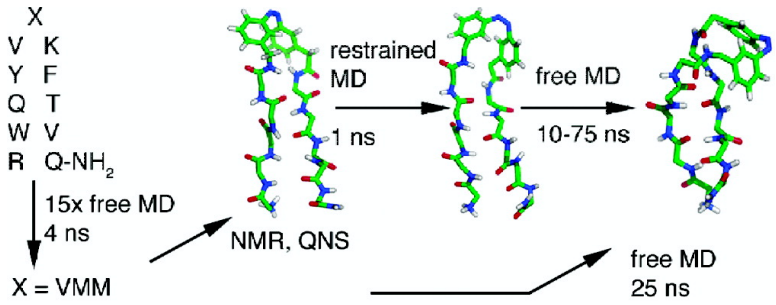
Article

Use of Molecular Dynamics in the Design and Structure Determination of a Photoinducible β-Hairpin

Vincent Krutler, Andreas Aemissegger, Philippe H. Hnenberger, Donald Hilvert, Tomas Hansson, and Wilfred F. van Gunsteren

J. Am. Chem. Soc., **2005**, 127 (13), 4935-4942• DOI: 10.1021/ja044253u • Publication Date (Web): 10 March 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 4 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 03/10/2005

Use of Molecular Dynamics in the Design and Structure Determination of a Photoinducible β-Hairpin

Vincent Kräutler,[†] Andreas Aemissegger,[‡] Philippe H. Hünenberger,[†] Donald Hilvert,[‡] Tomas Hansson,[†] and Wilfred F. van Gunsteren*,[†]

Contribution from the Laboratories of Physical Chemistry and Organic Chemistry, ETH Zürich, CH-8093 Zürich, Switzerland

Received September 21, 2004; E-mail: wfvgn@igc.phys.chem.ethz.ch

Abstract: The study presented here consists of three parts. In the first, the ability of a set of differently substituted diazobenzene-based linkers to act as photoswitchable β-turn building blocks was assessed. A 12-residue peptide known to form β-hairpins was taken as the basis for the modeling process. The central (β-turn) residue pair was successively replaced by six symmetrically ((*o*,*o*), (*m*,*m*), or (*p*,*p*)) substituted (aminomethyl/carboxymethyl or aminoethyl/carboxyethyl) diazobenzene derivatives leading to a set of peptides with a photoswitchable backbone conformation. The folding behavior of each peptide was then investigated by performing molecular dynamics simulations in water (4 ns) and in methanol (10 ns) at room temperature. The simulations suggest that (*o*,*o*)- and (*m*,*m*)-substituted linkers with a single methylene spacer are significantly better suited to act as photoswitchable β-turn building blocks than the other linkers examined in this study. The peptide containing the (*m*,*m*)-substituted linker was synthesized and characterized by NMR in its cis configuration. In the second part of this study, the structure of this peptide was refined using explicit-solvent simulations and NOE distance restraints, employing a variety of refinement protocols (instantaneous and time-averaged restraining as well as unrestrained simulations). We show that for this type of systems, even short simulations provide a significant improvement in our understanding of their structure if physically meaningful force fields are employed. In the third part, unrestrained explicit-solvent simulations starting from either the NMR model structure (75 ns) or a fully extended structure (25 ns) are shown to converge to a stable β-hairpin. The resulting ensemble is in good agreement with experimental data, indicating successful structure prediction of the investigated hairpin by classical explicit-solvent molecular dynamics simulations.

Introduction

Molecular fragments that act as "switches" can be introduced into protein structures in order to trigger selective and reversible conformational changes. Moieties that undergo a structural change upon irradiation with monochromatic light, "photoswitches", are especially interesting in this regard.[1] A generally applicable strategy for the use of such photoswitches requires the conformational regulation of a common, well-defined, and compact structural building block in proteins, such as the β-turn.

Here, we investigate the replacement of the central amino acid pair of a β-turn by a diazobenzene-containing linker (Figure 1). Diazobenzene is known to exist in two conformations, trans and cis, the trans form being thermodynamically more stable. The cis isomer can be generated from the trans isomer by irradiation with monochromatic light in the high-frequency visible range, and then spontaneously reverts back via thermal isomerization with a half-life at room temperature ranging from seconds to days depending on the substitution pattern of the phenyl rings.[2]

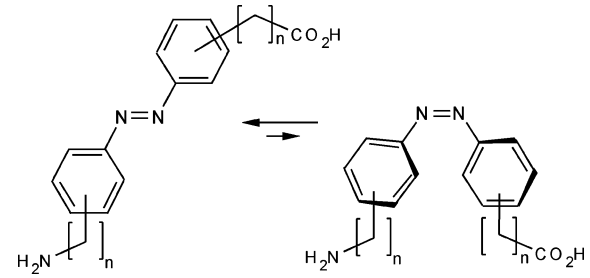


Figure 1. Structure and isomerization of the diazobenzene-based photoswitchable linkers considered in the present study. The linkers considered are symmetrically substituted in the *ortho*, *meta*, or *para* positions ((*o*,*o*)-, (*m*,*m*)-, or (*p*,*p*)-linkers) and involve either one or two methylene spacers between the phenyl ring and the corresponding functional group.

A number of diazobenzene-containing linkers have already been incorporated into peptide backbones and side chains, and their isomerization was found to be well-suited to the induction of reversible conformational changes in proteins.[3−5] The

[†] Laboratory of Physical Chemistry.
[‡] Laboratory of Organic Chemistry.

(1) Pieroni, O.; Fissi, A.; Angelini, N.; Lenci, F. *Acc. Chem. Res.* **2001**, *34*, 9−17.
(2) Griffiths, J. *Chem. Soc. Rev.* **1972**, *1*, 481−493.
(3) Luckner, U.; Cubillos, J.; Chmieliewski, J. *J. Am. Chem. Soc.* **1995**, *117*, 8466−8467.
(4) Behrendt, R.; Renner, C.; Schenk, M.; Wang, F.; Wachtveitl, J.; Oesterhelt, D.; Moroder, L. *Angew. Chem., Int. Ed.* **1999**, *38*, 2271−2274.
(5) Renner, C.; Behrendt, R.; Spörlein, S.; Wachtveitl, J.; Moroder, L. *Biopolymers* **2000**, *54*, 489−500 and 501−514.

structure of diazobenzene suggests that an appropriately substituted fragment involving the cis isomer could emulate the backbone of a β-turn. The β-turn-like structure would then be disrupted upon conversion to trans, leading to a photoswitchable turn.

