J Comput Aided Mol Des
DOI 10.1007/s10822-008-9241-4

A virtual library of constrained cyclic tetrapeptides that mimics all four side-chain orientations for over half the reverse turns in the protein data bank

Sage Arbor · Garland R. Marshall

Received: 30 April 2008 / Accepted: 2 September 2008

Abstract Reverse turns are often recognition sites for protein/protein interactions and, therefore, valuable potential targets for determining recognition motifs in development of potential therapeutics. A virtual combinatorial library of cyclic tetrapeptides (CTPs) was generated and the bonds in the low-energy structures were overlapped with canonical reverse-turn Cα–Cβ bonds (Tran et al., J Comput Aided Mol Des 19(8):551–566, 2005) to determine the utility of CTPs as reverse-turn peptidomimetics. All reverse turns in the Protein Data Bank (PDB) with a crystal structures resolution $\leq$3.0 Å were classified into the same known canonical reverse-turn Cα–Cβ bond clusters (Tran et al., J Comput Aided Mol Des 19(8):551–566, 2005). CTP reverse-turn mimics were compiled that mimicked both the relative orientations of three of the four as well as all four Cα–Cβ bonds in the reverse turns of the PDB. 54% of reverse turns represented in the PDB had eight or more CTPs structures that mimicked the orientation of all four of the Cα–Cβ bonds in the reverse turn.

Keywords Cyclic tetrapeptide · Reverse turn · Mimic · Conformational search

Introduction

Protein structures are comprised of the three ordered secondary structures—helices, sheets, and reverse turns. Of these, reverse turns are particularly interesting targets to mimic for therapeutics due to their location at the perimeter of globular proteins and their common role in macromolecular recognition, especially in ligands for G-protein coupled receptors [1, 2]. Reverse turns are usually defined by a four-residue sequence in which the main chain makes an $\sim$180° turn with the carbon alphas of the i and i + 3 residue within 7 Å. A subset of reverse turns, β-turns, has a hydrogen bond between carbonyl of the i and amide NH of the i + 3 residues. Turns have been found critical in a range of macromolecular recognition motifs, such as integrin or interferon-γ binding in biological systems ranging from cell adhesion [3] to antiviral agents [4], respectively. There has, therefore, been considerable effort to mimic reverse turns which has been accomplished by short linear peptides [5, 6], cyclic peptides [7–11], and small molecules [7–19] such as benzodiazepines [9, 12–16], β-turn dipeptides (BTD) [20], indolizinoindole β-turn mimetic (IBTM) [20], and γ and β-lactams [19]. The use of proline has long been known to help a peptide adopt a reverse-turn conformation [8]. For example, the classical type VI turn is defined as having a *cis*-amide bond between residue i + 1 and i + 2, which proline facilitates in the i + 2 position due to its substituted tertiary nitrogen. The sequence D-pro-L-Pro in particular has been found to adopt a reverse turn conformation [21–23], and an alternating D/L amino acid sequence has also been found to facilitate cyclization of small peptides [23]. In addition, cyclic tetrapeptides (CTPs) have been found to be inherently good mimics of reverse turns due to their $\sim$180° turn in all four of the sequential four-residue segments. Combination of alternating D and L prolines in CTPs have been found to be good reverse-turn mimetics [7, 24] and also more synthetically feasible to cyclize head-to-tail due to the alternating D- and L-proline propensity to stabilize reverse turns which causes the backbone of the linear peptide to turn back upon itself.

S. Arbor (✉) · G. R. Marshall
Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, MO 63110, USA
e-mail: sagearbor@gmail.com

As therapeutics cyclic tetrapeptides have the added benefit of potentially being bioavailable [25].

A bottleneck in developing therapeutics is the fact that the structure-activity relationship (SAR) investigations are often started de novo for each molecular interaction of interest, for example, by high-throughput screens. A more efficient approach to drug design involves classifying privileged chemical structures so that the search for mimetics can start from a set of scaffolds that correctly orient chemical substituents to reproduce one of the molecular surfaces involved in protein/protein recognition. For example, once a new turn of potential therapeutic interest is classified, then the set of preexisting turn mimics for that turn class can quickly be tested as a mimetic. In addition, unknown turn recognition motifs can more quickly be determined by screening against a small set of privileged scaffold compounds that represent all possible turn space as compared to a much larger library of compounds that cover conformational space less exhaustively. Historically, reverse turns have been classified by the ϕ and Ψ backbone torsional angles of the i + 1 and i + 2 residues. However, the biologically recognized moiety in reverse turns are the surfaces generated by the four side chains of the turn, not the conformation of the backbone chain which only indirectly impacts the recognition surface. This historical classification of turns does not specify the Ψ_i and ϕ_{i+3} torsional angles of the turn which dictates the relative orientations of the first and fourth side chain. It has recently been shown the relative orientation of all four side chains can more accurately be classified by grouping the four Cα–Cβ bonds in reverse turns into nine clusters [26]. This paper explores the conformational space of CTPs and classifies them as mimics for these nine β-turn clusters. All reverse turns in the recent PDB are then likewise classified into the same nine clusters; thereby, elucidating CTPs that can immediately be used as lead compounds in a SAR mimetic search of that given reverse turn.

