JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

pubs.acs.org/JACS

Unfolding the Conformational Behavior of Peptide Dendrimers: Insights from Molecular Dynamics Simulations

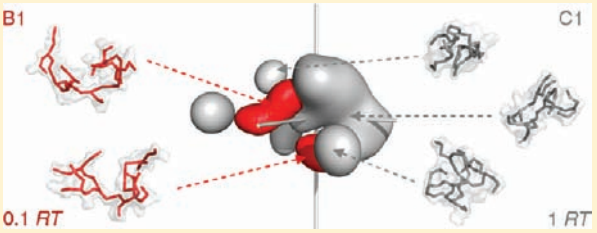
Luís C. S. Filipe,[†] Miguel Machuqueiro,[†,‡] and António M. Baptista*[,†]

[†]Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Av. da República, EAN, 2780-157 Oeiras, Portugal

[‡]Centro de Química e Bioquímica, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, C5, 1749-016 Lisboa, Portugal

Supporting Information

ABSTRACT: We present here the first comprehensive structural characterization of peptide dendrimers using molecular simulation methods. Multiple long molecular dynamics simulations are used to extensively sample the conformational preferences of five third-generation peptide dendrimers, including some known to bind aquacobalamine. We start by analyzing the compactness of the conformations thus sampled using their radius of gyration profiles. A more detailed analysis is then performed using dissimilarity measures, principal coordinate analysis, and free energy landscapes, with the aim of identifying groups of similar conformations. The results point to a high conformational flexibility of these molecules, with no clear "folded state", although two markedly distinct behaviors were found: one of the dendrimers displayed mostly compact conformations clustered into distinct basins (rough landscape), while the remaining dendrimers displayed mainly noncompact conformations with no significant clustering (downhill landscape). This study brings new insight into the conformational behavior of peptide dendrimers and may provide better routes for their functional design. In particular, we propose a yet unsynthesized peptide dendrimer that might exhibit enhanced ability to coordinate aquocobalamin.



1. INTRODUCTION

Dendrimers are a family of highly branched compounds that share a common layout where wedges emerge radially from a core by means of a regular branching pattern.[1−7] These structures are characterized by a combination of high end-group functionality and a precisely defined chemical composition containing three topologically distinct regions (core, branches, and periphery), where each region can manifest features that are modulated by the dendrimer as a whole.[1,3,7−11] Because of their particular architecture, these molecules have been used to combine miscellaneous building blocks, such as branched compounds,[1,10] proteinogenic and nonproteinogenic amino acids,[8,12,13] carbohydrates,[14−16] among others.[6] Nowadays, dendrimers constitute a class of molecules presenting a wide range of sizes, functionalities, and applications.[4−10,16−20]

Peptide dendrimers are a specific kind of dendrimers formed by alternating proteinogenic amino acids with branching diamino acids.[13,21,22] The peptide dendrimers obtained in this way have been continuously and systematically studied, providing models for different natural systems. Some examples include catalytic peptide dendrimers with esterolytic or aldolytic activity,[21−23] multivalent lectin binding dendrimers,[14,24] drug delivery compounds,[25] and peptide dendrimers for vitamin B_{12} transport.[26,27] Moreover, combinatorial synthesis methods allow for the sequential design of peptide dendrimers mimicking specific aspects of biological functions.[8,28,29] Peptide dendrimers have already been reported to exhibit biocompatibility,[14,16] enhanced resistance to proteolysis,[30] and acquisition of secondary structure motifs characteristic of folded proteins.[31]

It has been suggested that most peptide dendrimers are topologically constrained to adopt a globular shape.[8,22,23,28] Thus, multibranched molecules with protein-like structures could be shaped, and features such as molecular recognition would arise from cooperative interactions among the different amino acid constituents, in a manner similar to that of natural proteins.[8,22,23] Unfortunately, most studies of peptide dendrimers lack structural information at the molecular level. It has been proposed that peptide dendrimers exhibit relatively loose conformations similar to molten globule proteins,[23] which might explain the reason why it has not been possible to obtain detailed structural information using the most common experimental techniques.[32] Several experimental studies, such as diffusion NMR experiments on catalytic esterase peptide dendrimers, indicate that these molecules adopt roughly spherical structures,[8,23,32] while other studies, such as cobalamin-binding experiments, suggest the existence of more disordered states.[26,27,31]

Reliable information at the molecular level is therefore vital for understanding the functional properties of peptide dendrimers, and their structure in solution. The use of molecular mechanics/dynamics (MM/MD) methods, which are very well suited to study structural aspects of biomolecules, including large

Received: December 7, 2010
Published: March 15, 2011

conformational transitions like protein folding,[33−35] might be a valuable tool to investigate these issues, as these methods allow us to gather structural information about peptide dendrimers that would otherwise be inaccessible. This approach has been performed recently and for the first time to investigate catalytic peptide dendrimers with proven esterolytic activity, providing the first insights into these molecules structural behavior in solution.[32] In that work, Javor et al. adopted a simulated annealing procedure using multiple short MM/MD runs, finding that the dendrimers tended to adopt compact, roughly spherical structures.

Herein, we characterize five third-generation peptide dendrimers using multiple long MM/MD simulations and analyze their conformational details and structural preferences in solution. The purpose is to characterize the whole conformational space of these dendrimers, rather than simply finding their low-energy conformers.

The dendrimers investigated have an identical topology (Figure 1) but are composed of different amino acid residues (Table 1). Thus, we expect to analyze not only the structural tendencies inherent to each of these particular dendrimers, but also reach some conclusions concerning the relevance of topology on the preferentially acquired conformations. This approach is of interest since some authors advocate the idea that the acquisition of structure by peptide dendrimers is mainly topology driven.[8,23,28]

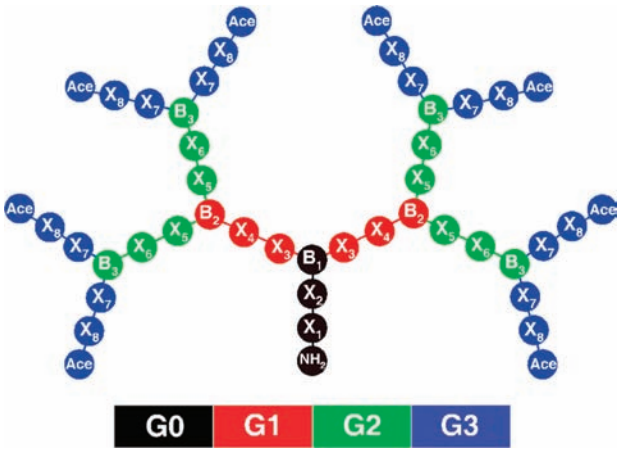


Figure 1. Schematic topology of the peptide dendrimers investigated. The X_n and B_n residues of each dendrimer are indicated in Table 1. Dendrimer generations (G0−G3) and their constituents are evidenced using different colors. The N- and C-terminal caps are, respectively, the acetyl group (Ace) and the amine group (NH_2).

The peptide dendrimers B1, B1H, and C1 have been synthesized and characterized using different experimental techniques,[26] having demonstrated ability to mimic the cobalt binding in vitamin B_{12} dependent enzymes or transport proteins. Their coordination with cobalamin is thought to be mediated by cysteine or histidine residues,[26] and their experimental cobalamin-binding affinities are also presented in Table 1. The intriguing question surrounding these three dendrimers is that the one presenting the higher number of potentially cobalamin-coordinating residues (C1, which presents two histidines) is not the one with the higher binding constant. Furthermore, the exchange of a single cysteine residue in B1 by a histidine one (producing B1H), leads to a marked decrease of the binding constant.[26] These counterintuitive results might be explained by the preferential conformations adopted by each of these dendrimers in solution.

The peptide dendrimers B1HH and B1HHH are both variants obtained from B1 by replacing with histidine the two 4-aminomethyl(benzoic) acid (Amb) residues (positions X_4 in Figure 1) and, in B1HHH, also the cysteine core residue (position X_2 in Figure 1). They have been included in the present study in order to examine the effect of incorporating two additional potentially cobalamin-coordinating residues into the overall dendritic topology under investigation. Because of the similarities in the amino acid residues composition between B1, B1H, B1HH, and B1HHH, they will be collectively designated as "B1-family".

