PROTEINS: Structure, Function, and Bioinformatics 55:152–168 (2004)

Vicinal Disulfide Bridge Conformers by Experimental Methods and by Ab Initio and DFT Molecular Computations

Ilona Hudáky,[1] Zoltán Gáspári,[1] Oliviero Carugo,[2] Maša Čemažar,[2] Sándor Pongor,[2] and András Perczel[1]*
[1]*Department of Organic Chemistry, Eötvös Loránd University, Budapest, Hungary*
[2]*International Centre for Genetic Engineering and Biotechnology, Trieste, Italy*

ABSTRACT **A systematic comparison is made between experimental and computational data gained on vicinal disulfide bridges in proteins and peptides. Structural and stability data of ab initio and density functional theory (DFT) calculations on the model compound 4,5-ditiaheptano-7-lactam and the model peptide HCO—*ox*-[Cys—Cys]—NH_2 at RHF/3-21G*, B3LYP/6-31+G(d), and B3LYP/6-311++G(d,p) levels of theory are presented. The data on Xxx—Cys—Cys—Yyy type amino acid sequence units retrieved from PDB SELECT, along with data on sequence units that have vicinal disulfide bridge, taken from the Brookhaven Protein Data Bank, are conformationally characterized. Amino acid backbone conformations, *cis-trans* isomerism of the amide bond between the two cysteine residues, and ring puckering are studied. Ring puckers are characterized by their relation to the conformers of the parent 4,5-ditiaheptano-7-lactam. Computational precision and accuracy are proved by frequency calculation and solvent model optimization on selected conformers. It is found that the *ox*-[Cys—Cys] unit is able to accept types I, II, VIa, VIb, and VIII β-turn structures. Proteins 2004;55: 152–168.** © 2003 Wiley-Liss, Inc.

Key words: vicinal disulfide bridge; Cys-Cys; conformer; *cis-trans* isomerism; ring puckering; ab initio and DFT; PCM; β-turn

INTRODUCTION

Biological Relevance of Vicinal Disulfide Bridge Investigations

Disulfide bridges are most commonly known to link sequentially distant parts of polypeptide chains. Thus, the presence of covalent bonds between neighboring cysteines in some native proteins is rather surprising. It could be expected that these rare structural motifs have an important structural/functional role. However, the biochemical relevance of vicinal disulfides is characterized in detail only for a few proteins.

The most remarkable example is that of the alpha subunit of nicotinic acetylcholine receptors (nAchRs), where a conserved disulfide bridge specific for acetylcholine-binding pentameric ligand-gated ion channels is involved in ligand and toxin binding. Specific types of these receptors can be found in neuromuscular junctions as well as in pre- and postsynaptic nerve terminals. The role of nAchRs in nicotine addiction and diseases such as Parkinson's and Alzheimer's makes them important targets of drug design.[1]

In quinoprotein alcohol dehydrogenases and atracotoxin J-ACTX-Hv1c, the vicinal disulfide is proved to be essential for biological activity by site-directed mutagenesis studies. In the former case, the vicinal disulfide is not directly involved in catalysis, suggesting a structural role,[2,3] while in the latter case, the Cys → Ser double mutant toxin retains its native-like conformation;[4] thus, the S—S bond itself can be part of the recognition site by its cognate partner molecules.

Conformational Background

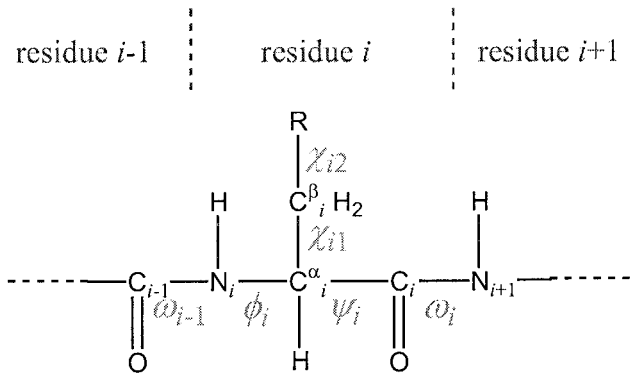
The conformational analysis of *ox*-[Cys—Cys] sequence units from experimentally determined protein and peptide structures and the systematic quantum chemical calculations on the HCO—*ox*-[Cys—Cys]—NH_2 model peptide are two different approaches to the description of vicinal disulfide bridges. Neither of them will be expected to deliver all possible conformers. On the one hand, the "experimental" approach can reveal only the energetically most favorable structures, thus reasonably populated minima. Therefore, certain conformers can escape detection simply because their statistical probability is too low. This approach is very sensitive to the size and the level of homology of the database used. On the other hand, computation can reveal all possible conformers available at the applied level of theory. However, a theoretical method ignores a set of factors associated with the molecular environment. In addition, the latter approach, applied on shorter model peptides, usually suggests a high stability for those conformers that incorporate locally preferred hydrogen-bonded systems. Meanwhile, it attributes higher energies to all other points of the conformational space,[5] and as a consequence, it may let a possible energy minimum converge into another minimum nearby in the

The Supplementary Materials Referred to in this article can be found at http://www.interscience.wiley.com/jpages/0887-3585/suppmat/index.html

Grant sponsor: Hungarian Scientific Research Foundation; Grant number: OTKA T 032486.

*Correspondence to: András Perczel, Department of Organic Chemistry, Eötvös Loránd University, 1117 Budapest P.O.B. 32, Hungary. E-mail: perczel@para.chem.elte.hu

Received 26 February 2003; Accepted 18 July 2003



Scheme 1. Selected torsional angles of the *i*th amino acid residue in a protein–peptide sequence: $\omega_{i-1}(C^{\alpha}_{i-1}—C_{i-1}—N_i—C^{\alpha}_i)$ is characteristic of amide bonds; ϕ_i ($C_{i-1}—N_i—C^{\alpha}_i—C_i$) and ψ_i ($N_i—C^{\alpha}_i—C_i—N_{i+1}$) reflect backbone conformation; χ_{i1} ($N_i—C^{\alpha}_i—C^{\beta}_i—C^{\gamma}_i$) and following side-chain torsional angles describe the orientation of the side-chain.

conformational space. Both facts have to be taken into account in the comparison of computational and statistical data, because most types of hydrogen bonds stabilizing proteins and longer peptides are formed between spatially close but sequentially remote, in our case, not vicinal, amino acid residues.

Every conformation of a given molecule is a point within its conformational space. This conformational space is defined by a set of independent conformational degrees of freedom. Peptide models are described by torsional angles of each amino acid residue[a] (see Scheme 1). A torsional angle of conformational relevance may accept two or more distinct values. An amide bond may be either *cis* or *trans* (i.e., $\omega \approx 0°$ or $\omega \approx 180°$, respectively). The backbone conformation of amino acid residues is conveniently represented by the Ramachandran map defined in terms of torsional angles ϕ and ψ, and thus may be labeled by a subscribed Greek letter as shown in Scheme 2. Side-chain torsional angles (e.g., χ_{i1}, χ_{i2}, etc.) usually may accept *gauche*− (or *g*−, ca. −60°), *gauche*+ (or *g*+, ca. 60°) or *anti* (or *a*, ca. 180°) conformation. The maximum number of conformers is derived by the multiplication of the numbers of possible minima at conformationally relevant and independent torsional angles.

In homoconformers, several amino acid residues of a sequence (e.g., *n* for the number of residues) adopt the same backbone conformation. An $(\alpha_L)_n$ homoconformer is a right-handed α-helix, a $(\beta_L)_n$ is a β-strand, while an $(\epsilon_L)_n$ is a polyproline II structure present in the triple helix of collagen. All these homoconformers have hydrogen bonds connecting residues spatially remoter than vicinal. Successful computational description of such homoconformers (i.e., energetical preference over other possible conformers) will be expected only when a model compound with sufficient length has been chosen, where the hydrogen bond pattern is repeated several times.[8] The numerous forms of turns, in contrast, are heteroconformers, because neighboring residues in a turn typically accept different backbone orientations. Sequences that have a tendency to fold into turns seem to be good subjects for computational investigation, because in their case, even a short model peptide may present the constraints and the hydrogen bonds determinant to the structure.[9] Such a sequence is Cys-Cys with a vicinal disulfide bridge, as it was reported to have a high preference[10] toward type VIII β-turn[11] conformation.

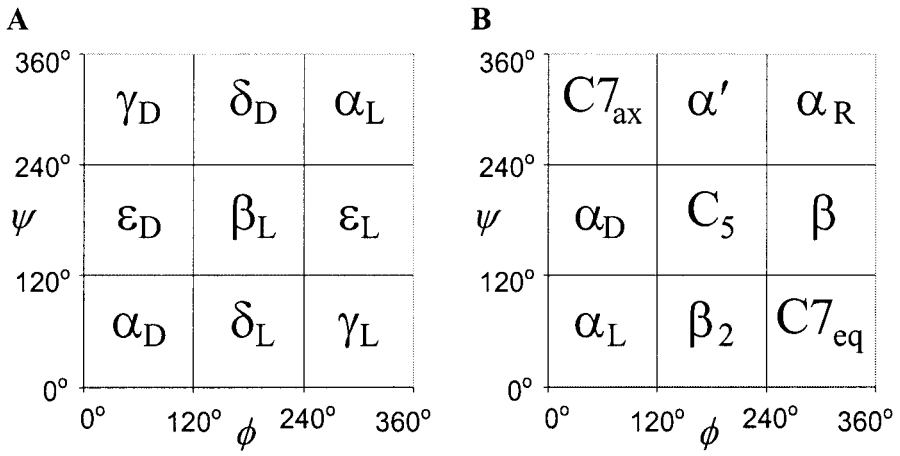
Different forms of β-turns are all involved in the reversal of the main-chain fold (e.g., β-sheet–β-turn–β-sheet motive) of peptides and proteins. Sequence units adopting a β-turn structure can be the target of several posttranslational modifications (glycosylation, phosphorylation), immune recognition, and other important biochemical processes. Hundreds of experimental and theoretical studies coupled to β-turns indicate that they are still of great importance.[12–14] A total of four residues are involved in the formation of β-turns, but those two ($i+1$ and $i+2$) located at the central part have a major role in the way reversal occurs. Different forms of β-turns (see Scheme 3) are most commonly distinguished: I, I′, II, II′, III, III′, VIa, VIb, VII, and VIII.[15] In globular proteins, the different types of β-turns exhibit very different natural abundances.[11,16–18]

Except for the two forms of type VI β-turn (see Scheme 3), the amide bond located between residue ($i+1$) and ($i+2$) is always *trans*. In types VIa and VIb, however, the same amide bond has a *cis* orientation. In proteins, amide bonds are typically in the *trans* conformation. According to a study on X-ray–determined protein structures,[22] 0.28% of the amide bonds are in the *cis* form. Among all these cases of a *cis* amide bond between residues Xxx and Yyy, in 90%, amino acid Yyy is Pro. In several proteins, a nonproline *cis* amide bond has been reported to lie near to the active site.[23–25]

The overall folding of a couple of consecutive amino acid residues can be described in a short and quantitative way after introducing parameters τ and d (see Scheme 4). An Xxx—Yyy sequence unit in a peptide or a protein is a β-turn if it reverses the backbone fold, which is a requirement likely to be fulfilled whenever $-90° \leq \tau \leq 90°$. A more strict definition of β-turns expects also that $d \leq 7$ Å.[11,19,26,27] An Xxx—Yyy sequence unit can be modeled in quantum chemical calculations by the compound For—Xxx—Yyy—NH_2. Considering that (1) such modeling replaces a C—C bond by a C—H bond and (2) the amide groups are nearly planar, a conformation of the model peptide that has the very same torsional angles as an Xxx—Yyy sequence unit in a longer polypeptide chain would result in almost the same dihedral angle τ and a somewhat shorter distance d.

This study aims at the computational description of the Cys—Cys sequence unit incorporating the vicinal disulfide bridge by ab initio and DFT methods. *Cis-trans* isomers, backbone folds, and ring puckers are all considered. Computational results are compared to relevant solvent model computations and experimental data. Both computation-

[a]In this article, the term *amino acid residue* refers to any amino acid preceded and followed by amide bonds, regardless of whether it is studied experimentally or computationally.



