FEBS Letters

FEBS PRESS science publishing by scientists

Computational investigation of retro-isomer equilibrium structures: Intrinsically disordered, foldable, and cyclic peptides

Gül H. Zerze[1], Frank H. Stillinger[2] and Pablo G. Debenedetti[1]

1 Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ, USA
2 Department of Chemistry, Princeton University, Princeton, NJ, USA

Correspondence
P. G. Debenedetti, Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ 08544, USA
Tel: +1 609 258 5480
E-mail: pdebene@princeton.edu

(Received 12 May 2019, revised 20 June 2019, accepted 26 July 2019, available online 9 August 2019)

doi:10.1002/1873-3468.13558

Edited by Michael Bubb

We use all-atom modeling and advanced-sampling molecular dynamics simulations to investigate quantitatively the effect of peptide bond directionality on the equilibrium structures of four linear (two foldable, two disordered) and two cyclic peptides. We find that the retro forms of cyclic and foldable linear peptides adopt distinctively different conformations compared to their parents. While the retro form of a linear intrinsically disordered peptide with transient secondary structure fails to reproduce a secondary structure content similar to that of its parent, the retro form of a shorter disordered linear peptide shows only minor differences compared to its parent.

Keywords: Aβ; cyclic peptides; intrinsically disordered peptides; p53; peptidomimetics; retro peptides

Peptide backbones are referred to as nonpalindromic, because there is a specific bond (and hence sequence) directionality from the N-terminus (amino group end) to the C-terminus (carboxy group end), which is not the same as that resulting from moving from the C-terminus to the N-terminus. Reversal of the peptide bond, with the resulting peptides referred to as retro peptides, has been an active subject of protein and peptide research, and has been considered in drug design, peptidomimetics, and molecular recognition feature (MoRF) prediction [1–5]. Earlier, it has been hypothesized that backward reading of the primary structure may result in mirror image protein structure [6], a view supported by lattice model simulations, which show that the retro form of the B domain of Staphylococcal protein A can fold into a three-helix bundle similar to its parent [7]. Later, however, CD and NMR experiments together with computational modeling showed evidence that the retro forms of the B domain of Staphylococcal Protein A, the SH3 domain of α-spectrin, and the B1 domain of Streptococcal Protein G are unfolded, unlike their parents [8]. Additionally, the resolved crystal structure of the retro form of GCN4 leucine zipper in retro form showed that it folds into a highly stable structure, distinct from the parent peptide's native structure [9]. These findings led to consensus on the view that peptide bond order has an intimate relationship with the folding characteristics of peptides. More recently, retro peptide structure has been studied by combining structure prediction algorithms and short molecular dynamic simulations using atomistic models [10], showing that although retro peptides may not fold as their parents, 'they retain secondary structure preferences similar to their parents'.

Amino acid order reversal has mostly been considered in the context of peptide drug design. Peptide-therapeutics have generally been found to have high binding affinity, specificity, and low toxicity, compared to other small-molecule analogs. However, in addition

Abbreviations
MoRF, molecular recognition feature; PDB, Protein Data Bank; PTWTE, parallel tempering in the well-tempered ensemble; RMSD, root mean square deviation.

to their superior properties, they have certain weaknesses that can limit their usage. One such weakness is their comparatively low stability against proteolytic degradation; they often have a short half-life in the body and are degraded before reaching their target. However, as opposed to levorotary [(L)-forms] peptides, dextrorotary [(D)-forms] peptides are much more stable against enzymatic degradation. The use of D-enantiomers of peptide drug candidates has, therefore, been proposed as a strategy to develop longer half-life peptide drugs. Nonetheless, because the D-peptide is the mirror image of the L-peptide [11,12], it does not preserve the side-chain geometry of the parent L-peptide, which plays a critical role in binding to the target [13]. Accordingly, retro isomerization has been brought into play here [1,14,15]: the same orientation of the side chains can be recovered by inverting the peptide bond sequence of the D-analog, as can be visualized in Fig. S1, resulting in the retro (inverted amino acid order)-inverso (inverted backbone chirality) peptide. Throughout this work, the term 'retro isomerization' is meant to denote backward reading of the amino acid sequence. It is not to be confused with 'retro-inverso peptide', which denotes a peptide whose amino acid sequence is read backwards (retro) and each of whose individual amino acids is replaced by its mirror-image enantiomer (opposite chirality). Retro-inverso peptides are not considered in this work, except for literature discussion, and pedagogical purposes (Fig. S1). We do not simulate retro-inverso peptides. In spite of above-described theoretical advantages, retro-inverso drug design has shown mixed success [16–18].

The retro-inverso design approach has typically failed for peptides with a relatively higher secondary structure content [19]. This failure has largely been attributed to structural differences arising from stereochemical inversion (inverso part of the retro-inverso concept) such as the difference between right- and left-handed helices, for the L- and D-forms of a helical peptide, respectively [19]. However, as mentioned above, a reversed amino acid sequence may fail to preserve the parent backbone structure. Side-chain geometry equivalence between parent and retro-inverso peptides is based on the assumption of backbone structure preservation. For example, the alanine peptides in Fig. S1 have the same backbone torsional angles ($\varphi = 180°$, $\psi = 180°$ for all peptide bonds) in both the normal and retro-inverso forms of the peptide. If the backbone structures of the normal and retro forms were different, there would not be side-chain geometry equivalence between them.