Methods

Conformational searches of cyclic tetrapeptides (CTPs)

All conformational searches were done using Macromodel 7.2 [27] software by 10,000 conformational steps using Monte-Carlo Multiple Minimum (MCMM) conformational search method [28, 29]. All conformations within 200 kJ of the lowest energy conformer were retained to verify sufficient sampling of CTPs conformational space. All 16 possible *cis*/*trans*-amide bond macrocyclic conformations were found multiple times which suggests that conformational space was adequately searched. Conformations were found that had identical *cis*/*trans* amide bond patterns but where an amide bond was flipped in one conformation relative to another conformation, thereby giving a different dipole. While these two conformations had the same Cα–Cβ orientations their differing dipoles could cause different binding affinities if a turn mimetic was created from them. Such instances were therefore screened for and both conformations considered unique when determining $\Delta\Delta G$ between the lowest energy conformer and the second lowest energy conformer, or listing the number of conformations found within 2 kcal of lowest energy structure (see detailed CTP library results at http://cmd.wustl.edu/sagearbor/CTPlib/). Calculations were done both in vacuo and in water using the GB/SA implicit solvation model [30]. During conformational searches, all 12 bonds in each CTPs macrocycle (including the four amide bonds) were allowed to rotate. For CTPs with proline or pipecolic rings, the low energy conformer of the macrocycle was then locked and one or two of the side-chain ring torsions, respectively, were then allowed to rotate in a second MCMM search. The puckering of the rings was not found to significantly affect the CTPs conformation of the backbone macrocycle (data not shown), but any preferred ring conformer of proline and pipecolic rings for a given macrocycle conformation was elucidated for those cases in which these side chain bonds were found to mimic Cα–Cβ reverse-turn bonds.

The Cambridge Structural Database (CSD) CTPs were conformationally searched using the seven different force fields: Amber*, Amber94, MM2*, MMFF94, MMFF94*, OPLS2001, and OPLS2005 as implemented by Macromodel 7.2 [27]. For each run, charges were applied using the corresponding force-field methodology. Calculations for the CSD CTPs were done in vacuo, chloroform, and water (Fig. 1a). The CTP combinatorial library was made by every non-redundant combination of amino acids: glycine, L-alanine, D-alanine, L-proline, D-proline, *N*-methyl-L-Alanine, *N*-methyl-D-alanine, L-pipecolic acid, and D-pipecolic acid. This resulted in 1,665 unique sequences of CTPs. Both vacuum and water MCMM conformational searches were run on each of the 1,665 unique CTP sequences resulting in a library of 3,330 sets of low-energy confomers.

Extraction of PDB turns

All crystal structures in the PDB as of August 2007 with a resolution of less than 3 Å were analyzed for reverse-turn conformers. Extraction of turns from the PDB was done with POSSE software [31] and a script created 'extractturns', which can be attained in current POSSE distribution (http://ccb.wustl.edu/~ewelsh/). Turns were specified as four residues with the first and fourth Cα's within 7 Å, i and i + 1 Cα's within 4.4 Å (to avoid finding chain breaks that when put together resembled turns), and where the