In this article, we report on the use of molecular dynamics (MD) simulations at different stages throughout the design and characterization process of a photoswitchable diazobenzene-containing β-hairpin. The study consists of three parts. In the first part, MD simulations are used to assess the relative abilities of a set of six differently substituted linkers to act as β-turn replacements (Figure 1). The linkers considered consist of a diazobenzene moiety which is symmetrically ((*o*,*o*), (*m*,*m*), or (*p*,*p*)) substituted with either aminomethyl/carboxymethyl or aminoethyl/carboxyethyl groups. For conciseness, the different linker geometries will be abbreviated as VOO, VMM, VPP, WOO, WMM, and WPP, where V and W indicate the presence of one (V) or two (W) methylene spacer units and the last two letters specify the phenyl ring substitution pattern. In the peptide sequence RWQYV^{D}PGKFTVQ-NH_2 (where ^{D}P stands for D-proline), which is known to form β-hairpins in aqueous solution,[6−8] the two central residues of the β-turn segment (^{D}PG) were replaced by the different linkers in a cis conformation. The six resulting peptides were subjected to unrestrained molecular dynamics simulations in water and in methanol at room temperature, starting from extended initial structures. The relative tendencies of the different peptides to assume a β-hairpin-like conformation were then assessed by analyzing the hydrogen-bonding and conformational patterns adopted during the simulations.

These simulations indicated that the peptides involving the VOO and VMM linkers in the cis conformation possessed a high propensity for hairpin formation. The VMM-containing peptide was synthesized and characterized[9] by NMR, confirming the prediction that this linker could serve as a suitable β-turn building block.

In the second part of this study, molecular dynamics simulations were used in the refinement of the preliminary structure obtained for the VMM-containing peptide using only NOE distance restraints, bonded interactions, and repulsive van der Waals interactions.

Typically, in the MD refinement of structures using NMR data, ^3J coupling constant restraints and NOE distance restraints are used in two separate artificial potential energy functions that restrain the structure so as to satisfy the experimental input data during the simulation.[10] Unfortunately, such potential energy functions usually model the solution structure as a single conformation, with little concession to the fact that NMR measurements really reflect a time average. This is particularly troubling in the case of small biopolymers, where the global fold is not necessarily sufficiently determined by sets of local restraints or where the structure may be sufficiently flexible to allow for sets of NOE restraints that may not be met at the same time, but only as time averages.

An alternative approach involves a restraining function which only requires the restraints to be fulfilled over the course of a predefined averaging time (usually a few picoseconds).[11] The ability of this method to sample a larger conformational space while satisfying the experimental data has highlighted the importance of analyzing whole MD trajectories, rather than single structures.[12] This method, however, comes with a significant drawback, which is the appearance of large, spurious forces along restraints that have not been satisfied over the course of a whole averaging period.[13] Here, we briefly evaluate different restraining protocols (instantaneous restraining and time-averaged restraining with two different averaging times) for the structure refinement of the VMM-containing peptide and attempt to provide additional insight concerning the effect of the respective restraining potential energy terms by using the end points of the restrained MD trajectories as starting points for free simulations.

In the third part of the present study, we present long, unrestrained simulations starting from either extended or NMR-derived model structures. Here, experimentally derived NOE restraints are used only to validate the generated trajectories, thereby avoiding the ambiguities brought about by the use of nonphysical potential energy functions. After 20 ns, both simulations express the same canonical β-hairpin, showing that the system of interest to this study can be described in atomic detail solely by MD simulation using a physically meaningful, internally consistent force field.

Methods

Simulations. All simulations were carried out using the GROMOS96[14,15] package of programs with the GROMOS 43A1 united atom force field.[14,16] The force field parameters of the linkers were derived by analogy with the α-amino acid parameters of the GROMOS96 force field, together with some experimental data[17] (see Supporting Information, Table S.I). Water was modeled using the simple-point-charge (SPC) model,[18] and for methanol, a GROMOS-compatible model was used.[19]

In all simulations, solute and solvent degrees of freedom were independently coupled to a heat bath at 300 K, with a relaxation time of 0.1 ps.[20] The box size was coupled to a pressure bath[20] at 1 atm, with a relaxation time of 0.5 ps and an isothermal compressibility of 45.75×10^{-5} (kJ mol^{-1} nm^{-3})$^{-1}$. All bond lengths were constrained using the SHAKE algorithm[21] with a relative geometric tolerance of

(6) Stanger, H.; Gellman, S. *J. Am. Chem. Soc.* **1998**, *120*, 4236−4237.
(7) Espinosa, J. F.; Gellman, S. *Angew. Chem., Int. Ed.* **2000**, *39*, 2330−2333.
(8) Fisk, J. D.; Gellman, S. *J. Am. Chem. Soc.* **2001**, *123*, 343−344.
(9) Aemissegger, A.; Kräutler, V.; van Gunsteren, W. F.; Hilvert, D. *J. Am. Chem. Soc.* **2005**, *127*, 2929−2936.
(10) Kaptein, R.; Zuiderweg, E. R. P.; Scheek, R. M.; Boelens, R.; van Gunsteren, W. F *J. Mol. Biol.* **1985**, *182*, 179−182.
(11) Torda, A. E.; Scheek, R. M.; van Gunsteren, W. F *Chem. Phys. Lett.* **1989**, *157*, 289−294.
(12) Torda, A. E.; Scheek, R. M.; van Gunsteren, W. F. *J. Mol. Biol.* **1990**, *214*, 223−235.
(13) Scott, W. R. P.; Mark, A. E.; van Gunsteren, W. F. *J. Biomol. NMR* **1998**, *12*, 501−508.
(14) van Gunsteren, W. F.; Krüger, P.; Billeter, S. R.; Mark, A. E.; Eising, A. A.; Scott, W. R. P.; Hünenberger, P. H.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; vdf Hochschulverlag AG an der ETH Zürich: Zurich, 1996.
(15) Scott, W. R. P.; Hünenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krüger, P.; van Gunsteren, W. F. *J. Phys. Chem. A* **1999**, *103*, 3596−3607.
(16) Daura, X.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **1998**, *19*, 535−547.
(17) Hampson, G. C.; Robertson, J. M. *J. Chem. Soc.* **1941**, 409−413.
(18) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, The Netherlands, 1981; pp 331−342.
(19) Walser, R.; Mark, A. E.; van Gunsteren, W. F.; Lauterbach, M.; Wipff, G. *J. Chem. Phys.* **2000**, *112*, 10450−10459.
(20) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(21) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.