While it has been established that reversed sequences of folded proteins are not able to adopt the same structure as their parents, it is not presently understood how retro sequences of nonfoldable peptides behave, compared to their parents. Here, we include two special classes of peptides, linear intrinsically disordered peptides and cyclic peptides. Retro inversion has been used with some success for drug design involving intrinsically disordered proteins [20,21] and in MoRF prediction algorithms [5]. Intrinsically disordered proteins or regions do not fold into well-defined, three-dimensional structures but they can show some, often transient, structural order at the secondary level. While some intrinsically disordered peptides only adopt coil-like extended structures [22,23], others can populate transiently formed local secondary structures [24–27]. In order to address systematically the effect of peptide bond sequence (retro isomerization) on backbone structure, we consider two linear foldable, two linear intrinsically disordered, and two cyclic peptides, both in their normal and reversed amino acid order. To our knowledge, this is the first atomistically detailed simulation study that systematically tests the effect of retro isomerization on peptide structure by studying the equilibrium ensembles. We find that while the retro forms of cyclic and linear foldable peptides are distinctly different from their parents, the effects in the case of intrinsically disordered peptides depend on the degree of structural order in the parent peptide. The retro form of one of the disordered peptides, which has partial secondary structure content, fails to reproduce parent's structure, whereas the retro form of the other disordered peptide considered here shows high structural similarity to its parent to the extent that structural order can be quantified.

Methods

We investigate the effect of retro isomerization for four linear and two cyclic peptides. The primary structures of the studied peptides are shown in Table 1. Initial coordinates of the linear peptides are generated using the CHARMM program [28] as fully extended configurations which are then energy-minimized using the steepest descent algorithm and relaxed in gas-phase simulations ($T = 300$ K) for 1 ps. The crystal structure of the cyc1 peptide is obtained from the Cambridge Structural Database (ID: BAMLIK) [29] and is used as the initial configuration for the simulation of cyc1 peptide. The Protein Data Bank (PDB) ID: 1JBL [30] is used for the initial coordinates of the 1JBL peptide. The coordinates of retro forms of the cyclic peptides are initially generated as linear peptides with the same backbone torsional angles as their parent peptide. Following that, the backbone is cyclized by forming the amide bond between N and C terminus, that is, by imposing the appropriate bond, angle, dihedral, and nonbonded 1–4 interaction

Table 1. Amino acid composition of the peptides studied in this work

Peptide	Primary Structure	Secondary Structure
GB1	GEWTYDDATKTFTVTE	β-hairpin
rGB1[a]	ETVTFTKTADDYTWEG	
Trp-cage	NLYIQWLKDGGPSSGRPPPS	Helix
rTrp-cage[a]	SPPPRGSSPGGDKLWQIYLN	
$p53_{15-29}$	SQETFSDLWKLLPEN	Intrinsically disordered
$rp53_{15-29}$[a]	NEPLLKWLDSFTEQS	
$A\beta_{16-20}$	KLVFF	Intrinsically disordered
$rA\beta_{16-20}$[a]	FFVLK	
cyc1	cyclo(-GHGAYG-)	
rcyc1[a]	cyclo(-GYAGHG-)	
1JBL[b]	cyclo(-GRCTKSIPPICFPD-)	
r1JBL[a,b]	cyclo(-DPFCIPPISKYCRG-)	

[a] Lower-case letter '*r*' in front of the peptide's name denotes 'retro'. [b] We note that there is a disulfide bond between cysteines of 1JBL (and r1JBL).

parameters, and running a steepest decent energy minimization. The coordinates of cyclized peptides are then relaxed in gas-phase simulations for 1 ps.

Trp-cage, $A\beta_{16-20}$, $p53_{15-29}$, cyc1, and 1JBL peptides are modeled using the Amber03w protein force field combined with the TIP4P/2005 water model [31,32], while the GB1 peptides are modeled using the Amber03* protein force field combined with the TIP3P water model [33]. Force fields are chosen according to the data in the literature; while the Amber03w (TIP4P/2005) force field has been shown to better represent disordered proteins and Trp-cage [31,34,35]; the Amber03* (TIP3P) force field has been shown to produce folding thermodynamics in close agreement with the experimental data for β-hairpins [36,37]. Linear peptides (both the normal and retro forms), Trp-cage, $A\beta_{16-20}$, $p53_{15-29,}$ and GB1 are solvated in a truncated octahedron box containing 2688, 1325, 2040, and 2718 water molecules, respectively. Cyclic peptides (both the normal and retro forms), cyc1 and 1JBL are also solvated in a truncated octahedron box containing 1388 and 1931 water molecules, respectively. Initial coordinates of the solvated peptides are energy-minimized and equilibrated in an NVT (T = 300 K, $\rho_{water} \approx 1$ kg·L^{-1}) simulation, followed by an NPT (P = 1 bar, T = 300 K) simulation, each for 100 ps. Further production simulations are performed in the NPT ensemble. Temperature is maintained using the Nosé–Hoover thermostat [38,39] with a 1 ps time constant and pressure is maintained at 1 bar using a Parrinello–Rahman barostat with isotropic coupling using a time constant of 2 ps.

We perform parallel tempering in the well-tempered ensemble (PTWTE) simulations [40–42] of the peptides for at least 300 ns/replica. The well-tempered ensemble amplifies the fluctuations of the potential energy (maintaining the same average potential energy), thereby increasing the potential energy distribution overlap between adjacent replicas and significantly reducing the number of replicas required to achieve an adequate average exchange acceptance [41]. Temperature ranges, number of replicas, resulting average replica-exchange acceptance rate and energy biasing parameters are given in the SI (Table S1). Convergence is monitored by following the cumulative average of the radius of gyration for each peptide at 300K, as shown in Figs S2 and S3. A certain amount of initial simulation data is treated as equilibration and excluded from the analysis, as also shown in Figs S2 and S3.

We simulate our systems using the GROMACS 2016.3 [43,44] MD engine and PLUMED 2.3.1 for metadynamics calculations [45]. Systems are propagated using the leap-frog algorithm with a 2 fs time step. The temperature of each replica is maintained using the Nosé–Hoover thermostat [38,39] with a 1 ps time constant. Electrostatic interactions are calculated using the particle-mesh Ewald method [46] with a real space cutoff distance of 1 nm. A 1-nm cutoff distance is used for the van der Waals interactions.