Scheme 2. (**A**) Labeling of the 9 ideal backbone conformers of an amino acid diamide according to their locations on a Ramachandran map.[6] (**B**) An alternative labeling applied in crystallography.[7]

Type of β-turn	Alternative structural code[19]	Dihedral angles (°)			
		ϕ_{i+1}	ψ_{i+1}	ϕ_{i+2}	ψ_{i+2}
I	$\alpha_L\delta_L$, $\alpha_L\alpha_L$, $\alpha_L\gamma_L$	−60	−30	−90	0
I′	$\alpha_D\delta_D$, $\alpha_D\alpha_D$, $\alpha_D\gamma_D$	60	30	90	0
II	$\varepsilon_L\alpha_D$, $\varepsilon_L\gamma_D$, $\varepsilon_L\delta_D$	−60	120	80	0
II′	$\varepsilon_D\alpha_L$, $\varepsilon_D\gamma_L$, $\varepsilon_D\delta_L$	60	−120	−80	0
III	$\alpha_L\alpha_L$	−60	−30	−60	−30
III′	$\alpha_D\alpha_D$	60	30	60	30
IV	Ambiguously defined	Types I–III′ β-turns with two or more torsional angles deviating more than 40° form the ideal values of Venkatachalam[20]			
V	$\gamma_L\gamma_D$	−80	80	80	−80
V′	$\gamma_D\gamma_L$	80	−80	−80	80
VIa*	$\varepsilon_L\alpha_L$, $\varepsilon_L\gamma_L$, $\varepsilon_L\delta_L$	−60	120	−90	0
VIb*	$\beta_L\alpha_L$, $\beta_L\gamma_L$, $\beta_L\delta_L$, $\beta_L\varepsilon_L$	−120	120	−60	0[11] or 150[21]
VII[13]	$\varepsilon_L\alpha_L$, $\varepsilon_L\gamma_L$, $\varepsilon_L\delta_L$	−90	140	−65	0
VIII	$\alpha_L\beta_L$	−60	−30	−120	120

Scheme 3. Major backbone conformational parameters of the most abundant β-turns, with traditional labels used in experimental protein and peptide science (column 1), together with the systematic labels used mainly in computational approaches (column 2). * represents *cis* amide bond between residues *i*+1 and *i*+2.

ally and experimentally favored conformers are assigned to traditional β-turn structures.

METHODS

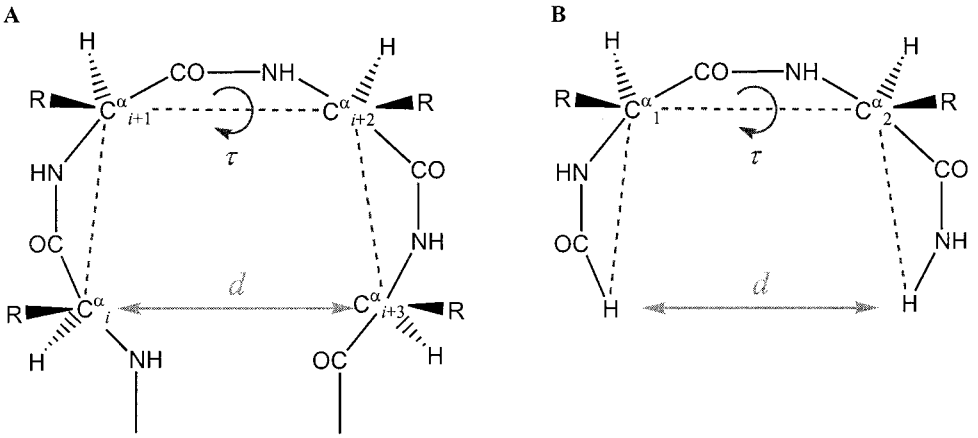
Ab Initio and DFT Molecular Computations on 4,5-Ditiaheptano-7-Lactam

4,5-Ditiaheptano-7-lactam was used as the simplest model for the conformational investigation of the 8-membered ring present in vicinal disulfide bridges (see Scheme 5). An exhaustive study was carried out at the RHF/3-21G* level of theory in order to discover all possible ring conformers. Each conformer was further optimized, first at B3LYP/6-31+G(d), and second at B3LYP/6-311++G(d,p) levels of theory. All calculations were carried out using the GAUSSIAN98 program package[28] with the GDIIS algorithm.[29–31]

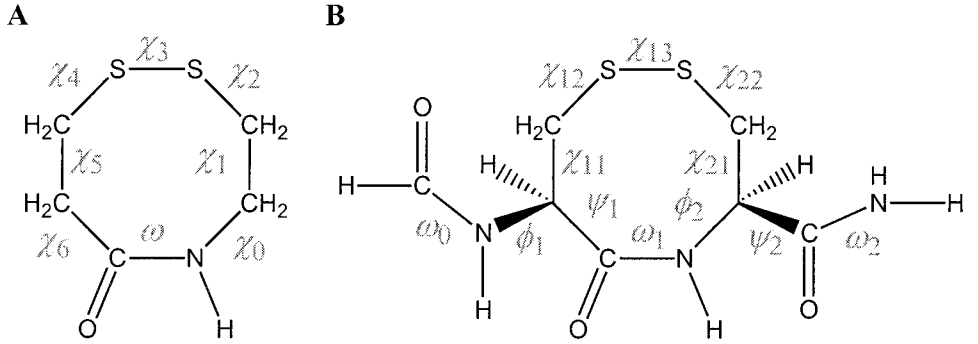
Conformers were labeled by a two-character code. The first letter is either *c* or *t*, reflecting whether the endocyclic amide bond is *cis* or *trans*. The second character stands for a simple numbering of the overall ring conformation (e.g., *c*5, *c*6, etc., see Table I). This ring is not planar; instead it is puckered in order to prevent eclipsing of bonds (see Fig. 1). Due to nonplanarity, the molecule loses its internal plane of symmetry; therefore, none of the conformers is identical with its mirror image. To emphasize the mirror image relationships between the conformers, Table I gives an explicit reference by providing an alternative label to each presented conformer: the label of the mirror image followed by a dash (e.g., *c*5 is identical with the mirror image of *c*1; hence, the label *c*1′ is derived as an alternative to *c*5, etc.).

Ab Initio and DFT Molecular Computations on the HCO—*ox*-[Cys—Cys]—NH_2 Model Peptide

At the RHF/3-21G* level of theory, all expected conformers of the HCO—*ox*-[Cys—Cys]—NH_2 model peptide were



Scheme 4. Definitions of dihedral angle τ and distance *d*. (**A**) For protein/peptide chains, dihedral angle τ is defined by four consecutive C^{α} atoms (C_i^{α}—C_{i+1}^{α}—C_{i+2}^{α}—C_{i+3}^{α}) of the polypeptide chain, given in degree. Parameter *d* is the atomic distance between C_i^{α} and C_{i+3}^{α}, given in angstrom units. (**B**) In the case of model peptide For—Xxx—Yyy—NH_2, both C_i^{α} and C_{i+3}^{α} are replaced by a hydrogen atom; thus, parameter τ is defined as dihedral angle H—C_1^{α}—C_2^{α}—H, while *d* is the atomic distance between these two hydrogen atoms, given in angstrom units.



Scheme 5. (**A**) Selected torsional angles of 4,5-ditiaheptano-7-lactam. All torsional angles are defined by ring atoms. (**B**) Selected torsional angles of the HCO—*ox*-[Cys—Cys]—NH_2 model peptide.

optimized. All ring conformers of 4,5-ditiaheptano-7-lactam were considered. While mirror image ring conformers (e.g., *c*6 and *c*2) have the same energy, hence, the same expected population in 4,5-ditiaheptano-7-lactam, they have to be examined independently in the model peptide. Adding the two pendant moieties (HCONH— and —$CONH_2$), the helical chirality of the ring is combined with the point chirality at the C^{α} atoms of the L-cysteine molecules. In the cases of torsional angles ϕ_1 and ψ_2, which are both independent of the ring, three distinct minima (i.e., *gauche*−, *gauche*+ and *anti*) were considered. All low-energy conformers, together with some higher energy conformers, were further optimized at the B3LYP/6-31+G(d) level of theory.

In order to investigate accuracy and precision obtained by simple minimization at the B3LYP/6-31+G(d) level of theory, frequency calculations as well as Polarizable Continuum Model (PCM)[33] single-point energy calculations were carried out for all low-energy conformers at the same level of theory. Each low-energy conformer was further optimized at the B3LYP/6-311++G(d,p) level of theory without solvent model, as well as at the B3LYP/6-31+G(d) level of theory with the application of PCM. The resultant thermodynamic parameters, relevant partly to gas phase and partly to aqueous solution, were compared to B3LYP/6-31+G(d) data. Most calculations were carried out using the GAUSSIAN98 program package[28] and the GDIIS algorithm.[29–32] PCM optimizations were performed using the GAUSSIAN03 program package and the GDIIS algorithm.[31]

A code is introduced to label the *ox*-[Cys—Cys] sequence unit irrespective of whether studied within a protein/peptide structure or within the HCO—*ox*-[Cys—Cys]—NH_2 model peptide. The set of torsional angles ψ_1, ω_1, ϕ_2, χ_{11}, χ_{13}, and χ_{21} describes unambiguously the ring; thus, the peptide conformers can be ascribed to the conformers of 4,5-ditiaheptano-7-lactam (see Scheme 5 and Table I). Note that torsional angles χ_{12}, χ_{13}, χ_{22}, χ_{21}, and ω_1 of the model peptide correspond to parameters χ_4, χ_3, χ_2, χ_1, and ω in the 8-membered ring compound. The definition of χ_{11}, ψ_1, and ϕ_2, conventional in peptide/protein nomenclature (see Scheme 1), does not allow strict correspondence to ring torsional angles, which are defined exclusively by ring atoms, though they are undoubtedly related to χ_5, χ_6, and χ_0, respectively. The label of the ring conformer (e.g., *c*5, *c*6,

TABLE I. Conformers of 4,5-Ditiaheptano-7-Lactam Optimized at Three Different Levels of Theory

RHF/3-21Gd*	E_h	ΔE	χ_0	χ_1	χ_2	χ_3	χ_4	χ_5	χ_6	ω
*c*5 (*c*1′)	−1113.438088	0.01	84.95	−80.19	76.75	−95.68	49.61	57.95	−107.96	0.93
*c*6 (*c*2′)	−1113.438111	0.00	96.71	−52.95	−52.68	96.44	−75.83	83.47	−94.41	12.13
*c*7 (*c*3′)	−1113.425905	7.66	78.43	−58.16	−55.05	78.21	24.98	−83.52	12.45	9.10
*c*8 (*c*4′)	−1113.423068	9.44	−45.31	91.07	−19.31	−79.83	52.21	64.64	−94.07	25.14
*t*5 (*t*1′)	−1113.428615	5.96	−45.96	−57.80	32.86	66.58	−97.69	61.97	−81.24	142.55
*t*6 (*t*2′)	−1113.437700	0.26	−59.10	−48.94	91.52	−88.40	93.91	−45.95	−56.37	149.58
*t*7 (*t*3′)	−1113.425481	7.93	−89.38	48.03	−90.75	81.72	20.03	−57.27	−25.51	140.11
*t*8 (*t*4′)	−1113.437359	0.47	−100.58	56.06	−78.33	102.48	−86.09	64.23	−88.28	146.72
B3LYP/6-31 + G(d)										
*c*5 (*c*1′)	−1122.335805	0.25	81.79	−81.29	76.65	−92.79	48.96	56.90	−108.14	4.43
*c*6 (*c*2′)	−1122.336202	0.00	91.20	−54.38	−51.10	94.29	−73.67	81.49	−97.14	18.97
*c*7 (*c*3′)	−1122.326197	6.28	71.99	−58.54	−54.11	76.94	24.02	−78.03	1.88	21.66
*c*8 (*c*4′)	−1122.324337	7.44	−52.55	92.23	−17.55	−79.16	51.28	64.28	−96.01	31.08
*t*5 (*t*1′)	−1122.326353	6.18	−45.66	−57.37	33.26	64.66	−95.25	61.47	−80.95	139.89
*t*6 (*t*2′)	−1122.332390	2.39	−60.61	−47.87	90.67	−86.58	91.20	−46.19	−53.73	145.09
*t*7 (*t*3′)	−1122.323703	7.84	−87.79	48.54	−89.43	81.54	17.41	−57.24	−24.36	136.92
*t*8 (*t*4′)	−1122.334458	1.09	−98.12	58.12	−78.33	99.08	−83.94	64.65	−90.20	142.73
B3LYP/6-311++G(d, p)										
*c*5 (*c*1′)	−1122.472673	0.28	81.65	−81.75	77.13	−93.10	48.94	57.40	−108.73	4.43
*c*6 (*c*2′)	−1122.473118	0.00	91.70	−54.57	−51.24	94.57	−74.28	82.22	−97.01	18.67
*c*7 (*c*3′)	−1122.463055	6.32	72.02	−58.96	−53.83	76.93	24.09	−78.91	2.65	21.69
*c*8 (*c*4′)	−1122.461283	7.43	−53.07	92.83	−17.27	−79.20	50.82	65.05	−95.95	30.55
*t*5 (*t*1′)	−1122.463366	6.12	−46.57	−58.27	34.88	63.53	−95.91	62.34	−80.16	139.78
*t*6 (*t*2′)	−1122.469531	2.25	−61.15	−48.14	90.68	−86.42	91.58	−46.84	−52.95	145.28
*t*7 (*t*3′)	−1122.460996	7.61	−87.93	48.67	−89.48	81.27	18.11	−58.06	−23.91	137.01
*t*8 (*t*4′)	−1122.471517	1.00	−98.76	58.21	−78.38	99.34	−84.33	64.99	−89.70	142.89

Conformer: labeled according to Methods section. Only one of each mirror image conformer pair is presented. Conformers *c*1, *c*2, *c*3, *c*4, *t*1, *t*2, *t*3, and *t*4 are mirror images of conformers of *c*5, *c*6, *c*7, *c*8, *t*5, *t*6, *t*7 and *t*8, respectively. This relationship is emphasized by the alternative labels in parentheses. The torsional angles of mirror image conformers are the same in absolute value but different in sign. E_h in Hartree; ΔE in kcal · mol^{-1}; torsional angles in degrees.

etc.) is followed by a pair of subscripted Greek letters (e.g., $\beta_L\beta_L$, $\beta_L\gamma_D$, etc.), which characterize the backbone conformation of the two adjacent cysteine residues according to Scheme 2(A). This code (e.g., *c*5_$\beta_L\beta_L$, *t*5_$\beta_L\gamma_D$, etc.) is redundant in the sense that both the first part referring to the ring and the second part reflecting the peptide backbone conformation imply the knowledge about torsional angles ϕ_1 and ψ_2 (see Table II).

