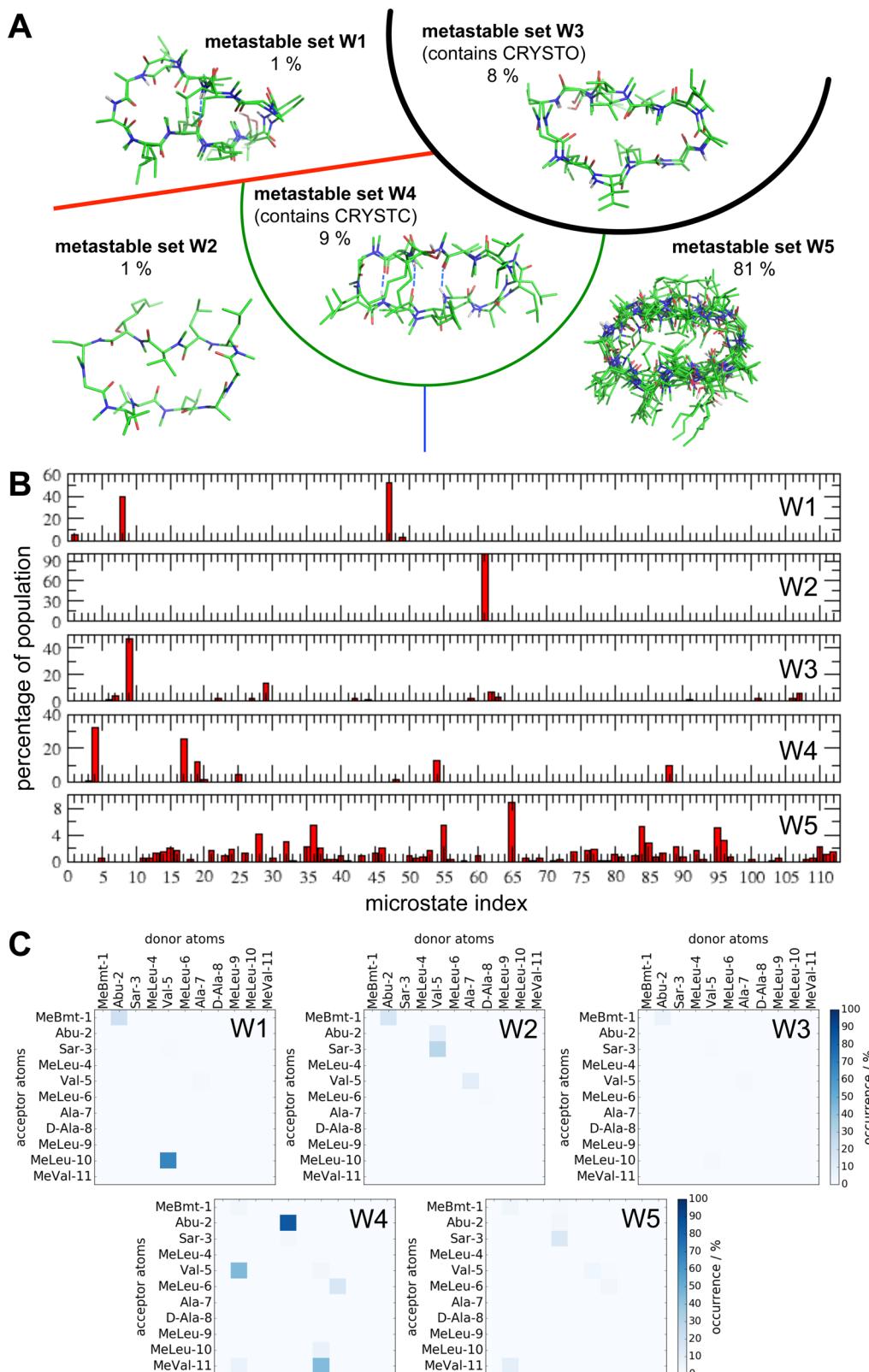
C1 and C2 are energetically similar (within the accuracy of the force field), but C2 is more long-lived because the interconversion process to any other metastable set is the slowest.

