Propensity for local folding induced by the urea fragment in short-chain oligomers†

Lucile Fischer,[a] Claude Didierjean,[b] Franck Jolibois,[c] Vincent Semetey,‡[a] Jose Manuel Lozano,[d] Jean-Paul Briand,[a] Michel Marraud,[e] Romuald Poteau[c] and Gilles Guichard*[a]

Examination of local folding and H-bonding patterns in model compounds can be extremely informative to gain insight into the propensity of longer-chain oligomers to adopt specific folding patterns (*i.e.* foldamers) based on remote interactions. Using a combination of experimental techniques (*i.e.* X-ray diffraction, FT-IR absorption and NMR spectroscopy) and theoretical calculations at the density functional theory (DFT) level, we have examined the local folding patterns induced by the urea fragment in short-chain aza analogues of β- and γ-amino acid derivatives. We found that the urea-turn, a robust C_8 conformation based on $1 \leftarrow 3$ H-bond interaction, is largely populated in model ureidopeptides (**I–IV**) obtained by replacing the α-carbon of a β-amino acid by a nitrogen. This H-bonding scheme is likely to compete with remote H-bond interactions, thus preventing the formation of secondary structures based on remote intrastrand interactions in longer oligomers. In related oligomers obtained by the addition of a methylene in the main chain (**V–VIII**), nearest-neighbour H-bonded interactions are unfavourable *i.e.* the corresponding C_9 folding pattern is hardly populated. In this series, folding based on remote intrastrand interactions becomes possible for longer oligomers. We present spectroscopic evidence that tetraurea **VIII** is likely to be the smallest unit capable of reproducing the H-bonded motif found in 2.5-helical *N,N′*-linked oligoureas.

Introduction

Non-natural oligoamides built from homologated amino acid residues (*e.g.* β- and γ-peptides) are the quintessential peptidomimetic foldamers.[1–7] Substituting heteroatoms for the carbon atoms in the backbone of ω-amino acid constituents of aliphatic oligoamides to generate non-amide linkages (*e.g.* *N*-oxyamide,[8,9] hydrazide,[10,11] urea[12,13]) can dramatically alter the pattern of intrastrand H-bond interactions, and thus represent a promising strategy to design new foldamers.[5] In particular, the urea group shares a number of interesting features with the amide linkage, *i.e.* rigidity, planarity, polarity, and hydrogen bonding capacity,

and thus represents an interesting surrogate. We are interested in determining the structure of urea-based oligomers belonging to the β- and γ-peptide lineages, namely compounds of type **A** and **B** (Fig. 1).[12,13]
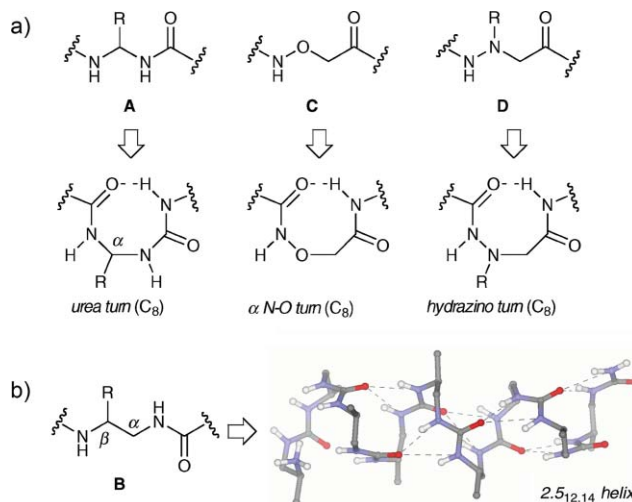
Fig. 1 a) The $1 \leftarrow 3$ H-bonding scheme in β-peptide isosteres: ureidopeptides of type **A**, α-aminoxy acid **C**, N^α-substituted hydrazinoacetic acid **D** derivatives; b) the 2.5-helical fold of *N,N′*-linked oligoureas of type **B**.

Previous work suggests that estimation of nearest-neighbour interactions in model compounds can be extremely useful to discriminate between backbones and identify those that may preferentially adopt compact folding patterns with long range order. For example, the original demonstration by Gellman and

[a] CNRS, Institut de Biologie Moléculaire et Cellulaire, Immunologie et Chimie Thérapeutique, 15, rue Descartes, F-67000, Strasbourg, France. E-mail: g.guichard@ibmc.u-strasbg.fr
[b] LCM3B, UMR-CNRS 7036, Groupe Biocristallographie, Université Henri Poincaré, BP 239, 54506, Vandœuvre, France
[c] Université de Toulouse; INSA, UPS, CNRS; LPCNO, 135 avenue de Rangueil, F-31077, Toulouse, France
[d] Fundacion Instituto de Immunologia de Colombia (FIDIC), and Universidad Nacional de Colombia, Bogotá, Colombia
[e] LCPM, UMR CNRS-INPL 7568, ENSIC-INPL, BP 451, F-54001, Nancy, France

† Electronic supplementary information (ESI) available: Influence of DMSO-d_6 content in CDCl$_3$–DMSO-d_6 mixtures on the NH proton resonances for **II.3**, **II.4**, **V.1a** and **VI.1**. NOESY correlations for **III.1** in CDCl$_3$; influence of the concentration and the solvent on the N–H stretching absorption in compound **V** series; formulae of model compounds **I.M** and **V.M** used for DFT calculations; calculated energetics, backbone torsion angles and hydrogen bond parameters for **I.M**, **I.3a**, **V.M**, and **V.2a**; crystal structure data. See DOI: 10.1039/b801139g
‡ Present address: Laboratoire Physico-Chimie Curie, CNRS UMR 168, Institut Curie, Paris, France.

Table 1 Urea–peptide derivatives **I–IV** containing a *gem*-diamino residue

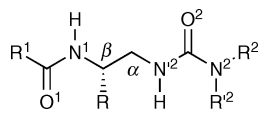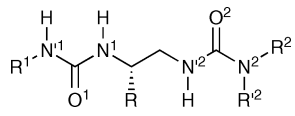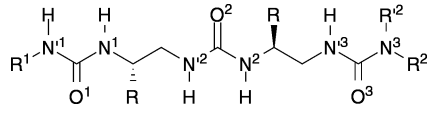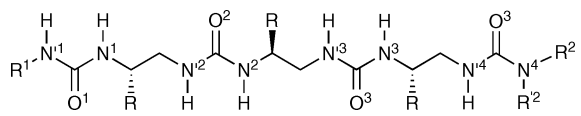| Compound	Code	R^1	R	R'	R^2	R'^2
	I.1a	*t*BuO	*i*Bu	H	Me	H
	I.1b	*t*BuO	*i*Bu	H	*i*Pr	H
	I.1″b	*t*Bu	*i*Bu	H	*i*Pr	H
	I.1c	*t*BuO	*i*Bu	H	Me	Me
	I.2a	*t*BuO	Bn	H	Me	H
	I.2b	*t*BuO	Bn	H	*i*Pr	H
	I.2c	*t*BuO	Bn	H	Me	Me
	I.3a	*t*BuO	CH$_2$OBn	H	Me	H
	I.3b	*t*BuO	CH$_2$OBn	H	*i*Pr	H
	I.3c	*t*BuO	CH$_2$OBn	H	Me	Me
	I.4a	*t*BuO	(CH$_2$)$_3$		Me	H
	I.4b	*t*BuO	(CH$_2$)$_3$		*i*Pr	H
	I.4c	*t*BuO	(CH$_2$)$_3$		Me	Me
	I.5[a]	Me	Bn	H	Me	H
	I.6[b]	Me	*i*Pr	H	Me	H
	II.1a	*t*BuO	*i*Bu	—	Me	H
	II.1′a	*t*Bu	*i*Bu	—	Me	H
	II.1b	*t*BuO	*i*Bu	—	*i*Pr	H
	II.1′c	*t*Bu	*i*Bu	—	Me	Me
	II.2	*t*BuO	*i*Bu	—	CH(Bn)CO$_2$Me	H
	II.3	BnO	Me	H	*i*Pr	H
	II.4	BnO	H	Me	*i*Pr	H
	III.1	*t*BuO	H	*i*Bu	—	—
	III.2	*t*BuO	*i*Bu	H	—	—
	IV	—	—	—	—	—

[a] Ref. 16*a*. [b] Ref. 16*b*.

coworkers that intramolecular H-bonding between neighbouring amide groups is not a favourable process in simple diamides derived from β-alanine led to the proposal that folding based on remote intrastrand interactions would be favoured in oligomers composed of β-amino acids.[14]

Preliminary results from our laboratories suggest that substituting a nitrogen for the α-carbon in β-amino acid derivatives, to give ureido compounds of type **A**, makes nearest-neighbour interactions more favourable than in β-amino acids counterparts.[13a] In non-polar solvents, type **A** *N*-acyl-*N'*-carbamoyl-*gem*-diaminoalkyl derivatives have been found to populate a C_8 conformation with a 1←3 H-bonding pattern, reminiscent of the C_7 α-peptide γ-turn. A particular feature of this folding pattern (*i.e.* the urea-turn) is that the urea adopts a characteristic *cis,trans* (*E*,*Z*) geometry.[13a] It is noteworthy that very similar turn conformations have been observed in isosteric systems

such as α-aminoxy acid **C** and N^α-substituted hydrazinoacetic acid **D** derivatives (*i.e.* the N–O[8] and hydrazino[10,11] (or N–N) turns, respectively) (Fig. 1a).[15]

In contrast, ureido monomers of type **B** possessing an extra carbon in their backbone are not expected to favor intramolecular hydrogen bonding between adjacent residues. This is supported by the finding that homochiral *N,N'*-disubstituted oligoureas of type **B** form robust helical structures stabilized by remote H-bond interactions closing 12 and 14-membered pseudorings (Fig. 1b).[12] To further delineate the propensity for local folding induced by the urea fragment in compounds of type **A** and **B**, we have now undertaken a detailed conformational investigation of short-chain ureido peptidomimetics **I–VIII** (see Tables 1 and 2) using a combination of experimental techniques (*i.e.* X-ray diffraction, FT-IR absorption and NMR spectroscopy) and theoretical calculations using density functional theory (DFT).

Table 2 Urea–peptide derivatives **V–VIII** having two sp^3 carbons in the main chain

Compound	Code	R^1	R	R^2	R$^{\prime 2}$
(structure)	**V.1a**	*t*BuO	Bn	Me	H
	V.1b	*t*BuO	Bn	*i*Pr	H
	V.1c	*t*BuO	Bn	Me	Me
	V.2a	*t*BuO	CH(Me)OBn	Me	H
	V.2c	*t*BuO	CH(Me)OBn	Me	Me
(structure)	**VI.1**	*t*Bu	CH(Me)OBn	Me	H
	VI.2	Bn	Bn	Me	H
(structure)	**VII**	Bn	Bn	Me	H
(structure)	**VIII**	Bn	Bn	Me	H

Results and discussion

Synthesis of urea-based compounds

The urea-containing peptides and oligoureas **I–VIII** (Tables 1 and 2) are divided in two families depending on the number of sp^3 carbons between the nitrogen atoms in consecutive urea or amide groups.

Compounds of type **A** (**I–IV**, containing a 1,1-diamino alkyl residue) were synthesized as previously described[13b] by coupling succinimidyl {1-{[(alkyloxy)carbonyl]amino}-1-X-methyl}carbamates or succinimidyl [1-(acylamino)-1-X-methyl] carbamates (derived from *N*-protected α-amino acid and dipeptides, respectively) to simple amines or α-amino acid esters. Ureido compounds of type **B** (**V–VIII** containing a 1,2-diamino alkyl residue) were prepared from succinimidyl {2-{[(*tert*-butoxy)carbonyl]amino}-2-X-ethyl}carbamates as previously reported.[17] Compounds **IV**,[13b] **VII** and **VIII** were obtained by repetitive urea formation with appropriate succinimidyl carbamates.

Comparative spectroscopic study of ureido compounds of type A and type B

The relative propensity for local folding induced by the urea fragment in model ureas with a 1,1-diaminoalkyl residue (**I**, type **A**) and a 1,2-diaminoalkyl residue (**V** and **VI**, type **B**) was studied in solution using a combination of FT-IR and NMR spectroscopies.