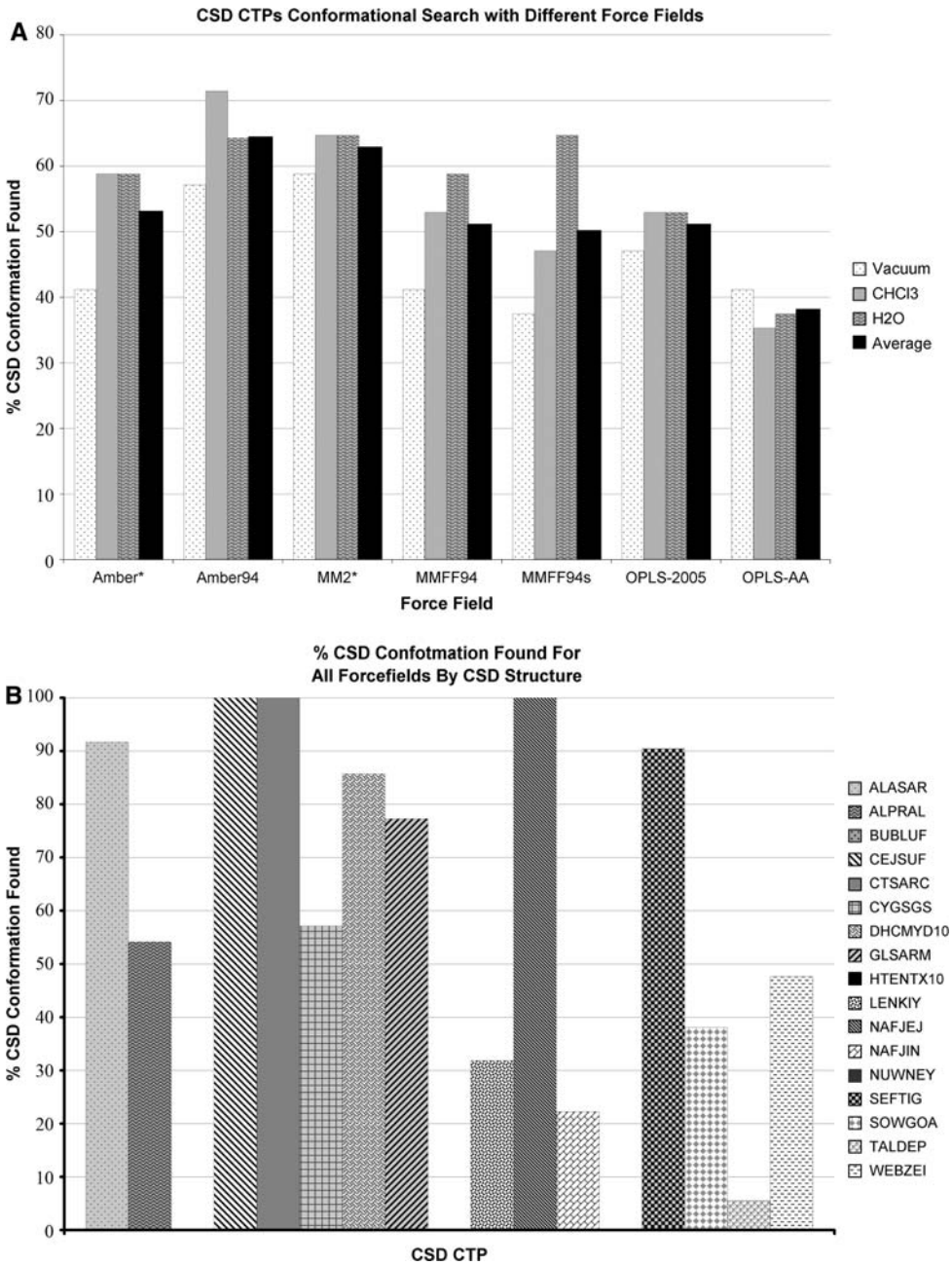

Fig. 1 Cyclic tetrapeptides (CTPs) with known crystal structures deposited in the Cambridge Structural Database (CSD) were conformational explored to measure the predictive capability of various molecular mechanics force fields. Calculations were done with vacuum, chloroform, and water solvation models, with the average results shown as % correct for each force field (**a**). The combined predictive capability for each CSD CTP averaging all force field results was also examined (**b**)

second and third residues were classified as turns according to POSSE [31] (thus avoiding finding helices). Each turn had its four Cα–Cβ bonds saved as its own PDB file and then converted to a mol2 file using the program Babel [32]. These mol2 files were then used with the program FOUNDATION [33] and converted into a database to determine which of the nine Tran et al. [26] reverse-turn Cα–Cβ clusters best overlapped the four Cα–Cβ vectors.

Overlaps

Tran et al. [26] established nine clusters of four vectors that were found to represent 90% of high-resolution reverse turns for non-homologous proteins in the PDB. PDB turns were mapped to these nine clusters by overlapping and assigning each turn to a cluster based on the lowest RMSD. CTPs had all of their bonds searched for low RMSD overlaps with the four vectors of these nine clusters in order to establish them as possible reverse-turn mimetics. Two methods were used for overlapping with the nine Tran et al. reverse-turn clusters. The Cα–Cβ in residue i, i + 1, i + 2, and i + 3 of the PDB reverse turns were overlapped with vector 1, 2, 3, and 4, respectively, in each Tran et al. cluster using Sybyl 7.3 [34] and assigned to the class with the lowest RMSD or, following precedent, left unassigned if the lowest RMSD was greater than 0.65 Å. All CTPs

were overlapped with the nine reverse turn Cα–Cβ Tran et al. clusters using the program FOUNDATION [33] at a 0.3, 0.5, and 0.7 Å RMSD cutoff. In all cases the nine Tran clusters were used as queries. The original nine Tran reverse-turn clusters, which had a carbon–carbon length (1.52 Å) were combinatorially modified to have carbon–hydrogen lengths (1.10 Å) as well as retaining the original carbon–carbon lengths. This yielded 16 groups of four vectors for each of the nine original Tran et al. clusters. All 144 of these sets of four vectors were used as queries in the program FOUNDATION to overlap with the database of the 3,330 low-energy CTP structures in order to elucidate CTPs that could serve as reverse-turn mimetics. This enabled elucidation of both C–H and C–C as bonds in the CTP molecule that could be modified to mimic the Cα–Cβ in a reverse turn. Searches were done in which all the bonds in the CTPs were overlapped with three of four as well as with all four of the Cα–Cβ Tran et al. cluster bonds. In all these methods each amino acid in the CTPs could be found to act as an overlap for any of the reverse turn Cα–Cβ bonds (from i, i + 1, i + 2, or i + 3 residue).