Experimental Data Analysis of Xxx—Cys—Cys—Yyy and Xxx—*ox*-[Cys—Cys]—Yyy Peptide Sequences

Xxx—Cys—Cys—Yyy type amino acid sequences were searched for in PDB SELECT[35,36] (2002 April update). PDB SELECT was further filtered, as entries with ambiguous residue numbering (non-numeric characters in the residue number field) and alternative conformers (at least one protein atom with more than one coordinate set) were not considered in the analysis. Altogether, 120 sequence units were found that build up a homology-filtered set of the Cys—Cys unit. Of these, only 3 contain a vicinal disulfide bridge. Because of this low occurrence, and in order to obtain a set of *ox*-[Cys—Cys] units acceptable for statistical analysis, sequences with vicinal disulfide bridges were searched for in the entire Brookhaven PDB[37] (Rel. 103, January 2003). The 48 sequence units found there represent an unfiltered set of the Xxx—*ox*-[Cys—Cys]—Yyy unit from proteins with determined structure. Both sets were conformationally characterized.

RESULTS

Ab Initio and DFT Molecular Computations on 4,5-Ditiaheptano-7-Lactam

A thorough ab initio investigation at the RHF/3-21G* level of theory found 16 different conformers of 4,5-ditiaheptano-7-lactam (see Table I). All conformers exist also at B3LYP/6-31+G(d) and B3LYP/6-311++G(d,p) levels of theory. Eight conformers have the *cis* amide bond, and the other 8 have the *trans* amide bond. Each conformer has a nonidentical mirror image. The energies of the mirror image conformer pairs are equal (i.e., $\Delta E^{c1} = \Delta E^{c5}$, etc.), while their torsional angles are the same in absolute value but different in sign ($\chi_0^{c1} = -\chi_0^{c5}$, etc.).

Most torsion angles are "well predicted" by the RHF/3−21G* level of theory, though a few differ by more than 10° from those values obtained by DFT methods. The largest deviation of ΔE obtained with the ab initio method from corresponding DFT data is 2 kcal · mol^{-1}. Data gained at the two DFT levels correlate even better, thus proving B3LYP/6-311+G(d) level calculations good enough. The greatest deviation of ΔE is 0.24 kcal · mol^{-1}, that of the torsional angles is 1.6°. Considering only energetically

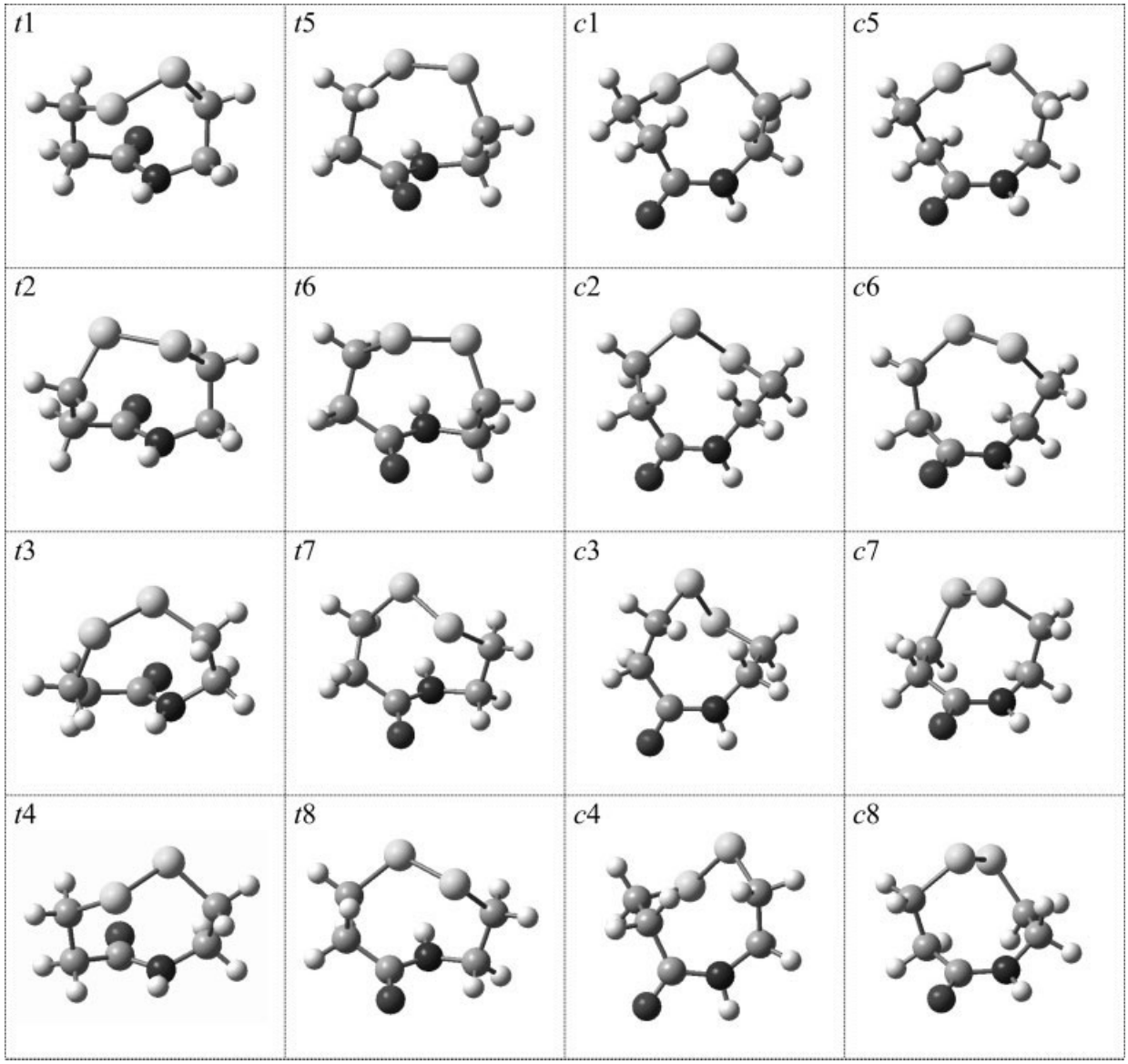


Fig. 1. Ring conformers of 4,5-ditiaheptano-7-lactam oriented according to Scheme 5(A). Conformers *c*1, *c*2, *c*3, *c*4, *t*1, *t*2, *t*3, and *t*4 are mirror images of conformers *c*5, *c*6, *c*7, *c*8, *t*5, *t*6, *t*7, and *t*8, respectively. Conformer pairs, without emphasis on the mirror image relationship, are captured from slightly different viewpoints.

favored conformers, the corresponding values are as low as 0.14 kcal · mol^{-1} and 0.8°.

The global minima are mirror image conformers *c*2 and *c*6 (or alternatively denoted as *c*2′ to emphasize that *c*6 is the mirror image of *c*2) at all three levels of theory. Conformers were divided into two groups according to their relative energies, ΔE, over the global minimum, *c*2 or *c*6. Energetically favored conformers have ΔE that is not more than 2.5 kcal · mol^{-1}. Conformers *c*1, *c*2, *c*5, *c*6, *t*2, *t*4, *t*6, and *t*8 are energetically favored. Energetically disfavored conformers, *c*3, *c*4, *c*7, *c*8, *t*1, *t*3, *t*5, and *t*7, have relative energies, ΔE, between 5 and 10 kcal · mol^{-1}.

Though linear peptides prefer *trans* to *cis* amide bonds, this 8-membered lactam ring adopts more easily a *cis* bond, while it is more constrained with a *trans* bond. The average deviation of parameter ω from the ideal values (i.e., 0° and 180°) is 12–19° in the case of *cis* bonds and 35–39° in the case of *trans* bonds.

In order to form a ring with a *trans* bond, both —CH_2—CH_2—S— moieties have to be on the same side of the amide plane. In conformers *t*1, *t*2, *t*3, and *t*4, this requirement is achieved by a positive χ_0, ω of about −145°, and a positive χ_6. Thus, ω may be viewed as one of the "governing coordinates" with two possible minima (e.g., −145° or +145°), while χ_0 and χ_6 depend on ω. χ_1 and χ_5 appear to be the other two "governing coordinates." Neither of them can be *anti*, as the ring system of this size seems to allow only one torsional angle, namely ω, to approach *anti* conformation. Both χ_1 and χ_5 are either *gauche*+ or *gauche*−, and the conformers represent their

TABLE II. Low-Energy Conformers of HCO—*ox*-[Cys—Cys]—NH_2 Optimized at the B3LYP/6-31+G(d) Level of Theory

Conformer	β-turn	ΔE	ΔE′	ϕ_1	ψ_1	ω_1	ϕ_2	ψ_2	χ_{11}	χ_{13}	χ_{21}	τ	*d*
*c*5_$\beta_L\beta_L$	ud.	0.98		−122.40	135.25	−2.99	−151.88	149.77	173.93	−94.16	−80.45	1.10	7.94
*c*5_$\beta_L\delta_L$	VIb	2.12		−127.64	134.93	3.62	−147.57	27.67	169.16	−94.25	−82.27	−35.23	7.38
*c*5_$\varepsilon_L\beta_L$	ud.	0.98		−114.06	133.66	−2.44	−152.49	148.84	175.13	−94.10	−80.34	2.25	7.79
*c*6_$\beta_L\beta_L$	ud.	0.00		−154.28	150.16	8.78	−141.47	141.14	−160.11	95.42	−54.04	18.69	8.35
*c*6_$\beta_L\delta_L$	VIb	0.38		−155.45	146.66	16.63	−139.17	18.97	−162.87	96.42	−55.29	−13.09	7.73
*c*6_$\varepsilon_L\delta_L$	VIa	1.84		−80.67	136.58	14.96	−138.14	65.68	−155.83	94.15	−51.75	11.81	6.19
*t*2_$\gamma_D\beta_L$	ud.	7.00	6.14	69.92	−73.38	−147.42	−176.39	162.27	177.37	87.83	46.64	−28.38	7.88
*t*2_$\gamma_D\delta_L$	ud.	3.97	3.11	68.04	−75.95	−145.89	−167.83	23.55	175.09	89.89	43.64	−13.23	6.09
*t*4_$\alpha_L\beta_L$	VIII	3.54	2.68	−89.11	−25.95	−142.67	−145.87	113.83	58.84	−97.48	−60.37	52.02	7.45
*t*4_$\alpha_L\delta_L$	I	2.02	1.16	−77.81	−27.28	−141.29	−139.07	19.69	58.28	−96.87	−62.34	38.80	6.30
*t*5_$\beta_L\gamma_D$	ud.	4.10	3.25	−157.00	159.61	138.46	86.72	−41.74	178.21	70.65	−59.98	9.44	7.93
*t*6_$\beta_L\alpha_D$	ud.	2.20	1.34	−159.07	175.82	144.06	77.45	22.48	75.88	−86.66	−53.20	54.32	7.95
*t*6_$\beta_L\gamma_D$	ud.	0.86	0.00	−159.51	175.12	143.06	81.39	−39.60	76.38	−86.27	−55.76	33.40	7.89
*t*8_$\beta_L\alpha_D$	ud.	4.23	3.38	−156.90	148.45	141.26	46.41	7.78	−178.04	104.08	50.32	−34.67	7.88

Conformer: labeled according to Methods section. β-turn, type of β-turn according to Scheme 3; ud.; undefined; Δ*E,* relative energy in kcal · mol^{-1} above the global minimum (*c*6-$\beta_L\beta_L$, E_h = −1459.743551). Δ*E*′, relative energy in kcal · mol^{-1} above minimum *t*6_$\beta_L\gamma_D$ (E_h = −1459.742186). All torsional angles in degrees; *d* in Å.

four variations. The remaining three torsional angles, χ_2, χ_3, and χ_4, all depend on the conformation of ω, χ_1, and χ_5. In all four energetically favored *trans* type conformers (i.e., *t*2, *t*4, *t*6, and *t*8), the signs of torsional angles χ_1, χ_2, χ_3, χ_4, and χ_5 alternate; $\chi_1 \approx \chi_5$ and $\chi_2 \approx \chi_4$; hence, a digir is found in these molecules if the nonequivalence of atoms N and carbonyl C is neglected.