In water, CsA exhibits two slow interconversion processes as in chloroform ( $>20\ \text{ns}$  and  $>15\ \text{ns}$ ), which are separated from a continuum of fast processes by a spectral gap. However, in water the gap was found to be much smaller than in chloroform, and several fast processes could be resolved at a lag time of  $2.5\ \text{ns}$ . In water, the microstates were thus grouped into five metastable sets termed W1, W2, W3, W4, and W5 (where “W” stands for water) using the PCCA+ method<sup>33,62,66</sup> (Figure 9A). The metastable set W3 was found to be separated from the rest of the conformational space by the lowest interconversion rates to all of the other metastable sets, implying high free energy barriers (black line in Figure 7 right). Two microstates (nos. 9 and 29) constitute over 60% of the conformations in W3 (Figure 9B). The structures of the centroids of the two microstates are shown in Figure 9A. No backbone–backbone H-bonds were observed in W3 (Figure 9C), suggesting an open conformation. Microstate no. 27 that is present in W3 (comprising 2.4% of the conformations) corresponds to CRYSTO, showing the same H-bonding behavior. The average backbone-atom-positional RMSD of W3 relative to CRYSTO is thus small (i.e.,  $0.117\ \text{nm}$ ; Table S6). The metastable set with the next lowest interconversion rate is W1 (red line in Figure 7 right). W1 comprises four microstates, i.e., microstates no. 1, 8, 47, and 49 (Figure 9B). The structures of the two most abundant microstates (nos. 8 and 47) are shown in Figure 9A. W1 is characterized by one intercycle backbone–backbone H-bond between Val-5 and MeLeu-10 (Figure 9C). Both the microstate composition and the backbone H-bonding pattern are similar to those of the metastable set C2 in chloroform. The interconversion rates of the remaining three metastable sets in water are higher (Figure 7 right), with W4 having the lowest rate among them, followed

by W2. Five microstates (nos. 4, 17, 19, 54, and 88) constitute over 92% of the conformations in W4 (Figure 9B). Three of them (nos. 4, 17, and 88) also belong to the most abundant microstates in the metastable set C1 in chloroform. In line with this, the four backbone–backbone H-bonds observed in CRYSTC are also present in W4 (Figure 9C), and the average backbone-atom-positional RMSD relative to CRYSTC is  $0.084\ \text{nm}$  (Table S6). The metastable set W2 on the other hand comprises only a single microstate, no. 61 (Figure 9A), with no intercycle backbone–backbone H-bond (Figure 9C). In the metastable set W5, all of the remaining microstates are bundled (Figure 9B). This set was therefore not analyzed in more detail. The transition fluxes between W3 and W4 and between W3 and W1 were analyzed using transition path theory<sup>67</sup> (Figure S13). The most likely pathway is in both cases via the ensemble set W5, although different pathways are possible. As the system exhibits only a few metastable sets, the number of transition pathways is generally small. The relative magnitudes of the fluxes along these pathways correspond to what is expected from general structural considerations. In chloroform, for example, when going from a conformation in C1 with a specific H-bond pattern to a conformation in C2, which has a different H-bond pattern, the molecule likely passes through conformations in which no or only few H-bonds are present, i.e., conformations in C3 (Figure S13).

From NMR studies in 1:1 water/methanol, it is known that the major conformation (or conformational ensemble) of chloroform is still present to about 20% in that more polar solvent mixture.<sup>27</sup> The metastable set W4 may represent this conformational ensemble in pure water. In line with this, the NOE distance bounds obtained in chloroform were generally fulfilled by this metastable set (lower panel in Figure 6B). The metastable sets W1, W2, and W3 may represent other distinct conformations of CsA with low interconversion rates on the NMR time scale, as observed in NMR measurements of CsA in dimethyl sulfoxide, methanol, and water/methanol.<sup>19,25,27,28</sup> The metastable set W3, where CsA forms H-bonds with the solvent, was found to be the most long-lived conformational state in water and also energetically the most favorable (within the accuracy of the force field) (Table 4). This observation agrees with the experimental finding that CsA binds in such an open conformation to cyclophilin in water. The relative populations of the metastable sets are 1% for W1 and W2, 8% for W3, 9% for W4, and 81% for W5 (Figure 9A). As not much is known from NMR measurements regarding the relative populations of the distinct conformations, a direct comparison is difficult. As for chloroform, the sampling time of  $10\ \mu\text{s}$  might still be too short to obtain accurate relative populations from the MSM.