Table 1. Summary of the Systems Considered in the Different Probing Simulations[a]

linker type	No. of solvent molecules	box edge length (nm)[b]	linker configuration
	Simulations in water		
VOO	2621	5.45	*cis*
VMM	2620	5.45	*cis*
VPP	2971	5.66	*cis*
VMM2	2608	5.47	*cis*
VMM3	3642	6.05	*cis*
VMMt	3217	5.81	*trans*
WOO	2784	5.57	*cis*
WMM	2618	5.45	*cis*
	Simulations in methanol		
VOO	2394	5.45	*cis*
VMM	2387	5.45	*cis*
VPP	2698	5.66	*cis*
WOO	2533	5.57	*cis*
WMM	2386	5.45	*cis*
WPP	2378	5.45	*cis*

[a] The length of each simulation is 4 ns (water) or 10 ns (methanol). [b] Box size refers to the initial edge length of the cube from which the truncated octahedral box is derived.

10^{-4}. The equations of motion were integrated using the leapfrog scheme and a time step size of 2 fs. The nonbonded interactions were handled using a twin-range cutoff scheme with cutoff radii of 0.8 and 1.4 nm[14] and a pairlist update frequency of 5 time steps. To account for the mean effect of electrostatic interactions beyond the long-range cutoff radius, a reaction-field force[22] with a relative dielectric permittivity of 54[23] was used. The center of mass motion was removed every 1000 steps. Coordinates (solute only) were saved for analysis every 0.2 ps, and energies were saved every 10 ps.

Structural Probing. The folding propensities of the six *cis*-diazobenzene-containing peptides were examined by performing molecular dynamics (MD) simulations in water (4 ns), as well as in methanol (10 ns). The latter solvent allows improved sampling at reduced computational costs due to its lower density and viscosity. Three additional simulations in water (4 ns) were carried out, involving the VMM-containing peptide with the linker in its cis configuration but different initial peptide conformations (VMM2, VMM3) and with the linker in its trans configuration (VMMt).

Initial solute coordinates were obtained by constructing fully extended peptides and by performing an energy minimization using the Insight II software package. The solute (bearing positive charges at the N-terminus and at the arginine and lysine side chains) was placed at the center of a truncated octahedral box with a minimum distance of 1.0 nm between any solute atom and the box wall. No counterions were included in order to avoid the very long relaxation times involved in equilibrating and sampling the counterion distribution. Moreover, the three (positive) charges present in the solute can be expected to be very effectively screened by the high dielectric solvents. The box was then filled with solvent molecules, preserving a minimum solute−solvent distance of 0.23 nm. The resulting solute−solvent systems (system sizes are reported in Table 1) were fully relaxed by a steepest-descent energy minimization with a convergence criterion of 0.1 kJ/mol, followed by 10 ps equilibration MD simulations.

Refinement of the NMR Model Structure. Since the structures obtained from the initial NMR refinement[9] showed no appreciable differences in terms of NOE violations, one of them was selected arbitrarily as the initial structure for the explicit-solvent MD simulations. The solute (charged at the N-terminus, as well as the arginine and lysine residues) was placed in a truncated octahedron box with a minimum solute-to-wall distance of 1.6 nm, and was hydrated by 3089 water molecules. To relax the system, a steepest-descent energy minimization with positionally constrained solute atoms was performed, followed by short MD runs (total 64 ps), with gradual increase of the temperature up to 300 K and decrease of the solute positional restraining force constant from 2.5×10^4 to 10 kJ mol^{-1} nm^{-2}.

A total of five simulations was performed. The first three simulations were initiated from the NMR model structure (see above) and performed using either (i) instantaneous distance restraints[10,14] with a force constant of 1000 kJ mol^{-1} nm^{-2} and a linearization cutoff of 1 nm (simulation *I*, 1 ns); (ii) time-averaged distance restraints[11] with the same force constant and linearization cutoff and an averaging time of 50 ps (simulation *T*, 1 ns); or (iii) time-averaged distance restraints with an averaging time of 300 ps (simulation T_L, 7 ns). The final configurations of simulations *I* and *T* were then used as starting points for two unrestrained simulations F_I, F_T (1 ns each). NOE-derived upper bound distances were obtained from the raw experimental data by applying the appropriate pseudo-atom and virtual-site corrections[14,24] (see Supporting Information, Table S.II).

Structure Prediction. Two MD simulations were performed. In the first (NMR), the final configuration of the simulation F_T was used as the starting configuration, and another 74 ns of simulation using conditions identical to those of F_T was performed, affording a total of 75 ns of unrestrained simulation. The second simulation (EXT) was started from an extended configuration generated and relaxed analogously to the ones used in the probing simulations affording a system with 4317 water molecules in a truncated octahedron based on a cubic box edge length of 6.46 nm. Here, a total of 25 ns of simulation was gathered.

Analysis. A conformational cluster analysis[25−27] of the trajectories was performed based on conformations sampled at 2 ps intervals, using a least-squares fit of the backbone atomic positions (C, C_α, N) and a cutoff criterion of 0.13 nm. Interstrand backbone hydrogen bonds were monitored, the existence of a hydrogen bond being defined by a maximum distance of 0.25 nm between the hydrogen and the acceptor atoms and a minimum angle of 135° between the donor, hydrogen, and acceptor. For the refinement and the long, unrestrained simulations, NOE distance upper bound violations were calculated. Here, NOE-derived upper bound distances were obtained from the raw experimental data by applying the appropriate pseudo-atom and virtual-site corrections[14,24] (see Supporting Information, Table S.II). The average interatomic distances were calculated using a r^{-3} averaging for the refinement simulations and a r^{-6} averaging for the long, unrestrained simulations.

Results

The results of the simulations in terms of occurrences of hydrogen bonds and of representative structures and populations of the most populated conformational clusters are presented in Figures 2−5, 7, and 8.