The alternation of the signs of torsional angles is preferred in *cis* rings, too. In each of the energetically favored conformers, only one neighboring torsional angle pair has identical helicity (i.e., $\chi_4 \approx \chi_5$ in *c*1 and *c*5, but $\chi_1 \approx \chi_2$ in *c*2 and *c*6).

Ab Initio and DFT Molecular Computations on the HCO—*ox*-[Cys—Cys]—NH_2 Model Peptide

Out of the 16 × 3 × 3 = 144 ideal conformers (16 ring conformers, 3 orientations for both ϕ_1 and ψ_2) of the HCO—*ox*-[Cys—Cys]—NH_2 model peptide, 91 were found at the RHF/3-21G* level of theory (see Table S.I in the Supplementary Information and Fig. 2). Within the model peptide, the same ring conformers were energetically favored as in 4,5-ditiaheptano-7-lactam, except that *c*1 and *c*2 turned out to be disfavored, while *t*5 became favored. Distinction between energetically favored and disfavored peptide conformers is not as evident as in the case of 4,5-ditiaheptano-7-lactam, because preferences of the ring are combined by the completely independent preferences of the backbone folding (the latter aspect is well analyzed on the HCO—Ala—Ala—NH_2 model peptide).[27,38]

In the course of optimization, most of those initial geometries that did not deliver a genuine minimum underwent conformational changes involving mainly the backbone torsions, while retaining the original ring conformer. Only one *trans* ring conformer change was detected (namely *t*7 → *t*8, i.e., one ideal peptide conformer with a *t*7 ring converged to a genuine minimum with a *t*8 ring). As the ring itself is rather constrained with a *trans* amide bond, ring torsional angles in *trans* peptide conformers do not differ much from those in the relevant 4,5-ditiaheptano-7-lactam. *Cis* rings, however, seem more flexible; thus, their ring torsional angles tend to change considerably in order to accommodate the two pendant moieties, HCONH— and —$CONH_2$. Four obvious cases of ring alteration were detected for *cis* ring conformers (two cases for both *c*3 → *c*8 and *c*4 → *c*6), and even the ring torsional angles of peptide conformers with presumably the same ring differ to a great extent. Nevertheless, subdivision of *cis* ring conformers is unnecessary, as low energy conformers, *c*5 and *c*6, are still very well defined by their torsional angles.

As a general observation about the torsional angle patterns of the RHF/3-21G* conformers (see Fig. 2), it can be stated that *cis* ring conformers are more flexible (i.e., their ring torsional angles tend to vary within wider ranges, while *trans* ring conformers are rather constrained, i.e. their ring torsional angles fall into narrower ranges). This is due to the relatively small ring, as discussed in the previous section, which is less constrained with a *cis* amide bond (although an unconstrained amide bond usually prefers *trans* conformation). Torsional angle ω_1 varies from −78° to +51° for *cis* bonds; but the deviation from the ideal value of 0° is less than 13° in the case of low-energy conformers. Deviation from the ideal *trans* value (i.e., 180°) is always greater than 26°; two groups of *trans* ω_1 values are formed at around 142° and −147°. Torsional angle ψ_1 may be either *gauche*− or *anti*, and parameter ϕ_2 is either *gauche*+ or *anti*; all four ψ_1/ϕ_2 combinations are found for *cis* rings, but only *anti*/*gauche*+ and *gauche*−/*anti* combinations are found for *trans* ring conformers.

All RHF/3-21G* conformers with relative energy, Δ*E*, over a global minimum (*t*6_$\beta_L\gamma_D$) not more than 7 kcal · mol^{-1}, together with selected higher energy conformers, were further optimized at B3LYP/6-31+G(d) level of theory (see Table S.II in the Supplementary Information) in order to establish correlation between ab initio and DFT calculations. Relative energies obtained by the two different methods correlate well for *cis* conformers and poorly for

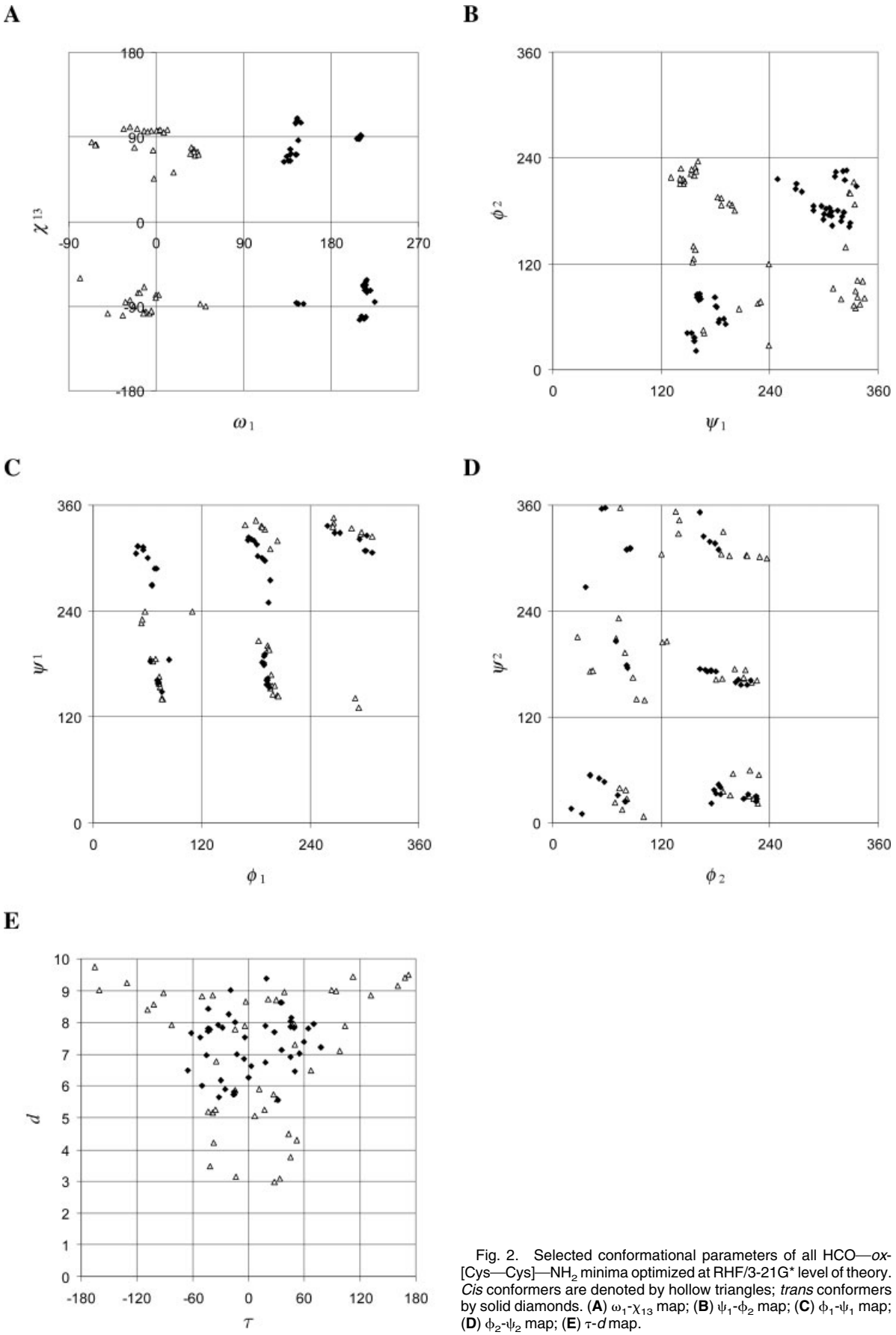


Fig. 2. Selected conformational parameters of all HCO—*ox*-[Cys—Cys]—NH_2 minima optimized at RHF/3-21G* level of theory. *Cis* conformers are denoted by hollow triangles; *trans* conformers by solid diamonds. (**A**) ω_1-χ_{13} map; (**B**) ψ_1-ϕ_2 map; (**C**) ϕ_1-ψ_1 map; (**D**) ϕ_2-ψ_2 map; (**E**) τ-*d* map.

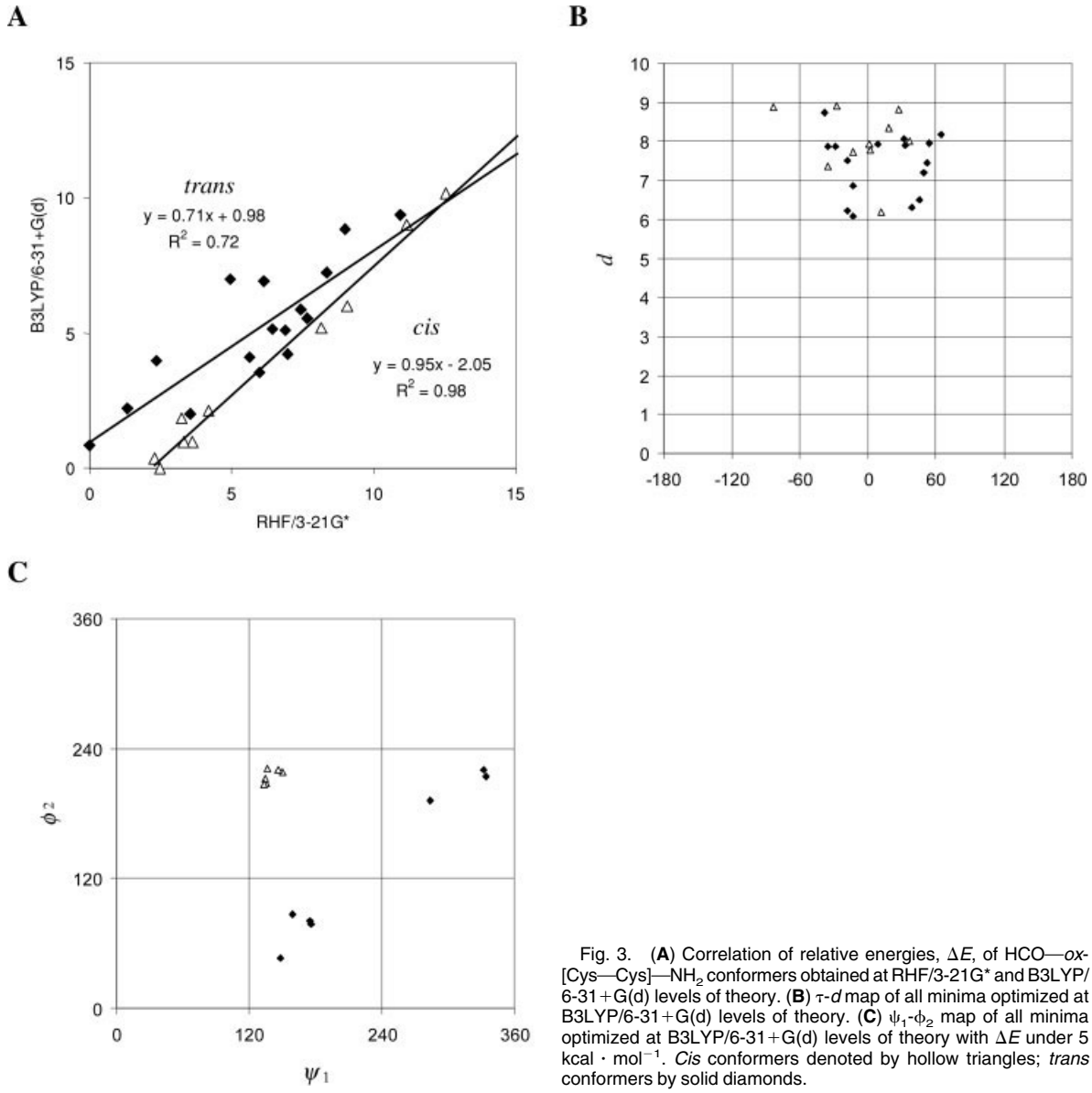


Fig. 3. (**A**) Correlation of relative energies, ΔE, of HCO—*ox*-[Cys—Cys]—NH_2 conformers obtained at RHF/3-21G* and B3LYP/6-31+G(d) levels of theory. (**B**) τ-*d* map of all minima optimized at B3LYP/6-31+G(d) levels of theory. (**C**) ψ_1-ϕ_2 map of all minima optimized at B3LYP/6-31+G(d) levels of theory with ΔE under 5 kcal · mol^{-1}. *Cis* conformers denoted by hollow triangles; *trans* conformers by solid diamonds.

trans conformers [see Fig. 3(A)]. By the DFT method, the global minimum turned to be conformer $c6\_\beta_L\beta_L$. It is important to note that conformer $c6\_\beta_L\beta_L$ has no stabilizing hydrogen bond, while conformer $t6\_\beta_L\gamma_D$ is stabilized by a γ-turn (hydrogen bond between an amide hydrogen and the carbonyl oxygen of the previous amide group, causing a seven-membered pseudoring).