The IR data (Table 3) of compounds **I** with R$^{\prime 2}$ = H are not significantly sensitive to dilution below 10 mM in CH$_2$Cl$_2$ and 2 mM in CCl$_4$, indicating a similar behaviour to analogous peptide models.[18,19]

The broad absorption in the 3300–3400 cm^{-1} region, which is assigned to H-bonded NHs, shifts to low frequencies when R^2 = Me is changed into *i*Pr, and disappears for R$^{\prime 2}$ = Me (Fig. 2a, Table 3). Therefore it may be assigned to the H-bonded N^{2}H
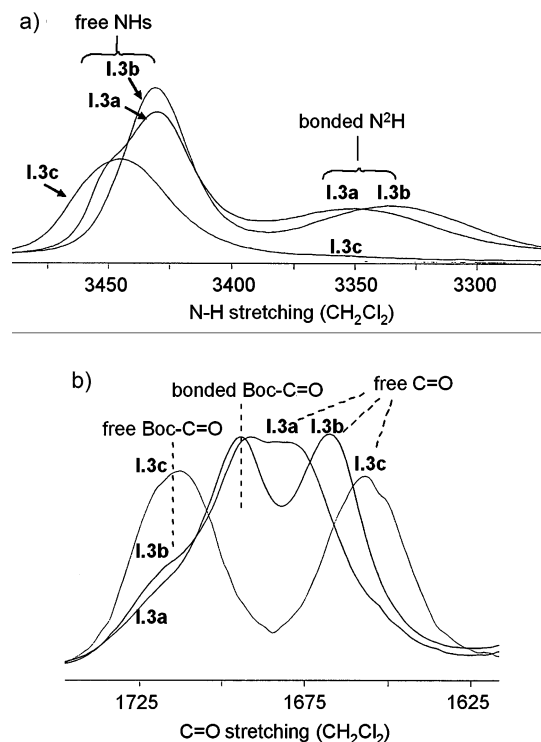
Fig. 2 Superimposition of a) NH stretch and b) C=O stretch region FT-IR data for 2 mM compounds **I.3a**, **I.3b** and **I.3c** in CH$_2$Cl$_2$.

vibrator.[19] On the other hand, the stretching absorption of the CO^1 carbonyl is shifted from about 1695 cm^{-1} for R$^{\prime 2}$ = H to about 1715 cm^{-1} for R$^{\prime 2}$ = Me (Fig. 2b, Table 3). We may then conclude that part of molecules **I** with R$^{\prime 2}$ = H are folded by an N^{2}H-to-CO^1 intramolecular H-bond closing an 8-membered ring (Fig. 3a).

Table 3 N–H and C=O stretching frequencies for the urea–peptide models **I** and **V**

| Cmpd	Solvent	N²–H + N'²–H Free	N²–H Free	N²–H Bonded	C=O¹ Free	C=O¹ Bonded	C=O² Free
I.1a	CH₂Cl₂	3453[sh]	3431	3350	1718[sh]	1695	1676
I.1b	CH₂Cl₂	——————3430——————		3339[m,br]	1718[sh]	1695	1667
I.1″b	CH₂Cl₂	——————3437——————		3286[m,br]	——————1662[br]——————		
I.3a	CH₂Cl₂	3453	3431	3363[w,br]	n.v.	1698	1679
I.3b	CH₂Cl₂	——————3431——————		3344[w,br]	n.v.	1698	1669
I.2a	CH₂Cl₂	3431	3450[sh]	3352[m,br]	1718[sh]	1692	1677
	CCl₄	3462/3439	—	3346[m,br]	1718[sh]	1693	1689
I.2b	CH₂Cl₂	——————3431——————		3336[m,br]	1718[sh]	1694	1667
	CCl₄	3462/3439	—	3346[m,br]	1714[sh]	1702	1682
I.2c	CH₂Cl₂	3445	—	—	1713	—	1657
I.4a	CH₂Cl₂	3429	3451[sh]	3312[br]	1696[sh]	——————1673——————	
I.4b	CH₂Cl₂	——————3431——————		3292[br]	1696[sh]	——————1670——————	
I.4c	CH₂Cl₂	3455	—	—	1693	—	1656
V.1a	CH₂Cl₂	3432	3452	3366[w,br]	1703	n.v.	1677
	CCl₄	3442	3459	3356[br]	1707	n.v.	1690
V.1b	CCl₄	3436	3448[sh]	3367[br]	1707	n.v.	1682
V.1c	CH₂Cl₂	3472/3432	—	—	1704	n.v.	1648
	CCl₄	3480/3442	—	—	1707	n.v.	1658
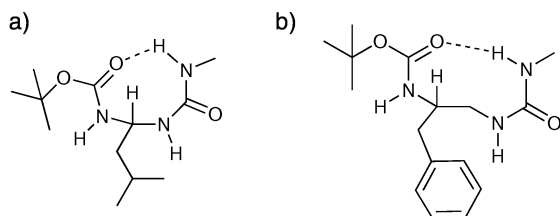
Fig. 3 Schematic representation of (a) the C_8 γ-like folded structure induced by the urea–peptide motif in **I.1a**, and (b) the minor folded C_9 structure induced by the urea–peptide motif in **V.1a**, showing the *cis,trans* conformation of the *N,N'*-disubstituted urea fragment.
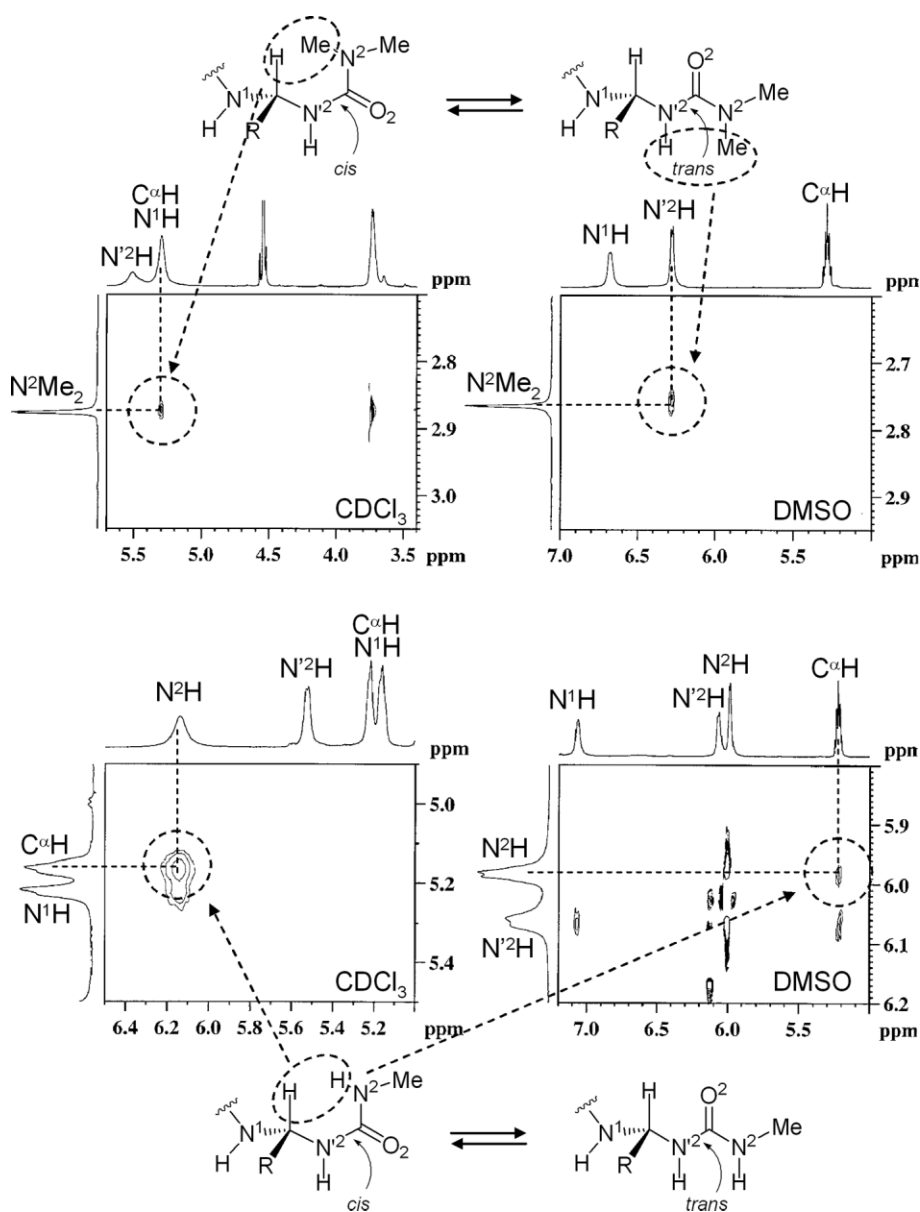
The N¹H and N'²H bonds are free from any intramolecular interaction, as demonstrated by NMR DMSO-d_6 titration experiments in CDCl₃. By progressive addition of DMSO-d_6 in CDCl₃, their proton resonances actually experience a higher shift ($\Delta\delta$: 1.6–2.1 ppm for N¹H and 0.83–1.22 ppm for N'²H in compounds **I.1a**–**I.3a**) to low fields than N²H ($\Delta\delta$: −0.22 to 0.42 ppm in

compounds **I.1a**–**I.3a**) (see Fig. 4 (heavy line), Fig. 6 (left panel) and Fig. 7 (left panel)).

The *trans,trans* conformation of the urea fragment is not compatible with the above N²H–CO¹ H-bond, and a *cis,trans* conformation is geometrically required. This point was confirmed by NOESY experiments in solution in CDCl₃ (Fig. 5).

Molecule **I.1a** in CDCl₃ actually exhibits a CᵅH/N²H NOE correlation, denoting a short interproton distance typical of the *cis,trans* conformation, but no N²H/N'²H NOE correlation that would denote the *trans,trans* conformation (Fig. 5, lower panel). For molecules **I.1a** and **I.1c**, we verified that the *cis,trans* conformation of the urea fragment is not strictly dictated by the occurrence of the N²H–CO¹ intramolecular H-bond. The *cis,trans* conformation is effectively retained by **I.1a** in DMSO, where solvation breaks this H-bond. In **I.1c** (R² = R'² = Me), where the single resonance for both methyl groups indicates a rapid rotation of the CO²–N²Me₂ bond, the Me/CᵅH (in CDCl₃) and Me/N'²H (in DMSO-d_6) NOE correlations denote a solvent-induced transition from the *cis* (CDCl₃) to the *trans* (DMSO-d_6) conformation of the N'²–CO² bond (Fig. 5, upper part).
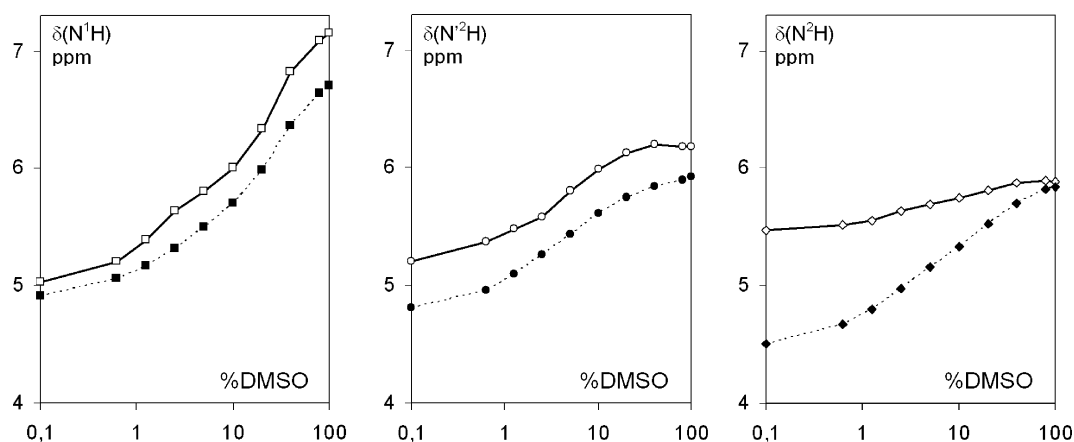
Fig. 4 Influence of DMSO-d_6 content in CDCl₃–DMSO-d_6 mixtures on the NH proton resonances for **I.2a** (solid line) and **V.1a** (dotted line).

This journal is © The Royal Society of Chemistry 2008

Org. Biomol. Chem., 2008, **6**, 2596–2610 | **2599**

Fig. 5 NOESY correlations demonstrating the *cis,trans* conformation of the *N,N'*-disubstituted urea fragment for **I.1a** in CDCl₃ and DMSO-d_6 (bottom) and the transition from the *cis* (CDCl₃) to the *trans* (DMSO-d_6) conformation for the urea N'²–CO² bond in **I.1c** (top).

Qualitative analyses of IR data reveal a similar behavior between compounds **I** and **V** (Table 3), but the intensity of the broad contribution in the 3300–3400 cm⁻¹ region is much weaker in CH₂Cl₂ for the latter, indicating a smaller percentage of folded molecules (see Fig. S1†).

In CCl₄, the intensity of the low frequency component is highly sensitive to the concentration down to 0.2 mM, which denotes a great tendency to molecular aggregation (see Fig. S1†). The residual absorption at 0.04 mM reflects the existence of a minor percentage of folded molecules with an N²H–CO¹ intramolecular H-bond closing an 9-membered ring (Fig. 3b).