For later comparison to the simulation results, the hydrogen-bonding patterns expected for a canonical β-hairpin consist of a pairing of either the 2W/10V, 4Y/8F, and 6X → 6X residues (X denoting the linker), termed *A*-type, or of the 1R/11Q, 3Q/9T, and 5V/7K residues termed *B*-type (see Figure 9). Here, a slash denotes the formation of a β-type hydrogen-bond pair, while an arrow denotes a single hydrogen bond.

(22) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 5451−5459.
(23) Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1994**, *100*, 3169−3174.
(24) Fletcher, C. M.; Jones, D. N. M.; Diamond, R.; Neuhaus, D. *J. Biomol. NMR* **1996**, *8*, 292−310.
(25) Daura, X.; van Gunsteren, W. F.; Mark, A. E. *Proteins* **1999**, *34*, 269−280.
(26) Peter, C.; Daura, X.; van Gunsteren, W. F. *J. Am. Chem. Soc.* **2000**, *122*, 7461.
(27) Hamprecht, F. A.; Peter, C.; Daura, X. *J. Chem. Phys.* **2001**, *114*, 2079−2089.
(28) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14.1*, 33−38.

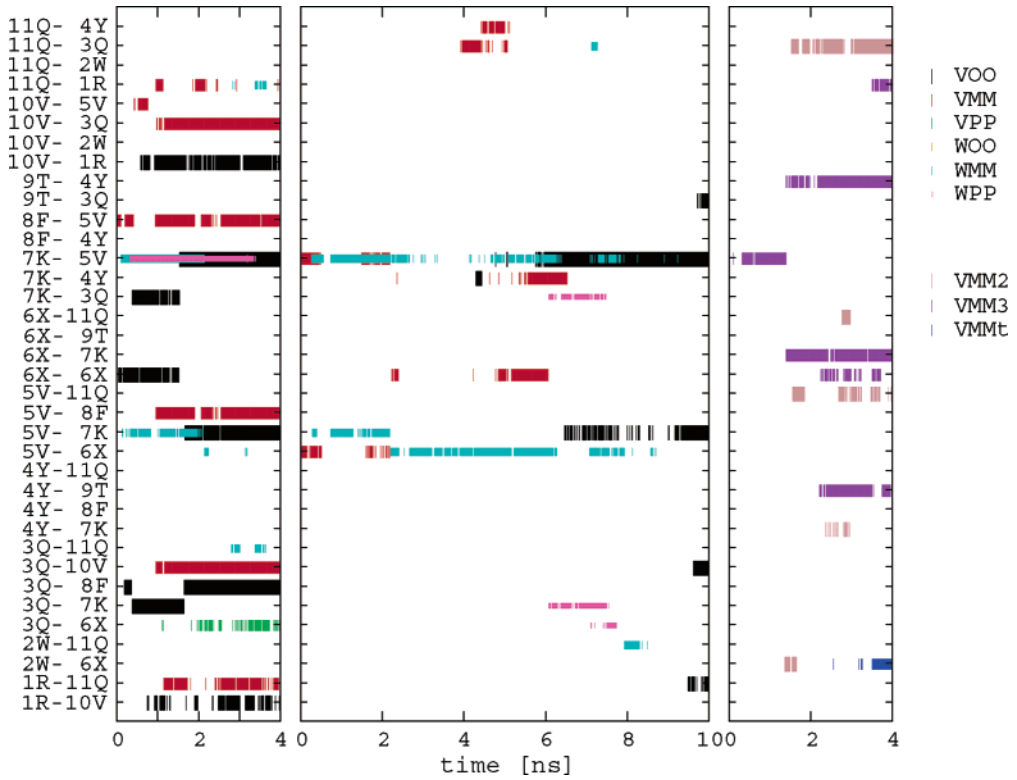


Figure 2. Interstrand backbone hydrogen-bond formation as a function of time for the different probing simulations (V, methylene spacer; W, ethylene spacer; O/M/P, substitution pattern). Only hydrogen bonds occurring in at least 1% of the configurations during the corresponding simulation are reported. The labels of individual hydrogen bonds indicate a hydrogen bond between the NH group of the first residue and the CO group of the second one. The three graphs correspond to simulations in water (left), simulations in methanol (center), and control simulations in water (right). The latter simulations correspond to different initial structures (VMM2, VMM3) or a linker in the trans configuration (VMMt).

Structural Probing. The VOO simulation is characterized by the formation of β-hairpin-like structures both in water and in methanol. In water, a kinked hairpin (1R/10V, 3Q → 8F, and 5V/7K) is formed after about 1.5 ns and appears to be very stable (most populated cluster with 63% occupancy along the trajectory). In methanol, a hydrogen-bonding pattern closer to that of a canonical β-hairpin (1R → 11Q, 3Q → 10V, 5V/7K, and 9T → 3Q) is found at the very end of the simulation. These configurations belong to the most populated cluster with 20% occupancy along the trajectory.

Among the simulations with the VMM linker in a cis configuration in water, the peptide is found to fold into stable (>40% along the trajectories) β-hairpin-like structures in two out of three cases (simulations VMM and VMM3) and a globular structure in one simulation (VMM2). Both VMM (1R/11Q, 3Q/10V, and 5V/8F) and VMM3 (1R ← 11Q, 4Y/9T, and alternative formation of 6X → 6X or 6X → 7K) show out-of-register hydrogen-bonding patterns compared to those of a canonical β-hairpin. In methanol, the same peptide does not form a stable structure, although a few transient hydrogen bonds are observed. The simulation involving the VMM linker in a trans configuration in water (VMMt) is characterized by disordered structures possessing, at most, one interstrand hydrogen bond (2W → 6X), but lacking any hairpin-like shape.

The peptide containing the VPP linker in water quickly folds into a stable structure, involving a single, noncanonical hydrogen bond (3Q → 6X). In methanol, the peptide does not fold at all.

The WMM simulations are characterized by a persistent canonical hydrogen bond pair (5V/7K in water; in methanol, 5V → 7K is replaced by 5V → 6X after about 2 ns), but only occasional hydrogen bonding (2W → 11Q, 11Q → 3Q) is observed further away from the linker. The WPP simulations show a single persistent hydrogen bond in water (7K → 5V) and the transient formation of a hydrogen-bonded pair (3Q/7K) in methanol, concurrent with the sampling of its most populated cluster. The WOO simulations do not show any persistent hydrogen bonds. None of the peptides with two methylene spacers in the linker fold into a stable hairpin structure. In many of these simulations, the C-terminal strand of the peptide samples a wide variety of secondary structure elements, including β-turns (Figures 3 and 4).

Refinement of the NMR Model Structure. As can be seen in Figure 5, all simulations sample similar regions of backbone conformational space corresponding to β-hairpin motifs. A pairwise combined clustering analysis (data not shown) confirms this observation, indicating a very significant overlap in the sampled configurations. In addition, Figure 6 and Table 2 show that all simulations are characterized by limited violations of the 53 NOE distance restraints (see Supporting Information, Table S.II). The sum of positive NOE violations increases in the order T (0.485 nm) < I (0.591 nm) < T_L (0.688 nm) < F_T (1.317 nm) < F_I (2.079 nm).

A closer look at the violated NOE distances (Table 2) reveals two interesting features. First, one-half of the violated NOEs involve the 5V residue, with three out of these four violations involving the atom H_β. The fact that these restraints are violated even in the strongly restrained T and I runs suggests that these restraints cannot be fulfilled simultaneously with each other and with the other restraints, indicating considerable mobility around this residue. Second, one should note the large violation found

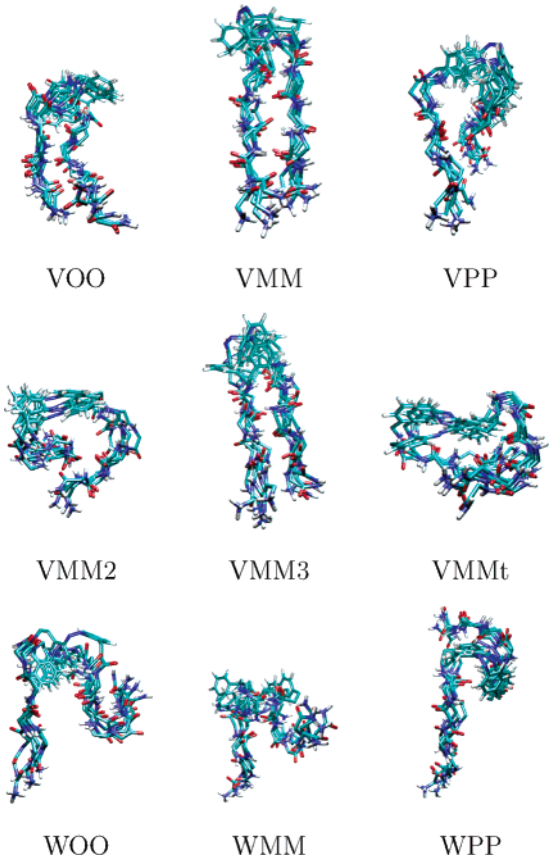


Figure 3. Backbone structures of the five most different members belonging to the primary (most populated) cluster (maximum backbone RMSD of 0.13 nm) for the different probing simulations in water (for labeling, see caption of Figure 2). All shown structures were superimposed using the VMDRMSD program[28] and a backbone (C, C_α, N) least-squares fit to enforce a similar orientation. The occurrence of the primary cluster during the different simulations is VOO (63%), VMM (45%), VPP (53%), VMM2 (30%), VMM3 (43%), VMMt (25%), WOO (20%), WMM (27%), and WPP (9%).

Table 2. NOE Distance Upper Bounds (see Supporting Information Table S.II for the definitions of the 53 NOE indices) that are Violated by More than 0.1 nm during the Refinement and Long, Unrestrained Simulations[a]

NOE	NOE upper bound violation (nm)								
index	T[b]	F_T[c]	T_L[b]	I[d]	F_I[c]	NMR[e]	EXT[f]	residues/atoms	
2					0.195			$R1_{HE}$	$Q11_H$
8		0.163				0.309	0.275	$W2_{HE3}$	$F8_H$
15		0.265	0.133			0.223	0.229	$Q3_H$	$F8_H$
31		0.132	0.114		0.273		0.129	$Y4_{HB+}$	$T9_{HB}$
33		0.126	0.107		0.331		0.149	$V5_{HA}$	$X6_{H10}$
36					0.176			$V5_{HB}$	$X6_{H14+}$
37	0.110	0.153		0.130	0.217			$V5_{HB}$	$K7_{HB+}$
38	0.108	0.135		0.104	0.322			$V5_{HB}$	$F8_H$

[a] Average interatomic distances were calculated using an r^{-3} averaging for the refinement simulations and r^{-6} averaging for the long, unrestrained simulations. [b] T indicates a simulation with time-averaged restraints (the L subscript indicates a longer simulation with an averaging time of 300 ps instead of 50 ps). [c] F is a free simulation (the I or T subscript indicates a simulation started from the final configuration of the corresponding restrained run). [d] I indicates a simulation with instantaneous distance restraints. [e] NMR indicates a long, unrestrained 75 ns simulation starting from the final configuration of T. [f] EXT is a long, unrestrained 25 ns simulation starting from an extended configuration.

for restraint 15 in F_T, which implies that the formation of a canonical *A*-type β-hairpin is apparently incompatible with this restraint.

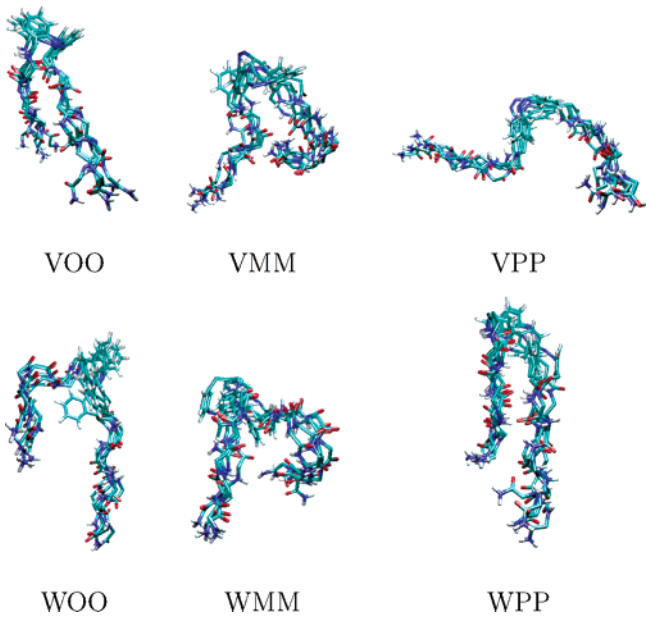


Figure 4. Backbone structures of the five most different members belonging to the primary (most populated) cluster (maximum backbone RMSD of 0.13 nm) for the different probing simulations in methanol (for labeling, see caption of Figure 2). All shown structures were superimposed using the VMDRMSD program[28] and a backbone (C, C_α, N) least-squares fit to enforce a similar orientation. The occurrence of the primary cluster during the different simulations is VOO (20%), VMM (9%), VPP (2%), WOO (7%), WMM (20%), and WPP (9%).

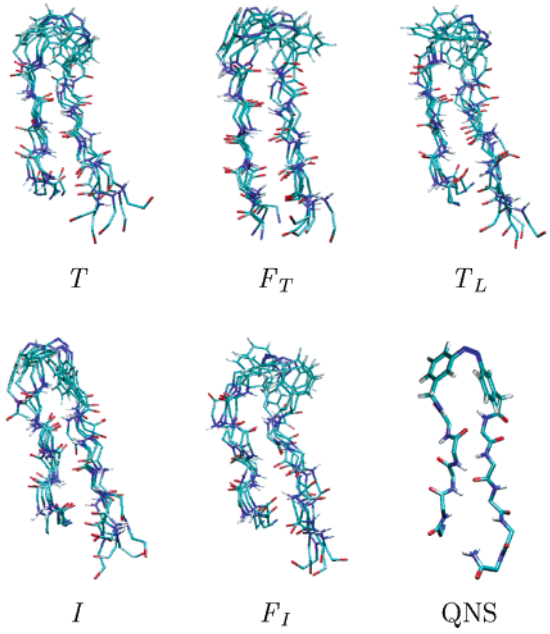


Figure 5. Backbone structures of the five most different members belonging to the primary (most populated) cluster (maximum backbone RMSD of 0.13 nm) for the different refinement simulations. All shown structures were superimposed using the VMDRMSD program[28] and a backbone (C, C_α, N) least-squares fit to enforce a similar orientation. The occurrence of the primary cluster during the different simulations is T (74%), F_T (55%), T_L (27%), I (98%), and F_I (76%). (Bottom right): Starting structure (derived using QNS) employed in the refinement simulations.

The hydrogen-bonding patterns (Table 3) vary widely between the different simulations, and a number of different hairpin-like conformations are actually visited (Figure 5). Still, of the two canonical hydrogen-bonding patterns, the *A*-type one is clearly dominating, with the pair 4Y/8F present in about 50%

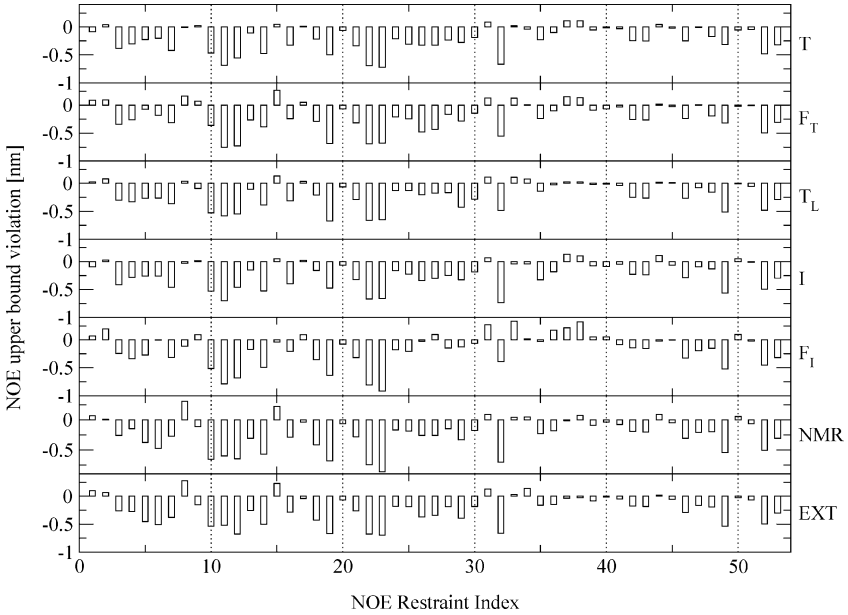


Figure 6. NOE violations during the refinement and long, unrestrained simulations. The average interatomic distances were calculated using r^{-3} averaging for the refinement simulations and r^{-6} averaging for the long, unrestrained simulations. The results are summarized in Table 2. For the definitions of the NOE indices, see Table S.II (Supporting Information).

Table 3. Occurrence of Interstrand Backbone Hydrogen Bonds during the Refinement and Long, Unrestrained Simulations[a]

Hairpin	Donor → Acceptor	Occurrence (%) of Hydrogen Bonds in the Simulations						
type	NH → CO	*T*	F_T	T_L	*I*	F_I	NMR	EXT
	2W → 9T	10	13					
A	2W → 10V		46	6			40	16
	2W → 11Q			22			30	
	3Q → 8F	6				57		
B	3Q → 9T		11	18				
	3Q → 10V			17				
A	4Y → 8F	71	58	56	37	19	86	67
A	6X → 6X			7			66	69
	7K → 4Y	21	5		39	27		
B	7K → 5V		37	49				
	8F → 3Q	6			42			
A	8F → 4Y	77	52		48	7	83	75
	8F → 6X		26					
	10V → 1R		16		22	16		
A	10V → 2W		49				58	29
	11Q → 2W						18	

[a] Only hydrogen bonds occurring in at least 5% of the configurations of the corresponding simulation are reported. Here, a hydrogen bond is defined by a maximum hydrogen−acceptor distance of 0.25 nm and a minimum donor−hydrogen−acceptor angle of 135°. A canonical anti-parallel β-hairpin is expected to show the pairing of either 2W/10V, 4Y/8F, and 6X → 6X (β-hairpin *A*-type) or 1R/11Q, 3Q/9T, and 5V/7K (β-hairpin *B*-type), depending on the hydrogen-bonding pattern in the turn region.

of all sampled configurations. In particular, this pattern is present in the configurations of the primary cluster of simulation F_T, with only the 6X → 6X hydrogen bond being replaced by 7K → 5V.

Structure Prediction. The unrestrained simulation starting from the NMR model structure[9] maintains a stable *A*-type β-hairpin over more than 90% of the trajectory (Figures 7 and 8, Table 3), with only minor fluctuations in the terminal region, where the W2/Q11 hydrogen-bonding pair is found to be competing with the canonical W2/V10 (Figure 7). While the number of violated NOE upper bounds as compared to that of the starting ensemble, F_T, decreases from six to two, the magnitude of the remaining violations increases significantly (Figure 6, Table 2). It is noteworthy that these restraints are obviously not compatible with a canonical β-hairpin (Supporting Information, Table S.II).

The unrestrained simulation starting from an extended structure undergoes a rapid (6 ns) initial folding process (Figure 7) affording a short-lived (0.5 ns) *A*-type β-hairpin. This is followed by a rearrangement of the C-terminal strand, which folds back onto itself forming the simulation's primary cluster (Figure 8) while maintaining canonical *A*-type hydrogen bonding in the linker region (Figure 7). This type of back-bending of the C-terminal strand is also found in the WOO and WMM simulations in water (Figure 3). After 13 ns, the simulation converges with the simulation starting from the NMR model structure (Figure 7) and shows comparable NOE violations, as well (Figure 6, Table 2).

Discussion

Structural Probing. Assuming that the tendency of a peptide to quickly fold into a hairpin-like structure is correlated with the stability of the corresponding hairpin, the molecular dynamics simulations presented here reveal some clear trends in the relative abilities of diazobenzene-containing linkers to act as β-turn replacements.

In the original peptide sequence, a hydrophobic cluster formed by the side chains (2W, 4Y, 8F, and 10V) is believed to play a crucial role in the stabilization of the hairpin structure.[6,8] Formation of a hairpin structure is, therefore, to be expected if two conditions are satisfied: (a) a thermodynamically stable linker geometry allows hairpin nucleation; (b) solvent and linker permit the formation of the hydrophobic cluster.

A number of observations made in the present simulations can be rationalized based on the first condition (a). First, linkers with ethylene spacers do not lead to stable hairpins, even though one of them (WMM) exhibits substantial interstrand hydrogen bonding in the expected turn region. This may be due to a comparatively greater flexibility of the linker, which involves eight rotatable bonds in the W-linkers compared to six in the

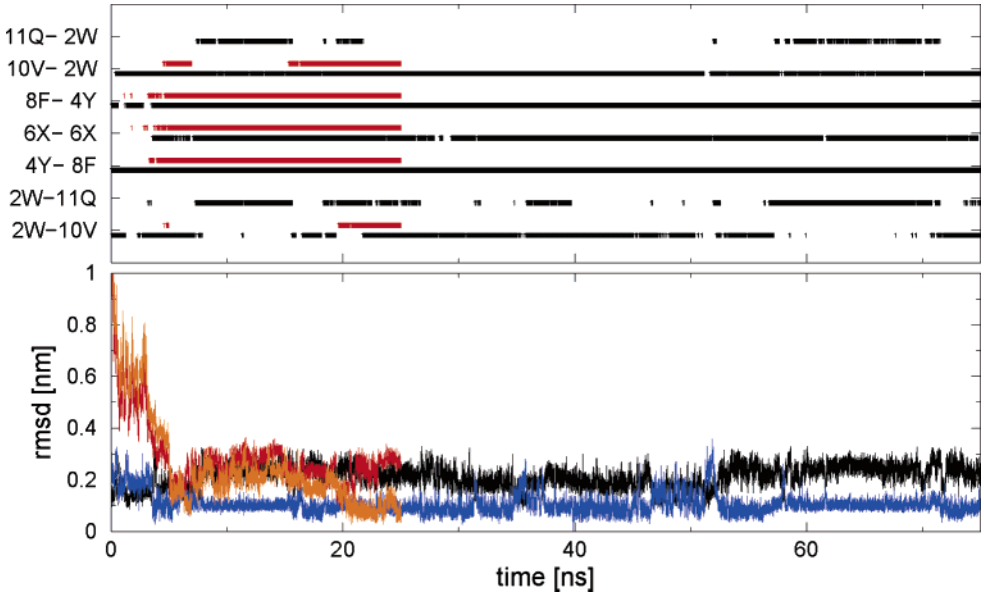


Figure 7. Time-dependent properties of the two long, unrestrained simulations starting from the NMR model structure[27] (NMR, 75 ns) or from an extended configuration (EXT, 25 ns). (Top): Canonical interstrand backbone hydrogen-bond formation (black = NMR; red = EXT). Only hydrogen bonds occurring in at least 1% of the configurations during the corresponding simulation are reported. The labels of individual hydrogen bonds indicate a hydrogen bond between the NH group of the first residue and the CO group of the second one. (Bottom): Backbone atom positional (C, C_α, N) RMSD. Black: NMR, RMSD from the NMR model structure. Blue: NMR, RMSD from the central structure of its primary cluster. Red: EXT, RMSD from the NMR model structure. Orange: EXT, RMSD from the central structure of the primary cluster of the NMR simulation.

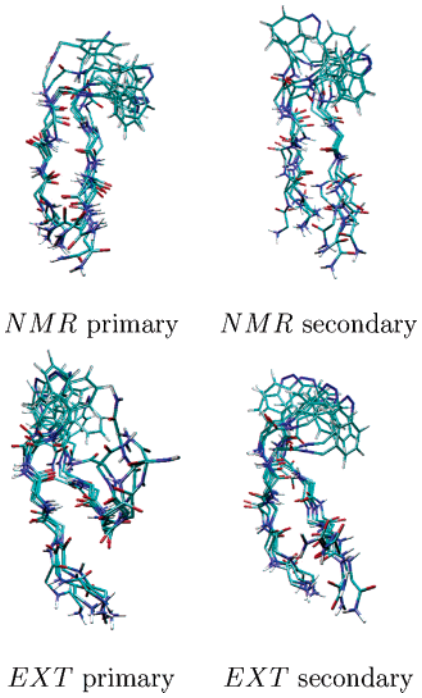


Figure 8. Backbone structures of the five most different members belonging to the primary (most populated) and secondary clusters (maximum backbone RMSD of 0.13 nm) for the two long, unrestrained simulations. All shown structures were superimposed using the VMDRMSD program[28] and a backbone (C, C_α, N) least-squares fit to enforce a similar orientation. The occurrence of the respective clusters during the different simulations is: NMR primary (87%), NMR secondary (7%); EXT primary (41%), EXT secondary (25%). The central configuration of the primary cluster of the simulation starting from an extended structure is sampled after 12.5 ns.

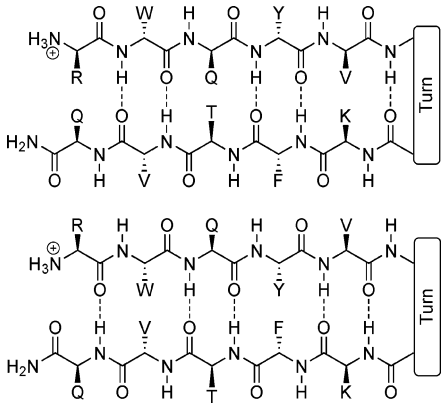


Figure 9. The hydrogen-bonding pattern expected for a VMM-containing canonical β-hairpin consists of a pairing of either 2W/10V, 4Y/8F, and 6X → 6X (top), termed *A*-type β-hairpin, or 1R/11Q, 3Q/9T, and 5V/7K (bottom), termed *B*-type β-hairpin.

V-linkers but only four in the original ^{D}PG residue pair. Second, *para*-substituted diazobenzenes are not good candidates to emulate β-turn building blocks. For these linkers, only the VPP simulation in water showed the formation of a stable structure, which remained fairly globular and lacked significant hydrogen bonding. With this substitution pattern, hydrogen bonds in the turn region could, in principle, only be formed if linkers with ethylene spacers are used. Third, the simulation of a VMM-containing peptide in the trans configuration did not sample hairpin-like structures, suggesting that the cis−trans isomerization could indeed be a suitable conformational switch for this system.

A number of other observations can also be rationalized based on the second condition (b). First, the stability of hairpin-like structures is substantially reduced in the simulations in methanol, which is likely due to destabilization of the hydrophobic cluster in a less polar solvent. While in these simulations the effect of methylene versus ethylene spacers on hairpin stability can no longer be distinguished, a *para*,*para*-substitution of the linkers is still clearly unfavorable. Second, additional interactions between the hydrophobic cluster and the aromatic rings of the linker are probably partially responsible for the lack of hairpin

formation in the simulations involving linkers with ethylene spacers in water. This phenomenon would also explain the lack of hairpin formation in the WMM simulation in methanol, where persistent interstrand hydrogen bonding in the turn region is observed.

Among the linkers examined in the present study, the VOO- and VMM-substituted diazobenzenes are the most likely to form stable β-hairpins (in water). Characterization by NMR of the VMM-containing peptide in its cis and trans configurations[9] confirmed the above prediction, leading to a photoswitchable peptide that forms a highly flexible canonical β-hairpin with the linker in its cis configuration, while forming an insoluble aggregate in the trans configuration. Note, however, that only symmetrically substituted linkers have been considered here. Linkers with asymmetric ring substitution patterns or asymmetric numbers of methylene spacers might also be good candidates for inserting photoswitchable β-turns into peptides.

Refinement of the NMR Model Structure. All resulting trajectories fulfill the experimental restraints within reasonable boundaries, with violations mostly involving the V5 and 8F residues, indicating significant mobility in the turn region. Removal of the restraints (simulations F_T and F_I) results in a sudden increase in the number of NOE upper bound violations, indicating significant strain due to this nonphysical potential energy term.

For the generation of the initial NMR model structure, only covalent and repulsive van der Waals potential energy terms were used in addition to experimental distance restraints. An improved picture of the hydrogen-bonding patterns is, therefore, the main result of the restrained explicit-solvent MD simulations presented here. While considerable qualitative differences in the resulting hydrogen-bonding patterns are observed, an *A*-type hydrogen-bonding pattern is generally preferred. It is interesting to note that for the original sequence, Espinosa and Gellman[7] proposed a *B*-type hairpin, while in our simulations, hydrogen bonds of this type are only found between nonaromatic residues in the relatively unrestrained F_T and T_L simulations.

It is concluded that conformations corresponding to a canonical *A*-type β-hairpin may contribute significantly to the overall ensemble. However, it is also found that simulations of this time scale are not sufficiently converged to obtain a clear and unique picture of the system's conformational distribution and dynamics, and that the application of experimental restraints may have led to the sampling of overly strained and, therefore, unphysical conformations.

Structure Prediction. Both the simulation starting from the NMR-derived model structure[9] and the simulation starting from an extended structure quickly (within 20 ns) converge to a canonical *A*-type (interstrand hydrogen bonds 6X → 6X, 4Y/8F, and 10V/2W) β-hairpin which shows the typical right-hand twist and, once formed, is stable throughout the trajectory. The trajectory starting from the NMR model structure fulfills 51 out of the 53 experimental NOE distance bounds, indicating that the structure adopted by the examined sequence may indeed be classified as an *A*-type β-hairpin. The simulation starting from the extended structure quickly (after 6 ns) forms a short-lived (0.5 ns) *A*-type β-hairpin, which then rearranges into a stable intermediate (13 ns retention time) involving a back-bending of the C-terminal strand, before it again assumes the expected β-hairpin fold, which is stable throughout the rest of the simulation (5 ns).

Overall, the system may be characterized as a stable *A*-type β-hairpin. However, the violation of specific NOE distance upper bounds, the formation of a stable intermediate, and structural data gathered for the original peptide sequence[7] indicate that fairly stable alternative conformations exist which are not well sampled at the time scales reached in the simulation.

It seems that unrestrained molecular dynamics simulations using a physically meaningful and internally consistent force field form a straightforward and reliable tool to investigate the structure and dynamics of the examined sequence at an atomic level of detail.

Supporting Information Available: Force field parameters used for the different diazobenzene-based linkers. List of NOE distance restraints used in the simulations. Figure depicting nomenclature of the linker atom names. This material is available free of charge via the Internet at http://pubs.acs.org.

JA044253U