Those conformers that have relative energy, ΔE not more than 5 kcal · mol^{-1} over the relevant global minimum ($t6\_\beta_L\gamma_D$ or $c6\_\beta_L\beta_L$) at either RHF/3-21G* or B3LYP/6-31+G(d) level of theory, were considered low-energy minima [see Table II and Figs. 3(B) and 4] and were further investigated (see Discussion section). Considering the criterion $-90° \leq \tau \leq 90°$, each low-energy conformer can be considered a β-turn. Criterion $d \leq 7$ Å, however, is seldom fulfilled. Most low-energy minima can be labeled also by the traditional β-turn nomenclature (see Table II); thus, types I, VIa, VIb, and VIII are readily distinguished. Type II β-turn ($\epsilon_L\gamma_D$ or $\epsilon_L\alpha_D$) structure is approximated by conformers $t5\_\beta_L\gamma_D$, $t6\_\beta_L\alpha_D$, $t6\_\beta_L\gamma_D$, and $t8\_\beta_L\alpha_D$.

Experimental Data Analysis of Xxx—Cys—Cys—Yyy and Xxx—*ox*-[Cys—Cys]—Yyy Peptide Sequences

We found 120 Xxx—Cys—Cys—Yyy type sequence units in PDB SELECT[35,36] and subjected them to conformational analysis. Of these, only 3 contain a vicinal disulfide bridge. Because of the very low occurrence in PDB SELECT, Xxx—*ox*-[Cys—Cys]—Yyy sequence units were retrieved also from the Brookhaven PDB and analyzed below. Several characteristic torsional angles of the sequences from the two sets are visualized in Figure 5.

In 45% of the Xxx—Cys—Cys—Yyy type sequences from PDB SELECT, both cysteine backbones are extended-like [mainly β_L or ϵ_L; see Fig. 5(C, D)]. In one fourth of the

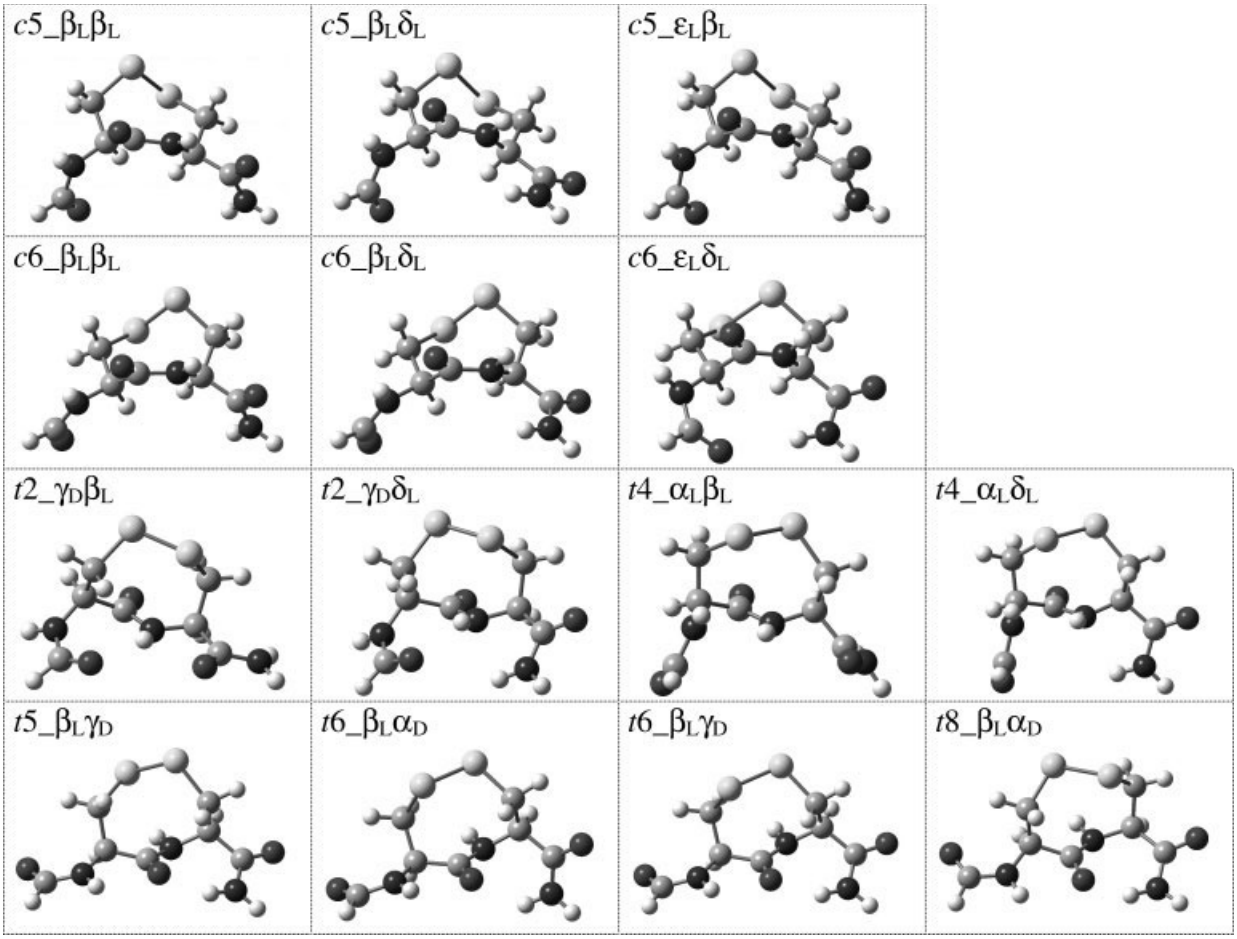


Fig. 4. Selected low-energy conformers of HCO—*ox*-[Cys—Cys]—NH_2 obtained at B3LYP/6-31+G(d) level of theory.

sequences, both cysteines adopt helix-like backbone conformation, whereas in 18%, all four residues are nearly helical; hence, a right-handed α-helix is formed. In 12%, an extended cysteine is followed by a helical one, in which case Xxx tends to be extended, while the conformation of Yyy varies. In another 12%, a helix-like cysteine is followed by an extended one, in which case neither Xxx nor Yyy show significant dependence except that rare backbone conformers (α_D, ϵ_D, γ_D, or δ_D) are not adopted. No *cis* amide bond was detected between the two cysteine residues.

As a vicinal disulfide bridge causes conformational constraints on torsional angles ψ_1 and ϕ_2, it is useful to plot these torsions against each other, although this does not result in a conventional Ramachandran map [see Fig. 5(A)]. Two populated areas can be defined: regions I and II. In region I sequences, the first cysteine residue is extended-like, while the second one is either extended-like or helix-like. It is very important to note that disulfide bridge–containing sequences are never found in region I. In region II sequences, the first cysteine is helix-like, while the second one is either helix-like (upper part of the region) or extended-like (lower part of the region, with several disulfide bridge–containing sequences). Region III is defined because it contains some sequences with a disulfide bridge, namely, those with a *t*6 or a *t*8 ring, hence, an $\epsilon_L\alpha_D$ backbone conformation of the two cysteine residues. Out of the 120 Xxx—Cys—Cys—Yyy type sequences, only two belong to region III, each with $\beta_L\alpha_D$ backbone conformation of the two cysteines.

Xxx—*ox*-[Cys—Cys]—Yyy sequence units were retrieved from the Brookhaven PDB.[37] The 48 hits were obtained from 31 different protein structures belonging to 11 families: *Alcaligenes* esterase 713; α-L-arabinanase; *Anguilla anguilla* agglutinin; acetylcholine-binding protein, with peptides corresponding to a segment of it (see below); quinoproteine alcohol dehydrogenases; palmytoil protein thioeseterases; alginate lyases; carboxipeptidase T; atracotoxin J-ACTX-Hv1c; human hepcidin; and a disulfide bond isomer of α-conotoxin. None of them had a *cis* amide bond connecting cysteine residues (see Table III). The numbers of sequence units with rings *t*2, *t*4, *t*6, and *t*8 are 2, 32, 11, and 2, respectively. One single-peptide unit could not be assigned to any of the ring conformers because it had torsion angle χ_{13} around 0°. Xxx tends to be an amino acid with a small (Gly, Ala) or even small hydrophilic side-chain (Ser, Thr). Yyy is mainly Asp, Asn, Arg or Pro.

A group of structures represents proteins and derived peptides from acetylcholine-binding proteins or receptors. In the acetylcholine-binding protein (AChBP), the ring

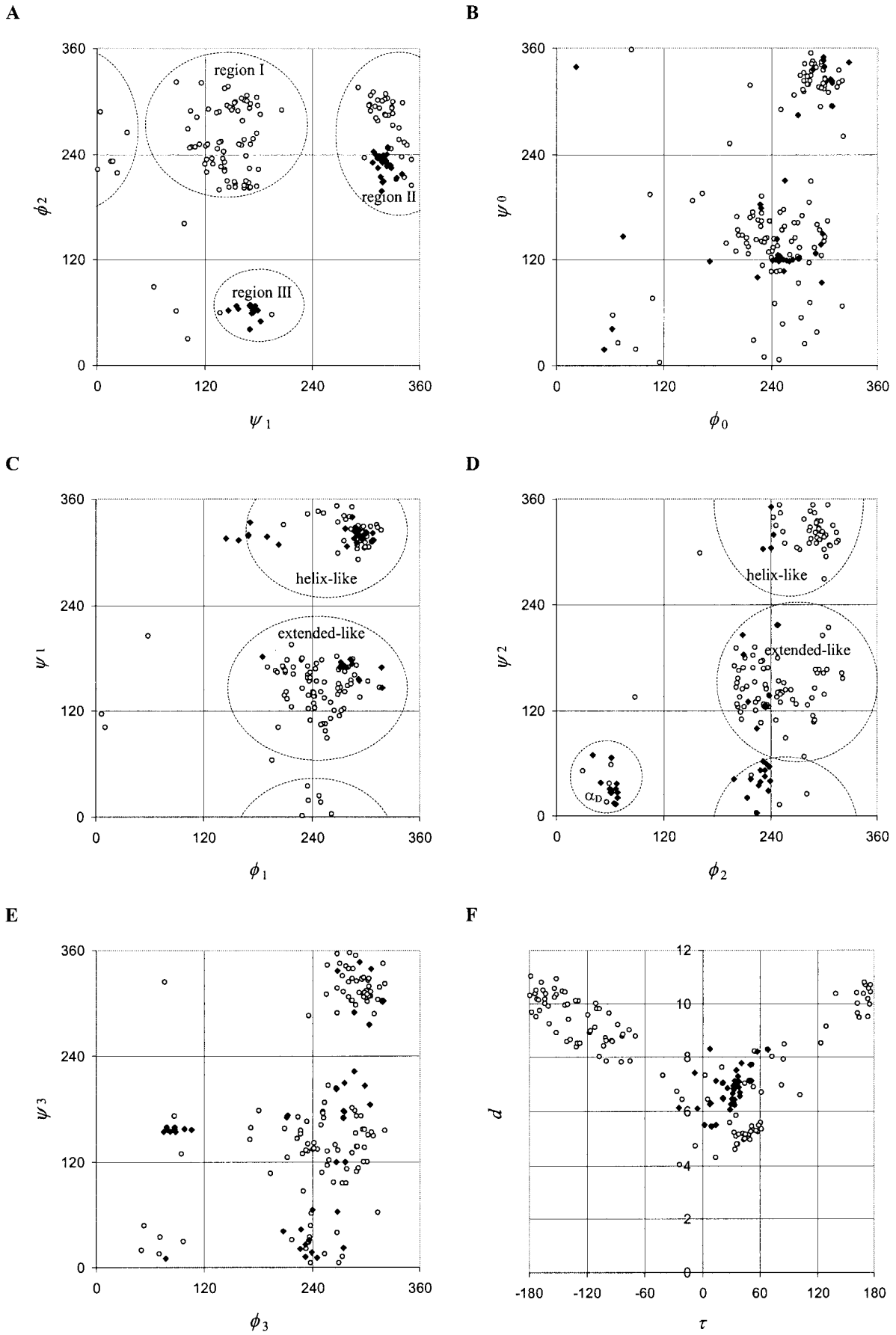


Fig. 5. Xxx—Cys—Cys—Yyy type sequence units found in PDB SELECT (hollow circle); Xxx—*ox*-[Cys—Cys]—Yyy sequence units found in PDB (solid diamond). (**A**) ψ_1-ϕ_2 map; (**B**) ϕ_0-ψ_0 map; (**C**) ϕ_1-ψ_1 map; (**D**) ϕ_2-ψ_2 map; (**E**) ϕ_3-ψ_3 map; (**F**) τ-*d* map.

