NXO beta structure mimicry: an ultrashort turn/hairpin mimic that folds in water

Constantin Rabong, Christoph Schuster, Tibor Liptaj, Nadezda Pronayova, Vassil B. Delchev, Ulrich Jordis and Jaywant Phopase

Linköping University Post Print

N.B.: When citing this work, cite the original article.

Original Publication:
http://dx.doi.org/10.1039/c4ra01210k
Copyright: Royal Society of Chemistry 
http://www.rsc.org/

Postprint available at: Linköping University Electronic Press 
http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-108819
NXO Beta Structure Mimicry: An Ultrashort Turn/Hairpin Mimic that Folds in Water

Constantin Rabong, Christoph Schuster, Tibor Liptaj, Nadežda Prónayová, Vassil B. Delchev, Ulrich Jordis, and Jaywant Phopase

We report the first application of NXO-pseudopeptides for β-turn mimicry. Incorporating the proline-derived NProO peptidomimetic building block, a minimal tetrapeptide β-hairpin mimic has been designed, synthesized and its solution structure elucidated. Emulating a natural proline-glycine β-turn, evidence from NMR, molecular modeling and CD suggests the formation of two rapidly interconverting hairpin folds in water, methanol and dimethyl-sulfoxide at room temperature, displaying the proline nitrogen amide bond in either cis or trans arrangement. The NProO-modified hairpin features peptidic backbone dihedrals Φ, Ψ characteristic of natural proline-containing turns composed of α-amino-acids only. Taken together, the observed folding behavior and inherently high designability render the NProO motif a building block for β-structure elaboration in aqueous medium.

1 Introduction

For a long time, peptidomimetics with predictable secondary structure characteristics have been in high demand in medically oriented synthetic chemistry. Combinatorial accessibility, straightforward synthetic diversification and stereochemical integrity during all manipulations are some of their prime features. In principle, all β-secondary structure mimicry relies on a turn-inducing motif with adjacent strands. Any turn mimic must, while suitably orienting the strand backbones to each other, arrange a reversal in peptide backbone direction. However, it has been recognized that amino acid side chain functionalities contribute to the stability of the generated fold, as well. Here, the intrinsic secondary structure propensities of sequenced amino acids and their cross-strand pairing propensities are the principal influences governing the native state ensemble.

The NXO concept represents an undertaking to create a synthetic foldamer scaffold amenable to elaboration of secondary structure found in peptides, in particular β-structure. Incorporating structure-guiding motifs at the repeating unit level, NXO-peptides mimic the dipole polarization of a natural amino acid backbone in β-sheet configuration (Fig. 1). Self-organizing within a mutually attractive hydrogen bonding pattern, NXO-modified peptides adopt extended folds. By virtue of their conformationally biased hydrazide (-NH-NH-CO-) “N”- and oxal-amide (-CO-CO-NH-) “O”-retrons, NXO-modified peptides sample the characteristic “X” amino acid dihedral angles Φ and Ψ in a conformational space characteristic of β-structure.

Selective restriction of conformational freedom and hence limitation of accessible molecular phase space has been a guiding theme in the development of new peptidomimetic scaffolds. Relying conceptionally on a restrained (often cyclic) turn-inducing template motif predisposing attractive strand-strand interactions, ingenious and sophisticated technologies have emerged. Enabling synthetic access to an ever-increasing number of peptide and foldamer sequences predictably adopting β-structure, the study of conformationally constrained peptidomimetics has been essential for an understanding of the principles underlying the generation of structure from sequence in natural proteins. Still, crafting a peptidomimetic foldamer intended for directed self-assembly in aqueous medium remains a daunting challenge.

At first sight, hydrogen bonding may not seem to be the top choice non-covalent interaction to stabilize secondary structure in the design of small molecule topomimetics. The main shortcomings of hydrogen bonds regarding technological exploitation as structure-guiding motifs are their intrinsic kinetic lability and reversibility of association. In protein β-sheets, cross-strand hydrogen bonding contributes most favorably when located in a hydrophobic region of the folded structure.
ever, the backbone of a small peptidomimetic must be regarded as virtually completely solvent exposed. A polar, protic medium like water will therefore effectively compete for hydrogen bonding interactions and relatively diminish intramolecular association. Notwithstanding, there are unique assets of hydrogen bonds: They are sequence-unspecific and feature strongly directional geometric restraints.

This paper describes the design, synthesis and analysis of a minimal β-turn mimic incorporating the L-proline derived NProO modification. We report our efforts towards expanding the frontiers in β-structure mimicry: Investigating peptidomimetic secondary structure motif-guided self-assembly in water at room temperature while, in an anticipated extension towards NXO-modified β-sheets, rendering the turn site functionally designable.