Because of the small concentration (0.2 mM) required to have non-aggregated molecules **V** in CCl₄, NOESY experiments could not be carried out in this case. However, the existence of the N²H–CO¹ H-bond implies that the *cis–trans* conformation of the urea fragment is populated (albeit to a low extent) in solution. The molecular flexibility of the urea models deriving from an 1,2-diaminoalkyl residue is corroborated by the magnetically equivalent main-chain CH₂ protons in compound **V** series. In addition, the larger chemical shifts experienced by N²H in the compounds **V** series compared to compound **I** series upon progressive addition of DMSO-d_6 in CDCl₃ indicates a higher solvent accessibility (see Fig. 4, right panel). The NH accessibility for diurea **VI.1** in CDCl₃–DMSO-d_6 mixtures revealed that N²H is a little less accessible in **VI.1** than in **V.1a** (see Fig. S2†), suggesting that the percentage of the N²H–CO¹ H-bond is a little higher, probably being related to the higher basicity of the urea carbonyl in **VI.1** compared with the urethane carbonyl in **V.1a**.
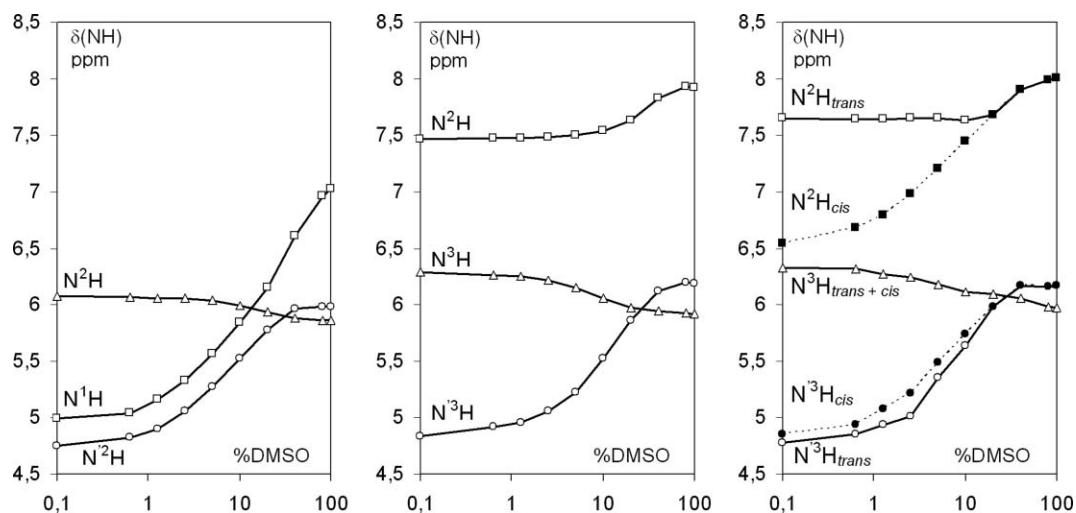
Fig. 6 Influence of DMSO-d_6 content in CDCl$_3$–DMSO-d_6 mixtures on the NH proton resonances for **I.1a** (left), **II.1′a** (middle) and **II.1a** (right).
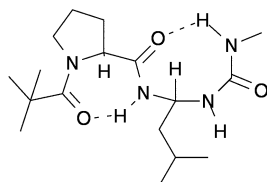
Fig. 7 Schematic representation of the C$_8$ urea-turn conformation associated with the γ-folded Pro residue in **II.1′a**.

Spectroscopic studies of type A amide/urea hybrids II–IV

An additional series of compounds of type **A** (**II–IV**, Table 1) have been prepared to evaluate the influence of i) the position of the urea moiety in the peptide chain (central in **III** versus terminal in **II**), and ii) the number of urea moieties (**IV**) on folding propensity and C$_8$ urea-turn formation.

When proline precedes the *gem*-diamino residue, the NH stretching for **II.1′a** exhibits a strong, broad absorption at 3327 cm^{-1} which considerably decreases for **II.1′c** (R$'^2$ = Me). The low solvent accessibility for the N^{3}H proton in **II.1′a** is quite similar to that for N^{2}H in **I.1a** (Fig. 6).

Substitution of Boc for Piv in **II.1a** allows the *cis–trans* equilibrium around the Boc–Pro bond, but has no influence on the N^{3}H solvent accessibility (Fig. 6). All these data indicate

that proline does not prevent the occurrence of the urea-turn. It is noteworthy that the high and low solvent accessibility for the N^{2}H proton contributions in **II.1a** depends on the *cis* and *trans* conformation of the Boc–Pro bond, respectively (Fig. 6, right panel). Combined with the persistence of a broad NH absorption at 3306 cm^{-1} for **II.1′c**, this observation indicates that N^{2}H in **II.1′a** is partly engaged in a 1←3 interaction with CO^1 to give overlapping γ- and urea-turns (Fig. 7).

It is noteworthy that the same trend *i.e.* γ-turn nucleation, albeit to a lower extent, was observed when alanine was substituted for proline (see Fig. S3†). The chemical shift variation observed for N^{2}H in **II.3** when increasing the DMSO concentration from 1.25% to 100% is larger than when a prolyl residue was present (0.92 ppm versus 0.37 ppm) but is significantly lower than that of a fully accessible amide NH (*ca.* 1.28 ppm for N^{3}H in compound **III** series; *vide infra*). Substituting D-Ala for L-Ala in **II.3** (*i.e.* heterochiral versus homochiral sequences) has no significant effect on the chemical shift variation of N^{2}H ($\Delta\delta$ = 0.94 ppm).

The incorporation of the urea group between two peptide bonds in **III.1** and **III.2** does not affect significantly the urea spectroscopic data. For example, the N^{2}H proton resonance is the less sensitive to solvation while those for N^{1}H and N^{3}H experience a large shift to low fields (Fig. 8).
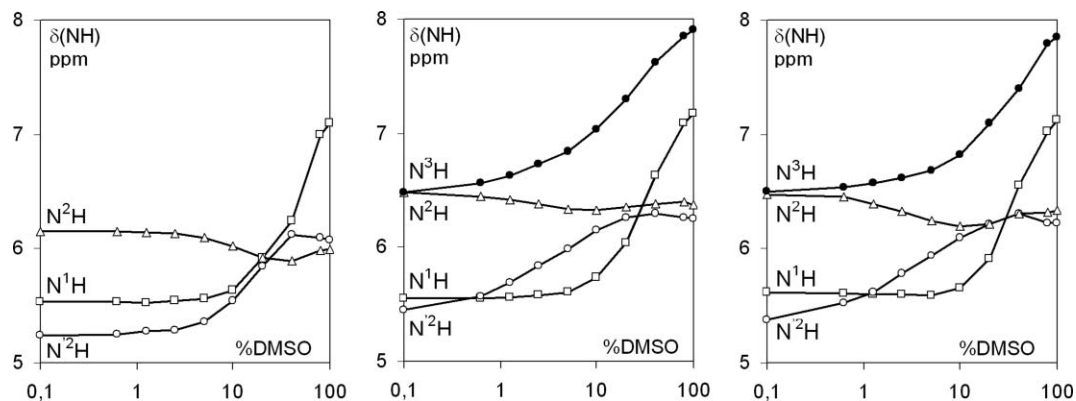
Fig. 8 Influence of DMSO-d_6 content in CDCl$_3$–DMSO-d_6 mixtures on the NH proton resonances for **I.3a** (left), **III.1** (middle) and **III.2** (right).

This journal is © The Royal Society of Chemistry 2008

Org. Biomol. Chem., 2008, **6**, 2596–2610 | **2601**

Moreover, the profiles of NH resonances for **III.1** and **III.2** are practically superimposed, indicating that the chirality of the sequence (homochiral *versus* heterochiral) has no influence on the conformational properties induced at a short distance by the urea group. The N²H/CαH NOE correlation again confirms the *cis,trans* conformation of the urea group (See Fig. S4†). This NOE pattern is conserved in longer hybrid oligomers made of alternating amide and urea bonds such as tetramer **IV** (Fig. 9).
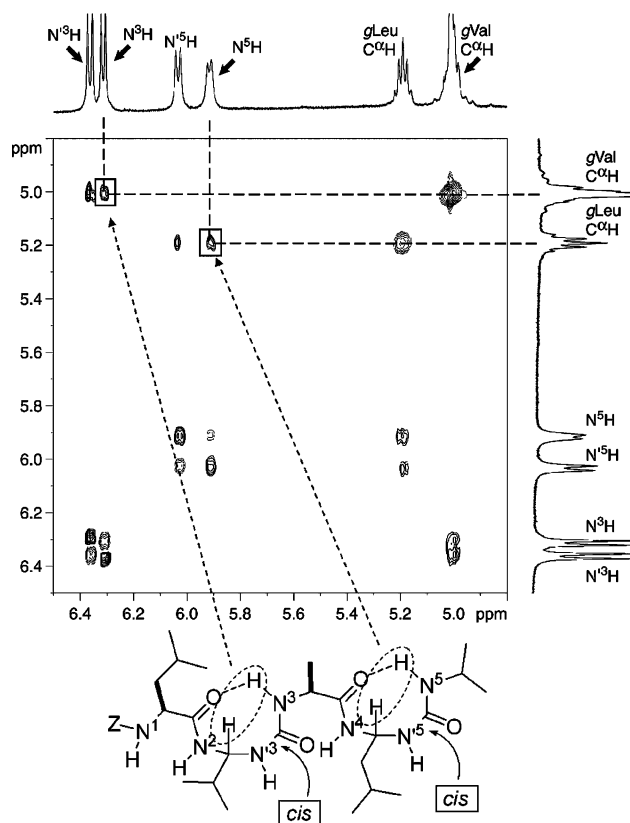
Fig. 9 NOESY correlations demonstrating the *cis,trans* conformation of N,N'-disubstituted urea fragments for **IV** in DMSO-d_6.

Theoretical calculations

To gain additional information on interactions between nearest neighbours in ureido compounds of type **A** and **B**, we have performed theoretical investigations using density functional theory (DFT) on model compounds **I** and **V**.

Before starting the exploration of potential energy surfaces of ureidopeptides, a brief comparison of the energy properties of N-methyl acetamide (NMA) and N,N'-dimethylurea (DMU), which are parent representatives of peptide and urea plaques, may be instructive. The lower thermodynamic stability of *cis*-NMA with respect to *trans*-NMA (by 2.6 kcal mol⁻¹ at the present level of calculation), restrictions in conformational space, non-covalent stabilization of *trans* amide bonds, and the so-called steric clash between successive side-chains in all-*cis* peptides, are invoked for explaining the extreme scarcity of *cis* peptide bonds in natural proteins. Rotation about urea bonds is less unfavourable, owing to
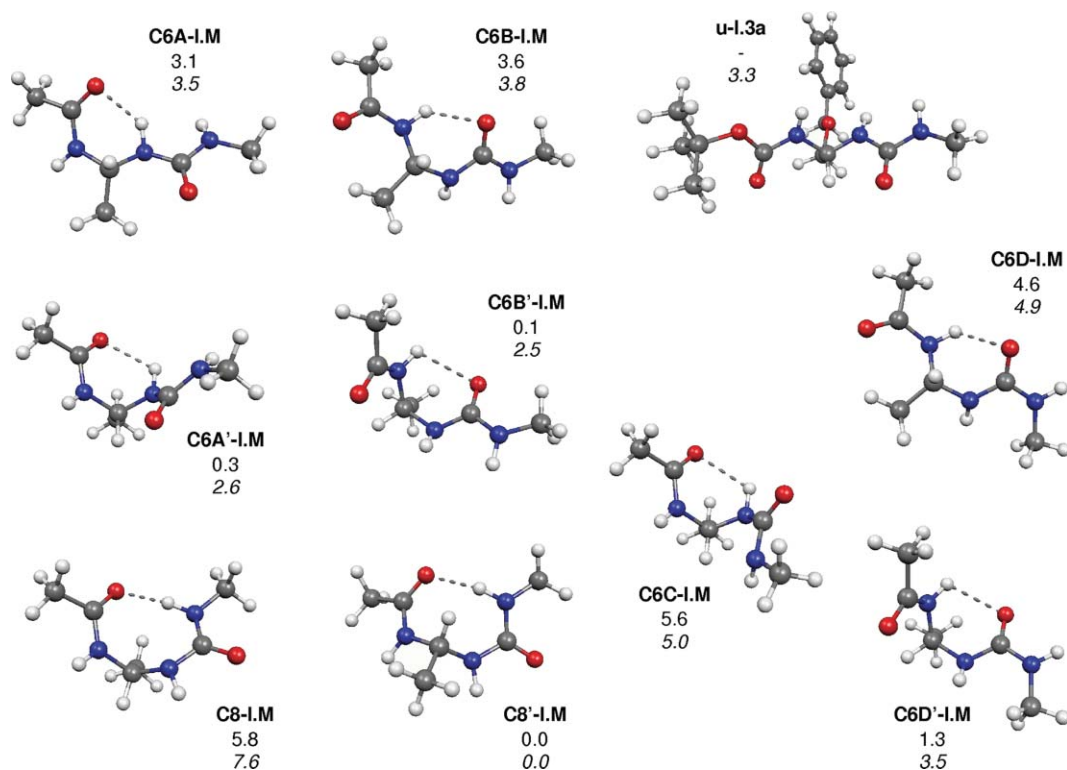
thermodynamic as well as to kinetic data. *cis,trans*-DMU actually lies at $\Delta_r G° = 1.3$ kcal mol⁻¹ above its *trans,trans*- counterpart, indicating that rotation about C(=O)–N bonds in ureas leads to a less unstable isomer. It should be noted that, according to theoretical calculations, some alkyl-substituted urea derivatives are expected to be more stable in their *cis* conformation.[20] In addition, while the lowest[21] free energy of isomerisation of NMA lies 19.6 kcal mol⁻¹ above the *trans* isomer, the transition state which connects *trans,trans* and *cis,trans* isomers of DMU lies only 9.1 kcal mol⁻¹ above the former geometry. This value is similar to those obtained by previous experimental[22] or theoretical[20] studies on ureas. In other words, rotation about an urea bond is less demanding than rotation about a peptide bond and *cis,trans* urea bonds should be significantly more populated than *cis* peptide bonds in oligomers.