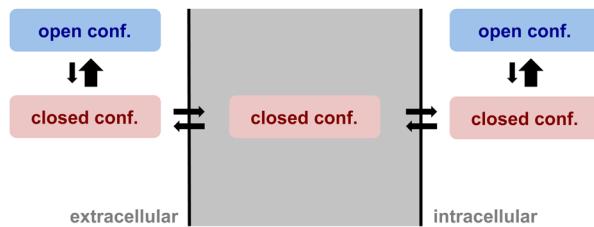
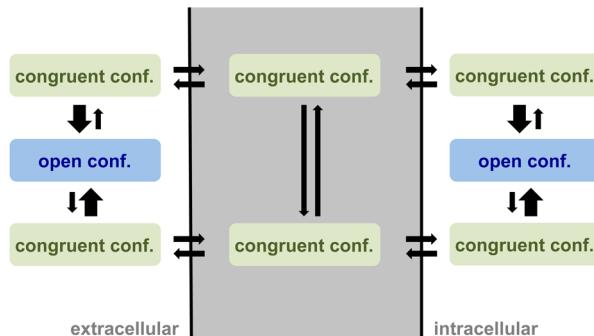
Interestingly, two pairs of largely similar conformational states were observed in the two solvents, i.e., C1/W4 and C2/W1, which we thus term “congruent” states. The current hypothesis for the passive diffusion of CsA through membranes is that there is a conformational equilibrium between the CRYSTO and CRYSTC conformations in water and that the peptide permeates the membrane in the CRYSTC conformation (Figure 10A).<sup>10</sup> In line with this, a conformational state that corresponds to CRYSTC was observed in water in the simulations. However, W4 was the energetically least favorable conformational ensemble (within the accuracy of the force field) (Table 4), and the relatively high interconversion rates to other metastable sets indicate that W4 may not be very long-lived. On the other hand, set W1 was found to be energetically



**Figure 9.** Metastable sets in water: W1, W2, W3, W4, and W5. (A) Schematic representation of the MSM. The colors of the dividing lines correspond to the implied time scale plots in the right panel of Figure 7. The structures shown are the centroids of the clusters corresponding to the two most abundant microstates in W1, W2, W3, and W4 and the nine most abundant microstates in W5 (generated using PyMOL<sup>45</sup>). The percentages indicate the relative populations obtained from the MSM. (B) Microstate distributions. (C) Patterns of backbone–backbone H-bonds.

more stable and more long-lived. At the same time, the corresponding sets in chloroform, C1 and C2, are energetically similar. These observations combined suggest that the

conformational state W1/C2 may be an alternative route for membrane permeation of CsA. On the basis of these findings, we propose to generalize the current hypothesis for passive

**A: Current hypothesis****B: Generalized hypothesis**

**Figure 10.** Schematic representation of the hypotheses concerning the origin of passive membrane permeability for peptides and peptidomimetics. (A) Current leading hypothesis: permeation occurs in a closed conformation similar to that observed in X-ray structures crystallized from apolar solvents, which is assumed to be significantly populated in the aqueous environment and in fast interconversion with open conformation(s).<sup>10</sup> (B) Generalized hypothesis proposed in this project: permeation is facilitated by congruent conformation(s), which are significantly populated in both the polar and the apolar environments; the closed conformation can be one of the congruent conformations.

diffusion of CsA and potentially cyclic peptides in general: membrane permeation could thus be facilitated whenever the peptide is able to adopt a “congruent” conformation in water, of which multiple may exist (Figure 10B). In the case of CsA, the CRYSTC conformation is one of two congruent conformations. The degree of permeability would then be correlated to the relative populations of these congruent conformations in water.