TABLE III. Xxx—*ox*—[Cys—Cys]—Yyy Sequence Units Retreived From PDB

PDB ID	Chain	RNo	Xxx residue	bb	ox-[Cys—Cys]	Yyy residue	bb	Method	Resolution [Å]	Type of β-turn	Description
1q1w	A, B	70	Gly	ε_L	*t*6_$\varepsilon_L\alpha_D$	Leu	β_L	diffraction	1.1	II	*Alcaligenes* esterase 713
1gyd	B	240	Leu	γ_L	*t*6_$\varepsilon_L\alpha_D$	Arg	δ_L	diffraction	2.05	II	α-L-arabinase
1gye	B	240	Leu	γ_L	*t*6_$\varepsilon_L\alpha_D$	Arg	δ_L	diffraction	2.5	II	
1gyh	A, B, C, D, F	240	Leu	ε_L	*t*6_$\varepsilon_L\alpha_D$	Arg	γ_L	diffraction	1.89	II	
1gyh	E	240	Leu	γ_L	*t*6_$\varepsilon_L\alpha_D$	Arg	δ_L	diffraction	1.89	II	
1kl2	A	81	Asp	α_L	*t*6_$\beta_L\alpha_D$	Gly	α_L	diffraction	1.9	ud.	*Anguilla anguilla* agglutinin
1idh	B	191	Thr	α_L	*t*8_$\varepsilon_L\alpha_D$	Pro	α_L	NMR		II	Peptides derived from α-Subunits of nicotinic acetylcholine receptors (nAChRs) complexed to toxins (α-bungarotoxin, cobratoxin)
1idg	B	191	Thr	α_L	*t*8_$\varepsilon_L\alpha_L$	Pro	ε_L	NMR		II	
1l4w	B	191	Thr	ε_L	*t*2_$\delta_D\delta_L$	Pro	α_L	NMR		ud.	
1ljz	B	191	Thr	ε_L	*t*2_$\delta_D\delta_L$	Pro	α_L	NMR		ud.	
1kc4	B	188	Glu	α_L	*t*4_$\delta_D\delta_L$	Lys	γ_L	NMR		ud.	
1kl8	B	188	Glu	α_L	*t*4_$\delta_D\delta_L$	Lys	γ_L	NMR		ud.	
1lxg	B	191	Thr	γ_L	*t*4_$\delta_D\beta_L$	Pro	ε_L	NMR		ud.	
1lxh	B	191	Thr	ε_D	*t*4_$\delta_D\beta_L$	Pro	ε_L	NMR		ud.	
1i9b*	A, B, C, D, E	186	Ser	α_L	*t*4_$\alpha_L\beta_L$	Pro	α_L	diffraction	2.7	VIII	Acetylcholine-binding protein
1h4i	A, C	102	Ala	ε_L	*t*4_$\alpha_L\beta_L$	Asp	ε_L	diffraction	1.9	VIII	Alcohol/methanol dehydrogenases (PQQ-containing quinoproteins or quinohemoproteins)
1h4j	A, C, E, D	102	Ala	γ_L	*t*4_$\alpha_L\delta_L$	Asp	ε_D	diffraction	3.0	I	
1g72	A, C	102	Met	γ_L	*t*4_$\alpha_L\delta_L$	Asp	ε_D	diffraction	1.9	I	
4aah	A, C	102	Met	γ_L	*t*4_$\alpha_L\delta_L$	Asp	ε_D	diffraction	2.4	I	
1kb0	A	115	Gly	ε_L	*t*4_$\alpha_L\delta_L$	Asp	ε_D	diffraction	1.4	I	
1kv9	A	104	Ser	δ_L	*t*4_$\alpha_L\delta_L$	Asp	ε_D	diffraction	1.9	I	
1flg	A, B	104	Pro	ε_L	*t*4_$\alpha_L\delta_L$	Asp	ε_D	diffraction	2.6	I	
1ei9	A	44	Ser	β_L	*t*4_$\alpha_L\delta_L$	Asn	ε_L	diffraction	2.3	I	Palmytoil protein thiesterases
1eh5	A	44	Ser	β_L	*t*4_$\alpha_L\delta_L$	Asn	ε_L	diffraction	2.5	I	
1exw	A	44	Ser	β_L	undefined	Asn	ε_L	diffraction	2.4	I	
1hv6*	A	187	Ser	α_L	*t*4_$\alpha_L\delta_L$	Asn	ε_L	diffraction	2.0	I	Alginate lyases
1qaz	A	187	Ser	α_L	*t*4_$\alpha_L\delta_D$	Asn	ε_L	diffraction	1.8	ud.	
1obr	a	154	Gly	α_D	*t*4_$\alpha_L\delta_L$	Gly	α_D	diffraction	2.3	I	Carboxypeptidase T
1d10*	A	12	Ala	α_L	*t*4_$\alpha_L\beta_L$	Pro	α_L	NMR		VIII	Atracotoxin
1m4e	A	7	Gly	α_D	*t*4_$\alpha_L\delta_D$	His	ε_L	NMR		ud.	Human hepcidin
1m4f	A	12	Gly	γ_L	*t*4_$\alpha_L\delta_D$	His	α_L	NMR		ud.	
1xgc	a	1	Glu	γ_D	*t*4_$\delta_D\delta_D$	Asn	δ_L	NMR		ud.	α-Conotoxin disulfide bond isomer

PDB ID, PDB reference code; *, present in PDB Select; RNo, Xxx residue number; bb, backbone conformational code [see Scheme 2(A)]; type of β-turn according to Scheme 3.

adopts *t*4 conformation (PDB code: 1i9b).[39] We also found four structures of complexes between peptides derived from the α-subunits of nicotinic acetylcholine receptors (nAChRs) and toxins: α-bungarotoxin (Bgtx) and α-cobratoxin (Cbtx) (PDB codes: 1idh, 1idg,[40] 1lxg and 1lxh,[41] 1kc4 and 1kl8,[42] 1l4w, and 1ljz[43], respectively). The observed ring conformers of the *ox*-[Cys—Cys] motif in the peptides are different from each other (Table III); thus, a conformational change in the vicinal disulfide upon toxin binding seems plausible, although it cannot be unambiguously characterized. We note that two ring conformers, *t*8 and *t*2, were found only in these complexes.

In *Alcaligenes* esterase 713, an α/β hydrolase, the *ox*-[Cys—Cys] ring is in the *t*6 conformation (1qlw).[44] The backbone nitrogen of Cys71 is involved in oxyanion hole formation, and the vicinal disulfide is believed to contribute to the structural stability of the corresponding β-turn.

Quinoproteine alcohol dehydrogenases (1flg,[45] 1h4i,[46] 1h4j,[2] 1g72,[47] 4aah,[48] 1kb0,[49] 1kv9[50]) contain the prosthetic group pyrroloquinoline quinone (PQQ). The conserved *ox*-[Cys—Cys] units are spatially close to PQQ, and despite not being involved directly in catalysis, replacement of any of the neighboring cysteines results in loss of enzyme activity, as judged by site-directed mutagenesis investigations on the methanol dehydrogenase of the methylotrophic bacterium *Methylobacterium extorquens*.[2,51] In all these sequences, Yyy is Asp and, with the exception of 1h4i, it accepts ϵ_D backbone conformation, which is extremely rare in the case of Cys-Cys units without disulfide bridge [see Fig. 5(E)].

In atracotoxin J-ACTX-Hv1c, the vicinal disulfide was shown to be essential for insecticidal activity, although the structure of the double mutant (Cys13Ser, Cys14Ser) appears to be essentially the same as that of the wild-type protein (1dl0).[4]

In the case of alginate lyases (1hv6,[52] 1qaz[53]) the *ox*-[Cys—Cys] unit is located in a solvent-exposed part of loop2 (L2). The disulfide is close to several conserved active site residues; thus, its role may be the maintainance of the active site conformation.

The Cys155–Cys156 bridge found in carboxypeptidase T (1obr[54]) is uncommon among carboxypeptidases. Its assumed role is to increase the rigidity of the corresponding surface loop to make it more stable against proteolysis.

In *A. anguilla* agglutinin (1k12[55]), the vicinal disulfide is solvent-exposed and is speculated to play a role in making nonspecific intermolecular contacts. It is also supposed to contribute to the conformation and rigidity of one of the loops responsible for ligand (fucose) binding.

The two predominant forms of human hepcidin, hepcidin-20 and hepcidin-25, have a disulfide bridge connecting adjacent cysteine residues in a loop between two β-strands.[56] Interestingly, the vicinal disulfide is involved in exchange process on the time scale of NMR measurements and represents the most flexible segment of the molecule.

The only non-native *ox*-[Cys—Cys] motif found in the PDB corresponds to an artificial disulfide bridge isomer of the α-conotoxin GI (1xgc[57]). Linking Cys2 to Cys3 leads to a family of at least two major and three minor conformers in contrast to the well-defined single structure of the native molecule. The only conformer of this isomer for which structure could be calculated diverges significantly from that of the wild-type toxin, as judged by root-mean-square deviation (RMSD) values of 4.2 Å for backbone and 7 Å for sulfur atoms. Interestingly, this isomer is only 10 times less potent biologically than the native form, while the other possible non-native isomers were found to be completely inactive.

In the case of palmytoil protein thioesterases (1ei9,[58] 1eh5,[58] 1exw[59]) and α-L-arabinase (PDB codes: 1gyd, 1gye, and 1gyh[60]), we found no indications for any structural or biological role of the vicinal disulfide bridge in the literature.

All 48 Xxx—*ox*-[Cys—Cys]—Yyy sequences retrieved from the Brookhaven PDB[37] fulfill the criterion of $-90° \leq \tau \leq 90°$ of the β-turns (parameter τ varies between −25° and 68°, with a weighted average of 26°). Distance *d* is within the interval of 5.4–8.3 Å and has a weighted average of 6.8 Å. All entries with *t*6 or *t*8 ring conformer, except for 1k12, have $\epsilon_L\alpha_D$ backbone conformation; thus, they compose type II β-turns (see Table III). The proteins with *t*2 ring conformer fold into $\delta_D\delta_L$ backbone conformation; therefore, they are not classical β-turns according to Scheme 3. Entries with the *t*4 ring, however, show greater variation in their backbone fold. Three entries in Table III (*t*4_$\alpha_L\beta_L$) belong to type VIII β-turns, 11 sequences ($\alpha_L\delta_L$, one of them with an undefined ring conformer) represent type I β-turn, while several backbone folds ($\delta_D\delta_L$, $\delta_D\beta_L$, $\alpha_L\delta_D$, $\delta_D\delta_D$) cannot be assigned to any of the traditional β-turns.

DISCUSSION

The significance of the *ox*-[Cys—Cys] sequence unit lies in the biological activity of the several proteins that contain a vicinal disulfide bridge. From the conformational point of view, such proteins were investigated by X-ray crystallography and NMR spectroscopy. In order to gain more detailed knowledge, the intriguing part of protein structure was modeled by various shorter sequences and examined either experimentally or computationally. This study aims to prove that the quantum chemical exploration of the conformational space of the shortest possible model peptide (i.e., that of HCO—*ox*-[Cys—Cys]—NH_2) provides an expedient grasp on the multifarious conformation of the *ox*-[Cys—Cys] sequence unit. This approach, however, cannot account for the experimental finding that long protein chains prefer different conformers than short peptides, a problem that biases abundance prediction but does not distort geometrical data.

The statistical analysis of vicinal disulfide bridges in proteins did not deliver any ring conformer with a *cis* bond. Experimental investigation of short, disulfide bridge–containing peptides, however, reported *cis* ring conformers in agreement with the ab initio and DFT calculations presented here. Boc—*ox*-[Cys—Cys]—OMe, determined by X-ray crystallography,[61] comprises a *c*6 ring. The fact that *cis* amide bonds are not found in as high abundance as

TABLE IV. Selected Geometric Parameters of Ring Conformers in Xxx—*ox*-[Cys—Cys]—Yyy Sequence Units From Experimentally Determined Protein Structures

Ring	n	ψ_1	ω_1	ϕ_2	χ_{11}	χ_{13}	χ_{21}	d_{S-S}
*t*2[a]	2	−45	−172	−141	135	106	11	2.05
*t*4	21	−40	−168	−131	72	−105	−41	2.05
*t*6[b]	6	172	161	62	82	−80	−35	2.09
*t*8[a]	2	158	169	52	157	109	16	2.05

n, number of different sequence units according to the entries of Table III; all torsional angles in degrees; d_{S-S}, distance of the two sulfur atoms in Å.
[a]NMR-determined sequence units.
[b]X-ray–determined sequence units (resolution: 1.1–2.0 Å).

predictable from either calculation or experimental investigation of shorter model peptides is not unique for the *ox*-[Cys—Cys] sequence unit. Indeed, we note, that independent of the amino acid sequence, *cis* amide bonds are underrepresented in experimentally determined protein structures when compared to abundances predicted from free enthalpy values.[22,62] The fact that no *cis* amide bonds were detected within as few as 48 observed cases of the *ox*-[Cys—Cys] unit, representing only 11 nonhomologous sets of natural proteins, may be the consequence of the uneven sampling of natural protein folds in the PDB, while another reason may be the low occurrence of *cis* protein amide bonds, which is independent of the sequence.