Preliminary studies on model compounds with R¹ = R = R² = Me ((1R)-N-[1-(3-methylureido)ethyl]-acetamide and (1S)-N-[1-methyl-2-(3-methylureido)ethyl]acetamide (See Fig. S5†), hereafter called **I.M** and **V.M**) have provided a first overview on the structures and energies of different folding patterns. In order to check the role of substituents and side-chains on the stability scale found for the models, we have also addressed two of the urea-containing derivatives studied in the experimental part of this work, namely **I.3a** and **V.2a**. All structures considered in this theoretical work are shown in Figs. 10 and 11 for compounds **I** and **V**, respectively.

The nomenclature first refers to unfolded patterns (**u**), or to the n-membered (**Cn**) H-bonded ring, followed by a capital letter when several geometrical patterns lead to the same ring size. Eight- and nine-membered rings are only observed in the case of urea-turns in **I** and **V** respectively (**C8–I** and **C9–V** types). Besides this folding pattern, for topological reasons hybrid amide–urea models such as **I** can only exhibit an unfolded structure (**u–I**) or 6-membered rings. Similarly, compounds **V** can be unfolded or may exhibit 7-membered pseudocycles. The stability of the conformations envisaged in this work is mainly ruled by three factors : i) the energy associated with the backbone, ii) the directionality and bond length of hydrogen bonds, iii) the energy cost associated to *trans→cis* inversion of urea bonds. Considering the intrinsic penalty of 2.5 kcal mol⁻¹ generated by *trans→cis* isomerisation of peptide bonds, we only considered peptide fragments in their *trans* conformation.

Urea models containing a *gem*-diamino residue (type A). All the results are collected in Table S1†. Relative Gibbs free energies ($\Delta_r G°$) differ from ΔE by less than 1 kcal mol⁻¹, thus showing that entropic effects do not play a significant role in the energies. Except for **C6C–I.M**, we systematically found two conformations associated with a given folding pattern. They differ by the torsion angles ϕ and ψ, and in most cases by the hydrogen bond parameters, *i.e.* the angle between N–H and H···O (θ) and the distance between H and O (r_{hb}). In the case of the urea-turn, it is interesting to notice that according to ϕ and ψ, structures of type **C8–I** and **C8'–I** are reminiscent of the C_7 α-peptide classic and inverse γ-turns. None of the geometry parameters not significantly differ between compounds **I.M** and **I.3a**.

The model compound in its unfolded geometry (**u–I.M**) is not a minimum on the potential energy surface, whereas we have found two stable geometries for the "real" system **I.3a**. Such a

2602 | *Org. Biomol. Chem.*, 2008, **6**, 2596–2610

This journal is © The Royal Society of Chemistry 2008

Fig. 10 Molecular conformations for compound **I.M** and selected conformation for compound **I.3a**. Calculated Gibbs free energies ($\Delta_r G^\circ$, in kcal mol^{-1}) are also given for **I.M** (plain text) and similar conformations for **I.3a** (italic). Hydrogen bonds are systematically depicted, even though values for θ and r_{hb} indicate that they are very weak for **C6A′–I**, **C6B′–I**, **C6C–I**, **C6D′–I**. See the ESI† for additional details.

structure may thus exist in the gas phase, but it is significantly less stable than the urea-turn **C8′–I.3a** by 3.3 kcal mol^{-1}. The most important result is that one of the two structures which exhibits a C_8 conformation is the most stable one among all the isomers considered in this work. This is already true for the model compound **I.M**, and more marked for **I.3a**, since the next structures in the energy scale (**C6A′–I.3a** and **C6B′–I.3a**, with *trans* peptide and urea bonds) lie *ca.* 2.5 kcal mol^{-1} above. We have also undertaken the study of two conformations which exhibit *cis,trans* and *trans,cis* bonds in NH–CO–NH (**C6C–I** and **C6D–I**). These structures are relatively high in energy.

In summary, it is noteworthy that the stability of the urea-turn is important enough to ensure that compound **I.3a** preferentially adopts this conformation. The relative stability of **C8′–I.M** suggests that this preference for the urea-turn motif should resist to substitution.

Urea models with an additional methylene group (type B). In this case, we have only found a single configuration for each folding pattern. All the results are given in Table S2†. In both the model system **V.M** and the real system **V.2a**, the urea-turn does not appear viable, being less stable than the extended conformation **u–V** and all other folded structures. Interestingly, we have found an additional structure (**//–V**) in which the peptide and urea plaques are almost parallel. This structure is stabilized by a dipole–dipole interaction ($\downarrow\uparrow$) between the two chemical fragments. As a matter of fact, our calculations on NMA and DMU show that these compounds have similar dipole moments (*ca.* 3.6 Debye).

This interaction is allowed by the flexibility introduced by the additional methylene group. Although this structure could be favoured under specific conditions (solvent or substituents), it was not the lowest-energy configuration in the gas phase. The two most stable conformations are **C7A–V** and **C7B–V**, for both the model and real systems. The directionality of the hydrogen bond (θ) together with the r_{hb} distance are in the range of relatively strong hydrogen bonds. It is also the case for the urea-turn **C9–V**, which is probably unstabilized owing to the backbone. In opposition to compounds of type **I**, the two conformations which exhibit *cis,trans* and *trans,cis* bonds in NH–CO–NH (**C7C–V** and **C7D–V**) have low energy. Nevertheless, it is unlikely that the former could be observed since, according to $\Delta_r G^\circ$, it lies at least 3 kcal mol^{-1} above the lowest-energy structures **C7A–V** and **C7B–V**.

In summary, in the case of ureido compounds with an additional methyl group in the backbone, C_9 turn conformations are not expected to be observed. The effect of side-chains or substituents in **V.2a** does not improve the stability of the urea-turn with respect to its counterpart in the model compound **V.M**. In addition, several folded structures, as well as the unfolded conformation, compete in a narrow energy range and are more likely populated than the urea-turn.

Crystal structures

Whereas spectroscopic and theoretical studies above indicate that nearest-neighbour hydrogen interaction are favoured in ureido compounds of type **A**, the situation turned out to be different in
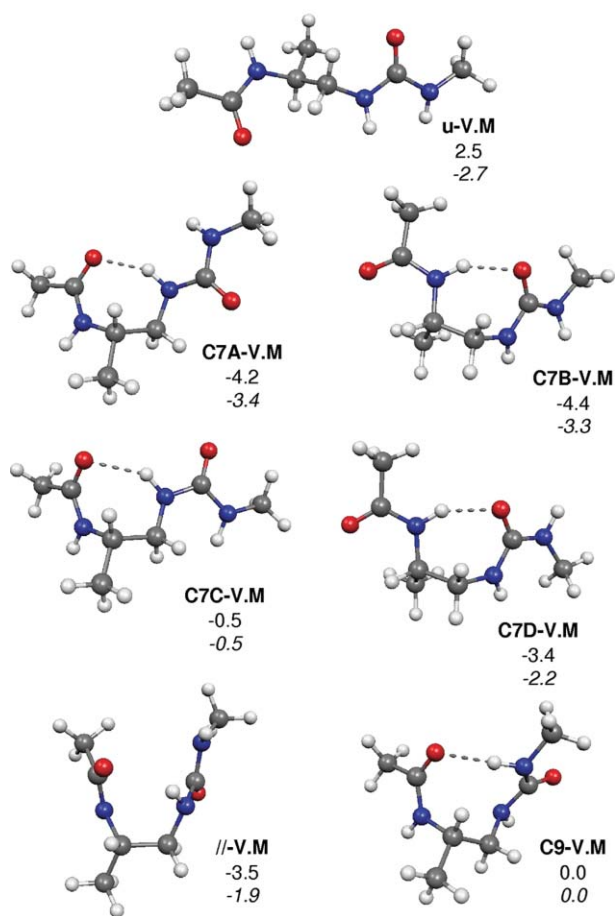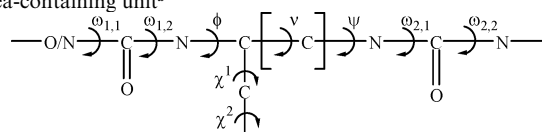
Fig. 11 Molecular conformations for compound **V.M**. Calculated Gibbs free energies ($\Delta_r G°$, in kcal mol^{-1}) are also given for **V.M** (plain text) and similar conformations for **V.2a** (italic). See the ESI† for additional details.

the solid state. We obtained crystals suitable for X-ray diffraction for eight of the urea-containing derivatives (*rac*-**I.1b**, **I.3c**, *rac*-**I.4b**, **II.1′a**, **II2**, **V.1a**, **V.1c**, **VI.2**, see Table S3†).

None of the crystal structures displays an intramolecular hydrogen bond. Except *rac*-**I.4b** and **II.1′a**, where proline logically induces a kink, the molecules assume more-or-less extended conformations, probably favoured by intermolecular interactions in the crystal (Fig. 12).

The main torsion angles for the urea-containing unit[23] are listed in Table 4.

In all cases, the NH–CO–NH urea fragment is nearly symmetrical and *trans,trans* planar.[24] One notes that, similarly to L-peptides, the ϕ angle is negative ($-133°$ to $-89°$). The pseudo "ψ" angle generally adopts a positive value ($84°$ to $137°$), except for **V.1a** where the negative value is associated with a skewed conformation of the 1,2-diamino ethane fragment. In all the cases except **II.2**, where the *i*Bu substituent is *trans* to the nitrogen ($\chi^1 = -173°$), the *i*Bu or Bn side substituent retains the same gauche-orientation ($\chi^1 = -71°$ to $-58°$). The *i*Pr substituent in **I.6** assumes the classical disposition, with $^\alpha$C–H and side-chain $^\beta$C–H bonds in the *trans* conformation.

All compounds form hydrogen-bonded ladder structures (Fig. 12). Successive molecules in the ladder are either parallel related by a translation for **I.5**, **V.1a**, **V.1c** and **VI.2** or antiparallel related by a 2_1 screw axis for *rac*-**I.1b**, **I.3c**, **I.6**, **II.1′a** and **II.2**. Compound *rac*-**I.4b** is different because adjacent molecules present the same chirality and are crystallographically independent. The intermolecular spacing (4.67–5.24 Å) is similar to that found for symmetrically disubstituted ureas.[25] The NH–CO–NH urea motif is involved in more-or-less complex contacts, with both urea NHs being engaged in an H-bond with i) the same amide carbonyl in **I.6**, **II.1′a** and **II.2**, ii) the urethane carbonyl in **I.1b**, or iii) the urea carbonyl in **I.4b**, **I.5**, **V.1a** and **VI.2**. The N · · · O distances (2.84–3.58 Å) may be above the limiting value for a classical H-bond.[26] The closest double contacts (2.91 Å and 2.97 Å) are observed for **VI.2**, and the loosest ones (3.51 Å and 3.58 Å) for **V.1a**. In one case (**II.1′a**), a urea NH is engaged in an H-bond network with three carbonyls, but the N · · · O distances are rather large (3.46, 3.48 and 3.50 Å). One also notes that the NHs of contiguous urea and/or amide groups in the same molecule may be oriented in the same (**I.1b**, **I.5**, **I.6**, **II.1′a** and **II.2**) or opposite (**V.1a** and **VI.2**) directions.