## CONCLUSIONS

The conformational landscapes of CsA in chloroform (as a substitute for the membrane interior) and in water have been investigated using multimicrosecond MD simulations and characterized by MSMs. Discrete microstates based on the backbone torsional angles were obtained and further kinetically clustered into three metastable sets in chloroform (C1, C2, and C3) and five metastable sets in water (W1, W2, W3, W4, and W5). The exhaustive sampling of CsA in chloroform and water showed complex conformational landscapes with significant similarities. The MSM of CsA in chloroform agrees qualitatively well with the available data from NMR measurements. The landscapes and metastable sets in the two solvents showed significant overlap. Two pairs of closely related metastable sets (termed “congruent” sets), C1/W4 and C2/W1, were found. The pair C1/W4 corresponds to the closed conformation observed in the small-molecule crystal structure of CsA (CRYSTC). The pair C2/W1 is described by a half-opened conformation with an intramolecular H-bond between Val-5 and MeLeu-10. C2 was the most long-lived set in chloroform and may correspond to the minor conformation observed in

NMR measurements. C2 and C1 were found to be energetically similar (within the accuracy of the force field). The most long-lived set in water was W3, which corresponds to the open conformation with no intramolecular H-bonds as it is observed in the crystal structure of the cyclophilin–CsA complex (CRYSTO). The existence of congruent conformational states in both solvents could explain the good membrane permeability of CsA. It has been hypothesized previously that the transition through the membrane occurs in the CRYSTC conformation, and the pair C1/W4 was indeed observed in the simulations. In addition, the pair C2/W1 may be an alternative conformational state for passive diffusion, especially as W1 was found to be more stable than W4 (within the accuracy of the force field). We have therefore proposed a generalized hypothesis for passive permeation in which congruent conformational states facilitate the transition between the polar and apolar environments, of which multiple may exist.

Our observations combined suggest that it is not sufficient to look at the apolar low-energy conformation alone in order to rationalize the membrane permeation of cyclic peptides. A workflow as proposed here could be used to identify the existence and accessibility of congruent conformational states for new peptides and compare them with those of known peptides. In the future, we want to apply the workflow to investigate and rationalize the differences in permeability of cyclosporin derivatives as well as other cyclic peptide families.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acs.jcim.6b00251](https://doi.org/10.1021/acs.jcim.6b00251).

Topology of CsA in the GROMOS 54A7 force field (ZIP)

Experimental procedures, Tables S1–S6, and Figures S1–S13 (PDF)

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [sriniker@ethz.ch](mailto:sriniker@ethz.ch).

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

J.W. thanks Jožica Dolenc for help with the analysis. J.W. and S.R. gratefully acknowledge financial support by the Swiss National Science Foundation (Grant 200021-159713) and by ETH Zürich (ETH-08 15-1). B.G.K. gratefully acknowledges financial support by the Fonds der Chemischen Industrie.

## REFERENCES

- (1) Driggers, E. M.; Hale, S. P.; Lee, J.; Terrett, N. K. The Exploration of Macrocycles for Drug Discovery - An Underexploited Structural Class. *Nat. Rev. Drug Discovery* **2008**, *7*, 608–624.
- (2) Horne, W. S. Peptide and Peptoid Foldamers in Medicinal Chemistry. *Expert Opin. Drug Discovery* **2011**, *6*, 1247–1262.
- (3) Marsault, E.; Peterson, M. L. Macrocycles Are Great Cycles: Applications, Opportunities and Challenges of Synthetic Macrocycles in Drug Discovery. *J. Med. Chem.* **2011**, *54*, 1961–2004.
- (4) Mallinson, J.; Collins, I. Macrocycles in New Drug Discovery. *Future Med. Chem.* **2012**, *4*, 1409–1438.
- (5) Villar, E. A.; Beglov, D.; Chennamdhavuni, S.; Porco, J. A., Jr.; Kozakov, D.; Vajda, S.; Whitty, A. How Proteins Bind Macrocycles. *Nat. Chem. Biol.* **2014**, *10*, 723–731.