To the best of our knowledge, these are the first long molecular dynamics simulations of peptide dendrimers and the first attempt to perform a comprehensive structural characterization of these molecules. We analyze several properties by resorting to histogram analysis and energy landscapes, which play a crucial role in rationalizing the conformational behavior of peptides and proteins.[33−40]

The present study can provide new insights into the structure−function relation in peptide dendrimers and, together with the available experimental results, contribute to an integrative understanding of their structure in solution and the development of novel applications.

2. COMPUTATIONAL METHODS

2.1. Dendrimer Building. We have used PyMOL[41] (version 0.99rc6) to obtain an initial set of 3D coordinates for each peptide dendrimer introduced in Table 1. All the dendrimers were built and simulated using the sequence with which they were synthesized (or would be in the case of B1HH and B1HHH),[26] namely, by attaching an amine group (NH_2) to

Table 1. Residue Composition and Cobalamin Binding Constant for Each Peptide Dendrimer

	residues at each position[a]											
dendrimer	X_1	X_2	B_1	X_3	X_4	B_2	X_5	X_6	B_3	X_7	X_8	binding constant,[b] 10^6 M^{-1}
B1	Asp	Cys	Dap	Tyr	Amb	Lys	Ala	Glu	Dap	Ser	Glu	5.0 ± 0.800
B1H	Asp	His	Dap	Tyr	Amb	Lys	Ala	Glu	Dap	Ser	Glu	0.083 ± 0.011
B1HH	Asp	Cys	Dap	Tyr	His	Lys	Ala	Glu	Dap	Ser	Glu	n.a.[c]
B1HHH	Asp	His	Dap	Tyr	His	Lys	Ala	Glu	Dap	Ser	Glu	n.a.[c]
C1	Ala	Arg	Dap	Thr	His	Dap	Tyr	Glu	Dap	Gly	Ser	0.022 ± 0.002

[a] Standard three-letter abbreviations are used for proteinogenic amino acids; Dap for s-2,3-diaminopropionic acid and Amb for 4-aminomethyl-(benzoic) acid. [b] Taken from ref 26. [c] Data not available.

dx.doi.org/10.1021/ja111001v | J. Am. Chem. Soc. 2011, 133, 5042–5052

the core C-terminus, and acetyl (Ace) groups to the peripherical N-termini (see Figure 1).

The starting structures obtained with PyMOL showed configurations resembling a random coil. To ensure that no compact conformational states were favored by these starting conformations, all structures were subject to energy minimization and initiation procedures previous to the production stage of the simulation (see section 2.2).

To facilitate the comparison among conformations of different peptide dendrimers we employ a color code to represent each of them along the article. Accordingly, B1 structures are colored in red, B1H in green, B1HH in blue, B1HHH in yellow, and C1 structures in gray.

2.2. MM/MD Settings. The GROMACS package,[42,43] version 4.0.2, and the GROMOS 53A6 force field[44] were employed to perform MD simulations on the five different peptide dendrimers. Most of the amino acid blocks considered for the topology of peptide dendrimers were already available in the GROMOS 53A6 set. Nevertheless, new topology blocks for the 4-aminomethyl(benzoic) acid (Amb), the s-2,3-diaminopropanoic acid (Dap) and the branching lysine residues, were constructed assuming the transferability of the force field (see Supporting Information). The charges assigned to each tritable residue were the ones typically present at pH 7. The suitability of the protonation states chosen was confirmed *a posteriori* by pK_a calculations over the conformation ensembles obtained from the simulations (data not shown).

The nonbonded interactions were treated with a twin-range cutoff of 8/14 Å and neighbor lists updated every 10 fs. The reaction-field method,[45] with a relative dielectric constant of 54.0,[46] was used for the long-range electrostatic interactions.

Molecular dynamics simulations were performed by integrating the equations of motion using the Verlet leapfrog algorithm with a time step of 2 fs, and the system coordinates (snapshots) were saved every 10 ps for further analysis. The LINCS algorithm[47] was employed to keep all bonds at their equilibrium values and the SETTLE algorithm[48] was used to maintain water molecules rigid.

Solvent and solute were separately coupled to temperature baths at 298.15 K, with Berendsen coupling[49] and a relaxation time of 0.1 ps. A Berendsen isotropic pressure coupling[49] was used at 1 bar, with a relaxation time of 0.5 ps and an isothermal compressibility of 4.5×10^{-5} bar^{-1}.

All simulations were done with explicit solvent, using 13619 (B1), 13608 (B1H), 9446 (B1HH), 9441 (B1HHH), and 9442 (C1) single point charge (SPC)[50] water molecules in rhombic dodecahedral unit-cells, while applying periodic boundary conditions. Suitable minimum distances between the peptide dendrimers and the end of the unit-cells were used to ensure that periodic images did not interact. The final systems contained about 28 to 40 $\times 10^3$ atoms.

All systems were energy minimized to remove excessive strain. This was done using the steepest descent algorithm, consisting of 12 000 steps with all dendrimer heavy atoms position restrained, followed by 15 000 steps without position restraints.

The initiation procedure consisted of seven different simulations. First a 20 ps simulation was performed with all dendrimer atoms restrained at a temperature equal to 298.15 K and assigning initial random velocities to the structure, thus generating the different replicates. This first step was followed by another simulation for 30 ps restraining only the α carbon atoms. A third simulation was performed for 50 ps without position restraints and heating the system up to 400 K. After that and to ensure the formation of extended conformations, the partial charges of each atom were changed from their reference values to a value of +0.1 *e* and heating continued during 100 ps up to 500 K. This brief change in all atomic electric charges promotes the repulsion among all dendrimer atoms, originating a generic initial structure that corresponds to the most "stretched" conformation of each dendrimer. This is, in our opinion, the less biased approach because it allows the simulations to depart from states that do not represent a conformational minimum and that are, in principle, the least favored ones.

Next, a simulation was run during 50 ps decreasing the systems temperature to 400 K, while keeping the positive charges on the atoms. Subsequently, all the atoms were kept positively charged throughout a 50 ps simulation, while cooling the system down to its reference temperature (298.15 K). Finally, all the charges were set to their original values and the initiation protocol was completed with a 20 ps MD run at 298.15 K with all dendrimer heavy atoms position restrained. The force constant used for the position restraints during all minimization and initiation steps was 1000 kJ mol^{-1} nm^{-2}.

Starting with these fully extended configurations, ten molecular dynamics simulations of 100 ns were performed for each of the five peptide dendrimers under study, amounting to a total of 1 μs per dendrimer. The replicates of each peptide dendrimer started from the same optimized system but with different sets of random velocities. The systems showed to be equilibrated at different time lengths (see section 3.1) and only the equilibrated trajectories were used for subsequent analysis, which was done using GROMACS[42,43] or in-house tools. This ensemble-dynamics approach, combining several replicates with long MD runs, allows us to obtain not only temporal averages but also ensemble averages. Despite this extensive sampling, we can not disregard the hypothesis that there may exist conformational events that are not accounted for, as the simulations may not have reached an ergodic stage.

2.3. Root Mean Square Deviation. The (dis)similarity among pairs of conformations was quantified using the root-mean-square deviation (rmsd) of their atomic Cartesian coordinates obtained after translational and rotational least-squares fitting. The rmsd between conformation *A* and conformation *B* of a given molecule was defined as the minimum of the function

$$\mathrm{rmsd}(A,B) = \left[\frac{1}{\sum_{i=1}^{N} \mathbf{m}_i} \sum_{i=1}^{N} \mathbf{m}_i \,\|\, \mathbf{r}_i(A) - \mathbf{r}_i(B)\|^2\right]^{1/2} \quad (1)$$

where *N* is the number of atoms in the summation, *i* is an index over these atoms, m_i is the mass of atom *i*, and $\mathbf{r}_i(A)$ and $\mathbf{r}_i(B)$ are the Cartesian coordinates of atom *i* in conformations *A* and *B*.[51] The minimum value of eq 1 is obtained by an optimal superposition of the two structures. The rmsd value quantifies how well the atoms of two molecular structures can be superimposed and fulfils the requirement of a metric,[52,53] meaning that it preserves the relative proximity between all the conformations.