2 Results and discussion

2.1 Design and Synthesis of an NXO β-turn/hairpin-mimic

Reverse turns are sites where the protein chain changes its direction, a prerequisite for the close packing seen in globular proteins. Often located at solvent-exposed protein surfaces, a good deal of molecular recognition and protein host-guest events known today are taking place here. Hence, a designable turn/hairpin peptidomimetic can be most useful. In proteins, proline is found at around 30% of turn i+1 positions. The design prerequisite of maintaining close topological analogy to the majority of L-proline containing turn sites of natural origin and previous studies motivated us to examine the turn-inducing capability of NXO-modified proline in the context of the extensively probed and proven two-residue proline-glycine turn scaffold (Fig. 2). We sought to mimic hairpin folding of a natural tetrapeptide sequence featuring an NXO-modified turn with L-proline at i+1. Incorporating the L-proline-derived NProO building block, the natural sequence would be modified accordingly: At the peptidic C-end, we introduced the conformationally more flexible, stronger solvent interacting hydrazide “N”-retron to mimic the turn glycine at i+2 whereas the anti-restrained oxalamide “O”-retron would constitute the i-position N-terminal.

Aiming to investigate the conformation of NXO-modified proline in water, β-turn mimic 1 was designed (Fig. 2). In order to maintain the structural topology of the β-turn a urea motif was chosen as a linker between the NProO hydrazide terminal at i+2 and i+3 glycine. Dimethylamino and methylamino functions were used for endcapping of the C- and N-terminals, respectively.

\[
\text{HCl}\cdot\text{H}_2\text{N} \quad \text{3}
\]
\[
\text{N} \quad \text{NH}_2 \quad \text{O} \quad \text{O} \\
\text{HCl}\cdot\text{H}_2\text{N} \quad \text{3}
\]
\[
\text{N} \quad \text{NH}_2 \quad \text{O} \quad \text{O} \\
\text{HCl}\cdot\text{H}_2\text{N} \quad \text{3}
\]
\[
\text{N} \quad \text{NH}_2 \quad \text{O} \quad \text{O} \\
\text{HCl}\cdot\text{H}_2\text{N} \quad \text{3}
\]
\[
\text{N} \quad \text{NH}_2 \quad \text{O} \quad \text{O} \\
\text{HCl}\cdot\text{H}_2\text{N} \quad \text{3}
\]
\[
\text{N} \quad \text{NH}_2 \quad \text{O} \quad \text{O} \\
\text{HCl}\cdot\text{H}_2\text{N} \quad \text{3}
\]
\[
\text{N} \quad \text{NH}_2 \quad \text{O} \quad \text{O} \\
\text{HCl}\cdot\text{H}_2\text{N} \quad \text{3}
\]

Scheme 1 Synthesis of NProO-modified β-turn mimic 1.

β-turn 1 was prepared starting from commercially available glycine ethyl ester hydrochloride 2 (Scheme 1). 2 was reacted with phosgene in the presence of DIPEA (ethylidisopropylamine) in dichloromethane to generate the corresponding isocyanate in situ, which was further reacted with Boc-hydrazide to give 3 in 48% overall yield. Compound 4 was generated in quantitative yield via direct amidation of ethyl ester 3 with dimethylamine in ethanol. Boc-deprotection by treatment of 4 with HCl-saturated ethyl acetate gave the product as the corresponding hydrochloride 5 in quantitative yield. 7 was prepared in two steps from 5: Firstly, 5 was reacted with Boc-L-proline using DCC (N,N'-dicyclohexylcarbodiimide) and HOBt (1-hydroxybenzotriazole) as coupling agents and DIPEA as base in dichloromethane to give 6 in 81% yield. Boc-deprotection using HCl-saturated ethyl acetate afforded 7 in quantitative yield. Compound 8 was found elusive towards isolation and attempts to react with oxalylchloride monoethyl or monomethyl ester using organic auxiliary bases (DIPEA, triethylamine) generated inseparable mixtures. Ultimately, 8 was prepared using a previously optimized protocol where 7 was treated with ethyl oxalylchloride in the presence of sodium bicarbonate in DMF at 0°C, affording crude 8 in 29% yield.
Without further purification, 8 was immediately reacted with a solution of methylamine in ethanol to afford 1 in 71% yield.