- (6) Yudin, A. K. Macrocycles: Lessons from the Distant Past, Recent Developments, and Future Directions. *Chem. Sci.* **2015**, *6*, 30–49.
- (7) Doak, B. C.; Zheng, J.; Dobritzsch, D.; Kihlberg, J. How Beyond Rule of 5 Drugs and Clinical Candidates Bind to Their Targets. *J. Med. Chem.* **2016**, *59*, 2312–2327.
- (8) Amidon, G. L.; Lee, H. J. Absorption of Peptide and Peptidomimetic Drugs. *Annu. Rev. Pharmacol. Toxicol.* **1994**, *34*, 321–341.
- (9) Rezai, T.; Yu, B.; Millhauser, G. L.; Jacobson, M. P.; Lokey, R. S. Testing the Conformational Hypothesis of Passive Membrane Permeability using Synthetic Cyclic Peptide Diastereomers. *J. Am. Chem. Soc.* **2006**, *128*, 2510–2511.
- (10) Rezai, T.; Bock, J. E.; Zhou, M. V.; Kalyanaraman, C.; Lokey, R. S.; Jacobson, M. P. Conformational Flexibility, Internal Hydrogen Bonding, and Passive Membrane Permeability: Successful In Silico Prediction of the Relative Permeabilities of Cyclic Peptides. *J. Am. Chem. Soc.* **2006**, *128*, 14073–14080.
- (11) Biron, E.; Chatterjee, J.; Ovadia, O.; Langenegger, D.; Brueggen, J.; Hoyer, D.; Schmid, H. A.; Jelinek, R.; Gilon, C.; Hoffman, A.; Kessler, H. Improving Oral Bioavailability of Peptides by Multiple N-Methylation: Somatostatin Analogs. *Angew. Chem., Int. Ed.* **2008**, *47*, 2595–2599.
- (12) White, T.; et al. On-Resin N-Methylation of Cyclic Peptides for Discovery of Orally Bioavailable Scaffolds. *Nat. Chem. Biol.* **2011**, *7*, 810–818.
- (13) Rüegger, A.; Kuhn, M.; Lichti, H.; Loosli, H.-R.; Huguenin, R.; Quiquerez, C.; von Wartburg, A. Cyclosporin A, ein immunsuppressiv wirksamer Peptidmetabolit aus *Trichoderma polysporum*. *Helv. Chim. Acta* **1976**, *59*, 1075–1092.
- (14) Wenger, R. M. Synthesis of Cyclosporine and Analogues: Structural Requirements for Immunosuppressive Activity. *Angew. Chem., Int. Ed. Engl.* **1985**, *24*, 77–85.
- (15) Hiestand, P. C.; Gubler, H. U. *Cyclosporins: Immunopharmacologic Properties of Natural Cyclosporins*; Springer: Berlin, 1988; pp 487–502.
- (16) Liu, J.; Farmer, J. D., Jr.; Lane, W. S.; Friedman, J.; Weissman, I.; Schreiber, S. L. Calcineurin is a Common Target of Cyclophilin-Cyclosporin A and FKBP-FK506 Complexes. *Cell* **1991**, *66*, 807–815.
- (17) Liu, J.; Albers, M. W.; Wandless, T. J.; Luan, S.; Alberg, D. G.; Belshaw, P. J.; Cohen, P.; MacKintosh, C.; Klee, C. B.; Schreiber, S. L. Inhibition of T-Cell Signaling by Immunophilin-Ligand Complexes Correlates with Loss of Calcineurin Phosphatase Activity. *Biochemistry* **1992**, *31*, 3896–3901.
- (18) Lipinski, C. A. Drug-Like Properties and the Causes of Poor Solubility and Poor Permeability. *J. Pharmacol. Toxicol. Methods* **2000**, *44*, 235–249.