Results and discussion

PDB reverse turns

Reverse turns have classically been described by the ϕ, ψ torsions of the i + 1 and i + 2 residues in the turn. However, Tran et al. [26] recently showed that the topography of all four side chains relevant to molecular recognition is better described by clustering reverse turns based on the orientation of their Cα–Cβ bonds. We screened the Protein Data Bank (PDB) for reverse turns as defined by a four-residue sequence that had the $C\alpha_i$ and $C\alpha_{i+3}$ within 7 Å of each other, as well as not being part of a helical region using secondary structure assignment portions of the program POSSE [31]. The Cα–Cβ bonds of the i, i + 1, i + 2, and i + 3 PDB residues were extracted and overlapped with corresponding four Cα–Cβ bonds in the nine Trans et al. clusters. Following precedent [26], each turn was thereby classified into the Tran et al. cluster with lowest RMSD, or tagged as unclassified if the RMSD >0.65 Å (Fig. 2). 54% of the turns in the PDB with a resolution ≤3.0 Å were clustered into one of the nine Tran et al. [26] clusters (Table 1 Row 1) which is significantly lower than the results of Tran et al. which found approximately 90% of reverse turns to be described by the one of the nine clusters. A sample of pdb structures were manually inspected but no trends, such as turn location relative to α-helix or length of loop containing turn, were seen which correlated with the probability of a good overlap between PDB turns and the nine Tran turn clusters.

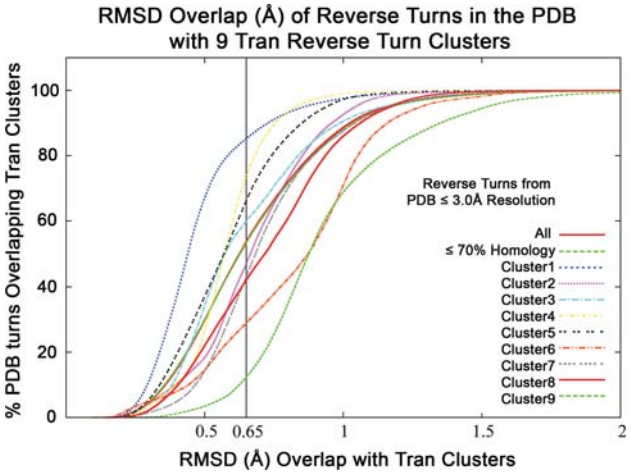

Fig. 2 All the PDB structures with a resolution <3Å had their reverse turns Cα–Cβ vectors overlapped with the nine Tran et al. reverse-turn Cα–Cβ clusters showing 54% PDB turn overlap within 0.65 Å RMSD. A subset of these turns with <70% homology was shown to give similar results. The PDB turns were assigned to one of the nine Tran et al. clusters by the best RMSD ≤0.65 Å or left unassigned. The reverse turns in the PDB are shown split into categories by which Tran clusters they overlap with best (1–9). The percent of PDB turns for each category that overlap a Tran Cluster at a given RMSD is shown. The PDB turns overlapped each turn cluster with varying tightness, cluster one clearly having the best overlap and cluster nine the worst

This discrepancy with Trans results may be due to the fact that in our analysis the lower 3 Å resolution turns were also screened as opposed to Trans more strict criteria of less than 2 Å resolution turns. If these nine clusters truly describe most of the allowed Cα–Cβ space of reverse turns in proteins then they could be used to improve the models of low quality crystal structure turns. For each PDB entry, the entire crystal lattice was screened for reverse turns. This meant that homodimers with identical monomer structures, for example, had each turn found twice and counted as separate turns. This skewed any statistical results by overweighting crystal structures with more monomers per unit cell assuming similar/identical monomer turn conformations. This was done because the desired result was a database in which any turn of scientific interest could be found in our database, whereas slight differences in monomer conformations in the crystal lattice might yield some turns that were mimicked and others not. Therefore, percentages reported in this paper are accurate for the set of turns in the PDB, or subsets discussed, but may be skewed by crystallographic sampling if considered canonically.

Validation of cyclic tetrapeptide conformational prediction

CTPs have been found useful as reverse-turn mimetics [7, 24]. In order to determine if the energy landscape of

Table 1 The clustering of PDB reverse turns into the nine Tran et al. clusters broken down by cluster (row 1)

	% of PDB turns or CTPs with RMSD ≤0.65 Å in cluster (# of PDB turns or CTPs with RMSD ≤0.65 Å in one of nine clusters)									
	1–9 All	1	2	3	4	5	6	7	8	9
PDB turns ≤3 Å Res (1,036,864)	54.44 (564,429)	16.00 (165,935)	3.88 (40,254)	4.40 (45,667)	6.75 (70,004)	8.73 (90,536)	3.07 (31,782)	5.79 (59,992)	4.56 (47,230)	1.26 (13,029)
CTP library 4 Vec OverLap all atoms <0.5 Å RMSD (3,330)	136.40 (4,542)	14.95 (498)	19.40 (646)	13.21 (440)	20.66 (688)	16.07 (535)	14.68 (489)	13.51 (450)	10.39 (346)	13.51 (450)
CTP library 4 Vec OverLap C–C & C–H atoms <0.5 Å RMSD (3,330)	13.18 (439)	0.24 (8)	0.78 (26)	2.28 (76)	2.64 (88)	1.62 (54)	0.72 (24)	3.18 (106)	1.41 (47)	0.30 (10)
CTP library 3 Vec OverLap all atoms <0.3 Å RMSD (3,330)	410.81 (13,680)	46.70 (1,555)	47.75 (1,590)	50.30 (1,675)	45.44 (1,513)	45.14 (1,503)	46.88 (1,561)	36.79 (1,225)	44.44 (1,480)	47.39 (1,578)
CTP library 3 Vec OverLap C–C & C–H atoms <0.3 Å RMSD (3,330)	70.06 (2,333)	2.91 (97)	7.66 (255)	12.91 (430)	11.71 (390)	4.77 (159)	9.93 (331)	10.00 (332)	8.26 (275)	1.92 (64)