Table 4 Main torsional angles (°) of the urea-containing unit^a

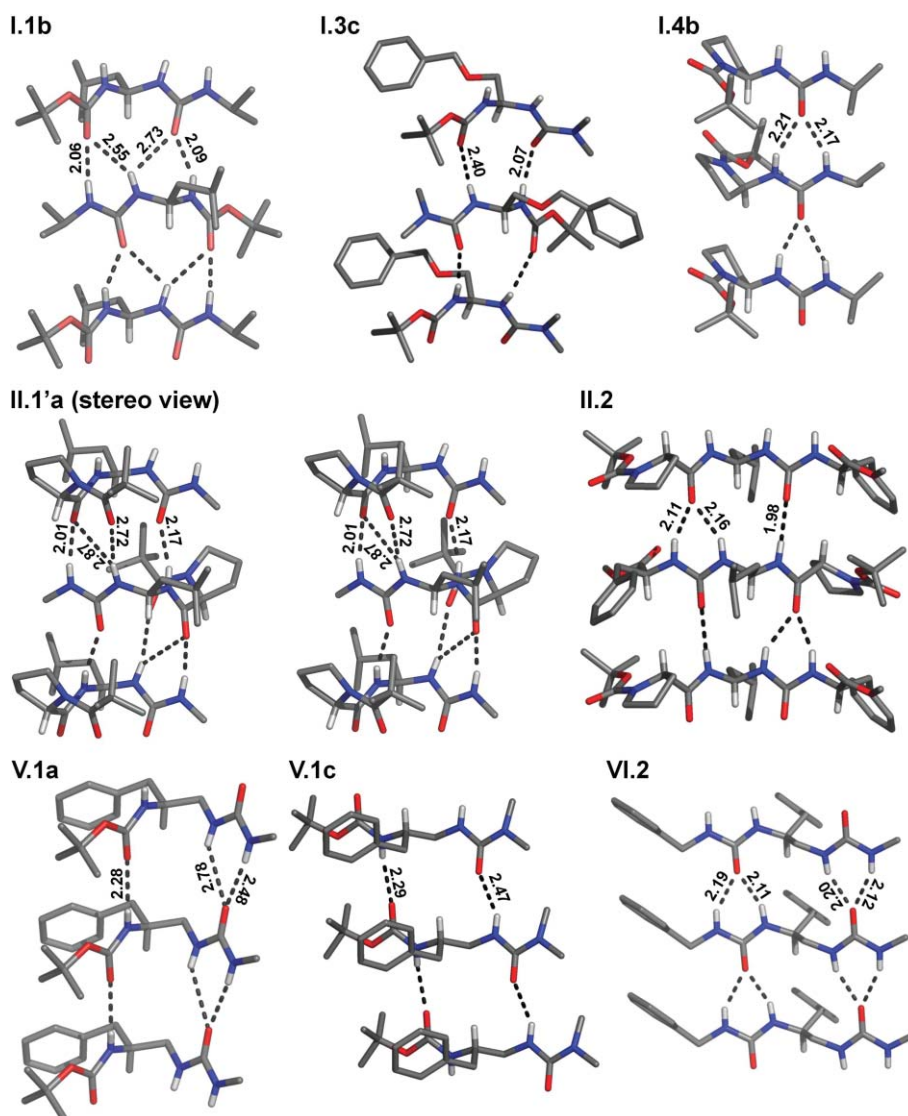| Compound^b	$\omega_{1,1}$	$\omega_{1,2}$	ϕ	ν	ψ	$\omega_{2,1}$	$\omega_{2,2}$	χ^1	χ^2
I.1b	171.4(2)	179.1(2)	−122.8(2)	—	−92.3(2)	169.3(2)	172.6(2)	−58.1(3)	−63.1(3)/173.2(2)
I.3c	−177.8(6)	−172.6(6)	−115.4(7)	—	129.8(6)	178.6(6)	−177.9(7)/15(1)	−67.0(8)	149(1)
I.4b	173.1(2)	3.5(3)	−76.3(2)	—	137.0(2)	−173.4(2)	−179.9(2)	30.6(2)	−39.6(2)
I.5^c	—	174	−111	—	93	−166	−174	−60	−89/91
I.6^d	—	176	−113	—	105	179	174	−58/180	—
II.1′a	—	178.8(3)	−125.5(3)	—	83.7(4)	−168.1(3)	−178.9(3)	−58.6(4)	−65.3(4)/171.4(3)
II.2	—	176.2(5)	−122.5(6)	—	118.5(5)	172.1(5)	−179.3(6)	−171.7(6)	69.8(8)/−166.1(6)
V.1a	176.8(6)	−167.0(5)	−123(1)	58(1)	−145(2)	−173(2)	−173(2)	−64.9(9)	−85.7(8)/94(1)
V.1c	177.7(2)	−169.1(2)	−133.0(2)	171.8(2)	112.9(3)	175.4(2)	175.4(3)/−22.4(4)	−62.6(3)	−87.7(2)/90.8(3)
VI.2	−179.4(5)	177.3(4)	−89.2(5)	177.1(4)	105.4(5)	175.7(4)	−179.3(5)	−71.1(5)	−100.5(7)/82.8(7)

^{a} Ref. 23. ^{b} The angles are reported for the molecule having the same absolute chirality as the starting L-amino acid. ^{c} Ref. 16a. ^{d} Ref. 16b.

Fig. 12 Molecular conformations and intermolecular interactions in the crystal structures of compounds **I.1b**, **I.3c**, **I.4b**, **II.1′a**, **II.2**, **V.1a**, **V.1c**, and **VI.2** (see Table S3 in ESI for crystal data and structure refinement and Table 4 for torsional angles of the urea-containing unit). Intermolecular NH···O distances in Å are indicated.

From unfavourable local folding to remote intrastrand interactions: ^{1}H NMR studies of short oligomers of type D

According to spectroscopic data collected on monomeric compounds **V** and **VI** and theoretical studies, intramolecular hydrogen bonding between nearest neighbours is not a favourable process for compounds of type **B**. This is consistent with previous finding from our laboratories showing that longer oligomers of type **B** (heptaurea and nonaurea) adopt a helical fold stabilized by remote intrastrand interactions.[12a,c] This 2.5-helical structure is stabilized by a double H-bonding scheme where each carbonyl is H-bonded to both NHs of the urea two residues ahead (See Fig. S6†).

Accordingly, one would expect a triurea to be the smallest molecule capable of reproducing a turn of the aforementioned helix. To investigate the minimum length required to nucleate a

helical conformation, we have investigated derivatives containing three (**VII**) and four (**VIII**) urea motifs.

Due to their low solubility in chloroform, compounds **VII** and **VIII** with three and four urea motifs were examined in DMSO-d_6 and in CDCl$_3$–30% DMSO-d_6 and compared to diurea **VI.2** (see Tables S4–S12†). Qualitative examination of ^{1}H NMR spectra revealed a number of interesting features. The NH regions of diurea **VI.2** and triurea **VII** remain poorly dispersed in both pure DMSO and CDCl$_3$–30% DMSO-d_6. In contrast, the urea NH signals of tetraurea **VIII** in CDCl$_3$–30% DMSO-d_6 are very well dispersed. The NH chemical shifts with DMSO solvation were generally smaller for **VIII** than for **VI.2**. In addition, the non-equivalence of the main-chain CH$_2$ protons spectacularly increased from diurea/triurea to tetraurea. In CDCl$_3$–30% DMSO-d_6, the chemical shift difference changes from 0.2 ppm for **VI.2** to

0.88–1.29 ppm for **VIII**, thus supporting the existence of a rigid structure for tetraurea **VIII**.

The present ^{1}H-NMR data obtained for **VIII** in CDCl$_3$–30% DMSO-d_6 strongly suggest that the minimum number of urea units required to initiate folding in N,N'-linked oligoureas is four. In particular, the large^3 J(N^{2}H,$^\beta$CH) values (9.7 Hz), the large chemical-shift differences ($\Delta\delta$) and strong differentiation of vicinal coupling constants for diastereotopic protons of the central residue ($\Delta\delta$ = 1.29 ppm; ^3J(N$^{'3}$H,$^\alpha$CH) = 3.2 Hz and 9.6 Hz for upfield and downfield $^\alpha$CH, respectively), as well as the $i/i+2$ NOE correlations (Fig. 13) observed in **VIII**, are all spectroscopic features previously associated with helical oligoureas.
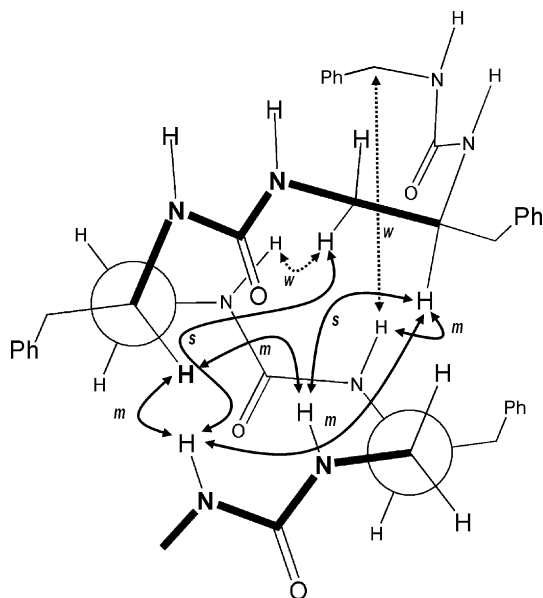
Fig. 13 Representative inter-residue NOE connectivities observed for tetraurea **VIII** in CDCl$_3$–DMSO-d_6 (70 : 30) at 298K. s = strong, m = medium, and w = weak. These NOE connectivities are consistent with **VIII** being the smallest unit capable of populating a 2.5-helical fold.

Conclusion

All the above spectroscopic data indicate that the N,N'-disubstituted urea is a flexible fragment. In the solid state, where it is systematically engaged in intermolecular H-bonds, the *trans,trans* conformer placing the N–H and C=O bonds in an *anti* orientation is observed in all cases. In solution, however, the urea motif recovers a conformational freedom, and the CO–N bond may assume the *cis* and *trans* conformations. In the case of ureido compounds of type **A**, it essentially populates the *cis,trans* conformation and gives rise to a short-distance interaction, with the preceding peptide carbonyl in a folded structure originally being term a urea-turn (Fig. 3a).[13a] Theoretical investigations using density functional theory (DFT) are in good agreement with experimental data. The H-bond closing an 8-membered ring in the urea-turn is not essential to the *cis–trans* conformation of the urea fragment.

The striking difference between the conformations observed in solution and those existing in the solid state is not unprecedented in the field of peptidomimetic foldamers. A similar dichotomy was noticed earlier by Yang and coworkers in the α- and β-aminoxy peptide series.[8,9b] Whereas in non-polar solvents short aminoxypeptides adopt helical conformations consisting of successive N–O turns, extended parallel and antiparallel sheet-like structures stabilized by intermolecular H-bonds have been observed in the solid state.

Interestingly, experimental NMR and IR data on (amide/urea) hybrids **II** support the formation of overlapping γ-turns and urea-turns (see Fig. 7). The comparison with type **III** urea/amide hybrids suggest that γ-turn nucleation could be favoured by the presence of a subsequent urea-turn. This conformational preference of oligo(amide/urea) hybrids in non-polar solvents parallels that observed in peptides composed of alternating α-aminoxy acids and α-amino acids.[27]

Examination of local folding and H-bonding patterns in model compounds can be extremely informative to gain insight into the propensity of longer-chain oligomers to adopt specific folding patterns based on remote interactions. Estimation of H-bonding between nearest-neighbour amide groups in simple β-alanine and γ-amino butyric acid derivatives was used by Gellman and coworkers as a criterion to estimate the relative propensity of β- and γ-peptide backbones to adopt compact and specific folding patterns.[14] By analogy, the 1←3 H-bonds that occur in model ureido compounds of type **C** and related oligo(urea/amide) hybrids is likely to compete with long-range order H-bonds, thus preventing the formation of secondary structures based on remote intrastrand interactions in longer oligomers. In contrast, the addition of a methylene in the main chain (*e.g.* type **B** residues) noticeably decreases the stability of the folded structure. Although the H-bond closing the 9-membered pseudocycle (Fig. 3b) is clearly visible in CCl$_4$, it is hardly populated in a slightly more polar solvent such as CH$_2$Cl$_2$. Theoretical calculations confirmed this trend and identified a number of alternative conformations that are more likely to be populated. Folding propensity does not increase significantly in corresponding diurea and triurea oligomers. However, the presence of four consecutive urea fragments in this series results in the appearance of a more rigid and folded structure, which probably corresponds to the 2.5-helical turn found in helical hepta- and nonaurea oligomers,[12] and which is reminiscent of the 14-helical structure of γ^4-peptides.[28]

Experimental

General

Amino acid derivatives were purchased from NeoMPS or Novabiochem. THF was freshly distilled from sodium/benzophenone under Ar. Toluene was distilled from P$_2$O$_5$ and stored over 4 Å molecular sieves. The reactions were carried out under an excess pressure of Ar. HPLC analysis was performed on a Nucleosil C$_{18}$ column (5 μM, 3.9 mm × 150 mm by using a linear gradient of A (0.1% TFA in H$_2$O) and B (0.08% TFA in MeCN) at a flow rate of 1.2 mL min^{-1} with UV detection at 214 nm. ^{1}H NMR and ^{13}C NMR spectra were recorded using Bruker Avance Apparatus DPX-300, ARX-300 and DRX-600. Chemical shifts (δ) are given in ppm, and J values are given in Hz. Optical rotations were obtained using a Perkin–Elmer polarimeter, with [a]$_D$ values being given in 10^{-1} deg cm^2 g^{-1}.

IR spectra were obtained in the Fourier transform mode on a Bruker IFS-25 apparatus. Matrix-assisted laser desorption ionization–time-of-flight (MALDI-TOF) mass analysis was performed on a linear MALDI-TOF instrument (Flex Control generated Xmass Data, 2000 Bruker Datoniks), using α-cyano-4-hydroxycinnamic acid as the matrix. Mass analyses in ESI (electrospray ionisation) mode were acquired on a LCQ Advantage MA MSn LCMS instrument (ThermoFischer Scientific).

General procedure for the preparation of ureas

To a stirred solution of the amine (1.1 equiv.) in 10 mL of MeCN or DMF were successively added succinimidyl carbamate (usually *ca.* 1 mmol) and Hunig's base (1.2 equiv.). After 10–30 min, the mixture was diluted with saturated NaHCO$_3$ and extracted with EtOAc. The organic layer was washed with 1 N KHSO$_4$, brine, saturated NaHCO$_3$ and brine, dried (Na$_2$SO$_4$), and evaporated. Flash chromatography and/or recrystallization afforded pure ureido peptidomimetics.