Secondary structure assignment of the GB1, Trp-cage, and $p53_{15-29}$ peptides is performed using the DSSP algorithm, which is based on hydrogen bonding patterns [47]. Details of β-sheet and α-helix assignments are provided by Kabsch and Sander [47]. The average β-sheet and α-helix fractions are calculated as the ensemble average of the above-described instantaneous quantity.

Per-residue turn fractions of the short cyclic peptide are calculated according to the φ and ψ angles sampled by the $i + 1$ and $i + 2$ residues. We analyze type I, II, I', and II' turns. We first calculate the φ and ψ angles for each residue and for each configuration in the ensemble. Then, for each pair of $i + 1$ and $i + 2$ residues, we search whether the dihedral angle criteria are satisfied. Ideal φ and ψ angles for $i + 1$ and $i + 2$ residues are taken from the work by Lin and coworkers [48] and ±30° is allowed for fluctuations around the ideal values for each angle. The angle criteria are visually represented in Fig. S4. If the angle criteria are satisfied by both residues, we then check whether a backbone hydrogen bond forms between residues i and i+3 [49]. If this is also satisfied, residues from i to $i + 4$ are assigned to be in a turn structure. We only find type I turn for the cyclic hexapeptide, and, therefore, we only report per-residue fraction of type I turn.

α-helix, β-strand, and ppII helix assignments for the longer cyclic peptide are performed according to the φ and ψ angles. The angle requirements for each structure type are shown in Fig. S5. In addition, we also impose a contiguity criterion of at least three consecutive residues to be found in the relevant basin of the φ-ψ space (α-basin, β-basin, and ppII basin, respectively for α-helix, β-strand, and ppII helix).

Clustering is performed based on structural similarity of the backbone heavy (i.e. non-hydrogen) atoms following the GROMOS algorithm [50] using a 0.15-nm root mean

square deviation (RMSD) cutoff distance. For each conformer in the ensemble, other conformers with respect to which the RMSD is 0.15 nm or less are found. The structure with the highest number of similar conformers (RMSD within 0.15 nm) forms the most populated cluster, together with the other conformers within the given cutoff distance. We report the most populated cluster with the percentage of the population (number of conformers in this cluster with respect to the total number of conformers).

Results and Discussion

Linear peptides

Four linear peptides are selected as representative of the typical structures that a peptide can assume. While Trp-cage [51] and GB1 [52] are well-known helix and β-hairpin folders, respectively, Aβ [22,23] and p53 [53] are classified as intrinsically disordered peptides. The p53 peptide is known to transiently form local helical structures, [53] whereas Aβ peptides are more similar to a random coil [22,23]. Retro-inverso forms of $A\beta_{16\text{-}20}$ and $p53_{15\text{-}29}$ fragments have previously been proposed as peptide-based drug candidates [20,54]. While the retro-inverso $A\beta_{16\text{-}20}$ fragment has been shown to have similar target binding as its parent [20,21], the retro-inverso form of the $p53_{15\text{-}29}$ fragment was ultimately found not to possess the same binding ability as the parent peptide [55].

Figure 1 shows the comparison of the peptides structures with their retro analogs at ambient conditions. We show the quantification of the structural observables that are most relevant to the structure of the parent peptide. The GB1 peptide folds into a β-hairpin [56], as we also observe here. However, the retro form of the peptide fails to fold into a stable β-structure (Fig. 1A). The most populated structure (see the Methods) of rGB1 contributes only 5% to its aqueous solution equilibrium ensemble, suggesting that this retro isomer does not fold into a stable alternative structure. Trp-cage is a helix folder with a α-helix near the N terminus and a 3_{10}-helix near the C terminus [57]. The retro form of Trp-cage, however, is only able to form an α-helix approximately five times less frequently in the same region as shown in Fig. 1B (N terminus for the parent peptide, C terminus for the retro peptide). Moreover, the 3_{10}-helix of the native Trp-cage near its C terminus is completely lost in the retro form of the peptide. Contributing only 6% of the entire population, the most populated cluster does not suggest the presence of an alternative stable structure, either. We find that rTrp-cage has a significant fraction of 3_{10}-helix between residues 10 and 15. In an earlier experimental study [58], rTrp-cage was found to be unfolded in aqueous solution. However, in 30% 2,2,2-trifluoroethanol (TFE) solution, rTrp-cage was found to stabilize a structure (published with a PDB ID: 2LUF), which contains an α-helix between residues 8 and 15. We note that this range of rTrp-cage residues encompasses the region that we find to form a relatively large fraction of 3_{10}-helix in aqueous solution. TFE is known to induce helix formation, therefore, it is possible that partially and transiently forming 3_{10}-helices in rTrp-cage in aqueous solution is stabilized and elongated upon addition of TFE [58].

In line with earlier findings [8], the retro forms of parent-foldable linear peptides investigated here are not foldable. They preserve only some residual secondary structure content, and they fail to stabilize a folded structure. These findings suggest that although the average hydropathy and net charge per residue indices alone work well to separate known foldable and disordered proteins [59], they are not sufficient criteria for predicting disorder in sequences [60].

