Theor Chem Acc (2007) 118:165–171
DOI 10.1007/s00214-007-0258-x

REGULAR ARTICLE

Intramolecular interactions and *cis* peptidic bonds

Pierre-François Loos · Xavier Assfeld · Jean-Louis Rivail

Received: 27 October 2006 / Accepted: 11 January 2007 / Published online: 15 February 2007

Abstract The *cis* Gly–Gly peptidic bond observed in dihydrofolate reductase, which takes place between a beta sheet and an alpha helix is chosen as an example to study the various factors which influence this conformation. The peptidic chain of 16 amino acids is studied by a QM/MM scheme (6-31+G* B3LYP—Amber 99). Both electronic and classical contributions to the total energy take part to the stabilization of the *cis* conformation. A special role is played by the alpha helix when it is bonded to the carbonyl group of the bond of interest.

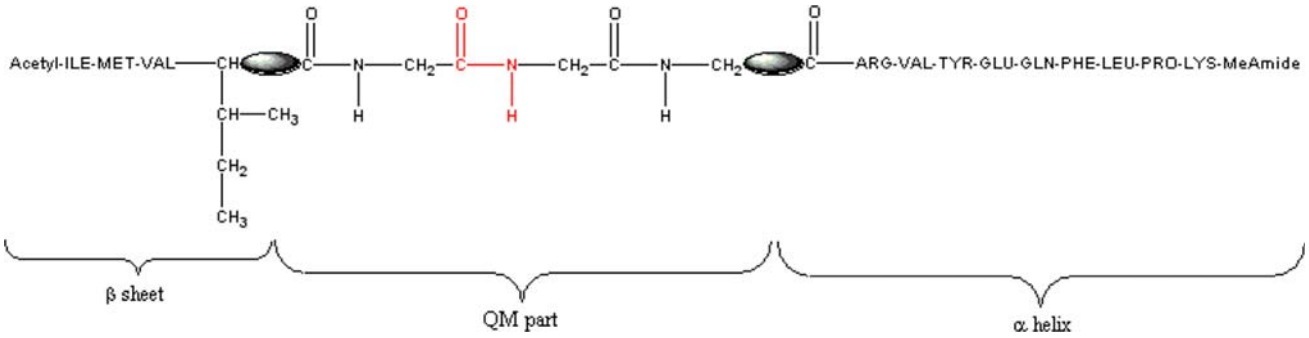
Keywords *Cis* peptidic bond · Alpha helix · Beta sheet · QM/MM

1 Introduction

In proteins, the great majority of the peptidic bonds adopt the *trans* conformation. However, it is not exceptional to find some bonds with the *cis* conformation. This often happens when one of the amino acids is proline and the reduced energy difference between the *cis* and the *trans* conformations has been related to the polarization of the alkyl chain attached to the nitrogen atom in this residue, which gives rise to a better induction energy in the *cis* conformation [1]. There are also some examples involving other amino acids (see Ref. 2(b) Protein Data Bank 1DRE.pdb) [2–4] and some attempts to explain these exceptions have been done with the help of quantum chemical studies [5–8]. An interesting example can be found in *Escherichia coli* dihydrofolate reductase in which a *cis* bond is observed between Gly95 and Gly96 [2]. It happens that this bond takes place between a beta sheet and an alpha helix and one may think that in this special instance the presence of these secondary structures plays an important role in the stabilization of the unusual *cis* conformation. In order to better understand the factors which may contribute to this kind of uncommon feature, we decided to analyse the influence of these neighbouring structures on the relative stability of the possible conformations. We focused our attention on the sequence from Ile 91 to Lys 106 which contains all the residues entering the beta sheet and the alpha helix. This 16 residue sequence is rather large for a full quantum chemical treatment at a reasonable level. Therefore, we chose a QM/MM approach in which the quantum part contains the two glycine residues involved in the *cis* bond and part of the amino acids bonded to them, in order to link the QM part to the MM ones by a C–C bond far enough from the bond of interest (Scheme 1). Our study involves comparison of the energies of a series of systems exhibiting the *cis* and the *trans* conformation of the Gly–Gly peptidic bond. We did not consider the possible transition states between these conformations because it is established that no equilibria take place in the protein. The information expected in this study does not pretend to give a definite answer to the complex problem of protein synthesis, but we hope to emphasize some factors which may play a role in this vital process. The systems that we considered are the whole peptide defined above and the structures which derive from this peptide by removing either the beta sheet, or the alpha helix or both and by exchanging the position of these sequences.