Analytical data for representative compounds I–VII

Boc-*g*Leu-CONH*i*Pr (I.1b). Boc-*g*Leu-COOSu (1.716 g, 5.00 mmol) was reacted with isopropylamine (511 μL, 6.00 mmol) according to the general procedure to yield **I.1b** (1.153 g, 80%): white solid, mp 123 °C. [a]$_D^{25}$ −13.7 (*c* 1.0 in DMF); δ$_H$(200 MHz, DMSO-d_6) 0.80 (6H, d, *J* 6.1, CH$_2$CH(C*H*$_3$)$_2$) 0.99 (6H, dd, *J* 6.5, 3.3, NHCH(C*H*$_3$)$_2$), 1.33–1.54 (10 H, m, C(C*H*$_3$)$_3$ + CH$_2$C*H*(CH$_3$)$_2$), 3.63 (1H, h, *J* 6.6, NHC*H*(CH$_3$)$_2$), 4.84–5.09 (1H, m, NHC*H*NH), 5.73–5.98 (2H, br m, N*H*CON*H*CH(CH$_3$)$_2$), 7.05 (1H, br d, *J* 6.9, N*H*COC(CH$_3$)$_3$); δ$_C$ (50 MHz, DMSO-d_6) 22.2 (CH$_3$), 23.1 (CH$_3$), 24.3 (CH), 25.1 (CH$_2$), 28.1 (CH$_3$), 40.7 (CH), 56.3 (CH), 77.6 (C), 154.7 (C), 156.1 (C); MALDI-TOF *m/z* 310.5 [M + Na]$^+$, 326.7 [M + K]$^+$.

Boc-*g*Phe-CONH*i*Pr (I.2b). Boc-*g*Phe-COOSu (100 mg, 0.26 mmol) was reacted with isopropylamine (68 μl, 0.79 mmol) according to the general procedure to yield **I.2b** (74 mg, 87%): white solid, mp 156 °C. [a]$_D^{25}$ −11.0 (*c* 1.1 in DMF); HPLC t_R = 12.81 min (linear gradient, 0–100% B, 20 min); δ$_H$(200 MHz, DMSO-d_6) 1.00 (3H, d, *J* 6.8, CH$_3$), 1.03 (3H, d, *J* 6.7, CH$_3$), 1.36 (9H, s, C(CH$_3$)$_3$), 2.88 (2H, br d, *J* 6.8, C*H*$_2$Ph), 3.65 (1H, m, C*H*(CH$_3$)$_2$), 5.14 (1H, m, NHC*H*NH), 5.95 (1H, br d, *J* 7.3, CON*H*CH), 6.09 (1H, br d, *J* 8.0, CON*H*CH), 7.16–7.32 (6H, m, arom. H + N*H*CO$_2$C(CH$_3$)$_3$); δ$_C$(50 Mhz, DMSO-d_6) 23.1 (CH$_3$), 28.1 (CH$_3$), 40.7 (CH), 40.9 (CH$_2$), 59.2 (CH), 77.7 (C), 126.0 (CH), 127.9 (CH), 129.1 (CH), 137.9 (C), 154.5 (C), 156.0 (C); MALDI-TOF *m/z* 344.9 [M + Na]$^+$, 360.6 [M + K]$^+$.

Boc-Pro-*g*Leu-CONHMe (II.1a). Boc-Pro-*g*Leu-COOSu (500 mg, 1.135 mmol) was reacted with HCl·NHMe (92 mg, 1.362 mmol) and DIEA (395 μL, 2.27 mmol) according to the general procedure to yield **II.1a** (320 mg, 79%): a white solid, mp 184 °C. [a]$_D^{25}$ −26.8 (*c* 1.0 in DMF); HPLC t_R = 5.93 min (linear gradient, 30–100% B, 20 min); δ$_H$(200 MHz, DMSO-d_6, signals of rotamers in italics) 0.83 (6H, d, *J* 5.8, CH(C*H*$_3$)$_2$), 1.30/*1.37* (10H, s, C(CH$_3$)$_3$ + C*H*(CH$_3$)$_2$), 1.40–1.58 (2H, m, C*H*$_2$CH(CH$_3$)$_2$), 1.64–1.85 (4H, m, CH$_2$C*H*$_2$C*H*$_2$CH), 2.50 (3H, d, *J* 4.6, NHC*H*$_3$), 3.16–3.43 (2H, m, NC*H*$_2$CH$_2$), 3.88–4.06 (1H, m, NC*H*CO), 5.19 (1H, br. t, *J* 7.9, NHC*H*NH), 5.97 (1H, br. d, *J* 4.4, NHCON*H*CH$_3$), 6.15 (1H, d, *J* 8.7, N*H*CONHCH$_3$),

7.99 (1H, d, *J* 7.5, N*H*COCH); δ$_C$(50 MHz, DMSO-d_6) 22.2 (CH$_3$) 22.3 (CH$_3$), 23.0 (CH$_3$), 24.2 (CH), 26.0 (CH$_3$), 27.8 (CH$_3$), 30.9 (CH$_2$), 43.7 (CH$_2$), 46.4 (CH$_2$), 55.3 (CH), 59.5 (CH), 78.3 (C), 153.4 (C), 157.3 (C), 171.7 (C); HRMS (ESI) *m/z* calcd for C$_{17}$H$_{33}$N$_4$O$_4$ [M + H]$^+$: 357.2496, found 357.2467.

Boc-Pro-*g*Leu-CO-Phe-OMe (II.2). Boc-Pro-*g*Leu-COOSu (500 mg, 1.135 mmol) was reacted with HCl·H-Phe-OMe (269 mg, 1.248 mmol) and DIEA (395 μL, 2.27 mmol) according to the general procedure to yield **II.2** (530 mg, 92%): a white solid, mp 168 °C. [a]$_D^{25}$ −17.7 (*c* 1.1 in DMF); HPLC t_R = 10.46 min (linear gradient, 30–100% B, 20 min); δ$_H$(200 MHz, DMSO-d_6) 0.69–0.88 (6H, m, CH$_3$), 1.21–1.54 (12H, s, C(C*H*$_3$)$_3$ + CH$_2$CH(CH$_3$)$_2$), 1.58–1.81 (3H, m, CH$_2$C*H*$_2$C*H*$_2$CH), 1.85–2.08 (3H, m, CH$_2$C*H*$_2$C*H*$_2$CH), 2.75–2.97 (2H, m, C*H*$_2$Ph), 3.15–3.41 (2H, m, NC*H*$_2$CH$_2$), 3.53 (3H, s, OC*H*$_3$), 3.85–4.01 (1H, m, NHC*H*CO), 4.25–4.44 (1H, m, NHC*H*CO), 5.02–5.24 (1H, m, NHC*H*NH), 6.36 (1H, br d, *J* 8.5, NHCON*H*), 6.45 (1H, br d, *J* 8.5, N*H*CONH), 8.97 (1H, br d, *J* 7.7, N*H*COCH); δ$_C$(50 MHz, DMSO-d_6) 22.3 (CH$_3$) 22.9 (CH$_2$), 27.9 (CH$_3$), 30.9 (CH$_2$), 37.8 (CH$_2$), 43.7 (CH$_2$), 46.4 (CH$_2$), 51.5 (CH), 51.5 (CH$_3$), 53.7 (CH), 55.2 (CH), 78.3 (C), 126.4 (CH), 128.1 (CH), 129.0 (CH), 136.9 (C), 153.3 (C), 155.9 (C), 171.6 (C), 172.6 (C); HRMS *m/z* calcd for C$_{26}$H$_{41}$N$_4$O$_6$ [M + H]$^+$: 505.3021, found 505.2730.

Cbz-Ala-*g*Leu-CONH*i*Pr (II.3). Cbz-Ala-*g*Leu-COOSu (2.65 g, 5.91 mmol) was reacted with isopropylamine (1.51 mL, 17.7 mmol) according to the general procedure to yield **II.3** (2.10 g, 91%): white solid, mp 170 °C. [a]$_D^{20}$ +1.5 (*c* 0.5 in DMF); HPLC t_R = 8.20 min (linear gradient, 30–100% B, 20 min); δ$_H$(300 MHz, DMSO-d_6) 0.84 (6H, d, *J* 5.9, CH$_2$CH(C*H*$_3$)$_2$) 0.98–1.02 (6H, m, NHCH(C*H*$_3$)$_2$), 1.17 (3H, d, *J* 7.1, CbzNHCHC*H*$_3$), 1.37-1.59 (3H, m, C*H*$_2$CH), 3.58–3.69 (1H, m, NHC*H*(CH$_3$)$_2$), 3.97 (1H, p, *J* 7.1, NHC*H*CO), 5.01 (1H, d, *J* 2.4, C*H*$_2$Ph), 5.14–5.24 (1H, m, NHC*H*NH), 5.94 (1H, d, *J* 7.5, N*H*iPr), 6.07 (1H, d, *J* 8.5, NHCHN*H*CONH), 7.29–7.36 (5H, m, arom. H), 7.39 (1H, d, *J* 7.8, CbzN*H*), 8.10 (1H, d, *J* 7.8, CHCON*H*); δ$_C$(75 MHz, DMSO-d_6) 23.4 (CH) 27.4 (CH$_3$), 27.6 (CH$_3$), 28.3 (CH$_3$), 28.4 (CH$_3$), 29.4 (CH), 46.0 (CH), 48.9 (CH$_2$), 55.2 (CH), 60.4 (CH), 70.5 (CH$_2$), 132.8 (CH), 132.9 (CH), 133.0 (CH), 133.5 (CH), 142.2 (C), 160.8 (C), 161.4 (C), 177.2 (C); HRMS (ESI) *m/z* calcd for C$_{20}$H$_{32}$LiN$_4$O$_4$ [M + Li]$^+$: 399.2579, found 399.2572.

Cbz-D-Ala-*g*Leu-CONH*i*Pr (II.4). Cbz-Ala-D-*g*Leu-COOSu (2.00 g, 4.46 mmol) was reacted with isopropylamine (1.14 mL, 13.38 mmol) according to the general procedure to yield **II.4** (1.32 g, 77%): white solid, mp 105 °C. [a]$_D^{20}$ −4.4 (*c* 0.5 in DMSO); HPLC t_R = 8.22 min (linear gradient, 30–100% B, 20 min); δ$_H$(300 MHz, DMSO-d_6) 0.84 (6H, d, *J* 3.3, CH$_2$CH(C*H*$_3$)$_2$) 0.98–1.02 (6H, m, NHCH(C*H*$_3$)$_2$), 1.17 (3H, d, *J* 7.1, CbzNHCHC*H*$_3$), 1.36–1.57 (3H, m, C*H*$_2$CH), 3.58–3.69 (1H, m, NHC*H*(CH$_3$)$_2$), 3.99 (1H, q, *J* 7.0, NHC*H*CO), 5.01 (1H, d, *J* 2.4, C*H*$_2$Ph), 5.15–5.24 (1H, m, NHC*H*NH), 5.99 (1H, d, *J* 6.8, N*H*iPr), 6.06 (1H, d, *J* 8.2, NHCHN*H*CONH), 7.28–7.37 (5H, m, arom. H), 7.35 (1H, d, *J* 7.4, CbzN*H*), 8.13 (1H, d, *J* 7.5, CHCON*H*); δ$_C$(75 MHz, DMSO-d_6) 18.8 (CH) 22.5 (CH$_3$), 22.9 (CH$_3$), 23.5 (CH$_3$), 23.7 (CH$_3$), 24.7 (CH), 41.3 (CH), 44.2 (CH$_2$), 50.5 (CH), 55.6 (CH), 65.8 (CH$_2$), 128.2 (CH), 128.8 (CH), 137.5 (C), 156.0 (C), 156.6 (C), 172.6 (C); HRMS (ESI) *m/z* calcd for C$_{20}$H$_{32}$N$_4$NaO$_4$ [M + Na]$^+$: 415.2316, found 415.2377.

This journal is © The Royal Society of Chemistry 2008

Org. Biomol. Chem., 2008, **6**, 2596–2610 | 2607

Boc-gSer(Bn)-CO-Leu-NHMe (III.1). White solid, mp 186 °C. HPLC t_R = 8.95 min (linear gradient, 30–100% B, 20 min); δ_H(300 MHz, DMSO-d_6) 0.83 (6H, d, J 6.5, CH(CH_3)$_2$) 0.85 (6H, d, J 6.5, CH(CH_3)$_2$), 1.37 (9H, s, tBuOCO), 1.41–1.23 (2H, m, CH_2CH(CH$_3$)$_2$), 1.58–1.45 (1H, m, CH_2CH(CH$_3$)$_2$), 2.56 (3H, d, J 4.4, NHCH_3), 3.43–3.40 (2H, m, CH_2OBn), 4.14–4.06 (1H, m, NCHCO), 4.48 (2H, s, OCH_2Ph), 5.28–5.18 (1H, m, NHCHNH), 6.25 (1H, d, J 8.6, NH), 6.37 (1H, d, J 8.6, NH), 7.17 (1H, d, J 7.8, tBuOCONH), 7.33–7.24 (5H, m, Ph), 7.90 (1H, q, J 4.8, NHCH$_3$); δ_C (75 MHz, DMSO-d_6), 22.4 (CH) 23.4 (CH$_3$), 24.7 (CH$_3$), 25.9 (CH$_3$), 28.6 (CH), 42.8 (CH$_2$), 51.8 (CH), 57.4 (CH), 71.4 (CH$_2$), 72.3 (CH$_2$), 78.4 (C), 127.8 (CH), 127.9 (CH), 128.6 (CH), 138.8 (C), 155.1 (C), 156.7 (C), 173.7 (C); HRMS (ESI) m/z calcd for C$_{22}$H$_{36}$LiN$_4$O$_5$ [M + Li]$^+$: 443.2841, found 443.2824.