- (19) Kessler, H.; Oschkinat, H.; Loosli, H. R. *Two-Dimensional NMR Spectroscopy*; Verlag Chemie: Weinheim, Germany, 1987; p 259.
- (20) Kessler, H.; Kock, M.; Wein, T.; Gehrke, M. Reinvestigation of the Conformation of Cyclosporin A in Chloroform. *Helv. Chim. Acta* **1990**, *73*, 1818–1831.
- (21) Efimov, S. V.; Karataeva, F. K.; Aganov, A. V.; Berger, S.; Klochkov, V. V. Spatial Structure of Cyclosporin A and Insight into its Flexibility. *J. Mol. Struct.* **2013**, *1036*, 298–304.
- (22) O'Donohue, M. F.; Burgess, A. W.; Walkinshaw, M. D.; Treutlein, H. R. Modeling Conformational Changes in Cyclosporin A. *Protein Sci.* **1995**, *4*, 2191–2202.
- (23) Loosli, H.-R.; Kessler, H.; Oschkinat, H.; Weber, H.-P.; Petcher, T. J.; Widmer, A. Peptide Conformations. Part 31. The Conformation of Cyclosporin A in the Crystal and in Solution. *Helv. Chim. Acta* **1985**, *68*, 682–704.
- (24) Lautz, J.; Kessler, H.; Kaptein, R.; van Gunsteren, W. F. Molecular Dynamics Simulations of Cyclosporin A: The Crystal Structure and Dynamic Modelling of a Structure in Apolar Solution Based on NMR Data. *J. Comput.-Aided Mol. Des.* **1987**, *1*, 219–241.
- (25) Lautz, J. K.; Kessler, H.; van Gunsteren, W. F.; Weber, H. P.; Wenger, R. M. On the Dependence of Molecular Conformation on the Type of Solvent Environment: A Molecular Dynamics study of Cyclosporin A. *Biopolymers* **1990**, *29*, 1669–1687.
- (26) Beck, J. G.; Chatterjee, J.; Laufer, B.; Kiran, M. U.; Frank, A. O.; Neubauer, S.; Ovadia, O.; Greenberg, S.; Gilon, C.; Hoffman, A.; Kessler, H. Intestinal Permeability of Cyclic Peptides: Common Key Backbone Motifs Identified. *J. Am. Chem. Soc.* **2012**, *134*, 12125–12133.
- (27) Ko, S. Y.; Dalvit, C. Conformation of Cyclosporin A in Polar Solvents. *Int. J. Pept. Protein Res.* **1992**, *40*, 380–382.
- (28) Shaw, R.; Mantsch, H.; Chowdhry, B. Solvent Influence on the Conformation of Cyclosporin. An FT-IR Study. *Can. J. Chem.* **1993**, *71*, 1334–1339.
- (29) Fesik, S. W.; Gampe, R. T., Jr.; Holzman, T. F.; Egan, D. A.; Edalji, R.; Luly, J. R.; Simmer, R.; Helfrich, R.; Kishore, V.; Rich, D. H. Isotope-Edited NMR of Cyclosporin A Bound to Cyclophilin: Evidence for a Trans 9,10 Amide Bond. *Science* **1990**, *250*, 1406–1409.
- (30) Weber, C.; Wider, G.; von Freyberg, B.; Traber, R.; Braun, W.; Widmer, H.; Wuethrich, K. NMR Structure of Cyclosporin A Bound to Cyclophilin in Aqueous Solution. *Biochemistry* **1991**, *30*, 6563–6574.
- (31) Fesik, S. W.; Neri, P.; Meadows, R.; Olejniczak, E. T.; Gemmecker, G. A Model of the Cyclophilin/Cyclosporin A (CSA) Complex from NMR and X-Ray Data Suggests that CSA Binds as a Transition-State Analog. *J. Am. Chem. Soc.* **1992**, *114*, 3165–3166.