Contribution to the Fernando Bernardi Memorial Issue.

P.-F. Loos · X. Assfeld · J.-L. Rivail (✉)
UMR 7565, Structure et Réactivité des Systèmes Moléculaires Complexes, CNRS-Universté Henri Poincaré, Nancy-Université, 54506, Nancy-Vandoeuvre Cedex, France
e-mail: Jean-Louis.Rivail@cbt.uhp-nancy.fr



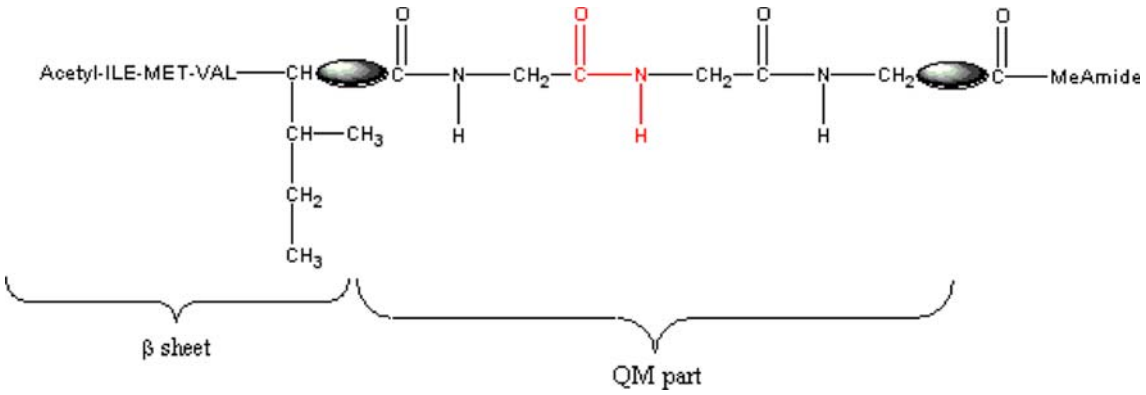
Scheme 1 *Beta-B-Alpha*

2 Computational methodology

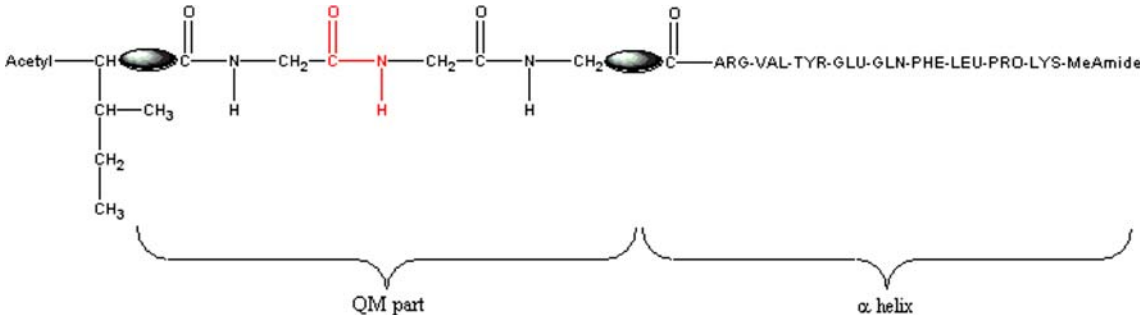
The QM/MM computations are performed by means of the LSCF method that we developed earlier [9–11] and which is implemented in our local Gaussian03 suite of programs [12]. In this method, the link between the QM and the MM parts is achieved by means of a strictly localized bond orbital (SLBO). In order to avoid some kind of spurious polarization of this bond, we usually manage to choose a carbon–carbon single bond. In the present case, the SLBOs are located between the α carbon and the carbonyl of an amino acid. This SLBO is derived from a computation on a model molecule containing the kind of bond of interest and with the same method and basis set as those used in study. In the present case, the computations are performed with the DFT formalism using a 6-31+G* basis set and the B3LYP functional [13,14]. The classical fragments are described with the help of the Amber 99 force field [15,16] in which the charges of the methyl groups added at both ends of the peptide have been slightly modified to ensure a neutral charge to these fragments. The van der Waals parameters for the atoms of the quantum fragment are given with the values defined in the force field for the classical atoms. The peptides considered in this study are made of a sequence of residues terminated by an acetyl group at the N terminus and an *N* methyl amide group at the C terminus. In all the systems that we considered in this study the quantum fragment is made of the carbonyl group of Ile 94, Gly 95, Gly 96 and the $NH–CH_2$ moiety of Gly 97 as indicated on Schemes 1, 2, 3, 4, 5, 6, 7. This fragment containing the peptidic bond of interest will be denoted by *B*. We shall denote the peptide derived from the dihydrofolate reductase structure, including the 16 residues from Ile 91 to Lys 106 as *Beta-B-Alpha*. We also considered the *Beta-B* (Scheme 2), the *B-Alpha* (Scheme 3), the *Alpha-B-Beta* built by exchanging the chains bonded to the central tetrapeptide (Scheme 4) the *Alpha-B* (Scheme 5) the *B-Beta* (Scheme 6) and finally the *Tetrapeptide* made of the sequence Ile-Gly-Gly-Gly, terminated as indicated above, in which the quantum fragment is the same as in the other structures (Scheme 7). According to the experimental structural determination performed at pH = 6 [2], we assumed that Arg 98 and Lys 116 are protonated, Glu 101 is deprotonated and Tyr 100 is neutral. These assumptions have been checked against the empirical prediction of pKa [17]. In order to avoid collapsing of the secondary structure of the beta sheet and of the alpha helix, the backbones in the classical fragments are fixed in the crystallographic structure. The side chains and the quantum fragment are fully optimized in all our computations.