Boc-gSer(Bn)-CO-D-Leu-NHMe (III.2). Boc-gSer(Bn)-COO-Su (1.15 g, 2.82 mmol) was reacted with CF$_3$COOH·H-D-Leu-NHMe (802 mg, 3.10 mmol) and DIEA (725 µL, 4.23 mmol) according to the general procedure to yield **III.2** (900 mg, 73%): white solid, mp 169 °C. $[a]_D^{20}$ −6.5 (c 1.0 in DMF); HPLC t_R = 8.90 min (linear gradient, 30–100% B, 20 min); δ_H(300 MHz, DMSO-d_6) 0.85 (6H, d, J 6.5, CH(CH_3)$_2$) 0.87 (6H, d, J 6.5, CH(CH_3)$_2$), 1.37 (9H, s, tBuOCO), 1.41–1.24 (2H, m, CH_2CH(CH$_3$)$_2$), 1.61–1.48 (1H, m, CH_2CH(CH$_3$)$_2$), 2.55 (3H, d, J 4.6, NHCH_3), 3.46–3.36 (2H, m, CH_2OBn), 4.12–4.04 (1H, m, NCHCO), 4.48 (2H, s, OCH_2Ph), 5.26–5.16 (1H, m, NHCHNH), 6.23 (1H, d, J 7.8, NH), 6.34 (1H, d, J 8.6, NH), 7.13 (1H, br, tBuOCONH), 7.36-7.24 (5H, m, Ph), 7.86 (1H, q, J 4.6, NHCH$_3$); δ_C (75 MHz, DMSO-d_6) 22.5 (CH) 23.4 (CH$_3$), 24.6 (CH$_3$), 25.9 (CH$_3$), 28.6 (CH), 42.8 (CH$_2$), 51.8 (CH), 57.5 (CH), 71.4 (CH$_2$), 72.3 (CH$_2$), 78.4 (C), 127.8 (CH), 127.9 (CH), 128.6 (CH), 138.7 (C), 155.0 (C), 156.7 (C), 173.7 (C); HRMS (ESI) m/z calcd for C$_{22}$H$_{36}$LiN$_4$O$_5$ [M + Li]$^+$: 443.2841, found 443.2830.

[[1-Benzyloxymethyl-2-(3-methylureido)ethyl]carbamic acid *tert*-butyl ester (V.2a). (S)-Succinimidyl-[(1-benzyloxymethyl-2-(*tert*-butoxycarbonylamino)ethyl]carbamate[17] (400 mg, 0.92 mmol) was reacted with methylamine hydrochloride (155 mg, 2.30 mmol) and DIEA (552 µL, 3.22 mmol) according to the general procedure to yield **V.2a** (340 mg, 98%): colourless oil. $[a]_D^{20}$ +3.9 (c 1.0, DMF); HPLC t_R = 8.58 min (linear gradient, 30–100% B, 20 min); δ_H(400 MHz, DMSO-d_6) 1.07 (3H, d, J 6.0, CHCH_3), 1.38 (9H, s, C(CH_3)$_3$), 2.53 (3H, d, J 4.5, NHCH_3), 2.96–2.03 (1H, m, NHCHCH_2), 3.19–3.26 (1H, m, NHCHCH_2), 3.52–3.59 (2H, m, NHCHCHOBn), 4.45 (1H, d, J 11.8, OCH_2Ph), 4.53 (1H, d, J 11.8, OCH_2Ph), 5.76 (1H, t, J 5.8, NH), 5.82–5.84 (1H, m, NH), 6.37 (1H, d, J 8.8, NH), 7.25–7.35 (5H, m Ph); δ_C (100 MHz, DMSO-d_6) 15.9 (CH$_3$) 26.7 (CH$_3$), 28.6 (3 CH$_3$), 40.7 (CH$_2$), 54.8 (CH), 70.3 (CH$_2$), 74.8 (CH), 78.0 (C), 127.6 (C), 127.8 (C), 128.5 (C), 139.3 (C), 156.0 (C), 159.1 (C); HRMS (ESI) m/z calcd for C$_{18}$H$_{29}$LiN$_3$O$_4$ [M + Li]$^+$: 358.2313, found 358.2303.

1-[3-Benzyloxy-2-(3-*tert*-butylureido)butyl]-3-methylurea (VI.1). Compound **V.2a** (240 mg, 0.69 mmol) was treated with TFA (2 mL) at 0 °C for 30 minutes. Concentration under reduced pressure and co-evaporation with cyclohexane left a residue which was dried under high vacuum. The resulting TFA salt was treated with *tert*-butyl isocyanate (91 µL, 0.87 mmol) and

DIEA (160 µL, 0.91 mmol) according to the general procedure to yield **VI.1** (220 mg, 92%): white solid, mp 156 °C. $[a]_D^{20}$ +3.4 (c 1.0 in DMF); HPLC t_R = 7.43 min (linear gradient, 30–100% B, 20 min); δ_H(400 MHz, DMSO-d_6) 1.07 (3H, d, J 6.2, CHCH_3), 1.20 (9H, s, C(CH_3)$_3$), 2.53 (3H, d, J 4.7, NHCH_3), 2.92–2.97 (1H, m, NHCH_2), 3.14–3.20 (1H, m, NHCH_2), 3.53–3.61 (1H, m, NHCH), 3.58–3.67 (1H, m, CHOBn), 4.41 (1H, d, J 11.5, OCH_2Ph), 4.55 (1H, d, J 11.6, OCH_2Ph), 5.50 (1H, d, J 9.1, NHCH), 5.81 (2H, t, J 5.3, NHCONHCH$_3$), 5.93 (1H, s, tBuNH), 7.26–7.38 (5H, m, Ph); δ_C(100 MHz, DMSO-d_6) 16.0 (CH$_3$) 26.4 (CH$_3$), 29.3 (3 CH$_3$), 41.5 (CH$_2$), 49.0 (C), 53.1 (CH), 70.2 (CH$_2$), 74.1 (CH), 127.3 (CH), 127.6 (CH), 128.1 (CH), 138.9 (C), 157.7 (C), 158.7 (C); HRMS (ESI) m/z calcd for C$_{18}$H$_{30}$LiN$_4$O$_3$ [M + Li]$^+$: 357.2473, found 357.2431.

1-[2-(3-Benzylureido)-3-phenylpropyl]-3-methylurea (VI.2). Compound **V.1a** (66 mg, 0.32 mmol) was treated with TFA (2 mL) at 0 °C for 30 minutes. Concentration under reduced pressure and co-evaporation with cyclohexane left a residue which was dried under high vacuum. The resulting TFA salt was treated with benzyl isocyanate (36 µL, 0.32 mmol) and DIEA (64 µL, 0.58 mmol) according to the general procedure to yield **VI.2** (77 mg, quant.): white solid, mp 197 °C. $[a]_D^{20}$ −0.140 (c 0.5 in CH$_2$Cl$_2$–MeOH 1 : 3); HPLC t_R = 5.72 min (linear gradient, 30–100% B, 20 min); ^{1}H NMR: See Tables S4–S6†; δ_C(75 MHz, DMSO-d_6) 31.6 (CH$_3$) 43.7 (CH$_2$), 48.0 (CH$_2$), 48.6 (CH$_2$), 56.8 (CH), 131.1 (CH), 131.7 (CH), 132.1 (CH), 133.3 (CH), 133.4 (CH), 134.4 (CH), 144.1 (C), 146.1 (C), 163.0 (C), 164.1 (C); HRMS (ESI) m/z calcd for C$_{19}$H$_{24}$LiN$_4$O$_2$ [M + Li]$^+$: 347.2054, found 347.2047.

Triurea VII. Compound **V.1a** (32 mg, 0.10 mmol) was treated with TFA (2 mL) at 0 °C for 30 minutes. Concentration under reduced pressure and co-evaporation with cyclohexane left a residue which was dried under high vacuum. The resulting TFA salt was treated with (S)-succinimidyl-[1-benzyl-2-(*tert*-butoxycarbonylamino)ethyl]carbamate[17] (30 mg, 0.10 mmol) and NMM (34 µL, 0.31 mmol) according to the general procedure to yield the corresponding diurea (40 mg, quant.). Treatment with TFA (2 mL) at 0 °C for 30 minutes followed by concentration under reduced pressure and co-evaporation with cyclohexane left a residue which was dried under high vacuum. The resulting TFA salt was treated with benzyl isocyanate (17 µL, 0.31 mmol) and DIEA (36 µL, 0.20 mmol) according to the general procedure to yield the triurea **VII** (40 mg, 97%): white solid, mp 195 °C. $[a]_D^{20}$ +0.046 (c 0.5 in CH$_2$Cl$_2$–MeOH 1 : 3); HPLC t_R = 8.72 min (linear gradient, 30–100% B, 20 min); ^{1}H NMR: See Tables S7–S9†; δ_C(75 MHz, DMSO-d_6) 31.6 (CH$_3$), 43.6 (CH$_2$), 43.7 (CH$_2$), 48.0 (CH$_2$), 48.5 (CH$_2$), 48.6 (CH$_2$), 56.5 (CH), 56.6 (CH), 131.1 (CH), 131.7 (CH), 132.0 (CH), 133.3 (CH), 134.4 (CH), 144.1 (2C), 146.0 (C), 163.1 (C), 163.2 (C), 164.1 (C); HRMS (ESI) m/z calcd for C$_{29}$H$_{36}$LiN$_6$O$_3$ [M + Li]$^+$: 523.3004, found 523.2969.

Tetraurea VIII. Compound **V.1a** (37 mg, 0.18 mmol) was treated with TFA (2 mL) at 0 °C for 30 minutes. Concentration under reduced pressure, co-evaporation with cyclohexane left a residue which was dried under high vacuum. The resulting TFA salt was treated with (S)-succinimidyl-[1-benzyl-2-(*tert*-butoxycarbonylamino)ethyl]carbamate[17] (52 mg, 0.18 mmol) and NMM (59 µL, 0.54 mmol) according to the general procedure to

2608 | *Org. Biomol. Chem.*, 2008, **6**, 2596–2610

This journal is © The Royal Society of Chemistry 2008

yield the corresponding diurea (70 mg, quant.). Deprotection and coupling steps were repeated to yield the corresponding triurea (97 mg, 80%). This triurea (40 mg, 0.06 mmol) was treated with TFA (2 mL) for 30 minutes then concentrated under reduced pressure. Co-evaporation with cyclohexane left a residue which was dried under high vacuum. The resulting TFA salt was treated with benzyl isocyanate (8 μL, 0.06 mmol) and DIEA (21 μL, 0.12 mmol) according to the general procedure to yield the tetraurea **VIII** (37 mg, 88%). HPLC t_R = 10.66 min (linear gradient, 30–100% B, 20 min); ^{1}H NMR: See Tables S10–S12†; δ_C(125 MHz, DMSO-d_6) 28.6 (CH$_3$), 40.6 (CH$_2$), 40.7 (CH$_2$), 40.9 (CH$_2$), 45.0 (CH$_2$), 45.2 (CH$_2$), 45.3 (CH$_2$), 46.0 (CH$_2$), 53.0 (CH), 53.4 (CH), 128.0 (CH), 128.1 (CH), 128.6 (CH), 128.8 (CH), 128.9 (CH), 129.2 (CH), 130.2 (CH), 130.3 (CH), 130.4 (CH), 131.3 (CH), 141.0 (C), 141.1 (C), 142.9 (C), 143.1 (C), 160.3 (C), 160.5 (C), 160.6 (C), 161.1 (C); HRMS (ESI) m/z calcd for C$_{39}$H$_{48}$N$_8$NaO$_4$ [M + Na]$^+$: 715.3691, found 715.3650.

Structural analyses in solution

^{1}H-NMR spectra were obtained with TMS as the internal reference. Concentrations used were 5 mM in CDCl$_3$ and DMSO-d_6. Proton resonances were assigned by COSY and NOESY experiments. The solvent sensitivity of the urea and amide NH protons, which is related to their free or hydrogen-bonded character, was investigated by considering the influence of DMSO-d_6 content in CDCl$_3$–DMSO-d_6 mixtures. The resonance of a solvent-exposed (free) proton is rapidly shifted to low fields whereas that of a hydrogen bonded (solvent-protected) proton is only weakly affected by DMSO-d_6 NH-solvation.[13a,29,30] IR spectra were obtained in the Fourier transform mode in order to investigate the NH (3200–3500 cm^{-1}) and CO (1580–1720 cm^{-1}) stretching frequencies in CCl$_4$, CH$_2$Cl$_2$ and DMSO. The concentration used depended both on the solvent and the nature of the compound. The IR spectra of compounds **I–III** with one 1,1-diaminoalkyl residue remained unchanged in DMSO whatever the concentration, but the concentration must be below 10 mM in CH$_2$Cl$_2$, and even 2 mM for compounds **I** in CCl$_4$, to exclude molecular aggregation. The same holds true for compounds **V** with one 1,2-diaminoalkyl residue in CH$_2$Cl$_2$, but the concentration must be reduced below 0.2 mM in CCl$_4$. In CH$_2$Cl$_2$ and CCl$_4$, a free secondary amide or urea group exhibits an NH absorption at 3400–3450 cm^{-1} and a CO absorption at 1650–1700 cm^{-1}. The above frequencies are reduced by 50–200 cm^{-1} and 10–20 cm^{-1}, respectively, when the NH and CO groups are hydrogen-bonded.[13a,30] The contributions of the residual water in the solvent, if any, were eliminated by correction in the 3500–3600 cm^{-1} region, where the peptide does not absorb.