2.2 Conformational analysis of 1 by stochastic conformational searching and solution phase Molecular Dynamics (MD) studies

Minimum energy conformation analysis and MD (Molecular Dynamics) were undertaken to assess folding and ensemble stability of 1.21 Stochastic conformational searching of >10⁶ ROESY (rotating-frame nuclear Overhauser effect correlation spectroscopy)-restrained, OPLS-AA-derived input geometries generated 956 conformations of which the 10 energy-lowest are shown RMSD (root-mean-square deviation)-superposed in Fig. 3, left model. All conformations are sampled in a reverse-turn characterized by hydrogen bond formation between acceptors and donors at (C3)O←HN17 and (C15)O←HN5, with (C15)O←HN8 providing additional stabilization. The interstrand distance C4·C16 (corresponding to Pos. i and i+3) measuring on average 5.4±0.5Å, 1 adopts a β-hairpin with a ten-membered β-turn.22 With the configuration of the turn-site proline given in the N²ProO-modification, the (C15)O←HN8 bonding motif can be formally assigned to be part of an inverse γ-turn,23 found frequently at the loop end in protein β-sheets.24

The lowest energy conformer generated was then freed of restraints and input to MD simulations with OPLS-AA parameters and implicitly treated solute-solvent interactions (Generalized Born model of solvation). The calculated trajectories further sustain the evidence that turn-site (i+1, i+2)-N²ProO incorporation stabilizes a hairpin conformation in a tetrapetide mimicking sequence (Fig. 3, right model).26 100ns trajectories in DMSO (medium relative permittivity ε=47Debye) at 297K and 327K display 1 as a β-turn adopting hairpin fold (see Fig. 2-4 and Table 2, supplementary information) in rapid equilibrium with an extended conformation throughout the entire simulation time. Again, the fold is mainly stabilized by dual hydrogen bonding at (C15)O←HN5 and (C3)O←HN17, with additional contributions stemming from (C15)O←HN8. 100ns MD simulations in water (ε=78.5Debye) give a qualitatively similar picture, that is, calculations do not show increased medium polarity detrimentally impacting the folding ability via destabilization of interstrand hydrogen bonding. The average hydrogen bonding lifetimes (see Table 3, supplementary information) are seen similar in both media, with (C3)O←HN17 and (C15)O←HN8, (C15)O←HN5 around 2.8ps and 2.1ps, respectively. Peptide backbone dihedrals representing the Nxo-modified turn site in 1 are displayed in the same regions of the Ramachandran plot also populated by natural peptides composed of α-amino acids.27 N²ProO at i+1 exhibits Φ (C15·N14·C10·C9; histogram data (see Fig. 5, supplementary information) around -72±8° while displaying two significant maxima for Ψ (N14·C10·C9·N8) at +109±18° and +154±20°, thus within the core β-region and characteristic of 1-proline at i+1 in a turn configuration.28 MD simulation gives the i-position ―O"- oxalamide constrained to ap (antiperiplanar) around 0(180°). The Nxo-modified hydrazide glycate at i+2 exhibits pertinent backbone dihedrals, sampling Φ (C9·N8·C7·C6) at +102±18 and -111±15° and Ψ (N8·N7·C6·N5) at (180)°±8°, mimicking glycine at i+2 in backbones consisting of α-amino acids.29 From MD simulations, N7H is seen pointing out of the turn cavity in the folded state. Instructively, facile rotation around the N8·N7-ax in the "N"-hydrazide motif, not attainable in the folded conformation, occurs in the extended conformation and during unfolding and refolding; H-N7·N8-H is displayed at +106±24°.30 Here, N8H is seen to provide crucial hydrogen bond donor capacity (H-N8·C9·C10 around +170±11°), acting as a nucleation site for folding. Hierarchically, after the innermost hydrogen bond (C15)O←HN8 is transiently established during incipient folding,31 association via interstrand binding (C15)O←HN5 sets the stage for establishing the prevailing hairpin conformation.32