Determining the rmsd value for a pair of chemically symmetrical or quasi-symmetrical molecules is a task hindered by the fact that such molecules have on their composition residues that are placed at different positions in the overall architecture but that are chemically equivalent. The peptide dendrimers studied here can be considered as chemically quasi-symmetrical

structures, as evidenced by Figure 1, because for any residue in a generation other than G0 there is at least one chemically equivalent residue. To overcome this problem we devised an automated procedure to construct all chemically equivalent permutations of a dendrimer, by iteratively swapping residue sections (pairs of spacer units) attached to the same branching units. There are seven branching residues ($1B_1 + 2B_2 + 4B_3$ in Figure 1) whose attached pair of spacer units may be in two different states (swapped or not), which amounts to a total of $2^7 = 128$ different permutations. (The Supporting Information explicitly shows the conformations thus obtained for the simpler case of a second-generation dendrimer, as well as an actual example involving a pair of conformations from the C1 simulations.) When a conformation is compared with some reference, the corresponding rmsd values are computed with all its 128 variants and the lowest one (the "true" rmsd) is selected. Because of the high computational cost of these calculations (128 times heavier than a standard rmsd calculation), we have performed them using only the conformations present at 0.2 ns intervals of the concatenated trajectories.

With the calculated rmsd values, we have built for each dendrimer the corresponding rmsd matrix, accounting for the dissimilarity between all pairs of structures. Following an approach described elsewhere,[40] we have used this matrix to determine the *central structure* that minimizes the dispersion of the rmsd values for each dendrimer.

2.4. Free Energy Landscapes. To characterize the conformational space of peptide dendrimers, we must identify the accessible energy basins and minima on the potential-energy surface. To achieve this, we have determined probability density functions, $P(\mathbf{r})$, in two different representation spaces. The first is a two-dimensional (2D) space using as coordinates the values of the radius of gyration (R_g) and rmsd to the central structure. The second uses the three-dimensional (3D) Euclidean coordinates obtained for a principal coordinate representation of the conformations (section 2.5). Both representation spaces were obtained using the concatenated trajectories at intervals of 0.2 ns. The probability density functions were estimated using a Gaussian kernel estimator,[54] while employing grids of $(0.009\ \text{Å})^3$ for the 2D energy surfaces, and $(0.05\ \text{Å})^3$ for C1 and $(0.08\ \text{Å})^3$ for the B1-family dendrimers 3D energy surfaces.

As done previously,[40] energy surfaces were computed from $P(\mathbf{r})$ according to

$$E(\mathbf{r}) = -RT \ln \frac{P(\mathbf{r})}{P_{\max}} \qquad (2)$$

where $P(\mathbf{r})$, $P_{\max}$, T, and R are, respectively, the probability density function, its maximum, the absolute temperature, and the ideal gas constant. For simplicity, we will hereafter refer to $E(\mathbf{r})$ as "energy", although it, in fact, represents a conditional free energy.[38−40]

We group the configurations expressed in the conformational spaces using as clustering condition the confinement within a common energy basin, while considering different energy cutoffs.[40]

2.5. Principal Coordinate Analysis. Creating useful representations of molecular conformation spaces is a task hindered by the high dimensionality of these spaces. Collective coordinate methods, such as *principal components analysis* or *principal coordinate analysis* (PCoorA), allow the projection of multidimensional data on low-dimensional subspaces.[55−57]

Table 2. Total Length of the Concatenated Equilibrium Simulations for Each Dendrimer

dendrimer	equilibrated simulation length (ns)[a]
B1	850.10
B1H	905.10
B1HH	905.10
B1HHH	880.10
C1	835.10

[a] This values refer to the sum of the equilibrated trajectories of 10 replicates.

PCoorA is a method that allows the mapping of objects into a system of coordinates (called principal coordinates), in such a manner that the Euclidean distances between objects in that space correspond (as well as possible) to their dissimilarities.[56−61]

We used the rmsd values between pairs of conformations as a dissimilarity measure and reduce the dimensionality of the data set to a 3D space.

3. RESULTS AND DISCUSSION

3.1. Equilibration and Validation. The equilibration time of the different replicates was determined by monitoring the radius of gyration (R_g),[62] which provides a rough measure of the compactness of a structure and is commonly used to study branched molecules.[63−68] More traditional properties like the potential and kinetic energies converge almost immediately. The replicates of all peptide dendrimers reach equilibrium at different times, ranging from 5 to 30 ns. By concatenating the trajectories of the different replicates, one obtains the total amount of production simulation in equilibrium conditions, whose corresponding times are presented in Table 2. For each peptide dendrimer, the subsequent analysis were performed over the corresponding group of ten equilibrated trajectories.

Although not many physical properties have been experimentally determined for the dendrimers studied here, a recent diffusion NMR spectroscopy study reported a diffusion coefficient of $1.10 \times 10^{-10}\ \text{m}^2\ \text{s}^{-1}$ for the B1 free form (without cobalamin) in aqueous solution.[27] This value can be used to assess the suitability of our approach and the use of GROMOS 53A6 force field for the simulation of these peptide dendrimers.

From the simulations, we have determined the diffusion coefficient for B1 in water by computing the mean-square displacement (MSD) and employing the Einstein relation.[69] For each of the ten replicates, we have determined the MSD of the dendrimer center of mass by considering only the last 80 ns; we discarded the first 20 ns of all replicates because that was the time the slowest replicate of B1 took to equilibrate. An average over these ten MSD curves originated an 80 ns average MSD curve (Figure 2), whose initial and final points were discarded to avoid the effects of ballistic motion and poor sampling.[70] The linear part of this average curve, ranging from 5 to 55 ns, was then subjected to a linear least-squares fit. The Einstein relation was then applied to the resulting straight line, yielding a diffusion coefficient of $2.85 \times 10^{-10}\ \text{m}^2\ \text{s}^{-1}$.

The statistical error associated with the diffusion coefficient was calculated as the difference between the diffusion coefficients obtained from fits over the two halves of the fit interval considered, in accordance with a methodology described elsewhere,[71] yielding a value of $0.43 \times 10^{-10}\ \text{m}^2\ \text{s}^{-1}$.

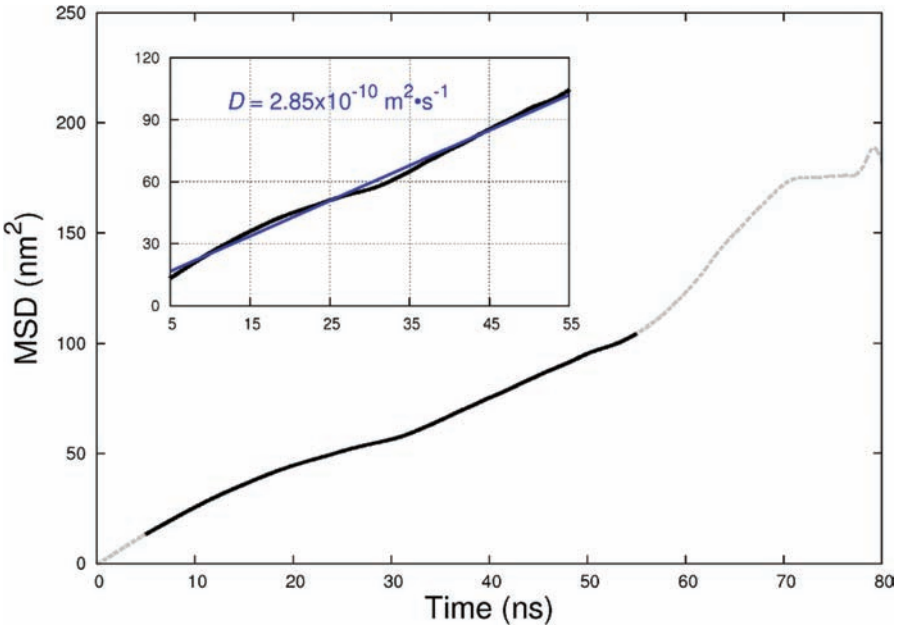


Figure 2. Average MSD curve for the B1 dendrimer. The dashed part of the curve represents the values that were neglected in the fit. The fit over the considered interval is presented in the inset plot, along with the diffusion coefficient obtained.

Considering the difficulties of estimating second-order time-dependent properties from simulations,[69] the computed diffusion coefficient ($(2.85 \pm 0.43) \times 10^{-10}$ m^2 s^{-1}) is in reasonably good agreement with the experimental value (1.10×10^{-10} m^2 s^{-1}).[27] Hence, the simulations seem to adequately reflect the experimental systems, supporting the suitability of the GROMOS 53A6 force field to model peptide dendrimers.[32]

3.2. Radius of Gyration. To investigate and characterize the existence of homogeneous conformational classes representative of each peptide dendrimer we have concatenated the equilibrated conformations sampled during the simulations and computed the R_g value for each of them. The resulting distributions of R_g values are shown in Figure 3, together with examples of structures associated with particular values.