General procedure for X-ray structure determination of compounds I.1b, I.3c, I.4b, II.1′a, II.2, V.1a, V.1c and VI.2. X-ray diffraction experiments for **I.1b**, **V.1c** and **VI.2** were carried out on a Nonius Mach3 four-circle diffractometer equipped with a graphite monochromator and a Cu rotating anode (Nonius FR591); for **I.3c**, **I.4b**, **II.1′a**, **II.2** and **V.1a** on a Bruker AXS Kappa CCD four-circle diffractometer equipped with a graphite monochromator and Mo sealed tube (Nonius FR590). The intensity data were corrected for Lp effects and no absorption correction was applied. The structures were solved by direct methods using the program SIR92.[31] Least-squares refinement

against F^2 was carried out on all non-hydrogen atoms using SHELXL97.[32] Hydrogen atoms were included by using a riding model (SHELXL97). The crystallographic data and the figures were prepared using WinGX[33] and Pymol,[34] respectively.

Pertinent crystallographic data are listed in Table S3. Crystallographic data can be found in the ESI.§

Theoretical calculations

All density functional theory (DFT) calculations were performed with the Gaussian03 suite of programs,[35] with the hybrid functional B3LYP. Although DFT approaches cannot be used to study systems whose intermolecular interactions are dominated by dispersion,[36] this method is reliable enough for the study of hydrogen bonds of the type N–H···O. Geometry optimisations were achieved in the gas phase without symmetry constraints by using the 6-31G(d,p) basis set. We have checked in some cases that this basis set yields results in close agreement with larger basis sets. For example, **C6A**–**I.M** lies 3.1 kcal mol^{-1} above **C8**–**I.M** in the 6-31G(d,p) basis set and is barely lowered in the 6-311++G(d,p) basis set, since it is found at $\Delta_r G$ = 2.9 kcal mol^{-1}. Moreover, geometry parameters, and among them the hydrogen bond feature, are not significantly modified. Several structural hypothesis were used as starting points. In addition, relaxed potential energy surface scans have been performed in order to probe the existence of alternative minima by varying ψ, ϕ or v. Calculation of vibrational frequencies was systematically done in order to characterize the nature of stationary points. Paths were traced from transition states to the corresponding minima using the Intrinsic Reaction Coordinate method.[37] Gibbs free energies $G°$ were calculated by means of the harmonic frequencies, *i.e.* by a straightforward application of the statistical thermodynamic equations.[38]

References

1 S. H. Gellman, *Acc. Chem. Res.*, 1998, **31**, 173–180.
2 *Foldamers: Structure, properties, and applications*, ed. S. Hecht and I. Huc, Wiley-VCH, Weinheim, 2007.
3 D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes and J. S. Moore, *Chem. Rev.*, 2001, **101**, 3893–4011.
4 C. M. Goodman, S. Choi, S. Shandler and W. F. DeGrado, *Nat. Chem. Biol.*, 2007, **3**, 252–262.

§ CCDC reference numbers are as follows: 668455 (**I.1b**), 668132 (**I.3c**), 668134 (**I.4b**), 668133 (**II.1′a**), 668136 (**II.2**), CCDC 668135 (**V.1a**), 668137 (**V.1c**) and 668138 (**VI.2**).

This journal is © The Royal Society of Chemistry 2008

Org. Biomol. Chem., 2008, **6**, 2596–2610 | 2609

5 (a) G. Guichard, in *Foldamers: Structure, properties and applications*, ed. S. Hecht and I. Huc, Wiley-VCH, Weinheim, 2007, pp. 35–74; (b) G. Guichard, in *Pseudopeptides in Drug Development*, ed. P.E. Nielsen, Wiley-VCH, Weinheim, 2004, pp. 33–120.

6 D. Seebach, A. K. Beck and D. J. Bierbaum, *Chem. Biodiversity*, 2004, **1**, 1111–1239.

7 I. Huc, *Eur. J. Org. Chem.*, 2004, 17–29.

8 X. Li and D. Yang, *Chem. Commun.*, 2006, 3367–3379.

9 (a) D. Yang, J. Qu, B. Li, F. F. Ng, X. C. Wang, K. K. Cheung, D. P. Wang and Y. D. Wu, *J. Am. Chem. Soc.*, 1999, **121**, 589–590; (b) D. Yang, J. Qu, W. Li, Y. Ren, D.-P. Wang and Y. D. Wu, *J. Am. Chem. Soc.*, 2003, **125**, 14452–14457; (c) D. Yang, Y. H. Zhang and N. Y. Zhu, *J. Am. Chem. Soc.*, 2002, **124**, 9966–9967; (d) D. Yang, Y. H. Zhang, B. Li, D. W. Zhang, J. C. Y. Chan, N. Y. Zhu, S. W. Luo and Y. D. Wu, *J. Am. Chem. Soc.*, 2004, **126**, 6956–6967; (e) D. Yang, D. W. Zhang, Y. Hao, Y. D. Wu, S. W. Luo and N. Y. Zhu, *Angew. Chem., Int. Ed.*, 2004, **43**, 6719–6722; (f) F. Chen, N. Y. Zhu and D. Yang, *J. Am. Chem. Soc.*, 2004, **126**, 15980–15981.

10 A. Aubry, J. P. Mangeot, J. Vidal, A. Collet, S. Zerkout and M. Marraud, *Int. J. Pept. Protein Res.*, 1994, **43**, 305–311.

11 (a) A. Cheguillaume, A. Salaün, S. Sinbandhit, M. Potel, P. Gall, M. Baudy-Floc'h and P. Le Grel, *J. Org. Chem.*, 2001, **66**, 4923–4929; (b) A. Salaün, M. Potel, T. Roisnel, P. Gall and P. Le Grel, *J. Org. Chem.*, 2005, **70**, 6499–6502; (c) A. Salaün, A. Favre, M. Le Grel, M. Potel and P. Le Grel, *J. Org. Chem.*, 2006, **71**, 150–158.

12 (a) V. Semetey, D. Rognan, C. Hemmerlin, R. Graff, J.-P. Briand, M. Marraud and G. Guichard, *Angew. Chem., Int. Ed.*, 2002, **41**, 1893–1895; (b) C. Hemmerlin, M. Marraud, D. Rognan, R. Graff, V. Semetey, J. P. Briand and G. Guichard, *Helv. Chim. Acta*, 2002, **85**, 3692–3711; (c) A. Violette, M. C. Averlant-Petit, V. Semetey, C. Hemmerlin, R. Casimir, R. Graff, M. Marraud, J.-P. Briand, D. Rognan and G. Guichard, *J. Am. Chem. Soc.*, 2005, **127**, 2156–2164; (d) A. Violette, S. Fournel, K. Lamour, O. Chaloin, B. Frisch, J.-P. Briand, H. Monteil and G. Guichard, *Chem. Biol.*, 2006, **13**, 531–538; (e) M. T. Oakley, G. Guichard and J. D. Hirst, *J. Phys. Chem. B*, 2007, **111**, 3274–3279.

13 (a) V. Semetey, C. Hemmerlin, C. Didierjean, A. P. Schaffner, A. G. Giner, A. Aubry, J. P. Briand, M. Marraud and G. Guichard, *Org. Lett.*, 2001, **3**, 3843–3846; (b) L. Fischer, V. Semetey, J. M. Lozano, A. P. Schaffner, J. P. Briand, C. Didierjean and G. Guichard, *Eur. J. Org. Chem.*, 2007, 2511–2525.

14 G. P. Dado and S. H. Gellman, *J. Am. Chem. Soc.*, 1994, **116**, 1054–1062.

15 Although rare among β-peptides, C_8-based conformations have been reported for particular types of β-peptides consisting either of (2R,3S)-α-hydroxylated β^{2,3}-amino acid residues, 1-aminomethylcyclopropanecarboxylic acid residues or *trans*-oxabornene-β-aminoacid residues: (a) K. Gademann, A. Häne, M. Rueping, B. Jaun and D. Seebach, *Angew. Chem., Int. Ed.*, 2003, **42**, 1534–1537; (b) S. Abele, P. Seiler and D. Seebach, *Helv. Chim. Acta*, 1999, **82**, 1559–1571; (c) R. J. Doerksen, B. Chen, J. Yuan, J. D. Winkler and M. L. Klein, *Chem. Commun.*, 2003, 2534–2535.

16 (a) A. A. Kemme, Y. Y. Bleidelis, Y. E. Antsans and G. I. Chipens, *Zh. Strukt. Khim.*, 1976, **17**, 1132–1135; (b) A. F. Mishnev, Y. Y. Bleidelis, Y. E. Antsans and G. I. Chipens, *Zh. Strukt. Khim.*, 1979, **20**, 154–157.

17 G. Guichard, V. Semetey, C. Didierjean, A. Aubry, J. P. Briand and M. Rodriguez, *J. Org. Chem.*, 1999, **64**, 8702–8705.

18 M. T. Cung, M. Marraud, J. Néel and A. Aubry, *Biopolymers*, 1978, **17**, 1149–1173.

19 G. Boussard and M. Marraud, *J. Am. Chem. Soc.*, 1985, **107**, 1825–1828.

20 V. S. Bryantsev, T. K. Firman and B. P. Hay, *J. Phys. Chem. A*, 2005, **109**, 832–842.

21 V. Villani, G. Alagona and C. Ghio, *Mol. Eng.*, 1999, **8**, 135–153.

22 K. A. Haushalter, J. Lau and J. D. Roberts, *J. Am. Chem. Soc.*, 1996, **118**, 8891–8896.

23 The torsional angles are defined in the same way as for peptides, according to IUPAC-IUB Commission on Biochemical Nomenclature: *Pure Appl. Chem.*, 1984, **56**, 595–624.

24 A search in the Cambridge Structural Database (CSD; Version 5.28; determined with R < 0.075, without disorder or errors and only organics (F. H. Allen, *Acta Crystallogr., Sect. B*, 2002, **58**, 380–388) indicated the presence of 84 structures with a urea motif similar to the one in compounds of type **A** and **B**. In all these structures the NH–CO–NH urea fragment is *trans,trans*-planar. The bond lengths and bond angles of the N,N'-disubstituted urea motif in compounds of type **A** and **B** (see ESI†) do not differ significantly from the standard values found for the urea motif in the CSD compounds.

25 T. L. Nguyen, F. W. Fowler and J. W. Lauher, *J. Am. Chem. Soc.*, 2001, **123**, 11057–11064.

26 E. N. Baker and R. E. Hubbard, *Prog. Biophys. Mol. Biol.*, 1984, **44**, 97–179.

27 D. Yang, W. Li, J. Qu, S. W. Luo and Y. D. Wu, *J. Am. Chem. Soc.*, 2003, **125**, 13018–13019.

28 T. Hintermann, K. Gademann, B. Jaun and D. Seebach, *Helv. Chim. Acta*, 1998, **81**, 983–1002.

29 T. P. Pitner and D. W. Urry, *J. Am. Chem. Soc.*, 1972, **94**, 1399–1400.

30 A. Aubry, M. T. Cung and M. Marraud, *J. Am. Chem. Soc.*, 1985, **107**, 7640–7647.

31 A. Altomare, G. Cascarano, G. Giacovazzo, A. Guagliardi, M. C. Burla, G. Polidori and M. Camalli, 'SIR92 – A Program for Automatic Solution of Crystal Structures by Direct Methods', *J. Appl. Crystallogr.*, 1994, **27**, 435.

32 G. M. Sheldrick, *SHELXL-97, Program for refinement of crystal structures*, University of Göttingen, Germany, 1997.

33 L. J. Farrugia, *J. Appl. Crystallogr.*, 1999, **32**, 837–838.

34 W. L. DeLano, *The PyMOL Molecular Graphics System*, 2002, see http://www.pymol.org.

35 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, *GAUSSIAN 03 (Revision B.05)*, Gaussian, Inc., Wallingford, CT, 2004.

36 A. K. Rappé and E. R. Bernstein, *J. Phys. Chem. A*, 2000, **104**, 6117–6128.

37 C. Gonzalez and H. B. Schlegel, *J. Phys. Chem.*, 1990, **94**, 5523–5527.

38 D. A. McQuarrie and J. D. Simon, *Molecular Thermodynamics*, University Science Books, Sausalito, California, 1999.

2610 | *Org. Biomol. Chem.*, 2008, **6**, 2596–2610

This journal is © The Royal Society of Chemistry 2008