To date, the equilibrium structure of the retro form of an intrinsically disordered peptide has not been compared to its parent's structure under same solution conditions. In Fig. 1C,D we present a structural comparison of the normal and retro forms of fragments of two intrinsically disordered peptides. As the $A\beta_{16\text{-}20}$ peptide is too short to analyze hydrogen bonding patterns for secondary structure calculation, we show the Ramachandran map of each residue, except for terminal residues, since the latter do not have a complete set of φ and ψ angles (the peptide is uncapped). The angles mostly populate the top-left corner of the Ramachandran map, indicating that the residues are in an extended configuration. A visual inspection of the most populated clusters of the normal and retro forms of the peptide confirms this conclusion (Fig. 1C, right). There are only minor differences between the Ramachandran maps of the normal and retro forms of the $A\beta_{16\text{-}20}$ peptide and an extended backbone is observed in either form. On the other hand, for a partially structured disordered peptide ($p53_{15\text{-}29}$), retro isomerization causes major structural changes, as evident from the per-residue secondary structure analysis (Fig. 1D). The per-residue equilibrium fractional assignment to alpha helix configuration is larger and extends over a longer residue sequence in the $rp53_{15\text{-}29}$, suggesting that retro isomerization results in enhanced ordering, as measured by secondary structure propensities. In the work by Lu and coworkers [55], the retro-inverso form of a p53 fragment was found to have a distinct stable left-handed helical configuration in TFE solution, whereas the parent form of the

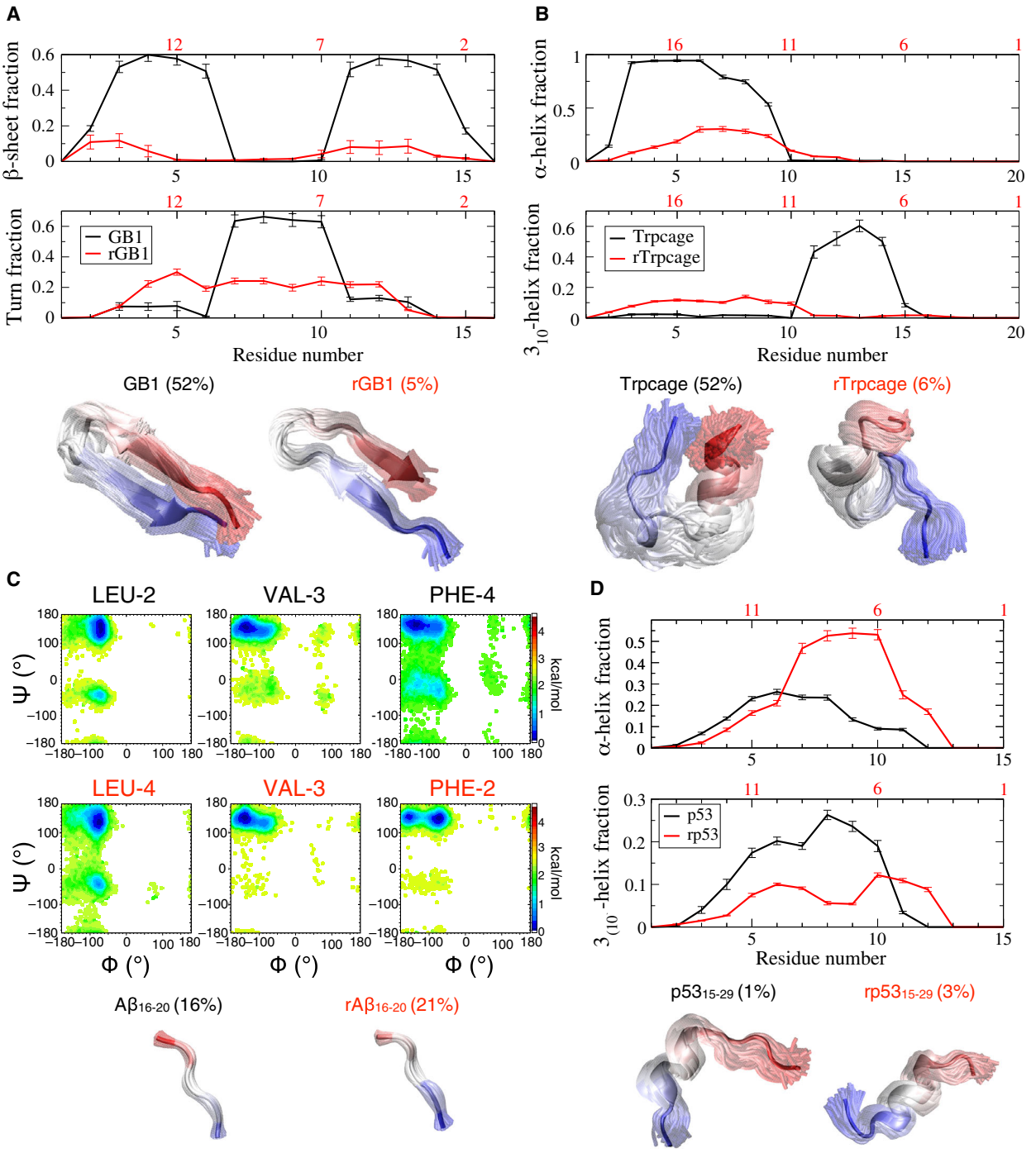

Fig. 1. Comparison between ensemble-average structure-related measures for the retro peptides and their parents (linear peptides). (A) Per-residue fractional assignments to turn and β-sheet configurations for β-hairpin former GB1 and its retro isomer. (B) Per-residue fractional assignments to α and 3_{10} helical configurations for the helical Trp-cage peptide and its retro isomer. (C) Free energy as a function of backbone torsion angles for the three nonterminal residues in the $A\beta_{16-20}$ peptide (top) and its retro isomer (bottom). (D) Per-residue fractional assignments to α- and 3_{10}-helices for the disordered $p53_{15-29}$ and its retro isomer. The most populated clusters of the peptides (both normal and retro, as labeled on top of each image) are illustrated at the bottom of each panel. Reported near the label of each cluster is the percent occurrence of the most populated cluster (see the Methods). Structures are colored according to the amino acid order. The amino acids from N terminus to C terminus (in the normal peptide) are colored from red to blue.

fragment was found to be only partially ordered in the same solution conditions, as measured by CD and NMR spectroscopy, although both the normal and retro-inverso forms were found to be predominantly disordered in an aqueous solution (phosphate buffer). Accordingly, these authors have concluded that an apparent increase in structural order occurs in the retro-inverso form of the peptide compared to the normal form [55]. Our findings showing increased structural order upon retro isomerization of p53 are, therefore, qualitatively consistent with Lu and coworkers' experimental results, insofar as quantitative measures of structural properties that are insensitive to chirality must be the same between chirally inverted peptides (such as retro and retro-inverso forms of the p53 fragment), while the ones sensitive to chirality must be exact mirror images of those of the parents [11,12,61]. Our results emphasize that secondary structure is a strong function of backbone directionality for sufficiently long peptides, even if the secondary structure is only local, as in the $p53_{15\text{-}29}$ peptide.