The first important observation is that the B1-family dendrimers exhibit R_g values that vary within similar ranges (approximately from 1 to 2 nm), with most conformations comprised in the interval between 1.3 and 1.5 nm. When compared to the B1-family dendrimers, the R_g values of C1 tend to be smaller and vary within a shorter range (from 0.8 to 1.4 nm). These different distributions are reflected on the average R_g values of each dendrimer (Table 3).

Moreover, C1 presents a number of conformations with low R_g values that are inaccessible to the B1-family dendrimers, implying that the combination of amino acid residues used to construct C1 promotes the formation of more compact structures. An analogous observation can be made for the conformations with high R_g values of the B1-family, which are inaccessible to C1.

An analysis of the conformations with high R_g values revealed a significant number of noncompact conformations of peptide dendrimers (e.g., B1 with R_g = 1.943 nm and C1 with R_g = 1.427 nm). The myriad of R_g values accessible to these dendrimers evidence the huge conformational variability that seems to characterize such structures. Still, from a detailed analysis of the conformations, we have verified that the most recurrent R_g values correspond to heterogeneous structures, sharing a common overall compactness but miscellaneous conformations, as illustrated by the images of structures collected in the vicinity of the most frequent R_g value of C1 (Figure 3). Therefore, although the histogram analysis performed allows the characterization of the overall compactness trends in the different peptide dendrimers, it does not satisfactory discriminates the conformations according to their structural homogeneity.

To further investigate the conformational preferences of these molecules by means of a single structural coordinate, we have performed an extensive analysis using the same histogram approach but employing different properties. We have studied the distribution of the conformations using as structural coordinates the following measures: total number of hydrogen bonds; solvent accessible surface area; sum of the distance between the branching residues α-carbons; asphericity;[72,73] acylindricity;[72,73] and the distance between the two farthest atoms in each structure. The histograms of all the aforementioned properties evidence similar tendencies to the ones displayed by the R_g histograms, with a clear distinction between the B1-family and the C1 dendrimers (data not shown). Although these properties further support the existence of more compact structures in C1 than in the B1-family, none of them provides a satisfactory detailed discrimination among the conformations of each dendrimer.

Additionally, we have computed the values for the dihedral angles of every snapshot of the equilibrium concatenated trajectories. The allowed φ and ψ angle values thus obtained are in accordance with the ones typically exhibited by proteins[74] (consult the Supporting Information for details). Therefore, the use of a dendritic architecture along with peptidic constituents does not seem to induce the occurrence of unusual (φ,ψ) combinations.

3.3. Residues Proximity. On the basis of what is currently known about the structure of peptide dendrimers in solution and by analogy with proteins, a "folded" dendrimer should, in principle, exhibit multiple close contacts between residues that lie far from each others on the overall topology.

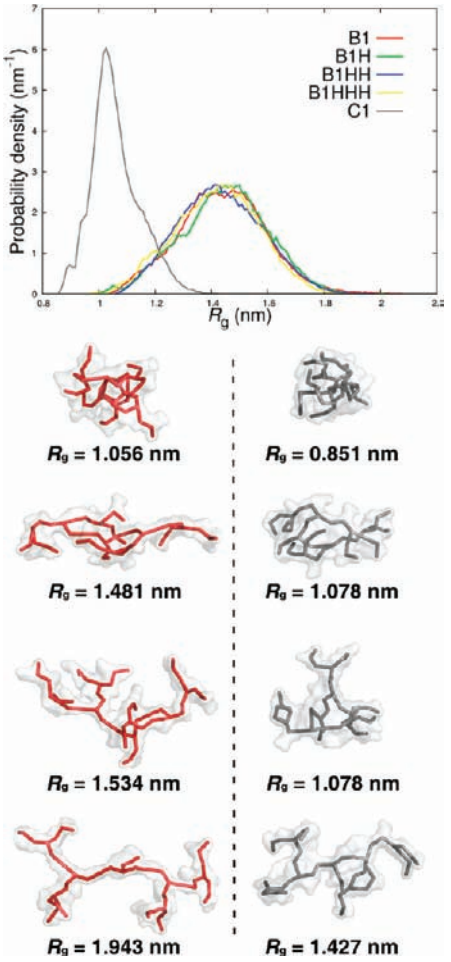


Figure 3. Radius of gyration probability density histograms for B1, B1H, B1HH, B1HHH, and C1. Some atomic configurations of B1 (red) and C1 (gray) are displayed, with the corresponding R_g values. The structures were represented considering only the α carbon atoms, and drawing lines between bonded residues. Also represented is the line connecting the nitrogen atom of the amino cap at the subsequent α carbon. Transparent atomic surfaces obtained using a 1.4 Å probe are also shown.

Table 3. Average Radius of Gyration of the Dendrimers

dendrimer	average R_g (nm)
B1	1.542
B1H	1.490
B1HH	1.502
B1HHH	1.419
C1	1.122

To evaluate the spatial proximity between all the residue pairs we employed an approach described elsewhere.[32] This analysis procedure implies plotting, for all residue pairs, the shortest atomic spatial distance between a pair versus their shortest topological distance in Figure 1, that is, the total number of residues along the shortest through-bond path connecting them.

In Figure 4, we have plotted the average shortest spatial distance for all possible pairs of residues of C1 and B1 versus the corresponding topological distances. The plots obtained for B1H, B1HH, and B1HHH are similar to the one of B1 and can be found in the Supporting Information.

Figure 4 shows that topologically distant residues tend to be spatially closer in C1 than in B1, indicating that in C1 there is the possibility of close contacts either among the different dendritic branches or among the branches and the dendrimer core. The trends displayed in the plot further support the results from the histogram analysis, highlighting once more the existence of two distinct structural behaviors, one characteristic of C1 and another of the B1-family.

3.4. 2D Energy Landscapes. Our analysis of the simulations using a single structural coordinate (section 3.2) could not identify distinct classes among the myriad of interconvertible dendrimer conformations. Therefore, we have used two structural coordinates, the radius of gyration (R_g) and the rmsd value to a central structure, in an attempt to adequately discriminate eventual conformational classes (Figure 5). The plots of B1H, B1HH, and B1HHH are very similar to those of B1 and are provided in the Supporting Information.

The scatter-plots presented at the top of Figure 5 suggest that there is no obvious relationship between the rmsd and R_g values, with similar values of rmsd accounting for very distinct conformations. Thus, the rmsd value obtained using as reference the central structure is not, by itself, a very good structural coordinate to analyze peptide dendrimers.

Some significant differences are visible between the B1-family and C1 scatter plots. While for the B1-family dendrimers the populated regions of R_g and rmsd values are very similar among them (consult the Supporting Information for B1H, B1HH, and B1HHH plots), C1 occupies completely distinct regions, characterized by lower R_g and rmsd values. Moreover, and contrary to the almost horizontal point cloud in the B1-family dendrimers, the scatter-plot of C1 has some visible point agglomeration indicating the existence of three main conformational groups.

Using the density of R_g and rmsd values, we have computed the corresponding 2D free energy profile. These profiles were used to investigate the existence of conformational classes and respective energy minima. The landscapes for B1 and C1 are displayed as contour maps in the center of Figure 5 (the contour maps for B1H, B1HH, and B1HHH are provided in the Supporting Information).

As evidenced by the contour figure, the landscape of B1 has a general funnel-like form and, if we draw an analogy with the landscapes of protein folding, we may consider that the dendrimer "folding" process is essentially downhill in terms of energy. Looking at some of the structures collected from the only existing energy basin in B1 (bottom left in Figure 5), one can corroborate that they share some features: all present a mildly loose structure, and none of them presents a well-defined structural nucleus. A clear distinction between the structures is still not obvious.

The energy landscapes of B1H, B1HH, and B1HHH (Supporting Information) suggest that the landscape features of all B1-family members are almost identical, with similar conformations accounting for the lowest-energy conformations. These findings support the idea that, in terms of energetic barriers, all B1-family dendrimers behave similarly, regardless of the amino acid differences that characterize each of them. It seems that, for these particular dendrimers, a small number of amino acid changes does not induces major transformations in the energy landscapes defined using R_g and rmsd.