The CH_3CO—*ox*-[Cys—Cys]—NH_2 model peptide was studied in aqueous solution by NMR spectroscopic methods.[10] Two *trans* amide conformers, together with two *cis* amide conformers, were determined in molar ratios 47:15:29:9. Ring torsional angles ω_1, ϕ_2, χ_{11}, and χ_{21} were determined experimentally, while ψ_1, together with χ_{12} and χ_{22} for cases of positive and negative χ_{13} values, were computed on the basis of a Monte Carlo conformational search. The ring conformers presented in that study (T−, T′−, C+, and C− with the original notation) can be unambiguously assigned to ring conformers *t*4, *t*6, *c*6, and *c*5. Other low-energy conformers found by the Monte Carlo conformational search, *t*2 and *t*8 (T+ and T′+) were not supported by NMR results. All six ring conformers delivered in the study on the CH_3CO—*ox*-[Cys—Cys]—NH_2 model peptide were predicted as low-energy ring conformers in ab initio and DFT investigations.

All four ring conformers found in proteins (i.e., *t*2, *t*4, *t*6, and *t*8) are energetically favored conformers of both 4,5-ditiaheptano-7-lactam and HCO—*ox*-[Cys—Cys]—NH_2. Data corresponding to the most reliable entries in Table III were averaged in order to assess experimentally determined torsional angles of these ring conformers (see Table IV). The comparison of experimental and computational data allows unambiguous identification of ring conformers. When comparing torsional angles of corresponding conformers found by both approaches, however, considerable numerical differences, around 10–20°, occur. Experimentally, parameter ω_1 deviates from the ideal value of a *trans* amide bond (180°) less, by only 11–21°, while computationally, this deviation tends to exceed 30°. The best correlation is in parameter χ_{13}, where the two methods differ only by a few degrees.

The quantum chemical investigation of the vicinal disulfide bridge using a short model peptide is promising, because the structure is inherently constrained to fold into a β-turn or to a similar backbone, while it cannot fit into a helix or a β-pleated sheet. Several of the computed conformers can be readily identified among protein sequence units (*t*4_$\alpha_L\delta_L$, *t*4_$\alpha_L\beta_L$, *t*6_$\beta_L\alpha_D$). Some other experimentally found conformers have calculated counterparts that differ mainly only in ϕ_1 (*t*2_$\delta_D\delta_L \rightarrow$ *t*2_$\gamma_D\delta_L$, *t*6_$\epsilon_L\alpha_D \rightarrow$ *t*6_$\beta_L\alpha_D$, *t*8_$\epsilon_L\alpha_D \rightarrow$ *t*8_$\beta_L\alpha_D$).

Apart from the comparison of computational and experimental data, another pertinent question is to ask whether the results obtained at relatively low levels [i.e., RHF/3-21G* and B3LYP/6-31+G(d)] could be consistent with more expensive calculations.

Frequency calculations on 14 selected low-level minima (see Table II) performed at B3LYP/6-31+G(d) level of theory proved that all these conformers are indeed minima. In only one case (*t*8_$\beta_L\alpha_D$), an imaginary frequency was detected and attributed to the flatness of the potential energy surface, because subsequent optimization with a "tight" convergence criterion resolved the ambiguity of the critical point. Furthermore, thermochemical parameters were gained from the frequency calculations and correlated to relative electronic energies obtainable from simple optimizations (see Table V). Most of the new energetic parameters (e.g., energy corrected with zero-point correction, ΔE_0; energy with thermal correction, ΔE_{therm}; enthalpy, ΔH) cause almost negligible changes in the relative order of conformers, and also the Gibbs free energy, ΔG, can be scaled to the simple electronic energy (see linear regression parameters in Table V).

In order to investigate the behavior of the model peptide in aqueous solution, PCM single-point energies were calculated at B3LYP/6-31+G(d) level of theory. The relative order of the conformers is changed by the solvent effect, the global minimum becomes *c*6_$\beta_L\delta_L$, and the correlation with gas phase relative energies is poor (see Table V). The newly gained data, however, do not match experimental data better than gas phase optimization results.

Conformers were further optimized at B3LYP/6-311++G(d,p) level of theory. Geometric as well as stability data (see Table V) correlate well with the B3LYP/6-31+G(d) results. Conformers *c*5_$\epsilon_L\beta_L$ and *t*4_$\alpha_L\beta_L$ migrated to *c*5_$\beta_L\beta_L$ and *t*4_$\alpha_L\delta_L$, respectively.

It is noted that changing the method from RHF to B3LYP, and the basis set from 3-21G to 6-31+G(d), and further to 6-311++G(d,p) is expected to keep the overall topology of the potential energy hypersurface (PEHS) and to annihilate a few conformers which converge to others.[63] Significant alteration of the PEHS may arise, however, when solvent effects are also considered (e.g., by the application of PCM method).[64] No scientific evidence published so far indicates that preselection of conformers in gas phase would not overlook some low energy minima of the PCM PEHS.

TABLE V. Selected Energetic Parameters of Low-Energy HCO—*ox*—[Cys—Cys]—NH_2 Conformers Optimized at B3LYP/6-31+G(d) Level of Theory and Their Correlation

Conformer	ΔE^a	$\Delta E_0{}^b$	$\Delta E_{therm}{}^b$	ΔH^b	ΔG^b	$\Delta E_{PCM}{}^c$	ΔE^d	$\Delta E_{PCM}{}^e$
$c5\_\beta_L\beta_L$	0.98	1.01	1.03	1.03	0.41	4.83	0.96	0.87
$c5\_\beta_L\delta_L$	2.12	2.16	2.13	2.13	1.83	3.66	2.01	1.59
$c5\_\varepsilon_L\beta_L$	0.98	1.05	1.07	1.07	0.61	4.27	not found	not found
$c6\_\beta_L\beta_L$	0.00	0.00	0.00	0.00	0.00	0.81	0.00	0.00
$c6\_\beta_L\delta_L$	0.38	0.48	0.37	0.38	0.70	0.00	0.36	0.23
$c6\_\varepsilon_L\delta_L$	1.84	2.04	2.00	2.00	1.58	6.65	1.76	not found
$t2\_\gamma_D\beta_L$	7.00	6.88	6.81	6.81	7.55	5.94	6.59	6.35
$t2\_\gamma_D\delta_L$	3.97	4.27	4.11	4.11	5.00	7.15	3.58	6.99
$t4\_\alpha_L\beta_L$	3.54	3.55	3.50	3.50	3.38	2.84	not found	1.43
$t4\_\alpha_L\delta_L$	2.02	2.02	1.90	1.90	2.39	2.60	1.92	1.27
$t5\_\beta_L\gamma_D$	4.10	4.62	4.38	4.38	5.16	8.42	4.00	7.26
$t6\_\beta_L\alpha_D$	2.20	2.32	2.11	2.11	2.80	1.48	1.87	2.21
$t6\_\beta_L\gamma_D$	0.86	1.20	0.91	0.91	1.91	2.77	0.66	3.33
$t8\_\beta_L\alpha_D$	4.23	4.78	4.45	4.45	5.24	6.47	3.97	6.63
Parameters of linear regressions								
m (ΔE^a)	1.00	1.01	1.00	1.00	1.13	0.84	0.94	1.10
b (ΔE^a)	0.00	0.12	0.05	0.05	−0.02	2.09	−0.03	0.29
R^2 (ΔE^a)	1.000	0.990	0.995	0.995	0.945	0.401	0.997	0.630
m (ΔG^b)	0.83	0.86	0.83	0.83	1.00	0.70	0.79	1.08
b (ΔG^b)	0.15	0.23	0.19	0.19	0.00	2.22	0.03	−0.10
R^2 (ΔG^b)	0.945	0.967	0.949	0.949	1.000	0.377	0.940	0.783

All relative energetic parameters are given in kcal · mol^{-1}. ΔE, electronic energy, directly gained from optimization; ΔE_0, energy corrected with zero-point correction; ΔE_{therm}; energy with thermal correction; ΔH, enthalpy; ΔG, Gibbs free energy; ΔE_{PCM}, energy with solvent effect correction.

[a]From optimization at B3LYP/6-31+G(d) (see Table II)

[b]From frequency calculation at B3LYP/6-31+G(d)

[c]From PCM single point energy calculation for aqueous solution at B3LYP/6-31+G(d)

[d]From further optimization of the conformer at B3LYP/6-311++G(d, p)

[e]From PCM optimization for aqueous solution at B3LYP/6-31+G(d)

m and b, parameters of linear regression equation $y = mx + b$ fitted to the energetic data of the same column, as y, and column ΔE^a or ΔG^b, as x.

R^2, linear correlation of the energetic data of the same column to those in column ΔE^a or ΔG^b, respectively.

Optimization of the 14 selected conformers at B3LYP/6-31+G(d) level of theory was performed also in water by applying PCM method. Conformers $c5\_\epsilon_L\beta_L$ and $c6\_\epsilon_L\delta_L$ migrated to $c5\_\beta_L\beta_L$ and $c6\_\beta_L\delta_L$, respectively. In the other 12 geometries, only minor structural changes occurred upon PCM optimization: on average, 2° at ring torsional angles and 7° at dihedrals ψ_1 and ϕ_2, with maximum alterations of 7° and 22°, respectively. Interesting to note is that parameter d, descriptive of the β-turn character of the conformers (see Scheme 4) never changed more than 0.35°. Relative energies (ΔE) correlate somewhat better to gas phase relative energies than single point PCM results (see Table V). The characteristic parameters of the linear correlation between PCM-optimized and PCM single point–calculated ΔE values are $m = 0.93$, $b = -0.47$, and $R^2 = 0.771$.

CONCLUSIONS

Exhaustive conformational search at RHF/3-21G* followed by subsequent optimization at B3LYP/6-31+G(d) delivered 14 low-energy conformers of the HCO—*ox*-[Cys—Cys]—NH_2 model peptide. Experimental data on the *ox*-[Cys—Cys] sequence unit from databases and literature were successfully related to the above computed conformers. It is found that the *ox*-[Cys—Cys] unit is able to adopt types I, II, VIa, VIb, and VIII β-turn structures, while without the vicinal disulfide bridge, the Cys—Cys sequence unit is not constrained to form a β-turn. The vicinal disulfide, despite its rare occurrence in proteins, is an important structural/functional motif in almost each case it was identified. We believe that vicinal disulfide bridges may stabilize otherwise high-energy β-turn structures, such as type VIII, as well as *cis* amide-containing types VIa and VIb. Several occurrences of the *ox*-[Cys—Cys] unit in proteins were found to be folded into type VIII β-turn, while none has been identified as type VIa or VIb β-turn.