Cyclic peptides

In addition to the above-discussed linear peptides, we also study cyclic peptides of two different lengths. We select a short cyclic peptide to represent a sequence free of prolines and noncanonical amino acids, as they may impose structural constraints on the backbone [62]. The longer peptide, on the other hand, includes prolines and a disulfide bond.

Ring strain due to backbone cyclization limits the number of configurations that can be adopted by cyclic peptides compared to linear peptides, the constraint being more pronounced for shorter cyclic peptides. Cyclic hexapeptides studied by experiments and simulations [48,63–65] exhibit primarily β-turn configurations. The conformational switches that they show are typically switches between different β-turn types. Accordingly, in this work, we analyze the per-residue fractions of type I, II, I', and II' turns. Turn types are decided according to Ramachandran angles occupied by residues $i + 1$ and $i + 2$ (4-residue turn, indexed from i to $i + 3$) (see Methods for further details). According to these angle criteria, we only find type I turn configurations for the cyclic hexapeptide studied here.

In Fig. 2A, we show a structural comparison between the normal and retro forms for the short cyclic peptide. We show Ramachandran maps for each residue in Fig. S6(A), and we find that the most populated basins are significantly different between the normal and retro forms of the cyc1 peptide. ALA-(4 for normal, 3 for retro), for example, switches from a compact configuration to a highly extended one (cyc1–rcyc1). While both the normal and retro forms of the hexapeptide show a significant fraction of type I turn per-residue (Fig. 2A), the regions involved in the turn differ, implying that retro isomerization does not yield a symmetric transformation for this cyclic peptide.

Similar observations apply to the longer peptide (1JBL): there are remarkable differences between the Ramachandran maps of the individual amino acids. As the peptide is longer, in this case, we calculate the per-residue fraction of secondary structures, α-helix, β-strand, and ppII helix (see the Methods for structural assignment details). We find that the fractions of locally forming secondary structures are quite different between normal and retro forms of the 1JBL, too (Fig. 2B). Our findings are consistent with previous NMR experimental work studying retro, inverso, and retro-inverso analogs of cyclic peptides, which show that the retro peptide conformations are significantly different from the conformations adopted by their parents [66].

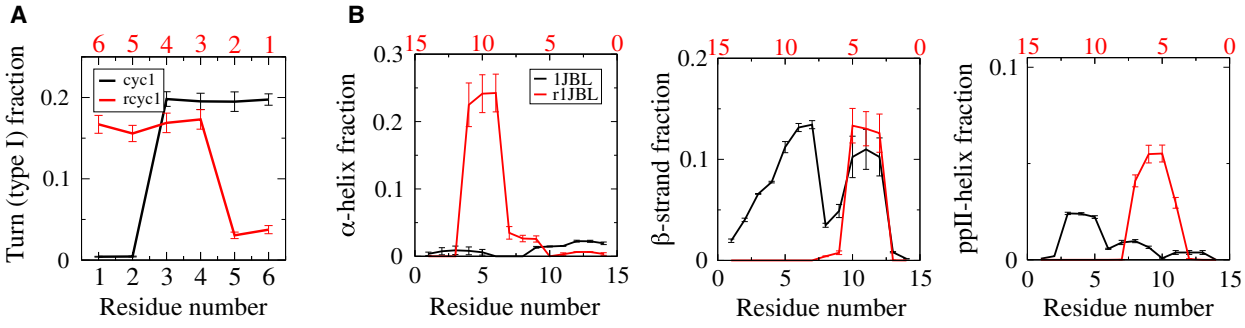

Fig. 2. Comparison between ensemble-average structure-related measures for the retro peptides and their parents (cyclic peptides). (A) Type I turn fraction per residue for the cyc1 peptide as calculated based on the torsion angles (see the text). (B) α-helix, β-strand, and ppII helix fractions per residue for 1JBL peptide as calculated based on the torsion angles (see text). Residue sequence is given from the N to the C terminus in every case.

Conclusion

Although it was initially thought that reversing peptide sequences would lead to a symmetry transformation resulting in comparable structures with respect to the parent peptide [6,7], subsequent evidence has shown that the directionality of the peptide bond is an important contributor to a peptide's unique structural properties [8,9]. Here, we study, for the first time, the equilibrium ensembles of retro peptides in atomistic detail with respect to their parents including linear foldable, linear intrinsically disordered, and cyclic peptides.

In agreement with previous findings [8,10,67], our results showed that reverse reading of the foldable peptides lead to a nonfoldable molecule. However, although these retro peptides fail to stabilize a three-dimensional structure, they preserve their intrinsic preference toward their parent's secondary structure. The transition from a foldable to a nonfoldable peptide is particularly striking, as the average hydrophobicity, charge composition, and even amino acid order are preserved in retro isomers (only the direction differs). These findings imply the need for consideration of backbone directionality for disorder prediction: although the average hydropathy and net charge per residue indices alone well separate known foldable and disordered proteins [59], and they are not the sufficient criteria for predicting disorder [60].