54% of turns in the PDB were clustered within 0.65 Å RMSD. Overlaps of the nine Tran et al. clusters of four Cα–Cβ bonds of reverse turns were compared to the CTP low-energy conformations. Searches were done, with a 0.5 Å RMSD cutoff, in which all four Cα–Cβ bonds were overlapped with all atoms in the CTPs (row 2) and where just C–C and C–H bonds were overlapped (row 3). Likewise three out of four Cα–Cβ bonds in the Tran et al. clusters were overlapped, with a more stringent 0.3 Å RMSD, with all atoms in the CTPs (row 4) and where just C–C and C–H bonds were overlapped (row 5). Percentages of PDB turns or CTPs in the library are shown, with absolute number of turns or CTP structures shown in parentheses

CTPs could be determined computationally with standard force fields, an analysis of known crystalline CTPs from the Cambridge Structural Database (CSD) was performed. CTPs conformations are largely determined by the *cis* or *trans* conformation of the four-amide bonds in their macrocyclic ring. Finding all four amide-bonds conformations in the CSD was used as a metric for the ability to correctly predict CTPs structures. Seventeen CTPs from the CSD were conformationally searched by the Monte Carlo Multiple Minimum (MCMM) method with the seven different force fields Amber*, Amber94, MM2*, MMFF94, MMFF94s, OPLS-2005, and OPLS-AA as implemented by Macromodel 7.2 [21] (Fig. 1a, b). The force fields ranged in their CTP predictive capability from 38 to 64% correct (Fig. 1a). Conformational searches were done in vacuum, chloroform, and water for all force fields. The water calculations most often predicted the correct conformation, with chloroform performing the second best, and vacuum calculations the worst (Fig. 1a). This is consistent with the crystal environment being a more hydrophilic environment. The amino acids makeup of the CTPs was also examined to see if specific amino acids coincided with the CTPs percentage conformational predictive capability. There were clearly some CTPs from the CSD that were 'easier' or 'harder' for all the force fields to accurately predict their conformations (Fig. 1b), and while some CTPs were predicted correctly by all force fields (e.g. CSD code CEJSUF) and others were not predicted correctly by any force fields (e.g. CSD code BULBUF), this did not correlate with inclusion of specific amino acids in the sequence. It is possible that certain dipeptide or tripeptide sequences were more and less difficult than others for current force fields to accurately assign relative conformational energies; however, this was not investigated in detail. TALDEP (CSD code), which was the only CTP in the CSD containing a pipecolic acid, only had its crystal structure predicted correctly as the low-energy conformation by one force field, OPLS2005. Because pipecolic acid was a building block in our combinatorial library, and the OPLS2005 force field was found to have a respectable 51% success rate on predicting the CSD CTPs conformation, this force field was used for predicting the low-energy conformations of our in silico-generated library of CTPs. The Amber94 force field which had the best all around predictive ability of 64% was not used because the correct conformation for the CTP TALDEP with a pipecolic acid residue was incorrectly found to be more than 15 kcal from the lowest energy structure.

Cyclic tetrapeptide combinatorial library

A cyclic tetrapeptide (CTP) library was created by combinatorially combining the nine amino acids: glycine, L-alanine, D-alanine, L-proline, D-proline, *N*-methyl-L-alanine, *N*-methyl-D-alanine, L-pipecolic acid, and D-pipecolic acid (Fig. 3). These residues were used due to their differing conformational tendencies in an attempt to cover as much CTP conformational space as possible. Glycine and proline are unique in their main chain structure, occupying unique phi/psi torsion space, compared to the other 18 natural amino acids which have a Ramachandran plot similar to alanine. This is due to glycines lack of substitution at the Cα and prolines exocyclic ring and tertiary substituted nitrogen. Pipecolic acid and *N*-methyl-alanine are more similar to proline than glycine or alanine in that they both have tertiary substituted nitrogens, but are unique enough due to relaxation of the five-membered exocylic ring to be included in the library; pipecolic acid has a more relaxed six-membered ring while *N*-methyl-alanine lacks this ring constraint altogether compared to proline. An analysis of the ϕ/Ψ space of each amino acid found the library to mainly adopt allowed conformational space for each residue (Fig. 3). Glycine being the most flexible of amino acids was expected to show a decreased sampling of conformational space due to cyclization of the tetrapeptide, and was found to have a much more limited range of ϕ torsions centered tightly around $-115°$ and $115°$, with two obvious main conformations centered around ϕ/Ψ of $-115°/85°$ and $115°/-85°$. The L- and D-stereoisomers for all other amino acids were found to have mirror-image Ramachandran (ϕ vs. Ψ) plots, as expected indicating that conformational space was equally sampled for both enantiomeric residues (Fig. 3).