The C1 energy landscape is more complex than the B1-family ones, clearly evidencing three distinct conformational regions, with the majority of conformations gathered in the region leading to the

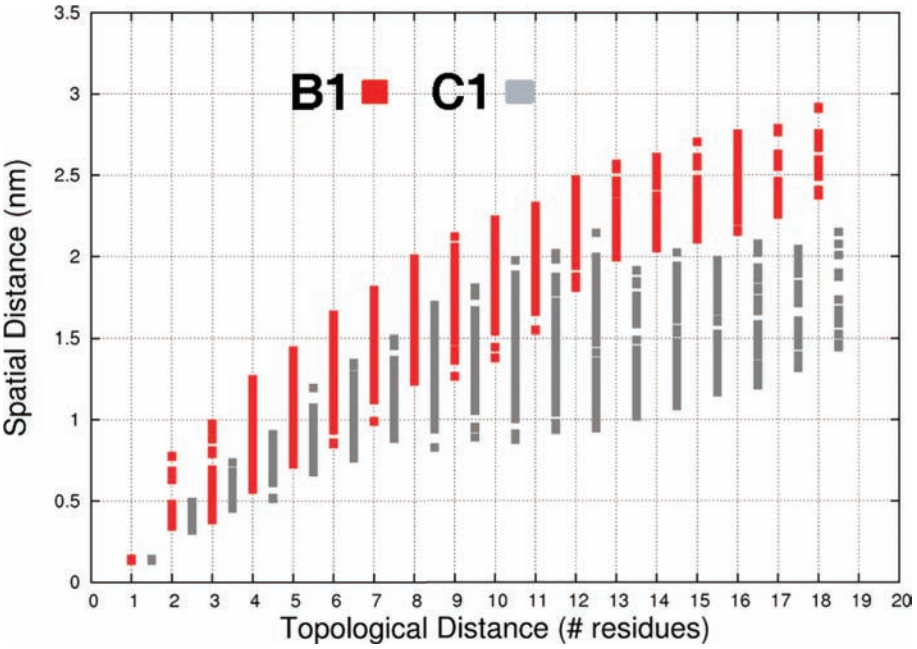


Figure 4. Shortest spatial distances of all residue pairs versus their topological distances. The spatial distance values used to construct the plots were obtained by averaging over 4251 (B1) and 4176 (C1) conformations, collected at 0.2 ns intervals of the total concatenated trajectories. To avoid the overlapping with B1 values, the C1 topological distances are shifted by +0.5 units.

global energy minimum, and where the two high-energy minima correspond to small portions of the total number of conformations.

The basin that incorporates the energy minimum (labeled as **1**) accounts for conformations that share some features but are structurally inhomogeneous. Nonetheless, their overall shape is a compact one. The basin labeled as **2** accounts for conformations with free energies of approximately 1.75*RT*, displaying compact structures with packed geometries. Finally, basin **3** accounts for conformations characterized by energy values varying between $5RT < E < 2.5RT$. Observing some of the conformations in this cluster, we conclude that they exhibit similar branching distributions. Furthermore, the energy landscape indicates that transitions between basin **3** and basin **1** would have to overcome high energy barriers.

It is noteworthy that the structures presented in the high energy basins (**2** and **3**), are usually more compact than the low energy ones. This means that the preferential configurations of C1 are not the most compact ones, but rather some mildly compact or molten-globule-like structures.

3.5. Principal Coordinate Analysis. An alternative set of free energy landscapes was computed in the 3D PCoorA spaces that better preserve the dissimilarities among all pairs of structures, as described in section 2.5. Figures 6−8 show those landscapes for the B1, B1H, and C1 dendrimers, respectively, displaying the isoenergetic surfaces for 2, 1, and 0.1 *RT*. The corresponding figures for B1HH and B1HHH are very similar to the one of B1 and can be found in the Supporting Information. Analogously to the isoenergetic lines in Figure 5, the isoenergetic surfaces delimit regions whose energy is below the indicated value, making possible to identify the landscape basins. The isoenergetic surfaces represented are those that better discriminate the existing basins.

If we analyze once more, the features of the energy landscapes making an analogy with protein folding, we can consider that the B1-family dendrimers exhibit very similar "folding" behaviors, with funnel-like landscapes and a clear downhill propensity. Furthermore, structures shown in Figures 6 and 7 (as well as those for B1HH and B1HHH in the Supporting Information) underline the similarities among the low energy conformations of the peptide dendrimers within this family. For B1, B1HH, and B1HHH, the only observable basins arise at very low energy values (approximately 0.1*RT*), corresponding to conformations that account for very small fractions of the entire conformational population.

The single exception to this downhill behavior in the B1-family is B1H (Figure 7), whose 3D landscapes reveals an additional basin around 2*RT*. This basin accounts for a small number of high energy structures (1.83%) that, despite being conformationally heterogeneous, generally display compact (or molten-globule like) conformations, with tightly packed residues.

The C1 landscape (Figure 8) differs markedly from those of the B1-family. In this dendrimer, we distinguish different basins and, as indicated by the 2D energy landscapes, we identify diverse conformational classes. Some of these basins account for substantial percentages of the total amount of conformations. Varying the isoenergetic surfaces reveals that the basins are connected in a nontrivial way, indicating that C1 has a complex and rough landscape. Furthermore, a comparison between the characteristic conformations depicted with each 3D landscape shows that, as observed with the 2D landscapes, the preferential structures populated by C1 are more compact than those favored by any of the other peptide dendrimers studied here.

Among the different analysis methodologies used in the present work, the conformational space defined using the coordinates provided by PCoorA is the one that more efficiently discriminates the different conformational preferences of peptide dendrimers, because, besides finding the features identified by the other methodologies, it also reveals new ones that were "hidden" in the complexity of the data. Nonetheless, the general topography (downhill/rough) of the energy landscapes of all peptide dendrimers seems to be captured by the 2D energy landscapes using R_g and rmsd as coordinates.