The disordered peptides studied here are ones whose retro-inverso isomers have previously been considered as drug candidates [20,54]. Retro-inverso isomers have been considered as topochemical equivalents to their parents in terms of providing identical side-chain orientation for target binding [1]. However, achieving identical side-chain orientation requires identical backbone structure (Fig. S1). Retro isomerization does not preserve the structure for foldable proteins. Therefore, retro-inverso drug candidates for foldable peptides are, therefore, not likely to be successful in reproducing the binding ability of the parent peptide. However, the extent of structure preservation upon reversing the peptide bond direction has remained an open question for disordered peptides. Here, we find that the retro isomer structure of a short (penta)peptide, $A\beta_{16\text{-}20}$, which has an extended backbone structure in its normal form, is not significantly different from that of its parent peptide. The retro-inverso isomer of this peptide has been found to have a binding affinity to its target similar to the affinity that the normal form of the peptide has toward the same target [20]. On the other hand, for the $p53_{15\text{-}29}$ fragment (although it is predominantly coil-like, it has a considerable fraction of helical structures), we find that the $rp53_{15\text{-}29}$ peptide does not reproduce the same or similar backbone structure. In particular, $rp53_{15\text{-}29}$ has a much longer and larger per-residue fraction of α-helices and, consistently, lower per-residue fraction of 3_{10} helix. We also note that retro-inverso $p53_{15\text{-}29}$ has been found to have much lower binding affinity to its target (MDM2 protein) compared to the binding affinity that $p53_{15\text{-}29}$ has to the same target [55].

It is well understood that total chiral inversion of an amino acid sequence yields mirror image structures; for example, right-handed helical structures become left handed in the same region. Accordingly, failure of retro-inverso drug candidates, when it happens, is attributed to the topological differences between the mirror image structures [19]. However, if the amino acid order reversal is already causing a major change in the structure, total chiral inversion only makes a mirror image of the changed structure, suggesting that retro isomerization in retro-inverso drug design might be the major contributor to the failure of these drugs. Our results, therefore, suggest that checking whether peptide structure is conserved upon retro isomerization would be a particularly useful first test when proposing a retro-inverso drug candidate. The use of peptides whose retro-inverso forms' biochemical efficacy has been previously evaluated experimentally [20,55] has been particularly useful to us in order to arrive at this conclusion.

Acknowledgements

GHZ and PGD acknowledge the support from Unilever R&D. The simulations presented in this work are performed on computational resources managed and supported by Princeton Research Computing, a consortium of groups including the Princeton Institute for Computational Science and Engineering (PICSciE) and the Office of Information Technology's High Performance Computing Center and Visualization Laboratory at Princeton University.

Author contributions

GHZ, FHS, and PGD designed research; GHZ performed research; GHZ, FHS, and PGD analyzed data; and GHZ. wrote the paper with assistance from FHS and PGD.

References

1 Goodman M and Chorev M (1979) On the concept of linear modified retro-peptide structures. *Acc Chem Res* **12**, 1–7.

2 Chorev M and Goodman M (1993) A dozen years of retro-inverso peptidomimetics. *Acc Chem Res* **26**, 266–273.
3 Chorev M and Goodman M (1995) Recent developments in retro peptides and proteinsan ongoing topochemical exploration. *Trends Biotechnol* **13**, 438–445.
4 Adessi C and Soto C (2002) Converting a peptide into a drug: strategies to improve stability and bioavailability. *Curr Med Chem* **9**, 963–978.
5 Xue B, Dunker AK and Uversky VN (2010) Retro-MoRFs: Identifying protein binding sites by normal and reverse alignment and intrinsic disorder prediction. *Int J Mol Sci* **11**, 3725–3747.
6 Guptasarma P (1992) Reversal of peptide backbone direction may result in the mirroring of protein structure. *FEBS Lett* **310**, 205–210.
7 Olszewski KA, Kolinski A and Skolnick J (1996) Does a backwardly read protein sequence have a unique native state? *Protein Eng Des Sel* **9**, 5–14.
8 Lacroix E, Viguera AR and Serrano L (1998) Reading protein sequences backwards. *Folding Des* **3**, 79–85.
9 Mittl PR, Deillon C, Sargent D, Liu N, Klauser S, Thomas RM, Gutte B and Grütter MG (2000) The retro-GCN4 leucine zipper sequence forms a stable three-dimensional structure. *Proc Natl Acad Sci USA* **97**, 2562–2566.
10 Zhang Y, Weber JK and Zhou R (2016) Folding and stabilization of native-sequence-reversed proteins. *Sci Rep* **6**, 25138.
11 Milton RD, Milton S and Kent S (1992) Total chemical synthesis of a D-enzyme: the enantiomers of HIV-1 protease show reciprocal chiral substrate specificity [corrected]. *Science* **256**, 1445–1448.
12 Zerze GH, Khan MN, Stillinger FH and Debenedetti PG (2018) Computational investigation of the effect of backbone chiral inversions on polypeptide structure. *J Phys Chem B* **122**, 6354–6363.
13 Moreira IS, Fernandes PA and Ramos MJ (2007) Hot spotsA review of the protein–protein interface determinant amino-acid residues. *Proteins Struct Funct Bioinf* **68**, 803–812.
14 Shemyakin M, Ovchinnikov YA, Ivanov V & Ryabova I (1967) Topochemical approach to the structure-activity relation. Retroenantio-Gly 5, 10-gramicidin S.. *Experientia* **23** 326.
15 Shemyakin M, Ovchinnikov YA and Ivanov V (1969) Topochemical investigations on peptide systems. *Angew Chem Int Ed Engl* **8**, 492–499.
16 Cheley S, Braha O, Lu X, Conlan S and Bayley H (1999) A functional protein pore with a retro transmembrane domain. *Protein Sci* **8**, 1257–1267.
17 Ben-Yedidia T, Beignon A-S, Partidos CD, Muller S and Arnon R (2002) A retro-inverso peptide analogue of inuenza virus hemagglutinin B-cell epitope 91-108 induces a strong mucosal and systemic immune response and confers protection in mice after intranasal immunization. *Mol Immunol* **39**, 323–331.
18 Li C, Zhan C, Zhao L, Chen X, Lu W-Y and Lu W (2013) Functional consequences of retro-inverso isomerization of a miniature protein inhibitor of the p53-MDM2 interaction. *Bioorg Med Chem* **21**, 4045–4050.
19 Garton M, Nim S, Stone TA, Wang KE, Deber CM and Kim PM (2018) Method to generate highly stable D-amino acid analogs of bioactive helical peptides using a mirror image of the entire PDB. *Proc Natl Acad Sci USA* **115**, 1505–1510.
20 Zhang G, Leibowitz MJ, Sinko PJ and Stein S (2003) Multiple-peptide conjugates for binding β-amyloid plaques of Alzheimer's disease. *Bioconjugate Chem* **14**, 86–92.
21 Taylor M, Moore S, Mayes J, Parkin E, Beeg M, Canovi M, Gobbi M, Mann DM and Allsop D (2010) Development of a proteolytically stable retro-inverso peptide inhibitor of β-amyloid oligomerization as a potential novel treatment for Alzheimers disease. *Biochemistry* **49**, 3261–3272.
22 Roche J, Shen Y, Lee JH, Ying J and Bax A (2016) Monomeric Aβ1–40 and Aβ1–42 peptides in solution adopt very similar Ramachandran map distributions that closely resemble random coil. *Biochemistry* **55**, 762–775.
23 Meng F, Bellaiche MM, Kim J-Y, Zerze GH, Best RB and Chung HS (2018) Highly disordered amyloid-β monomer probed by single-molecule FRET and MD simulation. *Biophysical J* **114**, 870–884.
24 Conicella AE, Zerze GH, Mittal J and Fawzi NL (2016) ALS mutations disrupt phase separation mediated by α-helical structure in the TDP-43 low-complexity C-terminal domain. *Structure* **24**, 1537–1549.
25 Miller C, Zerze GH and Mittal J (2013) Molecular simulations indicate marked differences in the structure of amylin mutants, correlated with known aggregation propensity. *J Phys Chem B* **117**, 16066–16075.
26 Wright PE and Dyson HJ (2009) Linking folding and binding. *Curr Opin Struct Biol* **19**, 31–38.
27 van der Lee R, Buljan M, Lang B, Weatheritt RJ, Daughdrill GW, Dunker AK, Fuxreiter M, Gough J, Gsponer J, Jones DT *et al.* (2014) Classification of intrinsically disordered regions and proteins. *Chem Rev* **114**, 6589–6631.
28 Brooks BR, Brooks CL 3rd, Mackerell AD Jr, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S *et al.* (2009) CHARMM: the biomolecular simulation program. *J Comput Chem* **30**, 1545–1614.
29 Yang C-H, Brown JN and Kopple KD (1981) Crystal structure and solution studies of the molecular