- (32) Kofron, J. L.; Kuzmic, P.; Kishore, V.; Gemmecker, G.; Fesik, S. W.; Rich, D. H. Lithium Chloride Perturbation of Cis-Trans Peptide Bond Equilibria: Effect on Conformational Equilibria in Cyclosporin A and on Time-Dependent Inhibition of Cyclophilin. *J. Am. Chem. Soc.* **1992**, *114*, 2670–2675.
- (33) Schütte, C.; Fischer, A.; Huisenga, W.; Deufelhard, P. A Direct Approach to Conformational Dynamics Based on Hybrid Monte Carlo. *J. Comput. Phys.* **1999**, *151*, 146–168.
- (34) Swope, W. C.; Pitera, J. W.; Suits, F. Describing Protein Folding Kinetics by Molecular Dynamics Simulations. 1. Theory. *J. Phys. Chem. B* **2004**, *108*, 6571–6581.
- (35) Keller, B.; Hünenberger, P.; van Gunsteren, W. F. An Analysis of the Validity of Markov State Models for Emulating the Dynamics of Classical Molecular Systems and Ensembles. *J. Chem. Theory Comput.* **2011**, *7*, 1032–1044.
- (36) Prinz, J.-H.; Held, M.; Smith, J. C.; Noé, F. Efficient Computation, Sensitivity, and Error Analysis of Committor Probabilities for Complex Dynamical Processes. *Multiscale Model. Simul.* **2011**, *9*, 545–567.
- (37) Voelz, V. A.; Bowman, G. R.; Beauchamp, K.; Pande, V. S. Molecular Simulation of Ab Initio Protein Folding for a Millisecond Folder NTL9 (1–39). *J. Am. Chem. Soc.* **2010**, *132*, 1526–1528.
- (38) Buch, I.; Giorgino, T.; De Fabritiis, G. Complete Reconstruction of an Enzyme-Inhibitor Binding Process by Molecular Dynamics Simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 10184–10189.
- (39) Qiao, Q.; Bowman, G. R.; Huang, X. Dynamics of an Intrinsically Disordered Protein Reveal Metastable Conformations that Potentially Seed Aggregation. *J. Am. Chem. Soc.* **2013**, *135*, 16092–16101.
- (40) Bowman, G. R.; Geissler, P. L. Equilibrium Fluctuations of a Single Folded Protein Reveal a Multitude of Potential Cryptic Allosteric Sites. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 11681–11686.
- (41) Malmstrom, R. D.; Kornev, A. P.; Taylor, S. S.; Amaro, R. E. Allostery Through the Computational Microscope: cAMP Activation of a Canonical Signalling Domain. *Nat. Commun.* **2015**, *6*, 7588.
- (42) Noé, F.; Doose, S.; Daidone, I.; Löllmann, M.; Sauer, M.; Chodera, J. D.; Smith, J. C. Dynamical Fingerprints for Probing Individual Relaxation Processes in Biomolecular Dynamics with Simulations and Kinetic Experiments. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 4822–4827.
- (43) Keller, B.; Prinz, J.-H.; Noé, F. Markov Models and Dynamical Fingerprints: Unraveling the Complexity of Molecular Kinetics. *Chem. Phys.* **2012**, *396*, 92–107.
- (44) Kajitani, K.; Fujihashi, M.; Kobayashi, Y.; Shimizu, S.; Tsujimoto, Y.; Miki, K. Crystal Structure of Human Cyclophilin D in Complex with Its Inhibitor, Cyclosporin A at 0.96-Å Resolution. *Proteins: Struct., Funct., Bioinf.* **2008**, *70*, 1635–1639.