3 Results and discussion

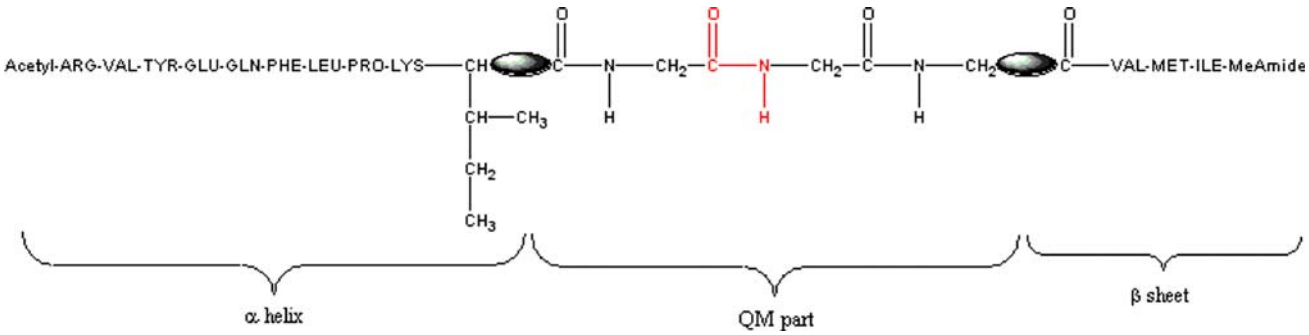
In Table 1 are collected the energies of the *cis* and *trans* conformations of the *Beta-B-Alpha* peptide and of the reduced peptides derived from this one. The total energy is also split into two contributions: the electronic contribution which is the sum of the energy of the quantum fragment and the electrostatic interaction of this fragment with the classical charges of the fragments described at the MM level on one hand, and the classical contribution which is the sum of the conformational energy of the classical fragments and the van der Waals interaction energy of these fragments with the quantum atoms on the other hand. Finally, we also report the differences of these quantities between the *cis* and *trans* conformations. From these results, it clearly appears that the *cis* conformation is energetically favoured in the whole peptide. If one looks at the influence of the various contributions to the total energy, one sees that the *cis* conformation is stabilized by the classical contribution which mainly arises from the interaction between the beta sheet and the alpha helix which are closer in the *cis* conformation than in the *trans* one. In the tetrapeptide, the *trans* conformation is the most stable, as expected. The presence of the alpha helix alone reduces the energy