conformation of the cyclic hexapeptide cyclo-(Gly-L-His-Gly-L-Ala-L-Tyr-Gly). *J Am Chem Soc* **103**, 1715–1719.
30 Korsinczky ML, Schirra HJ, Rosengren KJ, West J, Condie BA, Otvos L, Anderson MA and Craik DJ (2001) Solution structures by 1H NMR of the novel cyclic trypsin inhibitor SFTI-1 from sunower seeds and an acyclic permutant1. *J Mol Biol* **311**, 579–591.
31 Best RB and Mittal J (2010) Protein simulations with an optimized water model: cooperative helix formation and temperature-induced unfolded state collapse. *J Phys Chem B* **114**, 14916–14923.
32 Abascal JL and Vega C (2005) A general purpose model for the condensed phases of water: TIP4P/2005. *J Chem Phys* **123**, 234505.
33 Best R and Hummer G (2009) Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. *J Phys Chem B* **113**, 9004–9015.
34 Zerze GH, Best RB and Mittal J (2015) Sequence-and temperature-dependent properties of unfolded and disordered proteins from atomistic simulations. *J Phys Chem B* **119**, 14622–14630.
35 Kim SB, Palmer JC and Debenedetti PG (2016) Computational investigation of cold denaturation in the Trp-cage miniprotein. *Proc Natl Acad Sci USA* **113**, 8991–8996.
36 Best RB and Mittal J (2011) Free-energy landscape of the GB1 hairpin in all-atom explicit solvent simulations with different force fields: similarities and differences. *Proteins Struct Funct Bioinf* **79**, 1318–1328.
37 Zerze GH, Uz B and Mittal J (2015) Folding thermodynamics of β-hairpins studied by replicaexchange molecular dynamics simulations. *Proteins Struct Funct Bioinf* **83**, 1307–1315.
38 Nosé S (1984) A molecular dynamics method for simulations in the canonical ensemble. *Mol Phys* **52**, 255–268.
39 Hoover WG (1985) Canonical dynamics: equilibrium phase-space distributions. *Phys Rev A* **31**, 1695.
40 Sugita Y and Okamoto Y (1999) Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett* **314**, 141–151.
41 Bonomi M and Parrinello M (2010) Enhanced sampling in the well-tempered ensemble. *Phys Rev Lett* **104**, 190601.
42 Deighan M, Bonomi M and Pfaendtner J (2012) E_cient simulation of explicitly solvated proteins in the well-tempered ensemble. *J Chem Theory Comput* **8**, 2189–2192.
43 Berendsen HJ, van der Spoel D and van Drunen R (1995) GROMACS: A message-passing parallel molecular dynamics implementation. *Comput Phys Commun* **91**, 43–56.
44 Hess B, Kutzner C, Van Der Spoel D and Lindahl E (2008) GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* **4**, 435–447.
45 Bonomi M, Branduardi D, Bussi G, Camilloni C, Provasi D, Raiteri P, Donadio D, Marinelli F, Pietrucci F, Broglia R *et al.* (2009) PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Comput Phys Commun* **180**, 1961–1972.
46 Essmann U, Perera L, Berkowitz ML, Darden T, Lee H and Pedersen LG (1995) A smooth particle mesh Ewald method. *J Chem Phys* **103**, 8577–8593.
47 Kabsch W and Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637.
48 McHugh SM, Rogers JR, Yu H and Lin Y-S (2016) Insights into how cyclic peptides switch conformations. *J Chem Theory Comput* **12**, 2480–2488.
49 Gierasch LM, Deber CM, Madison V, Niu C-H and Blout ER (1981) Conformations of (X-L-Pro-Y)2 cyclic hexapeptides. Preferred β-turn conformers and implications for β-turns in proteins. *Biochemistry* **20**, 4730–4738.
50 Daura X, Gademann K, Jaun B, Seebach D, van Gunsteren WF and Mark AE (1999) Peptide folding: when simulation meets experiment. *Angew Chem Int Ed* **38**, 236–240.
51 Neidigh JW, Fesinmeyer RM and Andersen NH (2002) Designing a 20-residue protein. *Nat Struct Biol* **9**, 425.
52 Blanco FJ, Rivas G and Serrano L (1994) A short linear peptide that folds into a native stable β-hairpin in aqueous solution. *Nat Struct Biol* **1**, 584.
53 Mittal J, Yoo TH, Georgiou G and Truskett TM (2012) Structural ensemble of an intrinsically disordered polypeptide. *J Phys Chem B* **117**, 118–124.
54 Sakurai K, Chung HS and Kahne D (2004) Use of a retroinverso p53 peptide as an inhibitor of MDM2. *J Am Chem Soc* **126**, 16288–16289.
55 Li C, Pazgier M, Li J, Li C, Liu M, Zou G, Li Z, Chen J, Tarasov SG, Lu W-Y *et al.* (2010) Limitations of peptide retro-inverso isomerization in molecular mimicry. *J Biol Chem* **285**, 19572–19581.
56 Munoz V, Thompson PA, Hofrichter J and Eaton WA (1997) Folding dynamics and mechanism of β-hairpin formation. *Nature* **390**, 196.
57 Streicher WW and Makhatadze GI (2007) Unfolding thermodynamics of Trp-cage, a 20 residue miniprotein, studied by differential scanning calorimetry and circular dichroism spectroscopy. *Biochemistry* **46**, 2876–2880.
58 Vymětal J, Bathula SR, Cerný J, Chaloupková R, Zídek L, Sklenář V and Vondrášek J (2014) Retro operation on the Trp-cage miniprotein sequence produces an unstructured molecule capable of folding similar to the original only upon 2, 2, 2-triuoroethanol addition. *Protein Eng Des Sel* **27**, 463–472.

59 Uversky VN (2002) Natively unfolded proteins: a point where biology waits for physics. *Protein Sci* **11**, 739–756.
60 Kutyshenko VP, Prokhorov DA, Molochkov NV, Sharapov MG, Kolesnikov I and Uversky VN (2014) Dancing retro: solution structure and micelle interactions of the retro-SH3-domain, retro-SHH-Bergerac. *J Biomol Struct Dyn* **32**, 257–272.
61 Zerze GH, Stillinger FH and Debenedetti PG (2019) Effect of heterochiral inversions on the structure of a β-hairpin peptide. *Proteins: Struct, Funct, Bioinf* **87**, 569–578.
62 Kessler H, Anders U and Schudok M (1990) An unexpected cis peptide bond in the minor conformation of a cyclic hexapeptide containing only secondary amide bonds. *J Am Chem Soc* **112**, 5908–5916.
63 Matter H and Kessler H (1995) Structures, dynamics, and biological activities of 15 cyclic hexapeptide analogs of the. alpha.-Amylase Inhibitor Tendamistat (HOE 467) in Solution. *J Am Chem Soc* **117**, 3347–3359.
64 Kessler H, Gratias R and Hessler G (1996) Gurrath M.; Müller G. Conformation of cyclic peptides. Principle concepts and the design of selectivity and superactivity in bioactive sequences by 'spatial screening'. *Pure Appl Chem* **68**, 1201–1205.
65 McHugh SM, Yu H, Slough DP and Lin Y-S (2017) Mapping the sequence–structure relationships of simple cyclic hexapeptides. *Phys Chem Chem Phys* **19**, 3315–3324.
66 Wermuth J, Goodman S, Jonczyk A and Kessler H (1997) Stereoisomerism and biological activity of the selective and superactive αvβ3 integrin inhibitor cyclo(-RGDfV-) and its retro-inverso peptide. *J Am Chem Soc* **119**, 1328–1335.
67 English LR, Tischer A, Demeler AK, Demeler B and Whitten ST (2018) Sequence reversal prevents chain collapse and yields heat-sensitive intrinsic disorder. *Bio-Phys J* **115**, 328–340.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Details of parallel-tempering in well-tempered ensemble simulations of peptides.

Fig. S1. Side-chain orientation equivalence of an L-peptide, alanine-tetrapeptide (A) and the retro-inverso analog of the peptide (B).

Fig. S2. Radius of gyration of each linear peptide as a function of time (black) and cumulative average of radius of gyration as a function of time (red) at 300 K.

Fig. S3. Radius of gyration of each cyclic peptide as a function of time (black) and cumulative average of radius of gyration as a function of time (red) at 300 K.

Fig. S4. Ramachandran map showing the angle criteria (for i + 1 and i + 2 locations) of forming a type I turn.

Fig. S5. Ramachandran map showing the angle criteria to form α (blue), β (magenta) and ppII (green) structures.

Fig. S6. Free enery as a function of backbone torsion angles for the cyc1 peptide (A) and the 1JBL peptide (B) Turn, α-helix, β-strand and ppII helix fractions per residue (Figure 2) are calculated based on these torsion angles (see the main text).