The combination of these nine amino acids resulted in 1,665 unique CTPs (i.e., not identical by any symmetric operations). These CTPs were conformationally searched by MonteCarlo molecular mechanics in both vacuum and water (GBSA implicit solvation), yielding 3,330 sets of low-energy conformations (1,665 in vacuum and 1,665 in water). Both vacuum and water searches were performed in an effort to predict which conformers CTP scaffolds would adopt in binding to either a hydrophobic cleft, or a more hydrophilic surface site. Extrapolating from the CSD CTP conformational search results (see section above), half of these predicted conformations can be expected to be accurate. The in silico CTP library was compared to the known crystal structure CTPs from the CSD to see if a similar amide bond conformational distribution was found (Table 2). While crystal structure CTPs were found to adopt a alternating *cis–trans–cis–trans* conformation in 76.5% of cases, the in silico library only predicted this conformation 26.1% of the time. This may be due to molecular mechanics force fields high energetic penalty for a *cis* amide bond which is parameterized on a less strained linear peptide unit. Indeed, both the all trans and tttc conformations were found to be predicted more often than

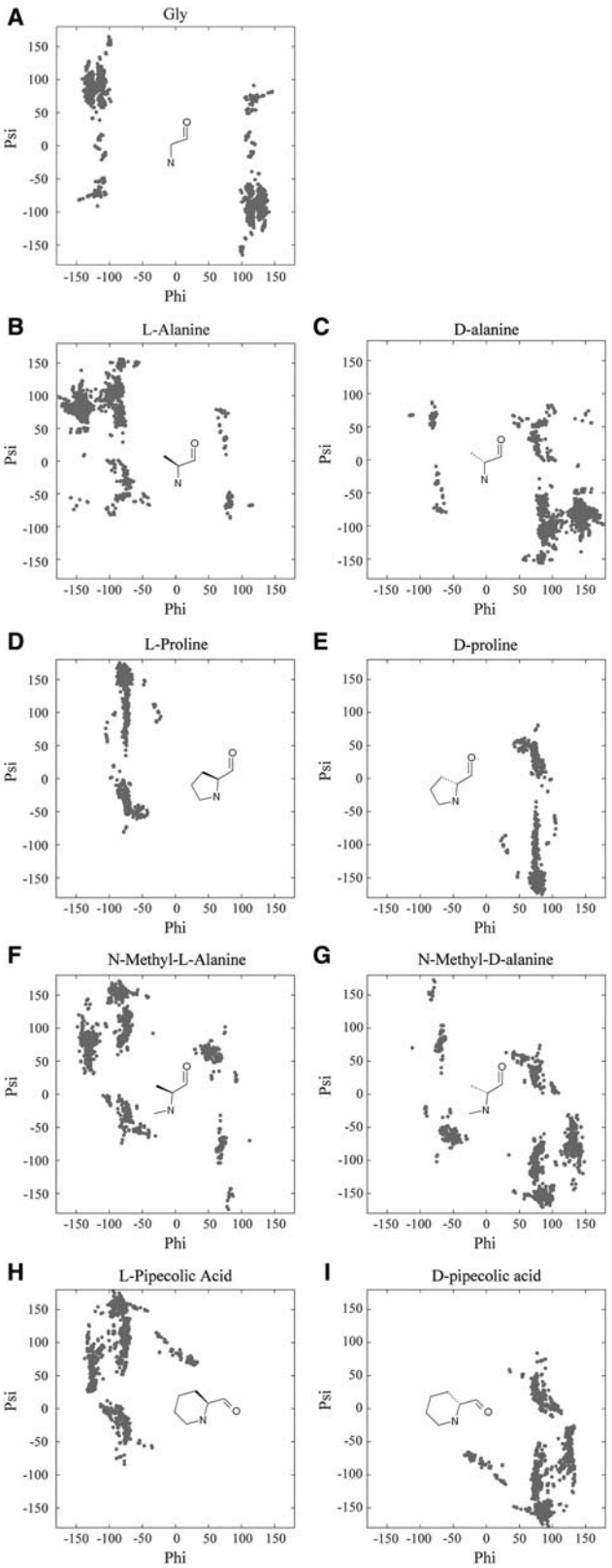

Fig. 3 Ramachandran plots showing the ϕ and Ψ values for each amino acid type Glycine (**a**), L-alanine (**b**), D-alanine (**c**), L-proline (**d**), D-proline (**e**), *N*-methyl-L-alanine (**f**), *N*-methyl-D-alanine (**g**), L-pipecolic acid (**h**), and D-pipecolic acid (**i**) in the cyclic tetrapeptide library

found in the limited CSD data set (Table 2). 56% (926) and 33% (549) of the CTPs in the library, in vacuum and water, respectively, were found to have a single low-energy conformation with more than a 2 Kcal/mol $\Delta\Delta G$ difference from the second lowest-energy conformation.