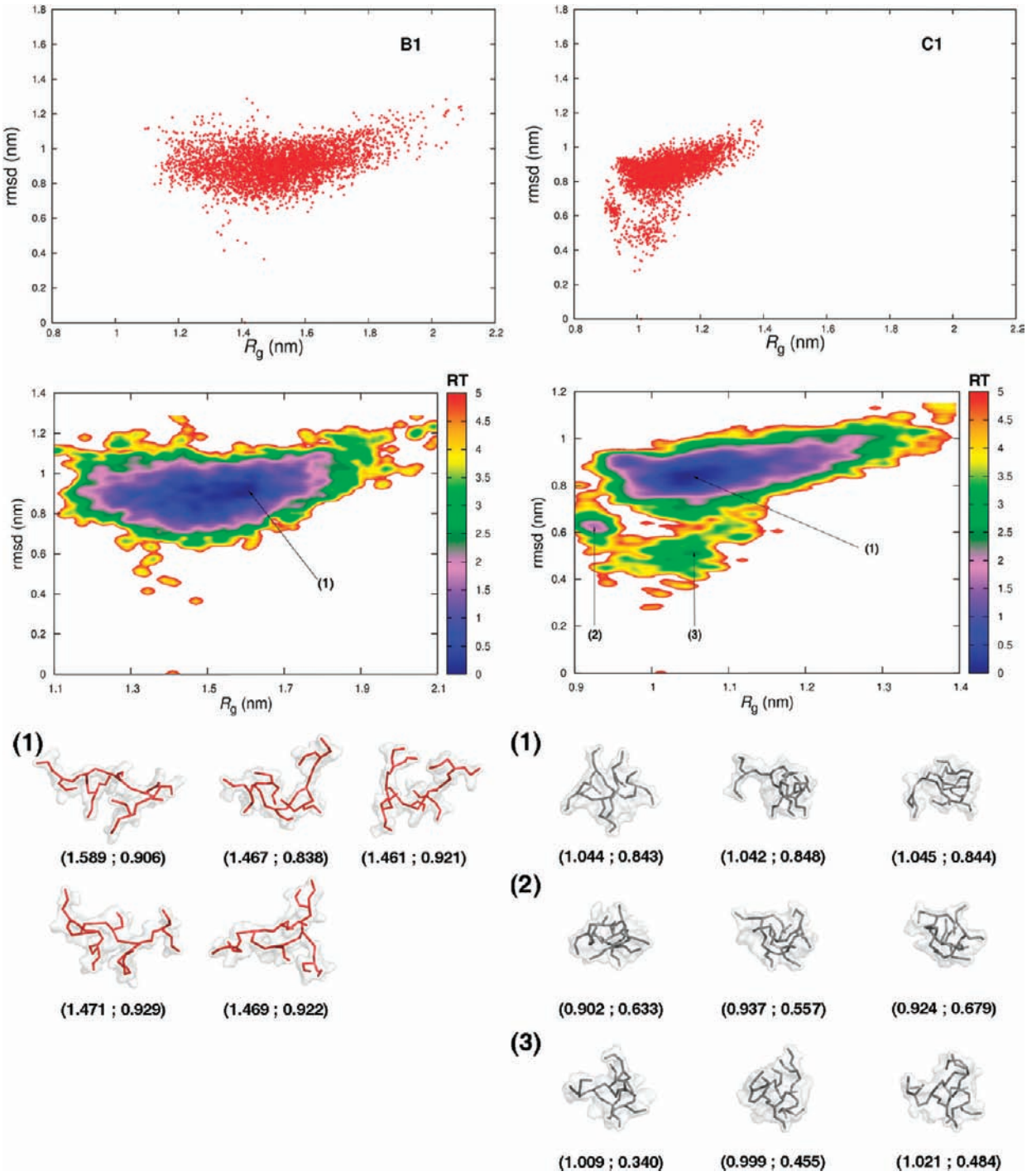


Figure 5. Scatter-plots (top) and free energy profiles (center) for B1 (left) and C1 (right), using R_g and rmsd as structural coordinates. The rmsd values were calculated relative to the central structure. A maximum cutoff of 5 *RT* is used for the free energy profiles. Examples of conformations from each basin are displayed, with the leftmost corresponding to the basin minimum.

4. CONCLUDING REMARKS

In the present work, we have extensively sampled the conformational space of peptide dendrimers with the first long time scale MD simulations of these systems, performing also a comprehensive analysis of their conformational preferences using histograms and free energy landscapes. Furthermore, we have shown that the diffusion coefficient experimentally measured for one of the dendrimers (B1) could be reasonably well reproduced by the MD simulations, supporting the suitability of the GROMOS 53A6 force field to model peptide dendrimers.

The present study supports the idea that some peptide dendrimers are indeed very flexible, with the B1-family dendrimers favoring loose and noncompact conformations where the nonbonded contacts between residues are scarce. Some more compact states are also observed in the B1-family, but they represent only a very small fraction of the total amount of accessible

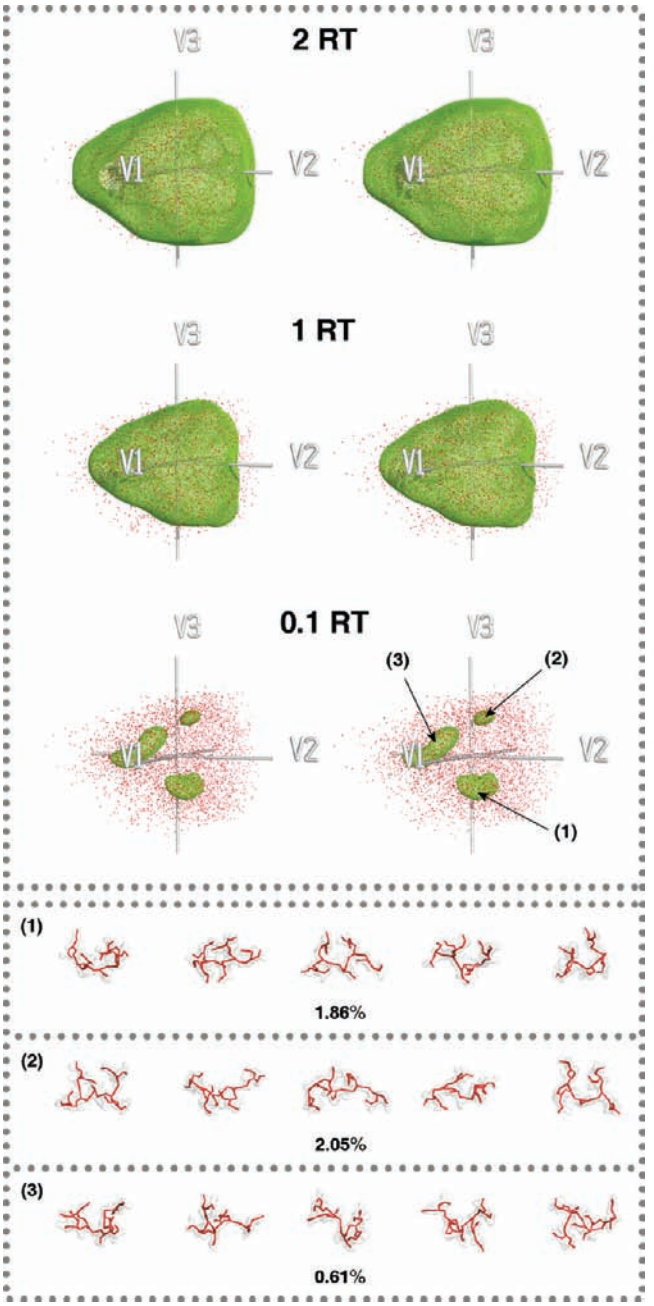


Figure 6. Cross-eye stereo images of B1 free energy landscapes. The structural coordinates employed are the principal coordinate vectors of a 3D space preserving the dissimilarities between all pairs of conformations (section 2.5). Isoenergetic surfaces of 2, 1, and 0.1 *RT* are represented in green, while the underlying density of points is shown as red dots. Examples of structures from each basin and relative populations are displayed, with the leftmost corresponding to the basin minimum. The basins are numbered according to their minimum energy values, with (1) corresponding to the lowest energy basin (high density peak).

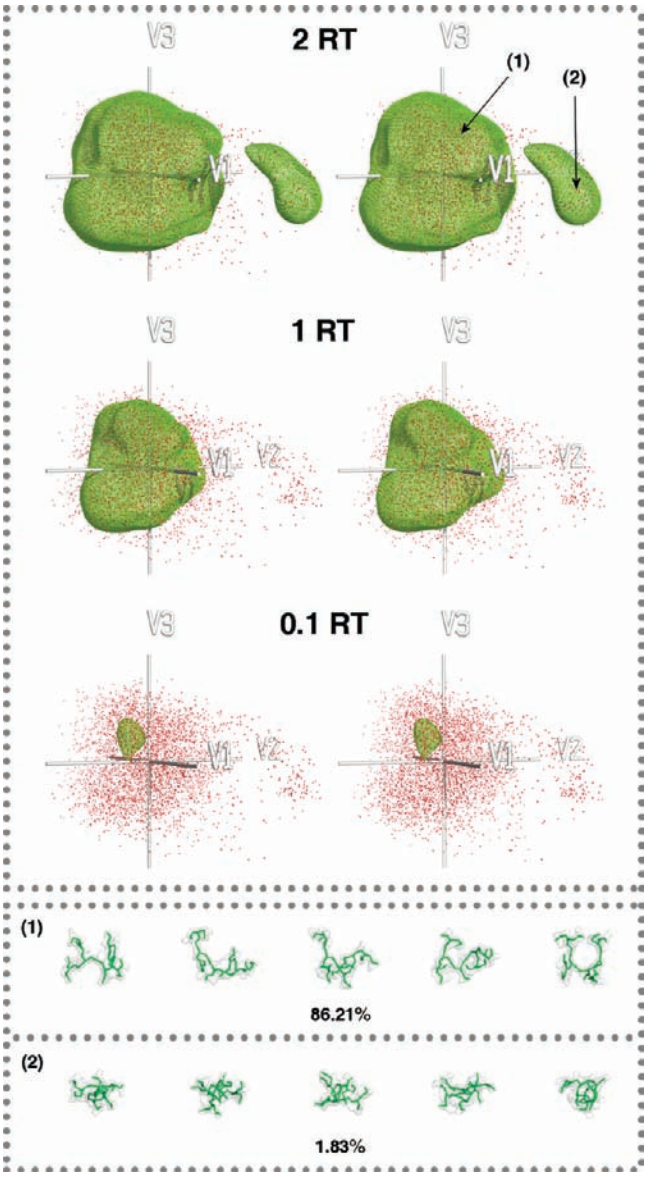


Figure 7. Cross-eye stereo images of B1H free energy landscapes. See the caption of Figure 6 for further details.

conformations. We have also identified a peptide dendrimer, C1, featuring tightly packed preferential conformations that may be regarded as molten globules. Therefore, the present results clearly support the existence of at least two distinct preferential conformational behaviors accessible to peptide dendrimers. These two distinct behaviors are also evident from the free energy landscapes: the landscapes of the B1-family dendrimers display funnel-like shapes with a clear downhill propensity, whereas C1 exhibits rough with multiple significantly populated basins.