2.3 NMR analysis of 1

Owing to the poor solubility of 1 in nonpolar solvents, NMR-experiments had to be restricted to water (H2O:D2O 9:1), DMSO-d₆ and MeOH-d₄. In DMSO-d₆, a first clue towards hydrogen bonding as a main factor contributing to fold stabilization were the shifts of the C5-urea and C17-oxalamide NH-signals appearing at δ=6.30ppm and 8.68ppm respectively, both well downfield from their resonances (δ=4.5 and 7.5ppm respectively) under conditions where no hydrogen bonding occurs.33 Given the observed chemical shift invariability upon varying sample concentration (between 2.5 to 25mM), this interaction was concluded to be intramolecular. In water, a fast exchange of NH protons with the solvent was seen; yet, N5H was observable even upon 1.5s of water pre saturation in the DPGSE-NOE (double pulsed field gradient spin echo-nuclear Overhauser effect spectroscopy) experiment,34 again likely so due to intramolecular hydrogen bonding.35 All experiments generated two distinctively different sets of signals in a ratio of ~3:2 (by integration of the respective C1 and C18 signals, see Fig. 4). Investigating the constitutional identity of these two species with COSY (correlation spectroscopy) and TOCSY (total correlation spectroscopy) experiments, their atom-connectivities were found similar; assignment of conformations was next to tackle. Bearing in mind the well described but generally slow cis-trans isomerism of N substituted prolines, we first envisaged a diastereomeric pair of N14·C15 amide rotamers.36 Given that shift differences between the two signal sets decreased proceeding towards the end of both strands from the proximate ring onward (by up to 0.6ppm for C10H, resonating at δ=4.94ppm and 4.34ppm in the major and minor set in DMSO-d₆ respectively), we concluded magnetic anisotropy around the stereogenic proline carbon to be the likely cause of splitting. That the two experimentally observed signal sets corresponded to the open and folded conformers seen in MD simulations (see Fig. 3, supplementary information) was excluded: Firstly, both observed sets displayed signals
characteristic of geometries with restricted backbone rotations. Secondly, the open-fold interconversion rate was simulated in the order of magnitude of $10^{-4}$ s at 290K, too fast to be discernable by the NMR experiments undertaken.

Fig. 4 600 MHz $^1$H NMR spectrum of 1 in MeOH-d$_6$ (298K, 5mM). Displayed in the box is the selective DPGSE-NOE irradiation of the C18 methyl group protons in both species 1a and 1b (at $\delta$=2.80ppm and 2.73ppm), showing NOE enhancement of the C1 methyl signals at $\delta$=2.93ppm and 3.02ppm, respectively.

Fig. 5 Key NOE interactions in 1, shown for both observed equilibrium geometries 1a and 1b. NOESY was carried out with a mixing time of 250ms at 290K, 297K and 298K in H$_2$O:D$_2$O=9:1, DMSO-d$_6$ and MeOH-d$_6$, respectively, at 5mM concentration.

Observing NOE (nuclear Overhauser effect spectroscopy) enhancements of methylolalamide C18H and dimethylamide CIH (Fig. 4 and 5) in both species, conformations with spatial proximity of strand-ends were indicated. Thus, the existence of two different sets of signals had to be explained in terms of two folds, rapidly equilibrating at NMR timescale.

Among envisaged structures were isomers exhibiting either a C9-C10 cis-amide or C15-C16 cis-oxalyl unit. Further, a cis-proline (N14-C15) conformer was contemplated, yet it remained unclear how such a fold would realize across-strand hydrogen bonding (experimentally indicated, though, by the downfield-shifting of the amide protons N5H and N17H observed in both signal sets). Ultimately, ab initio quantum mechanical calculation of NMR chemical shifts at the B3LYP/6-311G++(d,p) level of theory, including implicit solvation, clarified the issue: Among the minimum energy geometries located and substantiated by comparison to MD trajectories (see chapter 2.2), next to the previously observed hairpin (1a, Fig. 5, left structure and Fig. 6, blue model), a minimum geometry exhibiting the conspicuous downfield chemical shift of C10H was calculated (1b, Fig. 5, right structure and Fig. 6, red model). Overall in good agreement with experimental proton chemical shifts (Fig. 6 and 7, supplementary information), this geometry displayed both the amide linkage at the proline nitrogen and the strand-terminating methyl oxalamide unit in cis-arrangement. According to calculations, geometry 1b is 0.05-0.22kcalmol$^{-1}$ more stable than 1a. Although in isomer 1b the proline amide bond N14-C15 is in cis arrangement, concomitant rotation in the C16-N17 methylamide subunit allows 1b to project the terminal methyl C18 spatially similar to 1a while engaging in interstrand hydrogen bonding.

2.4 Studies of 1 by Circular Dichroism

While a delicate probe to study molecular conformation, CD (circular dichroism) cannot necessarily induce the structure characteristics of foldamers featuring novel functionalities. As a reference and conceptual augmentation, we prepared the N$^\beta$ProO-derivative 9, the enantiomer of 1, incorporating d-proline. The measurements shown in Fig. 7 provide evidence corroborating positive formation of a distinctive molecular conformation. In all three solvents of examination, 1 and 9 exhibit graphs characteristic of $\beta$-secondary structure: Indicating $\beta$-turn
adoption in water, 1 shows ellipticity $\theta$ maximal at $\lambda=231$ nm and a zero-transition at 209 nm, relatively red-shifted to the CD values for natural $\beta$-sheets.\textsuperscript{40} Folding propensities, taken from $\theta$ at 190 nm and from the local maxima in the 230-240 nm region, diminish on going from acetonitrile to methanol and water, correlating with solvent polarity. Thus, in agreement with NMR experiments, folding to a central well-defined spatial motif was observed in polar protic and aprotic medium alike.