REFERENCES

1. Paterson D, Nordberg A. Neuronal nicotinic receptors in the human brain. Prog Neurobiol 2000;61:75–111.
2. Afolabi PR, Mohammed F, Amaratunga K, Majekodunmi O, Dales SL, Gill R, Thompson D, Cooper JB, Wood SP, Goodwin PM, Anthony C. Site-directed mutagenesis and X-ray crystallography of the PPQ-containing quinoprotein methanol dehydrogenase and its electron acceptor, cytochrome c_L. Biochemistry 2001;40:9799–9809.
3. Avezoux A, Goodwin MG, Anthony C. The role of the novel

disulphide ring in the active site of the quinoprotein methanol dehydrogenase from *Methylobacterium extorquens*. Biochem J 1995;307:735–741.
4. Wang X, Connor M, Smith R, Maciejewski MW, Howden MEH, Nicholson GM, Christie MJ, King GF. Discovery and characterization of a family of insecticidal neurotoxins with a rare vicinal disulfide bridge. Nat Struct Biol 2000;7:505–513.
5. Csonka GI, Schubert GA, Perczel A, Sosa CP, Csizmadia IG. Ab initio conformational space study of model compounds of O-glycosides of serine diamide. Chem-Eur J 2002;8:4718–4733.
6. Perczel A, Ángyán JG, Kajtár M, Viviani W, Rivail JL, Marcoccia JF, Csizmadia IG. Peptide models 1: topology of selected peptide conformational potential-energy surfaces (glycine and alanine derivatives). J Am Chem Soc 1991;113:6256–6265.
7. Scarsdale JN, Van Alsenoy C, Klimkowski VJ, Schäfer L, Momany FA. Ab initio studies of molecular geometries 27: optimized molecular structures and conformational analysis of N^{α}-acetil-N-methylalaninamide and comparison with peptide crystal data and empirical calculations. J Am Chem Soc 1983;105:3438–3445.
8. Topol IA, Burt SK, Deretey E, Tang TH, Perczel A, Rashin A, Csizmadia IG. Alpha- and 3(10)-helix interconversion: a quantum-chemical study on polyalanine systems in the gas phase and in aqueous solvent. J Am Chem Soc 2001;123:6054–6060.
9. Perczel A, Jákli I, Csizmadia IG. Peptide models XXXV: relative stability of major types of β-turns as function of amino acid composition. Chem-Eur J 2003;9:2551-2566.
10. Creighton CJ, Reynolds CH, Lee DHS, Leo GC, Reitz AB. Conformational analysis of the eight-membered ring of the oxidized cysteinyl-cysteine unit implicated in nicotinic acetylcholine receptor ligand recognition. J Am Chem Soc 2001;123:12664–12669.
11. Wilmot CM, Thornton JM. β-turns and their distortions: a proposed new nomenclature. Protein Eng 1990;3:479–493.
12. Ferrara P, Caflisch A. Folding simulations of a three-stranded antiparallel beta-sheet peptide. Proc Natl Acad Sci USA 2000;97: 10780–10785.
13. Ashish, Grover A, Kishore R. Characterization of a novel type VII beta-turn conformation for a bio-active tetrapeptide rigin—a synergy between theoretical and experimental results. Eur J Biochem 2000;267:1455–1463.
14. Gunasekaran K, Gomathi L, Ramakrishnan C, Chandrasekhar J, Balaram P. Conformational interconversions in peptide beta-turns: analysis of turns in proteins and computational estimates of barriers. J Mol Biol 1998;284:1505–1516.
15. Lewis PN, Momany FA, Scheraga HA. Chain reversals in proteins. Biochim Biophys Acta 1973;303:211–229.
16. Chou PY, Fasman GD. Beta-turns in proteins. Biochemistry 1977;115:135–175.
17. Smith JA, Pease LG. Reverse turns in peptides and proteins. CRC Crit Rev Biochem 1980;8:315–399.
18. Perczel A, Jakli I, Foxman BM, Fasman GD. A search for the ideal type I beta-turn. Biopolymers 1996;38:723–732.
19. Perczel A, Mcallister MA, Császár P, Csizmadia IG. Peptide models 6. New beta-turn conformations from ab initio calculations confirmed by X-ray data of proteins. J Am Chem Soc 1993;115:4849-4858.
20. Venkatachalam CM. Stereochemical criteria for polypeptides and proteins. V. Conformation of a system of three linked peptide units. Biopolymers, 1968;6:1425–1436.
21. Lee HJ. Song JW, Choi YS, Park HM, Lee KB. A theoretical study of conformational properties of N-methyl azapeptide derivatives. J Am Chem Soc 2002;124:11881–11893.
22. Weiss MS, Jabs A, Hilgenfeld R. Peptides bonds revisited. Nat Struct Biol 1998;5:676.
23. Hennig M, Jansonius JN, Terwisscha van Scheltinga AC, Dijkstra BW, Schlesier B. Crystal structure of concanavalin B at 1.65 Å resolution: an inactivated chitinase from seeds of *Canavalia ensiformis*. J Mol Biol 1995;254:237–246.
24. Weiss MS, Metzner HJ, Hilgenfeld R. Two non-proline cis peptide bonds may be important for factor XIII function. FEBS Lett 1998;423:291–296.
25. Shima S, Warkentin E, Grabarse W, Sordel M, Wicke M, Thauer RK, Ermler U. Stucture of coenzyme F_{420} dependent methylenetetrahydromethanopterin reductase from two methanogenic archaea. J Mol Biol 2000;300:935–950.
26. Lewis PN, Momany FA, Scheraga HA. Chain reversal in proteins. Biochim Biophys Acta 1973;303:211–229.
27. Perczel A, McAlister MA, Császár P, Csizmadia IG. Peptide models 9: a complete conformational set of For-Ala-Ala-NH_2 from ab inito computations. Can J Chem 1994;72:2050–2070.
28. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, Montgomery JA, Stratmann RE, Burant JC, Dapprich S, Millam JM, Daniels AD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Gonzalez C, Challacombe M, Gill PMW, Johnson BG, Chen W, Wong MW, Andres JL, Head-Gordon M, Replogle ES, Pople JA. Gaussian 98 (Revision A.1). Pittsburgh, PA: Gaussian, Inc.; 1998.
29. Pulay P. Convergence acceleration of iterative sequences—the case of SCF iteration. Chem Phys Lett 1980;73:393–398.
30. Pulay P. Improved SCF convergence acceleration. J Comput Chem 1982;3:556–560.
31. Császár P, Pulay P. Geometry optimization by direct inversion in the iterative subspace. J Mol Struct (THEOCHEM) 1984;114:31–34.
32. Farkas Ö, Schlegel HB. Methods for optimizing large molecules: Part III. An improved algorithm for geometry optimization using direct inversion in the iterative subspace (GDIIS). Phys Chem Phys Lett 2002;4:11–15.
33. Amovilli C, Barone V, Cammi R, Cancés E, Cossi M, Mennucci B, Pomelli CS, Tomasi J. Recent advances in the description of solvent effects with the polarizable continuum model. Adv Quantum Chem 1999;32:227–261.
34. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA Jr, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, A Rabuck D, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA. Gaussian 03 (Revision B.01). Pittsburgh, PA: Gaussian, Inc.; 2003.
35. Hobohm U, Scharf M, Schneider R, Sander C. Selection of representative protein data sets. Protein Sci 1992;1:409–417.
36. Hobohm U, Sander C. Enlarged representative set of protein structures. Protein Sci 1994;3:522–524.
37. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucleic Acids Res 2000;28:235–242.
38. Ramek M,Yu CH, Schäer L. Ab initio conformational analysis of the model tripeptide N-formyl-L-alanyl-L-alanine amide. Can J Chem 1998;76:566–575.
39. Brejc K, van Dijk WJ, Klaassen RV, Schuurmans M, van der Oost J, Smit AB, Sixma TK. Crystal structure of an Ach-binding protein reveals the ligand-binding domain of nicotinic receptors. Nature 2001;411:269–276.
40. Zeng H, Moise L, Grant MA, Hawrot E. The solution structure of the complex formed between α-bungarotoxin and an 18-mer cognate peptide derived from the α1 subunit of the nicotinic acetylcholine receptor from the *Torpedo californica*. J Biol Chem 2001;276:22930–22940.
41. Zeng H, Hawrot E. NMR-based binding screen and structural analysis of the complex formed between α-cobratoxin and an 18-mer cognate peptide derived from the a1 subunit of the nicotinic acetylcholine receptor from *Torpedo californica*. J Biol Chem 2002;277:37439–37455.
42. Moise L, Piserchio A, Basus VJ, Hawrot E. NMR structural analysis of α-bungarotoxin and its complex with the principal α-neurotoxin-binding sequence on the α7 subunit of a neuronal nicotinic acetylcholine receptor. J Biol Chem 2002;277:12406–12417.
43. Samson AO, Scherf T, Eisenstein M, Chill JH, Anglister J. The

mechanism for acetylcholine receptor inhibition by α-neurotoxins and species-species resistance to α-bungarotoxin revealed by NMR. Neuron 2002;35:319–332.
44. Bourne PC, Isupov MN, Littlechild JA. The atomic-resolution structure of a novel bacterial esterase. Structure 2000;8:143–151.
45. Keitel T, Diehl A, Knaute T, Stezowski JJ, Höhne W, Görisch H. X-ray structure of the quinoprotein ethanol dehydrogenase from *Pseudomonas aeruginosa*: basis of substrate specificity. J Mol Biol 2000;297:961–974.
46. Ghosh M, Anthony C, Harlos K, Goodwin M, Blake C. The refined structure of the quinoprotein methanol dehydrogenase from *Methylobacterium extorquens* at 1.94 angstrom. Structure 1995;3:177–187.
47. Zheng YJ, Xia Z, Chen Z, Mathews FS, Bruice TC. Catalytic mechanism of quinoprotein methanol dehydrogenase: a theoretical and X-ray crystallographic investigation. Proc Natl Acad Sci USA 2001;98:432–434.
48. Xia Z, Dai W, Zhang Y, White SA, Boyd GD, Mathews FS. Determination of the gene sequence and the three-dimensional structure at 2.4 Å resolution of methanol dehydrogenase from Methylophilus W3A1. J Mol Biol 1996;259:480.
49. Oubrie A, Rozeboom HJ, Kalk KH, Huizinga EG, Dijkstra BW. Crystal structure of quinohemoprotein alcohol dehydrogenase from *Comamonas testosteroni*. J Biol Chem 2002;277:3727–3732.
50. Chen Z, Matsushita K, Yamashita T, Fujii T, Toyama H, Adachi O, Bellamy HD, Matthews FS. Structure at 1.9 Å resolution of a quinohemoprotein alcohol dehydrogenase from *Pseudomonas putida* HK5. Structure 2002;10:837–849.
51. Blake CCF, Ghosh M, Harlos K, Avezoux A, Anthony C. The active site of methanol dehydrogenase contains a disulphide bridge between adjacent cysteine residues. Nat Struct Biol 1994;1:102–103.
52. Yoon HJ, Hashimoto W, Miyake O, Murata K, Mikami B. Crystal structure of alginate lyase A1-III complexed with trisaccharide product at 2.0 Å resolution. J Mol Biol 2001;307:9–16.
53. Yoon HJ, Mikami B, Hashimoto W, Murata K. Crystal structure of alginate lyase A1-III from *Sphingomonas* species A1 at 1.78 Å resolution. J Mol Biol 1999;290:505–514.
54. Teplyakov A, Polyakov K, Obmolova G, Strokopytov B, Kuranova I, Osterman A, Grishin N, Smulevitch S, Zagnitko O, Galperina O, Matz M, Stepanov V. Crystal structure of carboxypeptidase T from *Thermoactinomecetes vulgaris*. Eur J Biochem 1992;208:281–288.
55. Bianchet MA, Odom EW, Vasta GR, Amzel M. A novel fucose recognition fold involved in innate immunity. Nat Struct Biol 2002;9:628–634.
56. Hunter HN, Fulton DB, Ganz T, Vogel HJ. The solution structure of human hepcidin, a peptide hormone with antimicrobial activity that is involved in iron uptake and hereditary hemochromatosis. J Biol Chem 2002;277:37597–37603.
57. Gehrmann J, Alewood PF, Craik DJ. Structure determination of the three disulfide bond isomers of α-conotoxin GI: a model for the role of disulfide bonds in structural stability. J Mol Biol 1998;278:401–415.
58. Bellizzi JJ, Widom J, Kemp C, Lu JY, Das AK, Hofmann DJ, Clardy J. The crystal structure of palmitoyl protein thioesterase 1 and the molecular basis of infantile neuronal ceroid lipofuscinosis. Proc Natl Acad Sci USA 2000;97:4573–4578.
59. Das AK, Bellizzi JJ, Tandel S, Biehl E, Clardy J, Hofmann SL. Structural basis for the insensitivity of a serine enzyme (palmitoyl-protein thioesterase) to phenylmethylsulfonyl fluoride. J Biol Chem 2000;275:23847–23851.
60. Nurizzo D, Turkenburg JP, Charnock SJ, Roberts SM, Dodson EJ, McKie VA, Taylor EJ, Gilbert HJ, Davies GJ. *Cellvibrio japonicus* α-L-arabinase 43A has a novel five-blade β-propeller fold. Nat Struct Biol 2002;9:665–668.
61. Hata Y, Matsuura Y, Tanaka N, Ashida T, Kakudo M. Tert-butyloxycarbonyl-L-cysteinyl-L-cysteine disulfide methyl-ester. Acta Crystallogr B 1977;33:3561–3564.
62. Jabs A, Weiss M, Hilgenfeld R. Non-proline cis peptide bonds in proteins. J Mol Biol 1999;286:291–304.
63. Perczel A, Farkas Ö, Jákli I, Topol IA, Csizmadia IG. Peptide models XXXIII. Extrapolation of low-level Hartree–Fock data of peptide conformation to large basis set SCF, MP2, DFT, and CCSD(T) results: the Ramachandran surface of alanine dipeptide computed at various levels of theory. J Comput Chem, 2003. Forthcoming.
64. Iwaoka M, Okada M, Tomoda S. Solvent effects on the ϕ-ψ potential surfaces of glycine and alanine dipeptides studied by PCM and I-PCM methods. J Mol Struct (THEOCHEM) 2002;586:111–124.