Each CTP had all of it bonds exhaustively searched to overlap the nine Tran et al. reverse-turn clusters using the program Foundation [33]. Overlaps were done searching all bonds in the CTPs or just C–C and C–H. Searching just the C–C and C–H bonds was attempted to limit overlaps with the macrocycle or carbonyl bonds for which side chain mimic modifications are less synthetically feasible. However, the overlaps that searched all atoms in the CTP molecule found synthetically useful *N*-methyl in *N*-methyl-alanine, N–Cδ in prolines, and N–Cϵ bonds in pipecolic acids that yielded many more overlaps. When just C–C and C–H bonds were overlapped, each of the nine reverse-turn clusters were mimicked ($\leq$0.5 Å RMSD) by at least eight CTPs, and by an average of 49 CTPs, in the CTP library (Table 1 Row 3). Assuming half have the correct low-energy conformation predicted, as the control CSD structures analysis revealed, this would correlate into a minimum of 4, and an average of 25, CTP mimics for each reverse-turn cluster. The number of CTPs reported in Table 1 that mimic a reverse-turn cluster are non-redundant, meaning that if one CTP can mimic the same Tran et al. cluster using different bonds, it is still only counted once in Table 1. Notice that the total number of CTPs that mimic the nine Tran et al. clusters when all bonds are overlapped, 4,542, exceeds 100% (136%) of the entire CTP library of 3,330 structures. A single CTP scaffold can mimic more than one Tran et al. cluster by overlapping different bonds in its structure to the four Cα–Cβ vectors of the reverse-turn bonds in the nine clusters. In addition, a single CTP conformer could overlap in between two of the Tran et al. turn clusters, but be within 0.5 Å RMSD of each, in which case it might not be a good choice as a unique mimic for either cluster. In Table 1, the theoretical maximum of 900% for the CTP rows would indicate each CTP overlapped with all nine clusters for the given Å RMSD, while 100% indicates each CTP overlapping, on average, just one of the nine clusters.

Many reverse turns contain a glycine that allows the main chain to fold back on itself due to glycines higher degree of flexibility. If mimicking one of these turns, it would only be important for the privileged scaffold to cover the tertiary space of the three Cα–Cβ bonds. With that in mind, a screen was done where only three of the four Cα–Cβ bonds in the nine Tran et al. clusters were required to overlap the bonds in the CTPs, but with a much tighter fit of 0.3 Å RMSD. Again screens were done where all bonds in the CTP were overlapped and just C–C and C–H bonds were searched. When just C–C and C–H bonds

Table 2 The four amide bond conformations from the in silico CTP library (3,330 CTPs) were compared to the amide bond conformations found in the CSD CTPs (17 CTPs) and were found to have a alternating *trans–cis–trans–cis* conformation predicted less often than observed in the crystal structures CTPs

	% of CTPs by amide bond conformation (t = *trans*, c = *cis*)					
	tttt	tctc/ctct	tttc	ttcc	tccc	cccc
Computational CTP library	27.4	26.1	19.2	5.8	18.5	3.0
CSD CTPS	11.7	76.5	5.9	0	5.9	0

were overlapped, 70% of the CTPs were found to mimic one of the nine Tran et al. clusters (Table 1 Row 5). Each of the clusters was overlapped by at least 64 structures and an average of 259 CTPs/cluster. When all the bonds in the CTPs were searched for overlap with three of the four Cα–Cβ bonds in the nine turn clusters each CTP was found to mimic, on average, half the nine clusters (Table 1, row 4). These results exemplify how inherently good CTPs serve as mimics of reverse turns. The inclusion of prolines and pipecolic acids at bonds in close proximity to reverse-turn space aid in making overlap of Cα–Cβ bonds in a reverse turn highly likely.

Conclusion

A conformational library of CTPs has been established which mimics known reverse-turn structure. Any protein of interest can be examined (http://cmd.wustl.edu/sagearbor/CTPlib/) by PDB entry code. All reverse turns from the PDB can be listed and downloaded in mol2 format as well as any corresponding CTP-turn mimics with information about conformational stability and bond overlaps with Cα–Cβs of the corresponding reverse turns. While this database represents the entire PDB as of August 2007, any future reverse turns could easily be overlapped by the nine Tran et al. clusters and this CTP library of mimetics used for them as well. This modular classification readily allows use as a lead scaffold for turn structures yet to be elucidated. Future work will include synthesis of the CTP mimics to verify conformation by NMR or X-ray, and methyl replacement of hydrogen on the CTP scaffolds to ascertain how side-chain substitution affects low-energy conformations.

Acknowledgements This work was funded in part by NIH Grant GM 68460. Sage Arbor acknowledges fellowship support from the Division of Biology and Biological Sciences, Washington University in St. Louis, Kauffman Cancer Research Pathway, and Washington University Center for Computational Biology Pathway.