- (45) *The PyMOL Molecular Graphics System*, version 1.8; Schrödinger, LLC: New York, 2015.
- (46) Herrmann, T.; Güntert, P.; Wüthrich, K. Protein NMR Structure Determination with Automated NOE-Identification in the NOESY Spectra Using the New Software ATNOS. *J. Biomol. NMR* **2002**, *24*, 171–189.
- (47) Herrmann, T.; Güntert, P.; Wüthrich, K. Protein NMR Structure Determination with Automated NOE Assignment using the New Software CANDID and the Torsion Angle Dynamics Algorithm DYANA. *J. Mol. Biol.* **2002**, *319*, 209–227.
- (48) Güntert, P. Automated NMR Structure Calculation with CYANA. *Methods Mol. Biol.* **2004**, *278*, 353–378.
- (49) Wagner, G.; Pardi, A.; Wuethrich, K. Hydrogen Bond Length and Proton NMR Chemical Shifts in Proteins. *J. Am. Chem. Soc.* **1983**, *105*, 5948–5949.
- (50) Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. Definition and Testing of the GROMOS Force-Field Version 54A7 and 54B7. *Eur. Biophys. J.* **2011**, *40*, 843–856.
- (51) Eichenberger, A. P.; Allison, J. R.; Dolenc, J.; Geerke, D. P.; Horta, B. A. C.; Meier, K.; Oostenbrink, C.; Schmid, N.; Steiner, D.; Wang, D.; van Gunsteren, W. F. GROMOS++ Software for the Analysis of Biomolecular Simulation Trajectories. *J. Chem. Theory Comput.* **2011**, *7*, 3379–3390.
- (52) Scott, W. R. P.; Huenenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krueger, P.; van Gunsteren, W. F. The GROMOS Biomolecular Simulation Program Package. *J. Phys. Chem. A* **1999**, *103*, 3596–3607.
- (53) Hockney, R. W. Potential Calculation and Some Applications. *Methods Comput. Phys.* **1970**, *9*, 135–211.
- (54) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. Interaction Models for Water in Relation to Protein Hydration. *Struct. Bonding (Berlin)* **1981**, *14*, 331–342.
- (55) Tironi, I. G.; van Gunsteren, W. F. A Molecular Dynamics Simulation Study of Chloroform. *Mol. Phys.* **1994**, *83*, 381–403.
- (56) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (57) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. A Generalized Reaction Field Method for Molecular Dynamics Simulations. *J. Chem. Phys.* **1995**, *102*, 5451–5459.
- (58) Heinz, T. N.; van Gunsteren, W. F.; Huenenberger, P. H. Comparison of Four Methods to Compute the Dielectric Permittivity of Liquids from Molecular Dynamics Simulations. *J. Chem. Phys.* **2001**, *115*, 1125–1136.
- (59) Tironi, I. G.; van Gunsteren, W. F. A Molecular Dynamics Simulation Study of Chloroform. *Mol. Phys.* **1994**, *83*, 381–403.
- (60) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.
- (61) Landrum, G. A. RDKit: Open-Source Cheminformatics Toolkit. <http://www.rdkit.org>.
- (62) Deuflhard, P.; Weber, M. Robust Perron Cluster Analysis in Conformation Dynamics. *Linear Algebra Appl.* **2005**, *398*, 161–184.
- (63) Chodera, J. D.; Singhal, N.; Pande, V. S.; Dill, K. A.; Swope, W. C. Automatic Discovery of Metastable States for the Construction of Markov Models of Macromolecular Conformational Dynamics. *J. Chem. Phys.* **2007**, *126*, 155101.
- (64) Keller, B.; Daura, X.; van Gunsteren, W. F. Comparing Geometric and Kinetic Cluster Algorithms for Molecular Simulation Data. *J. Chem. Phys.* **2010**, *132*, 074110.
- (65) Senne, M.; Trendelkamp-Schroer, B.; Mey, A. S.; Schütte, C.; Noé, F. EMMA: A Software Package for Markov Model Building and Analysis. *J. Chem. Theory Comput.* **2012**, *8*, 2223–2238.
- (66) Weber, M. *Improved Perron Cluster Analysis*; ZIB: Berlin, 2003.
- (67) Noé, F.; Schütte, C.; Vanden-Eijnden, E.; Reich, L.; Weikl, T. R. Constructing the Equilibrium Ensemble of Folding Pathways from Short Off-Equilibrium Simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 19011–19016.
- (68) Guentert, P.; Buchner, L. Combined Automated NOE Assignment and Structure Calculation with CYANA. *J. Biomol. NMR* **2015**, *62*, 453–471.
- (69) Vitalini, F.; Mey, A. S. J. S.; Noé, F.; Keller, B. G. Dynamic Properties of Force Fields. *J. Chem. Phys.* **2015**, *142*, 084101.
- (70) Sarich, M.; Noé, F.; Schütte, C. On the Approximation Quality of Markov State Models. *Multiscale Model. Simul.* **2010**, *8*, 1154–1177.
- (71) Razavi, A. M.; Wuest, W. M.; Voelz, V. A. Computational Screening and Selection of Cyclic Peptide Hairpin Mimetics by Molecular Simulation and Kinetic Network Models. *J. Chem. Inf. Model.* **2014**, *54*, 1425–1432.