**Fig. 7** CD graphs of 1 and corresponding $\delta$-proline enantiomer 9 in acetonitrile, methanol and water. Measurements were carried out at 295K and 0.1 mM concentration.

### 2.5 Discussion

The studies undertaken in the context of this work present the N\textsuperscript{2}ProO-modified model peptidomimetic 1 folding to a minimal $\beta$-sheet increment, displaying the aptitude of NNO-modified peptides to engage extensively in intramolecular hydrogen bonding. Subordinate to a fast (in the order of magnitude of $10^{-11}$ s at 290 K) open-fold equilibrium, 1 adopts a native state of two $\beta$-turn conformers in aqueous medium around room temperature. The trans-proline hairpin 1a and the cis-proline conformer 1b, the latter being slightly more abundant (ratio 1a:1b $\sim$ 2:3) and exhibiting a cis-proline amide bond and a cis-oxalyl unit in the oxalamide $O$-motif. 1a and 1b feature dual interstrand hydrogen bonding interactions. Shown in Fig. 8 are backbone-RMSD minimized overlays of the $\beta$-hairpin equilibrium geometry 1a, supported by spectroscopy and computation, with corresponding ideal $\beta$-turns (see caption Fig. 8). Both, the turn NProO-modified t-proline at $i+1$ and the hydrazide “N”-retron mimicking glycine at $i+2$ display their equivalent backbone dihedrals $\Phi$, $\Psi$ in the same region of the Ramachandran plot as their parent natural amino acids.\textsuperscript{41} 1 is seen as a versatile mimic of natural $\beta$-turns; ensemble conformer 1a is a close mimic of type II and type I$'$ $\beta$-turns. While in natural proteins, glycine at $i+2$ is frequently found in a type II turn configuration, type I$'$ turns were shown to be effective hairpin inducers.\textsuperscript{42}

In trying to rationalize the equilibrium ensemble composition of 1, a comparison of hydrogen bonding in NNO-modified peptides with analog tetrapeptide turns composed of $\alpha$-amino acids exclusively can be instructive.\textsuperscript{43} At given temperature, pressure and concentration, the state of a small peptide in solution is characterized by the Boltzmann-weighted ensemble of molecular microstates, that is, a multitude of low-energy structures in dynamic equilibrium.\textsuperscript{44} For a peptidomimetic, reproducing peptide secondary structure is an essential criterion; yet, in permitting technological application, the ability to represent the collectivity of ensemble properties is ultimately important.

**Fig. 8** Backbone RMSD-minimized superposition of the central MD conformer 1a from MD simulations (OPLS-AA in water at 290 K) with ideal turns. Ideal turn $i+1$ and $i+2$ torsions $\Phi$, $\Psi$ are taken from B. L. Sibanda, J. M. Thornton, Nature 1985, 316, 170. Turn types (model color, $i+1$ and $i+2$ backbone atom RMSD-difference to 1a in Å): Upper model: 1 (green, 0.26); 1' (yellow, 0.05); 11 (pink, 0.12); 11' (turquoise, 0.25). Lower model: III (orange, 0.62); III' (seagreen, 0.11); V (amber, 0.16); V' (purple, 2.02). Dotted lines indicate hydrogen bonding.

Understanding of the long time observed conformational plurality manifested by $\beta$-sheets and the high mobility of $\beta$-strands
due to inherent strand hydrogen bonding diversity was significantly advanced about a decade ago: Then, evidence was presented that in β-sheets, efficient cross-strand hydrogen bonding requires a slightly “skewed” geometry with a relative register shift of opposite strands to minimize electrostatically repulsive interactions, concomitant factors of intrastrand hydrogen bonding.\(^\text{45}\) These competing interactions cause the intrinsic conformational flexibility (in turn reflected by a shallow energy surface of Φ-torsions in the peptide backbone) of strands and, therefore, determine β-sheet ordering. This insight into β-structure conformational diversity led us to investigate whether it could be favorable to have an increased number of potential hydrogen bonding donor and acceptor moieties allocated in a peptidomimetic backbone. We envisaged a scenario of β-structure mimicry where, even in a polar medium like water, multiple hydrogen bonding capabilities constituted an asset rather than a liability.

To increase the number of potential interaction sites is advantageous from the perspective of statistical mechanics: Cross-strand hydrogen bond pairings that are laterally offset and hence inconsistent with formal sheet patterns are still entropically advantageous (and may “smear out” energetic separation between folded and completely unfolded states, the effect increasing with strand length). Structurally cooperative folding involving cross-strand backbone hydrogen bonding has been described before.\(^\text{46}\) Given that the backbone modification meets the requirement of matching hydrogen bond donor and acceptor functions in a mutually attractive fashion (see Fig. 1), incorporation of a backbone donor/acceptor site manifold draws upon this effect. Then, since in a partially unfolded structure the contiguous placement of folded and unfolded strand sections was shown to be thermodynamically unfavorable,\(^\text{47}\) presenting additional cross-strand hydrogen bonding options can drive the equilibrium towards folding by contributing positive entropy termi, diminishing the overall entropic cost of folding.\(^\text{48}\) The free energy of folding for an ensemble species \(i\) at temperature \(T\), \(\Delta G_{i}\), is given by

\[
\Delta G_i = \Delta H_i - T \Delta S_i \quad \text{Eq. (1)}
\]

where \(H_i\) and \(S_i\) are the associated enthalpy and entropy of folding. \(S_i\) depends upon the measure of ensemble microstates \(\Omega_i\)

\[
S_i = k_B \ln \Omega_i \quad \text{Eq. (2)}
\]

where \(k_B\) is Boltzmann’s constant. \(\Omega_i\) relates the number of realized microstates \(n_i\) to the number of accessible microstates \(N_i\)

\[
\Omega_i = N_i / n_i \quad \text{Eq. (3)}
\]

The relationship of \(\Omega_i\) with \(\Delta G_i\) is demonstrated for two folding ensemble entities, \(i = j, k\). As simplifications, it is assumed that both species \(j, k\) attain the native fold and derive their entire stabilization from the same number of reversible, independent and isoenergetic interactions. Other contributions to entropy are neglected so that Eq. (2) represents the entire conformational entropy of the molecule. The intrinsic difference in free energy of folding between \(j, k\) can be expressed as

\[
\Delta \Delta G_{jk} = \Delta \Delta H_{jk} - T (\Delta \Delta S_{jk}) \quad \text{Eq. (4)}
\]

The difference in the folding entropy term reads

\[
\Delta \Delta S_{jk} = \Delta S_j - \Delta S_k = k_B \ln \Omega_j / \Omega_k = k_B (\ln \Omega_j - \ln \Omega_k) = k_B \Delta \ln \Omega_{jk}
\]

\[
\text{Eq. (5)}
\]

Further, with Eq. (3),

\[
\Delta \Delta G_{jk} = k_B \left( \frac{N_j}{n_j} - \frac{N_k}{n_k} \right) = k_B \left( \Delta \ln N_{jk} - \Delta \ln n_{jk} \right)
\]

\[
\text{Eq. (6)}
\]

Thus, from entropy contributions alone and assuming \(H_j = H_k\), the difference in folding free energy between species \(j, k\) is related to the difference in population of microstates \(\Delta N_{jk}\) and \(\Delta n_{jk}\) by

\[
\Delta \Delta G_{jk} = -k_B T \Delta \ln \Omega_{jk} = -k_B T (\Delta \ln N_{jk} - \Delta \ln n_{jk})
\]

\[
\text{Eq. (7)}
\]

For \(N_j > N_k\) and \(n_j < n_k\), \(\Delta G_{jk} < 0\). It is significant that both, a reduction in the number of accessible microstates \(N_k\) and an increase in the number of folded states \(n_k\) are fold-stabilizing species \(k\) relative to \(j\). Whereas the favorability of reducing \(N_j\) (see Eq. (3)) is generally understood, foldamers exhibiting a plurality of interacting sites can capitalise on an increase of \(n_i\) to promote folding, as well.

In protein chemistry, the definition of higher-order structure is subject to rigorous characteristics. For structure classification, a set of formal criteria is consulted; structure elucidation focuses on the experimentally observable canonical ensemble member exclusively. Recently, facilitated by ever-increasing computing power, the analysis of protein dynamics has gained in importance. Significant advances in the understanding of ensemble dynamics and characterization of the molecular manifold have been made.\(^\text{49}\) It is here that a foldamer-inspired outlook on peptidomimetics can view structure formation, dynamics and ordering from a different perspective: Structural energetics suggest that it is the dynamics of interconversion between folded, unfolded and partially folded states together with a rigidified potential energy surface of backbone torsions that guide the mimic when presenting the functionalities lined on it and relate closely to the performance of the mimic in attaining tight interaction profiles with molecular targets.\(^\text{50}\) Consequently, RMSD difference as the criterion for comparing structure motifs gains in importance.

The folding event restructures a number of randomly ordered conformations through the formation of energetically stabilizing, non-covalent interactions. A reduction in the number of accessible ensemble conformations and accompanying loss of backbone and side chain conformational entropy is connected with a large energy penalty, known to be the single most unfavorable energetic factor in protein folding.\(^\text{51}\) By virtue of strand-planarization through short-range dipole ordering its constituent “N”- and “O”-motifs, the NXO-modification enhances the interaction strength of both interstrand and solute-solvent hydrogen bonding, enthalpic factors contributing to a large negative ESF (Electrostatic Solvation Free Energy) upon folding.\(^\text{52}\) Yet, incorporation of conformationally biased, “constrained” repeating motifs intrinsically predisposes the thermodynamic ensemble to undergo folding interactions. Likewise, decreasing accessibility to ensemble backbone conformations yields a lower conformational entropy penalty upon folding.\(^\text{53}\)

The native state of \(I\) featuring two distinct folds \(1a\) and \(1b\) in equilibrium presents an opportunity to study the behavior of
the NXO-modification in greater detail: It has been established that in absence of auxiliary stabilization, proline-N cis and trans amide isomers in proteins are nearly isoenergetic, the cis form usually being slightly disfavored.54 The energetic barrier of roughly 20 kcal mol\(^{-1}\) commonly associated with proline amide cis-trans isomerization in proteins can be expected to be significantly lowered in a short peptide sequence.55 Disregarding any further mechanism facilitating cis-trans isomerization (kinetics have been shown, e.g., to be accelerated by intramolecular assistance via hydrogen bonding to a proline nitrogen acceptor, effectively enhancing the sp\(^{3}\)-character of the proline ring nitrogen and hence decreasing amide double bond resonance\(^{39}\)), stabilization of the cis-proline fold 1b can be rationalized from evaluation of the intrinsic energetics in the “O”-oxalamide retron: \textit{Ab initio} quantum chemical energy profiling of relative configurations in oxalamides found the all-trans (ttt) conformer to be the minimum energy configuration in \(N, N\)-dimethyloxalamide, with cis-trans-trans (cct) and cis-skew-cis (csc) relatively destabilized by 6.2 kcal mol\(^{-1}\) and 12.7 kcal mol\(^{-1}\), respectively.57 In csc, an unfavorable steric interaction between the terminal NH-CH\(_2\) in the methylamide and the preceding carbonyl was found, the distance being 2.82 Å and thus slightly below the combined van der Waals radius.58

In 1, the difference between ttt and ctt is lowered, the proline ring nitrogen being a tertiary amide. In the solvated minimum geometries from \textit{ab initio} calculations (see Fig. 6), the hydrogend bond oxalamide dihedrals N14-C15-C16-N17 are indeed seen arranged skew (145° and 137° for 1a and 1b, respectively). The “O”-retron in 1 attains ttt and csc configuration; hence 1b avoids an unfavorable steric clash of the terminal methyl C18 with C15(O) (in 1b, the calculated distance C15(O)-C18 is 2.88 Å; the plane of the terminal methylamide, C16-N17-C18 is rotated 41° relative to the plane of C15-C16-O). Moreover, from Mulliken atomic charges, it was seen that upon hydration, the oxalamide carbonyls in \(N, N\)-dimethyloxalamide become stronger H-bond acceptors (by -0.1 electrons); here, the much larger dipole moment of the csc minimum (3.07 Debye in water) compared to both ctt and ttt (0.42 and ~0, respectively) allows for a stabilizing solvent-solute interaction, computed at 2.8 kcal mol\(^{-1}\) relative to the global ttt minimum.57a These results suggest the “O”-retron in 1 to gain upon stabilization through intrastrand hydrogen bonding; however, in 1b, this interaction can only be realized with the methylamide unit in cis configuration at C16-N17. Taken together, this arrangement then satisfies a skew oxalyl unit at C15-C16 and a second (C3(O)–H17) hydrogen bond at the cost of a further insignificant cis amide energy penalty.

3 Conclusions

In summary, employing NXO-peptidomimicry, we have created a minimal β-sheet increment featuring the N\(^{2}\)ProO-modification as novel β-hairpin nucleator. Simulation and spectroscopy account for folding to a turn/hairpin native state in polar solvents, including water, at room temperature. Thus, even in a minimal tetrapeptide mimic, NXO can impart β-structure characteristics and achieve secondary structure mimicry. As an outlook towards further studies, the high designability and ready diversification, together with excellent solubility properties in polar media, render NXO-derived β-structures valuable scaffolds for the elaboration of β-peptidomimetics in water.

Acknowledgements

JP is pleased to acknowledge financial support from the Swedish Research Council, Sweden (grant no. 2012-42315-94008-81).

Notes and references

\(^{a}\) Integrative Regenerative Medicine Centre (IGEN) & Department of Physics, Chemistry and Biology (IFM), Linköping University Campus Valla, 58183 Linköping (Sweden). E-mail: jayph@ifm.liu.se

\(^{b}\) Institute of Applied Synthetic Chemistry, Vienna University of Technology, Getreidemarkt 9/163, A-1060 Vienna (Austria).

\(^{c}\) Department of Environmental Geosciences, University of Vienna, Althanstrasse 14, A-1090 Vienna (Austria).

\(^{d}\) Department of NMR and Mass Spectrometry, Institute of Analytical Chemistry, Faculty of Chemical and Food Technology, Slovak University of Technology, Radlinského 9, 81237 Bratislava (Slovak Republic).

\(^{e}\) Department of Physical Chemistry, University of Plovdiv, Tzar Assen Street 24, 4000 Plovdiv (Bulgaria)

\(^{f}\) For information about NXO building blocks contact Ferentis at info@ferentis.eu

\(^{†}\) Electronic Supplementary Information (ESI) available: Experimental details of Synthesis and characterization details, Mass, NMR, CD spectra, computational data. See DOI: 10.1039/b000000x/


a) J. M. Word, S. C. Lovell, T. H. LaBean, H. C. Taylor, M. E. Zalis, B. K. Presley, J. S. Richardson, D. C. Richardson, J. Mol. Biol., 1999, 285, 1711; b) B. K. Ho, A. Thomas, R. Brasseur, Proc. Sci., 2003, 12, 2508; a minor peak sampling Ψ around -50° in the α-region and corresponding to the opened fold was observed, as well; see Fig. 5, supplementary information.


Formation of the γ-turn relative minimum conformer was calculated to be slightly exergonic with $\Delta G_{\text{TS}}^{\gamma} = -0.7 \pm 0.3 \text{ kcal/mol}$, at the B3LYP/6-311G+(d,p) level of theory, using solvation SCRF=water in the GAUSSIAN program package (M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, et al., Gaussian 09, Revision A.01, Gaussian, Inc., Wallingford CT, 2009). Relative to the extended conformation, the native fold exhibiting the γ-turn motif realizes energetic stabilization overtly from van der Waals (2.4 ±0.5 kcal/mol) and electrostatic termini (−6.1 kcal/mol); from a 100ns MD trajectory in water at 290K; see Table 2, supplementary information). Compare C. Guo, H. Levine, D. A. Kessler, Phys. Rev. Lett., 2000, 84, 3490.

The conformation with an extended backbone typically samples the hydrazide dihedral H-N7-N8-H around -120°, whereas in the fold 1 this angle is typically seen around +110°; thus, the relative populations of this torsion also reflect the population of the corresponding conformers (see Figures 2–4, supplementary information).


From calculated potential energies with SCRF=water solvation, at B3LYP with the 6-311G+(d,p) basis set. It is seen that isomer 1b gains stabilization of 1.3 kcal/mol relative to the a1 upon solvation from vacuum; see supplementary information for details.

The backbone dihedrals for 1b are, from the ab initio calculated lowest energy conformer, at i=1: $\Phi = -73^\circ, \Psi = +131^\circ$; at i=2: $\Phi = +92^\circ, \Psi = -11^\circ$.


Comparison of backbone-RMSD from the superposition of molecular fragments consisting of different constituent atoms as in Fig. 8 appears yet justified given the backbone pitch (measured as the distance between sequential α-carbons) in NXO-modified peptides being close to natural peptide sequences: For example, ab initio calculations gave the pitch between equivalent backbone atoms in a truncated NAlaO pseudopeptide, MeNHCO-N\(^2\)AlaO-NHMe, at 7.44±0.04Å, deviating less than 1.8% from the distance in a corresponding sheetlike l-alanine tripeptide, the latter calculated at 7.57±0.04Å for the distance between the terminal α-carbons.


It has been noted that the largest contribution to the free energy of folding came from stabilization through backbone hydrogen bonding and not from side chain interactions, see Ref. 8a and references cited therein; see also Ref. 43b and C. Zhang, J. L. Cornette, C. Delisi, Pro. Sci., 1997, 6, 1057.


58 For *N,N*-diacetylhydrazide, the *trans-gauche-trans* (tgt) isomer was found the most stable, with *tt* and *cgr* being 1.3kcalmol$^{-1}$ and 1.4kcalmol$^{-1}$ less stable, respectively. The NH-bond was again seen relatively deshielded (by +0.05 electrons) upon hydration; see Ref. 57a and compare analysis of the hydrazide "N"-retron above, chapter 2.2.