The existing experimental results on these systems suggest the possibility of miscellaneous behaviors,[8,23,26,31,32] where some peptide dendrimers are regarded as presenting unorganized flexible shapes, while others can display organized compact structures. In light of the present results, the idea that peptide dendrimers are topologically constrained to adopt a globular shape might need some revision.

The factors conditioning the preferential shapes adopted in solution by peptide dendrimers remain largely unstudied. Nonetheless, some hypotheses can be drawn from the present study. The clearly distinct conformational preferences observed here for the C1 and B1-family dendrimers probably result from a large difference in the composition of these two types of molecules. One such difference is the length of the side chain of the B_2

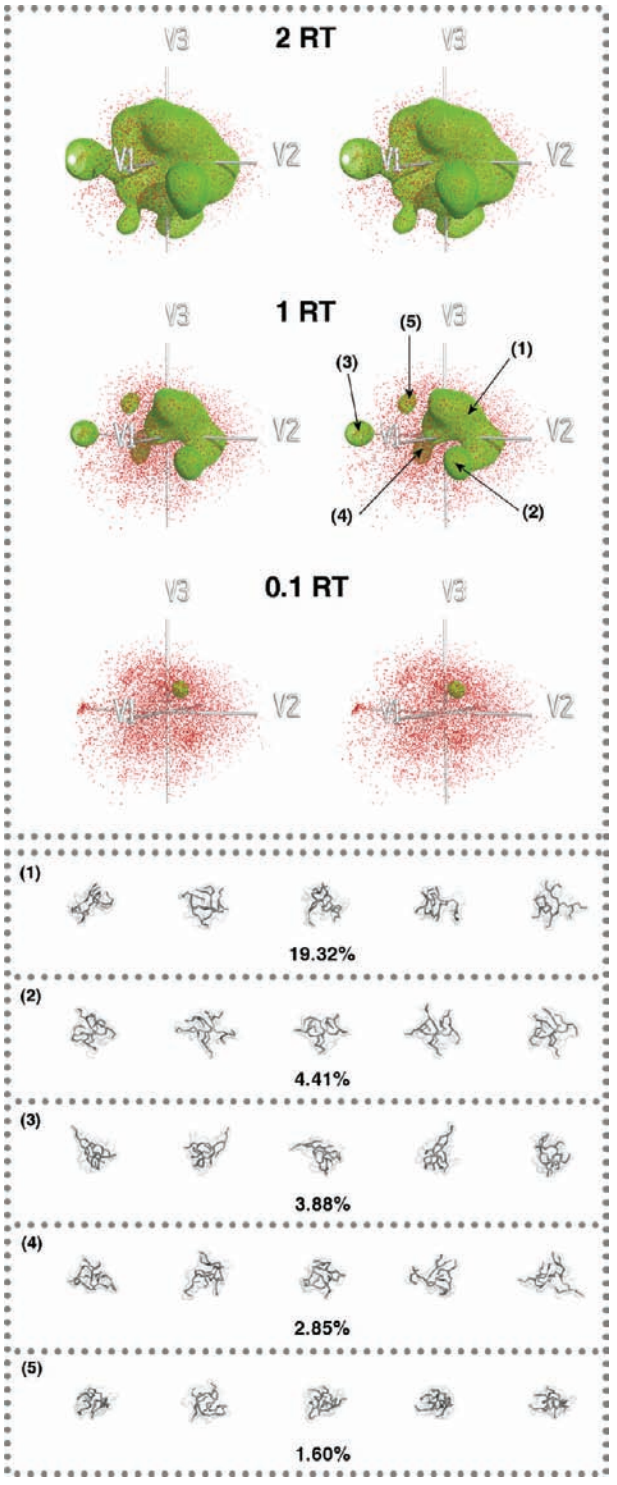


Figure 8. Cross-eye stereo images of C1 free energy landscapes. See the caption of Figure 6 for further details.

branching residues (see Figure 1 and Table 1). While those positions are occupied in C1 by Dap residues with a short side chains [$-CH_2NH_2$], in the B1-family they consist of Lys residues with much longer and freely rotating side chains [$-(CH_2)_4NH_2$], that may confer greater flexibility and lead to more disordered conformations. This hypothesis is consistent with a simulation study of third-generation peptide dendrimers with estereolytic activity, where compact molten globule structures were observed for dendrimers using Dap as the single branching residues.[32] On the other hand, a recent experimental study by Uhlich et al.[27] suggests that the size of peptide dendrimers increases with the number of negatively charged residues. This hypothesis could also explain the present results, since the B1-family contains a large number of negative Glu residues at the X_8 positions, where C1 contains neutral Ser residues (see Figure 1 and Table 1). Clearly, further and more systematic studies are needed to better understand the conformational determinants of peptide dendrimers. In any case, we found the conformational preferences within the B1-family to be remarkably robust to the replacement of a few residues, suggesting that it might be possible to predict the structure of a peptide dendrimer from others with similar composition.

The present work also provides some clues to enhance the functionalization of cobalamin-binding peptide dendrimers. The B1 and B1H dendrimers have very similar conformational behaviors, suggesting that their different binding constants for cobalamin simply reflect a better coordinating ability of Cys relative to His at position X_2 (see Figure 1 and Table 1). Furthermore, the C1 dendrimer has a lower affinity for cobalamin than B1H, despite having an additional His, indicating that its more closed conformations may hinder a proper interaction and formation of a tight complex. Therefore, although we have not conducted further experimental or computational cobalamin binding studies, we propose that the B1HH dendrimer, having a Cys residue, a chemical composition very similar to B1, and the open conformations characteristic of the B1-family, but with two additional His residues, may exhibit an even higher binding constant for cobalamin.

■ ASSOCIATED CONTENT

Supporting Information. GROMOS 53A6 topology blocks for Amb and Dap; details on dendrimer rmsd calculation; Ramachandran plots; additional 2D and 3D energy profiles; spatial distances versus topogical distances plots; complete ref 14. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
baptista@itqb.unl.pt

■ ACKNOWLEDGMENT

We thank Jean-Louis Reymond and Tamis Darbre for sharing experimental data ahead of publication, Paulo Martel for providing us the program to perform principal coordinate analysis and Cláudio M. Soares and Sara R.R. Campos for helpful discussions. We acknowledge the financial support from Fundação para a Ciência e a Tecnologia, Portugal, through project grant PTDC/QUI-QUI/100416/2008.

■ REFERENCES

(1) Smith, D. K.; Diederich, F. *Chem.—Eur. J.* **1998**, *4*, 1353–1361.
(2) Grayson, S. M.; Fréchet, J. M. J. *Chem. Rev.* **2001**, *101*, 3819–3867.
(3) Zeng, F.; Zimmerman, S. C. *Chem. Rev.* **1997**, *97*, 1681–1712.

(4) Gillies, E. R.; Fréchet, J. M. J. *Drug Discovery Today* **2005**, *10*, 35–43.
(5) Bosman, A. W.; Janssen, H. M.; Meijer, E. W. *Chem. Rev.* **1999**, *99*, 1665–1688.
(6) Medina, S. H.; El-Sayed, E. H. *Chem. Rev.* **2009**, *109*, 3141–3157.
(7) Lee, C. C.; Fréchet, J. M. J.; Szoka, F. C. *Nat. Biotechnol.* **2005**, *23*, 1517–1526.
(8) Darbre, T.; Reymond, J.-L. *Acc. Chem. Res.* **2006**, *39*, 925–934.
(9) Kofoed, J.; Reymond, J.-L. *Curr. Opin. Chem. Biol.* **2005**, *9*, 656–664.
(10) Helms, B.; Fréchet, J. M. J. *Adv. Synth. Catal.* **2006**, *348*, 1125–1148.
(11) Liang, C.; Fréchet, J. M. J. *Prog. Polym. Sci.* **2005**, *30*, 385–402.
(12) Crespo, L.; Sanclimens, G.; Pons, M.; Giralt, E.; Royo, M.; Albericio, F. *Chem. Rev.* **2005**, *105*, 1663–1681.
(13) Kim, Y; Zeng, F.; Zimmerman, S. C. *Chem.—Eur. J.* **1999**, *5*, 2133–2138.
(14) Johansson, E. M.; et al. *Chem. Biol.* **2008**, *15*, 1249–1257.
(15) Vepřek, P.; Ježek, J. *J. Pept. Sci.* **1999**, *5*, 203–220.
(16) Darbre, T.; Reymond, J.-L. *Curr. Top. Med. Chem.* **2008**, *8*, 1286–1293.
(17) Stiriba, S.-E.; Frey, H.; Haag, R. *Angew. Chem., Int. Ed.* **2002**, *41*, 1329–1334.
(18) Cloninger, M. J. *Curr. Opin. Chem. Biol.* **2002**, *6*, 742–748.
(19) Boas, U.; Heegaard, P. M. *Chem. Soc. Rev.* **2004**, *33*, 43–63.
(20) Sadler, K.; Tam, J. P. *Rev. Mol. Biotechnol.* **2002**, *90*, 195–229.
(21) Clouet, A.; Darbre, T.; Reymond, J.-L. *Adv. Synth. Catal.* **2004**, *346*, 1195–1204.
(22) Delort, E.; Nguyen-Trung, N.-Q.; Darbre, T.; Reymond, J.-L. *J. Org. Chem.* **2006**, *71*, 4468–4480.
(23) Javor, S.; Delort, E.; Darbre, T.; Reymond, J.-L. *J. Am. Chem. Soc.* **2007**, *129*, 13238–13246.
(24) Kolomiets, E.; Swiderska, M. A.; Kadam, R. U.; Johansson, E. M. V.; Jaeger, K.-E.; Darbre, T.; Reymond, J.-L. *ChemMedChem* **2009**, *4*, 562–569.
(25) Lagnoux, D.; Darbre, T.; Schmitz, M. L.; Reymond, J.-L. *Chem.—Eur. J.* **2005**, *11*, 3941–3950.
(26) Sommer, P.; Uhlich, N. A.; Reymond, J.-L.; Darbre, T. *ChemBioChem* **2008**, *9*, 689–693.
(27) Uhlich, N. A.; Natalello, A.; Kadam, R. U.; Doglia, S. M.; Reymond, J.-L.; Darbre, T. *ChemBioChem* **2010**, *11*, 358–365.
(28) Clouet, A.; Darbre, T.; Reymond, J.-L. *Biopolymers* **2006**, *84*, 114–123.
(29) Maillard, N.; Clouet, A.; Darbre, T.; Reymond, J.-L. *Nat. Protoc.* **2009**, *4*, 132–142.
(30) Sommer, P.; Fluxa, V. S.; Darbre, T.; Reymond, J.-L. *ChemBioChem* **2009**, *10*, 1527–1536.
(31) Javor, S.; Natalello, A.; Doglia, S. M.; Reymond, J.-L. *J. Am. Chem. Soc.* **2008**, *130*, 17248–17249.
(32) Javor, S.; Reymond, J.-L. *J. Org. Chem.* **2009**, *74*, 3665–3674.
(33) Shea, J.-E.; Brooks, C. L., III *Annu. Rev. Phys. Chem.* **2001**, *52*, 499–535.
(34) Brooks, C. L. *Acc. Chem. Res.* **2002**, *35*, 447–454.
(35) Muñoz, V. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 395–412.
(36) Plotkin, S. S.; Onuchic, J. N. *Q. Rev. Biophys.* **2002**, *35*, 111–167.
(37) Onuchic, J. N.; Luthey-Schulten, Z.; Wolynes, P. G. *Annu. Rev. Phys. Chem.* **1997**, *48*, 545–600.
(38) Bryngelson, J. D.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. G. *Proteins: Struct., Funct., Bioinf.* **1995**, *21*, 167–195.
(39) Maisuradze, G. G.; Liwo, A.; Scheraga, H. A. *J. Chem. Theory Comput.* **2010**, *6*, 583–595.
(40) Campos, S. R. R.; Baptista, A. M. *J. Phys. Chem. B* **2009**, *113*, 15930–16001.
(41) DeLano, W. L. *The PyMOL Molecular Graphics System*, version 0.99rc6; DeLano Scientific: San Carlos, CA, 2002.
(42) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.
(43) Berendsen, H. C.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Commun.* **1995**, *91*, 43–56.
(44) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **2004**, *25*, 1656–1676.
(45) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 2451–5459.
(46) Smith, P. S.; van Gunsteren, W. F. *J. Chem. Phys.* **1994**, *100*, 3169–3174.
(47) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
(48) Miyamoto, S.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 952–962.
(49) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
(50) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. *Intermolecular Forces*; Pullman, B., Ed.; D. Reidel Publishing Company, Dorbrecht, The Netherlands, 1981; pp 331−342.
(51) Cohen, F. E.; Sternberg, M. J. E. *J. Mol. Biol.* **1980**, *138*, 321–333.
(52) Kaindl, K.; Steipe, B. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **1997**, *53*, 809.
(53) Steipe, B. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **2002**, *58*, 506.
(54) Silverman, B. W. *Density Estimation for Statistics and Data Analysis*; Monographs on Statistics an Applied Probability; Chapman and Hall: New York, 1986.
(55) Maisuradze, G. G.; Liwo, A.; Scheraga, H. A. *J. Mol. Biol.* **2009**, *385*, 312–329.
(56) Gower, J. C. *Biometrika* **1966**, *53*, 325–338.
(57) Cox, T. F.; Cox, M. A. A. *Multidimensional Scaling*, 2nd ed.; Chapman & Hall/CRC: Boca Raton, FL, 2001.
(58) Becker, O. M. *J. Comput. Chem.* **1998**, *19*, 1255–1267.
(59) Elmaci, N.; Berry, R .S. *J. Chem. Phys.* **1999**, *110*, 10606–10622.
(60) Levy, Y.; Becker, O. M. *J. Chem. Phys.* **2001**, *114*, 993–1009.
(61) Hamprecht, F. A.; Peter, C.; Daura, X.; Thiel, W.; van Gunsteren, W. F. *J. Chem. Phys.* **2001**, *114*, 2079–2089.
(62) Cantor, C .R.; Schimmel, P. R. *Biophysical Chemistry, Part III: The Behavior of Biological Macromolecules*; W. H. Freeman and Company: New York, 1980.
(63) Lee, I.; Athey, B. D.; Wetzel, A. W.; Meixner, W.; Baker, J. R. *Macromolecules* **2002**, *35*, 4510–4520.
(64) Maiti, P. K.; Çağin, T.; Lin, S.-T.; Goddard, W. A., III *Macromolecules* **2005**, *38*, 979–991.
(65) Han, M.; Chen, P.; Yang, X. *Polymer* **2005**, *46*, 3481–3488.
(66) Tanis, I.; Tragoudaras, D.; Karatsos, K.; Anastasiadis, S. H. *J. Chem. Phys. B* **2009**, *113*, 5356–5368.
(67) Metullio, L.; Ferrone, M.; Coslanich, A.; Fuchs, S.; Fermeglia, M.; Paneni, M. S.; Pricl, S. *Biomacromolecules* **2004**, *5*, 1371–1378.
(68) Lee, H.; Larson, R. G. *J. Phys. Chem. B* **2009**, *113*, 13202–13207.
(69) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press, Inc.: New York, 1987.
(70) Chitra, R.; Yashonath, S. *J. Chem. Phys. B* **1997**, *101*, 5437–5445.
(71) Micaelo, N. M.; Baptista, A. M.; Soares, S. M. *J. Chem. Phys. B* **2006**, *110*, 14444–14451.
(72) Mattice, W. L.; Suter, U. W. *Conformational Theory of Large Molecules*; Wiley-Interscience: New York, 1994.
(73) Steinhauser, M. O. *J. Chem. Phys.* **2005**, *122*, 094901.
(74) Cantor, C. R.; Schimmel, P. R. *Biophysical Chemistry, Part I: The Conformation of Biological Molecules*; W. H. Freeman and Company: New York, 1980.