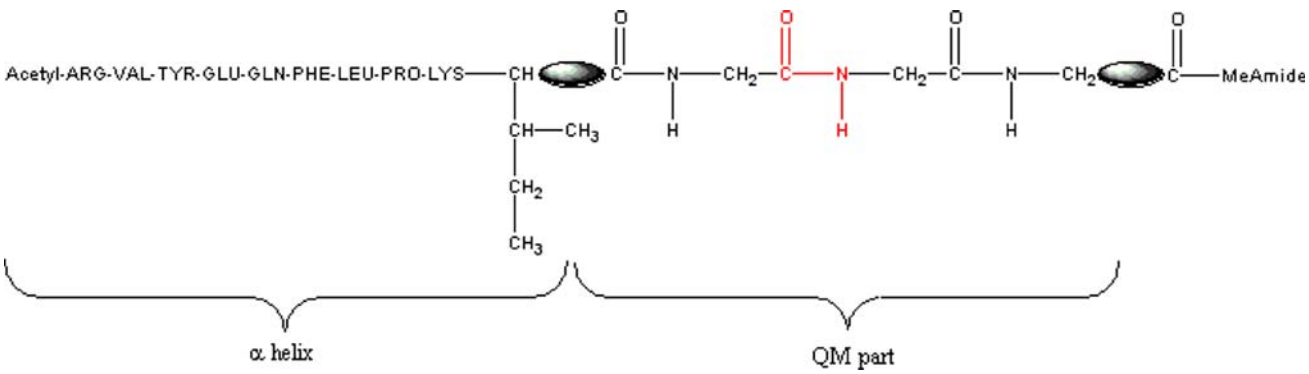
Scheme 2 *Beta-B*



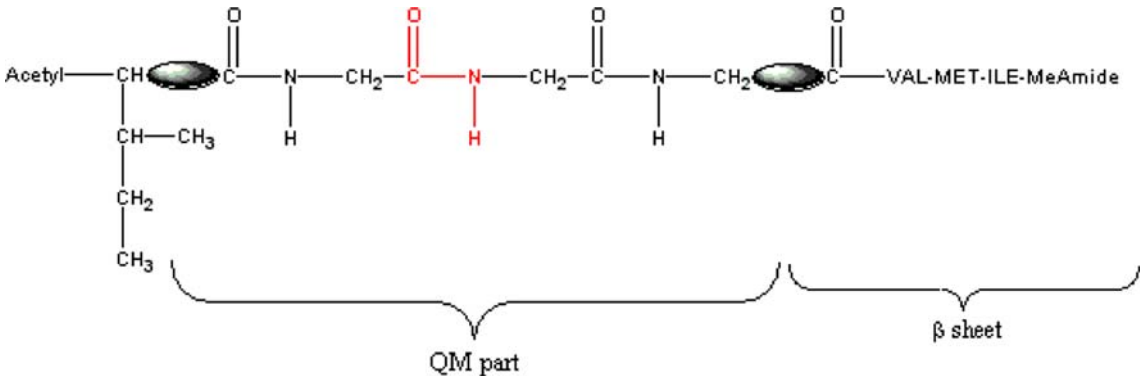
Scheme 3 *B-Alpha*



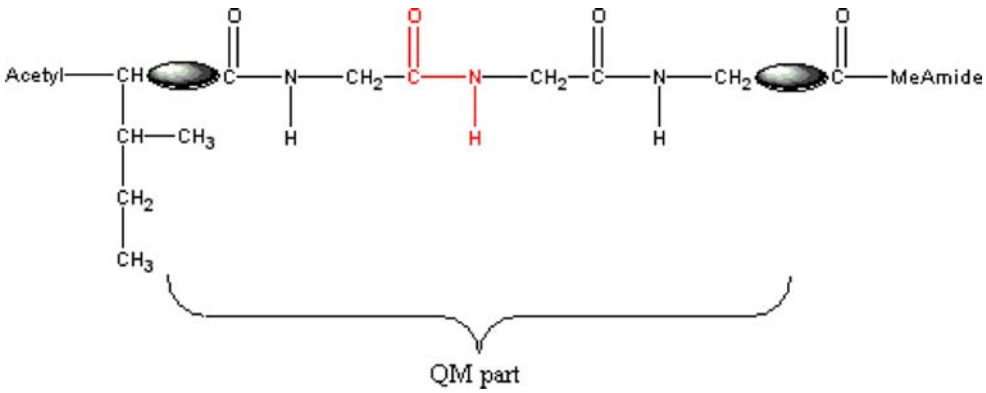
Scheme 4 *Alpha-B-Beta*



Scheme 5 *Alpha-B*



Scheme 6 *B-Beta*



Scheme 7 *Tetrapeptide*

Table 1 Energy data for the *Beta-B-Alpha* molecule and the derived systems

Energies	System	*Beta-B-Alpha*	*Tetrapeptide*	*B-Alpha*	*Beta-B*
Total (a.u.)	*Cis*	−625.29792	−624.78220	−625.30883	−624.77713
	Trans	−625.28663	−624.78826	−625.31130	−624.78338
Electronic (a.u.)	*Cis*	−624.71805	−624.71884	−624.72942	−624.71778
	Trans	−624.72389	−624.72531	−624.72218	−624.72417
Classical (a.u.)	*Cis*	−0.57986	−0.06336	−0.57942	−0.05935
	Trans	−0.56274	−0.06295	−0.58912	−0.05921
Cis-trans difference (kcal/mol)	Total	−7.08	3.80	1.55	3.92
	Electronic	3.66	4.06	−4.54	4.01
	Classical	−10.75	−0.26	6.09	−0.09

Table 2 Energy data for the *Alpha-B-Beta* molecule and the derived systems

Energies	System	*Alpha-B-Beta*	*Tetrapeptide*	*Alpha-B*	*B-Beta*
Total (a.u.)	*Cis*	−625.18519	−624.78220	−625.25033	−624.76490
	Trans	−625.21962	−624.78826	−625.25304	−624.77044
Electronic (a.u.)	*Cis*	−624.70399	−624.71884	−624.70909	−624.71137
	Trans	−624.71763	−624.72531	−624.71777	−624.72355
Classical (a.u.)	*Cis*	−0.48120	−0.06336	−0.54124	−0.05352
	Trans	−0.50198	−0.06295	−0.53527	−0.04689
Cis-trans difference (kcal/mol)	Total	21.60	3.80	1.70	3.48
	Electronic	8.56	4.06	5.45	7.64
	Classical	13.04	−0.26	−3.75	−4.16

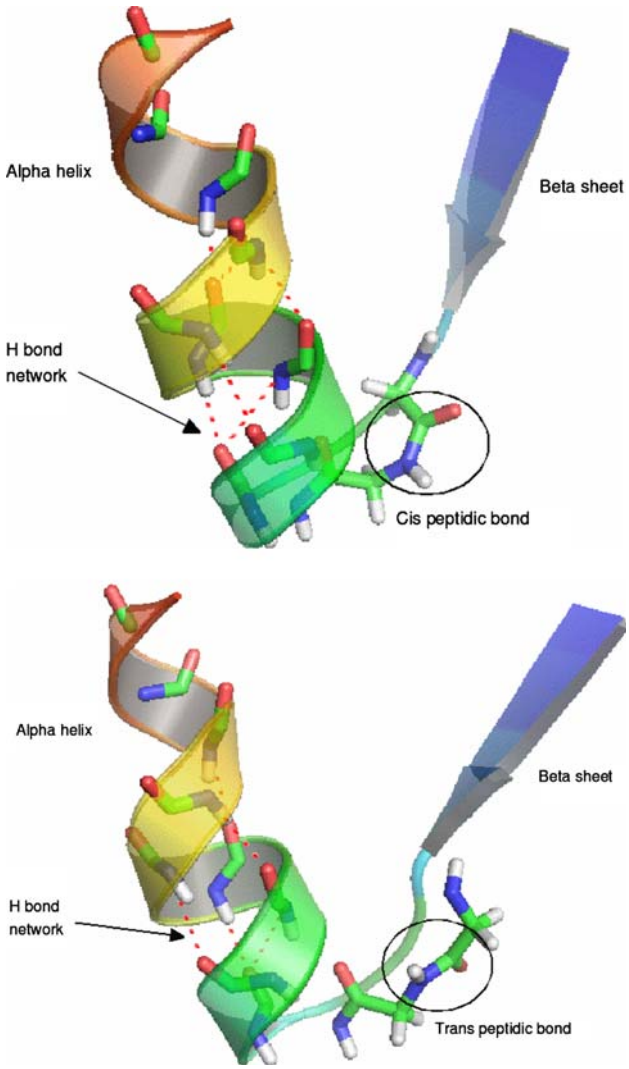


Fig. 1 **a** *Cis Beta-B-Alpha*. **b** *Trans Beta-B-Alpha*

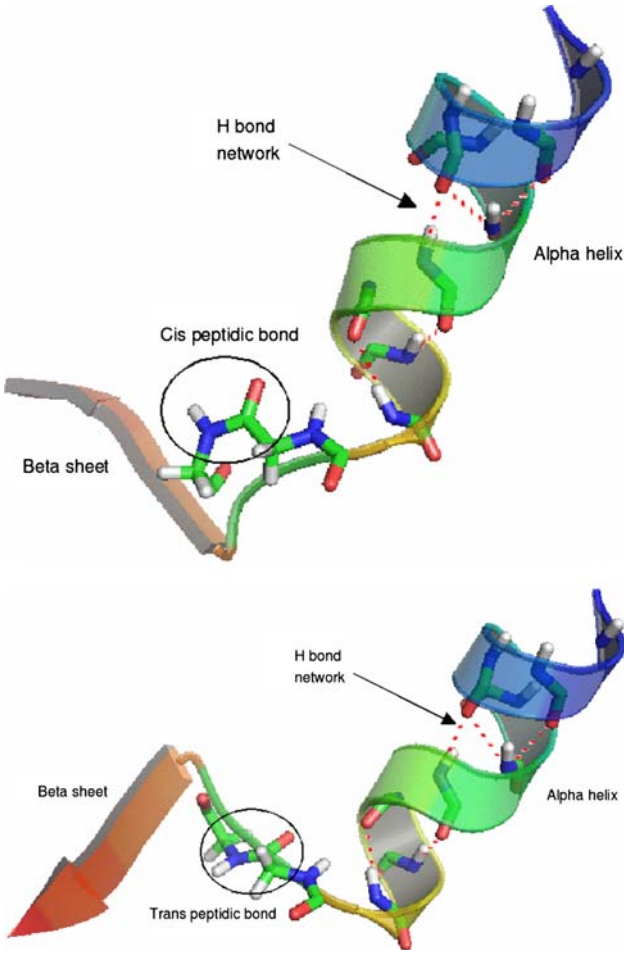


Fig. 2 **a** *Cis Alpha-B-Beta*. **b** *Trans Alpha-B-Beta*

difference, but the reduction is not enough to favour the *cis* conformation and in this case, it appears that the electronic contribution is in favour of the *cis* conformation, but it is counterbalanced by the classical term. The beta sheet alone has exactly the opposite effect. Finally, an examination of Tables 1 and 2 shows that the influence of both classical fragments is not additive. This comes obviously from the fact that the whole peptidic system and the truncated ones have been optimized separately by means of a QM/MM computation.

Exchanging the beta sheet and the alpha helix has a dramatic effect on the stability of the *cis* conformation as shown in Table 2. In particular, one notices in Fig. 1 that, contrary to the *Beta-B-Alpha* case in which the beta sheet and the alpha helix are close together in the *cis* conformation, so that the classical contribution to the total energy favours this conformation, in the present case, they are pushed apart probably because of unfavourable electrostatic interactions. Similarly, the alpha helix alone has now an electronic contribution to the total energy which is unfavourable to the *cis* conformation contrary to the previous case.

These remarks focus our attention on the alpha helix which is known to have a strong electrostatic influence on its neighbouring residues, thanks to its regular structure which keeps its peptidic bonds parallel to each other and generates a network of hydrogen bonds as shown in Figs. 1 and 2. A quite visible effect of this influence is the polarization of the peptidic bond which interacts with the helix [18] and which shows an increased electron transfer from the nitrogen lone pair to the carbonyl carbon atom so that the C–N bond order increases. The typical example of maximum electrostatic interaction can be found in the case of a molecule in solution in a solvent of high dielectric constant like water. In Table 3 we give the results of a computation on *trans* and *cis N* methylacetamide as a prototype of peptide [6]. These results are obtained at the same level of computation as in the

Table 3 Electronic polarization of *N* methyl formamide

	Vacuum			Aqueous solution		
	Relative energy (kcal/mol)	Dipole moment (D)	C–N bond order	Relative energy (kcal/mol)	Dipole moment (D)	C–N bond order
Trans	0	4.07	1.20	0	5.52	1.28
Cis	2.29	4.36	1.28	2.48	5.88	1.33

Table 4 Mayer bond orders of the Gly–Gly C–N bond

	Beta-B-Alpha	*Beta-B*	*B-Alpha*	*Alpha-B-Beta*	*Alpha-B*	*B-Beta*	*Tetrapeptide*
Cis	1.25	1.04	1.15	1.03	1.03	1.13	1.04
Trans	1.35	1.13	1.30	1.18	1.16	1.16	1.14

present study, with the help of the SCRF method developed in our group [19,20]. One notices that, although the dipole moments of both conformations have different orientations, polarization has the same consequence on C–N Mayer's bond order [21] and that there is a relationship between the polarization and the stabilization of the structures. The role of electron delocalization in amides has been analysed in detail by Gao et al. [22].

The results regarding the Gly–Gly bond in our models listed in Table 4 show some difference. In this case, the *trans* conformation in all systems has a greater C–N bond order than the *cis*, probably thanks to the substituents on the carbon atoms. But one sees that in the original *Beta-B-Alpha* system the influence of alpha helix is greater than that of the beta sheet. This can be understood by the fact that the helix is larger and better organized than the beta sheet. One also notices the difference of the effect of the alpha helix, whether it stands at the C terminus of the peptidic bond or at the N terminus. The effect is strong in the first case and negligible in the second one. To a smaller extent, the beta sheet shows a comparable effect in the *cis B-Beta* peptide. This may indicate an easier polarization by the group bonded to the C terminus of the peptidic bond. A simple explanation may be the change of the orientation of the polarizing field with respect to the bond of interest but there could also be some non-linear effects which may arise from the strong and inhomogeneous local electric field acting on this bond and would deserve a more detailed study.

4 Conclusion

The example of the *cis* Gly–Gly peptidic bond in dihydrofolate reductase shows that the stability of the *cis* conformation is the result of several contributions to the intramolecular interaction, which may act in opposite ways. Among them, the polarizing effect of the alpha helix plays a special role especially when the helix is bonded to the carbonyl end of the peptidic group. On the contrary, this effect is rather limited when the helix stands at the other end of the group. These remarks explain the rare occurrence of non-proline *cis* peptidic bonds in proteins, which require, to be formed, a conjunction of several favourable features.

References

1. Rivail J-L, Bouchy A, Loos P-F (2007) J Arg Chem Soc (in press)
2. Sawaya MR, Kraut J (1997) Biochemistry 36:586
3. Jabs A, Weiss MS, Hilgenfeld (1999) J Mol Biol 286:291 (review paper)
4. E Svensson A-K, O'Neill Jr JC, Matthews CR (2003) J Mol Biol 326:569
5. Zimmerman SS, Scheraga HA (1976) Macromolecules 9:408
6. Jorgensen WL, Gao J (1988) J Am Chem Soc 110:4212
7. Poteau R, Trinquier G (2005) J Am Chem Soc 127:13875
8. Enriz RD, Morales ME, Baldoni HA, (2005) J Mol Struct Theochem 731:177
9. Assfeld X, Rivail J-L (1996) Chem Phys Lett 263:100
10. Ferré N, Assfeld X, Rivail J-L (2002) J Comput Chem 23:610
11. Ferré N, Assfeld X (2002) J Chem Phys 117:4119
12. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery Jr JA, 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, Bakken V, 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, Rabuck AD, 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 (2004) Gaussian 03, Revision B.05, Gaussian, Inc., Wallingford

13. Lee C, Yang W, Parr RG (1988) Phys Rev B 37:785
14. Becke AD (1993) J Chem Phys 98:5648
15. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz Jr KM., Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA, (1995) J Am Chem Soc 117:5179
16. Wang J, Cieplak P, Kollman PA (2000) J Comput Chem 21:1049
17. Li H, Robertson AD, Jensen JH, (2005) Proteins 61:704
18. Assfeld X, Ferré N, Rivail J-L (2004) Theor Chem Acc 11:328
19. Rinaldi D, Bouchy A, Rivail J-L, Dillet V (2004) J Chem Phys 120:2343
20. Rinaldi D, Bouchy A, Rivail J-L (2006) Theor Chem Acc 116:664
21. Mayer I (1983) Chem Phys Lett 97:270
22. Mo Y, Schleyer PvR, Wu W, Lin M, Zhang Q, Gao J (2003) J Phys Chem A 107:10011