References

1. Fairlie DP, West ML, Wong AK (1998) Towards protein surface mimetics. Curr Med Chem 5(1):29–62
2. Tyndall JD, Pfeiffer B, Abbenante G, Fairlie DP (2005) Over one hundred peptide-activated G protein-coupled receptors recognize ligands with turn structure. Chem Rev 105(3):793–826
3. Li R, Hoess RH, Bennett JS, DeGrado WF (2003) Use of phage display to probe the evolution of binding specificity and affinity in integrins. Protein Eng 16(1):65–72
4. Hanessian S, McNaughtonsmith G, Lombart HG, Lubell WD (1997) Design and synthesis of conformationally constrained amino acids as versatile scaffolds and peptide mimetics. Tetrahedron 53(38):12789–12854
5. Marraud M, Aubry A (1996) Crystal structures of peptides and modified peptides [Review]. Biopolymers 40(1):45–83
6. Karle IL, Awasthi SK, Balaram P (1996) A designed beta-hairpin peptide in crystals. Proc Natl Acad Sci USA 93(16):8189–8193
7. Arbor SK, Kao J, Yun W, Marshall G (2007) c[D-pro-Pro-D-pro-*N*-methyl-Ala] adopts a rigid conformation that serves as a scaffold to mimic reverse-turn. Biopolym Peptide Sci (in press)
8. Rose GD, Gierasch LM, Smith JA (1985) Turns in peptides and proteins. Adv Protein Chem 37:1–109
9. Ripka WC, DeLucca GV, Bach AC II, Pottorf RS, Blaney JM (1993) Protein β-turn mimetics I: design, synthesis, and evaluation in model cyclic peptides. Tetrahedron 49(17):3593–3608
10. Muller G, Gurrath M, Kurz M, Kessler H (1993) Beta VI turns in peptides and proteins: a model peptide mimicry. Proteins 15(3):235–251
11. Reaka AJ, Ho CM, Marshall GR (2002) Metal complexes of chiral pentaazacrowns as conformational templates for beta-turn recognition. J Comput Aided Mol Des 16(8–9):585–600
12. Bunin B, Plunkett M, Ellman J (1994) The combinatorial synthesis and chemical and biological evaluation of a 1,4-benzodiazapine library. Proc Natl Acad Sci USA 91:4708–4712
13. Ripka WC, Lucca GVD, Bach ACII, Pottorf RS, Blaney JM (1993) Protein β-turn mimetics II: design, synthesis, and evaluation in the cyclic peptide gramicidin S. Tetrahedron 49(17):3609–3628
14. Horton DA, Bourne GT, Smythe ML (2003) The combinatorial synthesis of bicyclic privileged structures or privileged substructures. Chem Rev 103(3):893–930
15. Breinbauer R, Vetter IR, Waldmann H (2002) From protein domains to drug candidates-natural products as guiding principles in the design and synthesis of compound libraries. Angew Chem Int Ed Engl 41(16):2879–2890
16. Muller G (2003) Medicinal chemistry of target family-directed masterkeys. Drug Discov Today 8(15):681–691
17. Baca M, Alewood PF, Kent SB (1993) Structural engineering of the HIV–1 protease molecule with a beta-turn mimic of fixed geometry. Protein Sci 2(7):1085–1091
18. Blomberg D (2007) Synthesis of β-turn and pyridine based peptidomimetics. Intellecta DocuSys, Västra Frölunda
19. Palomo C, Aizpurua JM, Balentova E, Jimenez A, Oyarbide J, Fratila RM, Miranda JI (2007) Synthesis of beta-lactam scaffolds for ditopic peptidomimetics. Org Lett 9(1):101–104
20. Martin-Martinez M, De la Figuera N, Latorre M, Garcia-Lopez MT, Cenarruzabeitia E, Del Rio J, Gonzalez-Muniz R (2005)

Conformationally constrained CCK4 analogues incorporating IBTM and BTD beta-turn mimetics. J Med Chem 48(24):7667–7674
21. Chalmers DK, Marshall GR (1995) Pro-D-NMe-amino acid and D-Pro-NMe-amino acid: simple, efficient reverse-turn constraints. J Am Chem Soc 117:5927–5937
22. Takeuchi Y, Marshall GR (1998) Conformational analysis of reverse-turn constraints by *N*-methylation and *N*-hydroxylation of amide bonds in peptides and non-peptide mimetics. J Am Chem Soc 120(22):5363–5372
23. Spath J, Stuart F, Jiang L, Robinson JA (1998) Stabilization of a β-hairpin conformation in a cyclic peptide using the templating effect of a heterochiral diproline unit. Helv Chim Acta 81:1726–1738
24. Che Y, Marshall GR (2006) Engineering cyclic tetrapeptides containing chimeric amino acids as preferred reverse-turn scaffolds. J Med Chem 49(1):111–124
25. Rizo J, Gierasch LM (1992) Constrained peptides: models of bioactive peptides and protein substructures. Annu Rev Biochem 61:387–418
26. Tran TT, McKie J, Meutermans WD, Bourne GT, Andrews PR, Smythe M (2005) Topological side-chain classification of beta-turns: ideal motifs for peptidomimetic development. J Comput Aided Mol Des 19(8):551–566
27. Macromodel 7.2 (2001) Schrodinger, Inc, Portland
28. Chang G, Guida WC, Still WC (1989) An internal coordinate Monte-Carlo method for searching conformational space. J Am Chem Soc 111:4379–4386
29. Saunders M, Houk KN, Wu Y-D, Still WC, Lipton M, Chang G, Guida WC (1990) Conformations of cycloheptadecane. A comparison of methods for conformational searching. J Am Chem Soc 112(4):1419–1427
30. Still WC, Tempczyk A, Hawley RC, Hendrickson T (1990) Semianalytical treatment of solvation for molecular mechanics and dynamics. J Am Chem Soc 112(16):6127–6129
31. Welsh EA (2005) POSSE: a program for the automated pair-wise superposition and alignment of protein structures. Washington University, St. Louis
32. The Open Babel Package 2.02 (2006) http://openbabel.sourceforge.net
33. Ho CMW, Marshall GR (1993) FOUNDATION: a program to retrieve subsets of query elements, including active site region accessibility, from three-dimensional databases. J Comput Aided Mol Des 7:3–22
34. Sybyl 7.3 (2006) Tripos international: 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA