J Mol Model (2013) 19:193–203
DOI 10.1007/s00894-012-1531-z

ORIGINAL PAPER

Performance comparison of computational methods for modeling alpha-helical structures

Alexandru Lupan · Attila-Zsolt Kun · Francisco Carrascoza · Radu Silaghi-Dumitrescu

Received: 8 March 2012 / Accepted: 5 July 2012 / Published online: 31 July 2012

Abstract Geometry optimization results are reported for secondary structural elements of small proteins and polypeptides. Emphasis is placed on how well molecular mechanics as well as semiempirical, ab initio, and density functional methods describe α-helical and related structures in purely theoretical models (Gly_{10}, Ile_{10}) as well as in realistic models (an α-helical region of calmodulin, and the complete structure of a small protein). Many of the methods examined here were found to provide unsatisfactory descriptions of the hydrogen-bonding interactions within polypeptide-type structures, as the α-helical canonical secondary structure motif was not reproduced accurately. Ab initio and DFT methods provided reasonable results only when solvation models were included, although Hartree–Fock failed even with solvation in one of the test cases; among the semiempirical methods, one of the PM6 implementations performed very well.

Keywords Peptide · Protein · Alpha helix · Hydrogen bond · DFT · PM6

Introduction

Computational examinations of enzyme mechanisms utilizing methods such as density functional theory (DFT) or Hartree–Fock (HF) have traditionally been restricted—especially in metalloproteins containing transition metal ions—to smaller-size models of the active sites, due to the substantial computing resources required [1–5]. On the other hand, the reliability of empirical methods for describing protein structures has long been established [4, 6–9]. In recent years, the use of accurate solvation models and QM/MM-type techniques has allowed the influence of the protein environment to be accounted for more explicitly, although the less-relevant parts of the polypeptide are treated at levels of theory that are inferior to those applied to the active site itself [1–5]. Additionally, there have been constant efforts to improve the performance of lower levels of theory (semiempirical, molecular mechanics), so that they can be applied not just to the distant polypeptide regions but also to the actual active site [10–13]. At the current rate of evolution in computer performance, it may well be possible in the near future to apply DFT or (post-)HF methods to whole-protein enzyme models. Indeed, single-point energies have already occasionally been computed for large—and even complete—protein models at higher levels of theory, such as DFT [13–25]. It is in this context that the study described in this paper sought to gauge how well some of the commonly used computational models predicted the geometries of polypeptide chains. This was considered to be a particularly challenging situation for several of these models. First, polypeptide architectures tend to rely primarily on weak, non-covalent interactions (mainly hydrogen bonding, although recent results have suggested significantly more important contributions from other factors, such as van der Waals dispersion interactions [26–28]). Second, methods requiring parametrization may not necessarily perform efficiently if they have not been parameterized, especially when they are applied to proteins. The present study sought to examine the degree to which α-helical geometries can be described by computational methods; it did not aim to predict the relative stability of this type of secondary structure compared to others.

Electronic supplementary material The online version of this article (doi:10.1007/s00894-012-1531-z) contains supplementary material, which is available to authorized users.

A. Lupan · A.-Z. Kun · F. Carrascoza · R. Silaghi-Dumitrescu (✉)
Department of Chemistry and Chemical Engineering,
Babes-Bolyai University,
11 Arany Janos str,
Cluj-Napoca 400028, Romania
e-mail: rsilaghi@chem.ubbcluj.ro

Methods

The present study employed the following α-helical models: Gly_{10}, Gly_{50}, Ile_{10}, and Ile_{50}. Additionally, two experimentally known structures that were extracted from the Protein Data Bank (3CLN and 1ALG) were also examined computationally. These models were built within the Builder module of the Spartan software package [29], and were either capped at the N- and C-termini with hydrogen atoms (i.e., C-terminal COOH, N-terminal $–NH_2$) or left in ionized forms (carboxylate and ammonium). A distinctive feature of these structures is the regularity imposed by hydrogen bonds between the C=O and NH groups of different peptide bonds; the tables provided in this manuscript show the hydrogen-bond lengths extracted from these starting α-helical canonical structures.

Geometry optimizations were performed either in vacuum or with the CPCM continuum solvent model as implemented in the Gaussian09 software package [30]. Molecular mechanics calculations utilized the force fields Amber and UFF, as implemented in Gaussian09 [30] and Hyperchem [31]. The semiempirical PM6 method was employed as implemented in the Gaussian09 [30] and MOPAC [32] software packages, and the PM3 implementation from Hyperchem [31] was also used.

HF/3-21G* and density functional theory (M062X/6-31G**, M062/6-311+G**, BP86/6-31G**) computations were performed in Gaussian09 [31]. The CPCM solvation model was employed for HF and DFT calculations [30, 33]. Standard convergence criteria as defined in the respective software packages were employed.

Calculations performed with the Gaussian09 package utilized the default settings for the SCF cycles and geometry optimization, namely the fine grid (75,302) to numerically evaluate the integrals, a self-consistent field convergence of 10^{-8} hartrees, a maximum force of 0.000450 hartrees/bohr, an RMS force of 0.000300 hartrees/bohr, a maximum displacement of 0.001800 bohrs, and an RMS displacement of 0.001200 bohrs. The default SCF procedure uses a combination of EDIIS and CDIIS [34], with no damping or Fermi broadening. The nature of each stationary point after optimization was checked by calculating the harmonic vibrational frequencies to ensure that the stationary points found were genuine minima.

We recently reported a technical evaluation of the various optimization algorithms as employed with the PM6 method in the MOPAC2009 software package, using an α-helix of decaglycine as a test case [35]. Based on this extensive study, we chose to use the PM6-D2 Hamiltonian [36] and cpcm continuous solvent model (using a conductor-like screening model [37], a dielectric constant of water of 78.39, and a solvent radius of 1.3 Å) with the L-BFGS optimization method. The default optimizer in MOPAC2009 is Baker's eigenvector-following method [38]. The Davidon–Fletcher–Powel algorithm [39, 40] was the first quasi-Newton generalized method implemented in MOPAC; this was subsequently improved upon by the Broyden–Fletcher–Goldfarb–Shanno (BFGS) procedure [41–44]. An advanced variant of the BFGS optimizer is the "limited-memory BFGS" function minimizer, which calculates the inverse Hessian as needed [45–47], and is thus suitable for optimizing larger systems.

Note that, although we examined the geometries of α-helices, this does not imply that we expect them to be the only—or even the most stable—forms of secondary structure available for the peptides examined in the present study. It is, nevertheless, a basic biochemical principle that all simple peptides are able to form the two main types of secondary structure (assuming that less usual amino acids such as proline are not present), including the α helix. Geometry optimizations (M062x/6-31G**) reveal for instance that the 3_{10} and α-helix geometries are within ~5 kcal mol^{-1} of each other energetically, and are close in energy to other forms as well. These issues relating to the relative stabilities of secondary structural elements are, however, a topic for ongoing investigations, and are not discussed further here.

Results and discussion

Gly_{10}

Geometry optimization results obtained for the neutral Gly_{10} model are shown in Table 1 and Fig. 1. Among the molecular mechanics methods, UFF leads to a strongly distorted structure (cf. Fig. 1), where the length of the helix drops from 13.9 Å in the initial canonical structure to 10.8 Å; further data on this result are therefore not shown in Table 1. The Amber force field (Hyperchem implementation) preserves the helix length to within 1 Å of the starting structure, but yields average hydrogen bonds that are ~0.25 Å longer than in the starting canonical model. Among the semiempirical methods, by far the best performance is provided by the MOPAC implementation of PM6: the values of hydrogen bonds remain very close to each other (1.9–2.1 Å) throughout the model, in contrast to the Gaussian PM6 and PM3 methods, which yield values as high as 4.5. The MOPAC PM6 also appears to perform well in terms of the total length of the helix: it is within 0.6 Å of the canonical geometry.

Helix lengths predicted by ab initio and DFT methods fall within ±0.1 Å of the canonical geometry, with slightly better values obtained from solvated models. The lengths of the hydrogen bonds appear to be optimally modeled by HF in the solvated model: an average of 1.88 Å was obtained, compared to values of >2 Å obtained with the other methods

Table 1 CO···NH hydrogen-bond lengths and helix lengths for the neutral Gly_{10} models

Neutral Gly_{10} [a]	Initial	Amber Hyperchem	PM3 Hyperchem	PM6 Gaussian	PM6 MOPAC	HF/ 3-21G*	M062X/ 6-31G**	HF/ 3-21G*	M062X/ 6-31G**
	Vacuum	Vacuum	Vacuum	Vacuum	Water	Vacuum	Vacuum	Water	Water
1–5	1.74	2.13	1.85	2.18	1.99	2.21	2.51	1.96	2.12
2–6	1.74	1.97	1.89	2.04	1.92	1.93	2.13	1.81	1.97
3–7	1.74	2.04	1.89	3.45	1.94	2.03	2.15	1.9	2.02
4–8	1.74	2	1.86	2.12	1.93	2.05	2.23	1.86	1.99
5–9	1.74	2.05	3.14	2.09	1.99	1.92	2.68	1.87	2.11
6–10	1.74	2.1	4.56	4.47	2.12	2.12	3.08	1.86	2.26
$(CO\cdots HN)_{avg}$	1.74	2.05	2.53	2.72	1.98	2.04	2.46	1.88	2.08
Helix length	13.91	14.53	14.52	12.08	14.49	12.94	14.9	14.03	14.61

[a] Amino acids whose NH/CO groups engage in hydrogen bonds are indicated in this column; numbering starts from the N-terminus of the peptide

[b] Length of the helix, measured between the α-carbon atoms of the first and last amino acids, respectively

of its class. The M06 functional predicts two unusually long hydrogen bonds at the two termini of the helix in the unsolvated model but not in the solvated one. On the other hand, Table S1 of the "Electronic supplementary material" (ESM) shows that another density functional, BP86, completely fails to reproduce an α-helical structure in vacuum, even though it performs reasonably well when used together with a solvation model (average hydrogen bond length −1.95 Å).

One general observation is that all methods predict asymmetry in the helix, with the hydrogen bonds at the extremities of the chain displaying larger values than those in the middle of the chain. Furthermore, the lengths of the two hydrogen bonds at the two ends are not identical according to all of the methods employed here.

For the zwitterionic α-Gly_{10} structure, Fig. 2 and Table 2 illustrate the starting point and selected optimized geometries. The Amber force field leads to loss of the α-helical character of the structure, with hydrogen-bonding distances increasing to as much as 7.8 Å. The UFF force field was found to give similar results to Amber (data not shown).

The zwitterionic Gly_{10} HF structure obtained in vacuum is entirely nonhelical (cf. Fig. 2). Far better results are

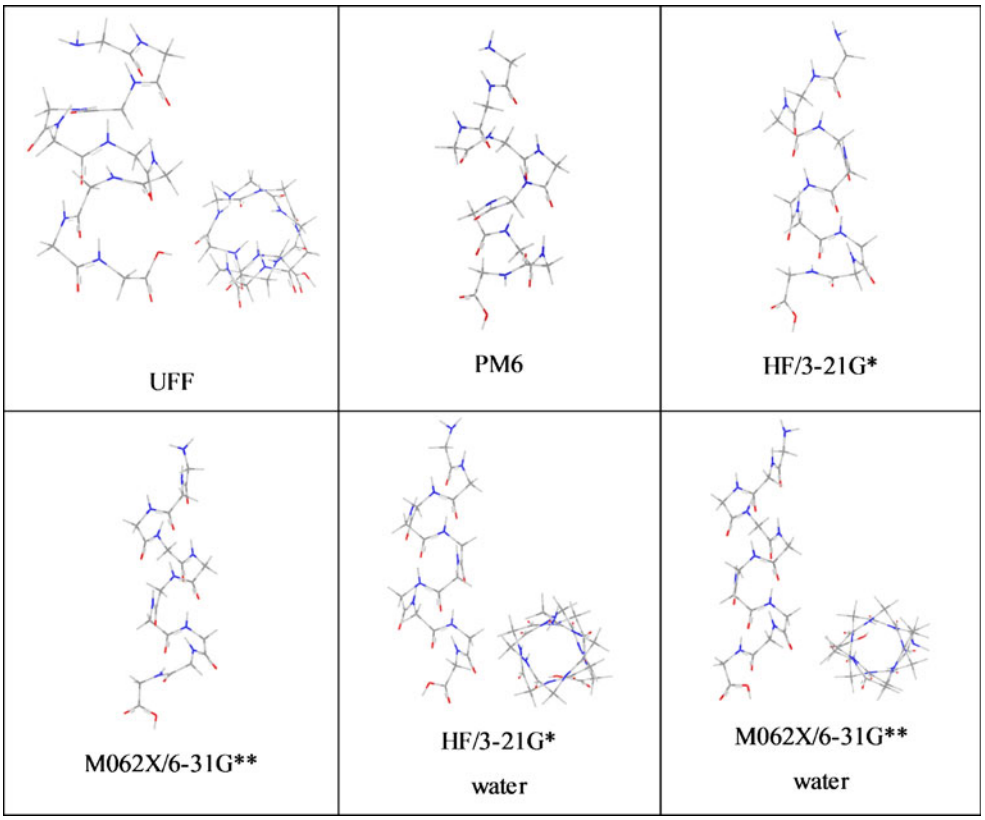

Fig. 1 Graphical representation of the optimized neutral Gly_{10} geometries provided by various methods employed in the present study

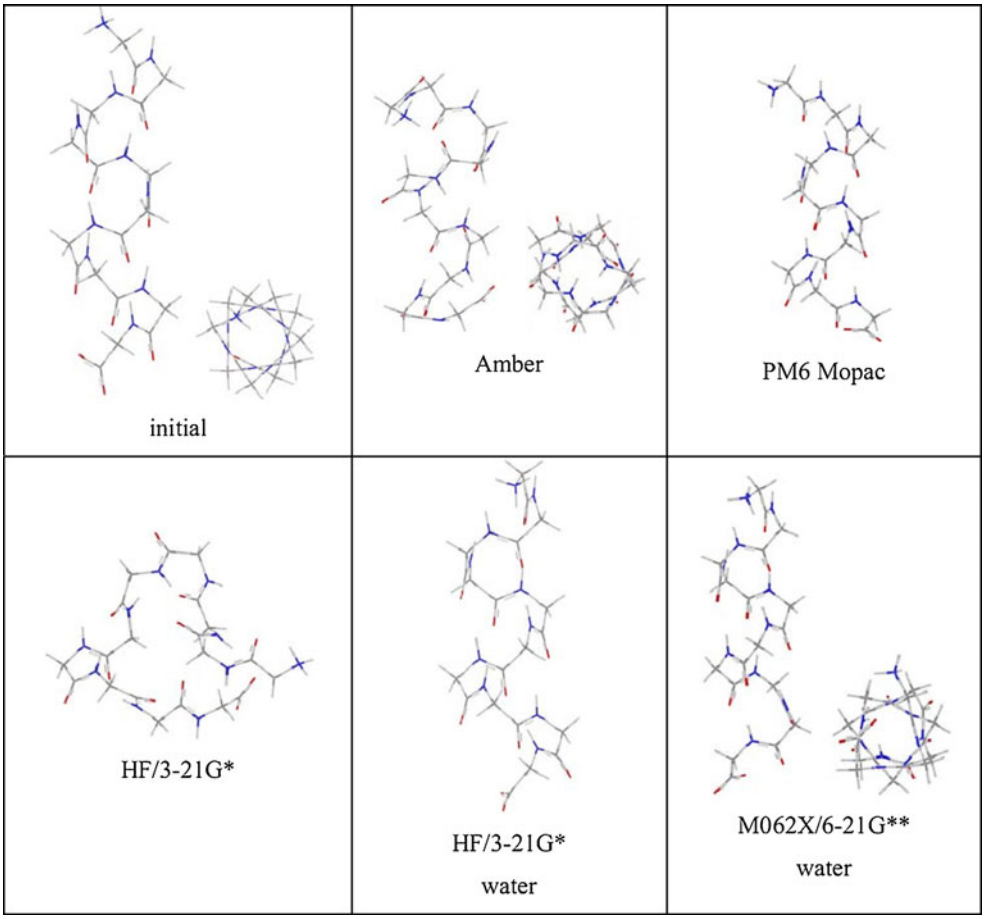

Fig. 2 Graphical representation of optimized zwitterionic Gly_{10} geometries provided by various methods employed in the present study

obtained when solvation is included with the HF and DFT methods. The M06 results also illustrate the effect of using a larger basis set on the results: hydrogen bonds contract by 0.1 Å upon changing from 6-31G** to 6-311+G**—a change that is expected, since the larger basis set leads to a more accurate description of weak interactions such as hydrogen bonds. In contrast to the M06 data shown in Table 2, DFT data obtained for the Gly10 zwitterion with the BP86 functional show that even solvation cannot produce an intact helical structure during geometry optimization, as the two terminal hydrogen bonds are elongated to more than 4 Å (Table S1 of the ESM).

Among the methods employed here, HF/3-21G* with water as solvent appeared to offer the best performance for the Gly_{10} zwitterion, as it maintained an α-helical structure that had regular hydrogen bonds with lengths close to those expected for the canonical α-helix. However, MOPAC PM6 gave a similar level of performance (the average of

Table 2 CO···NH hydrogen-bond lengths and helix lengths for the zwitterionic Gly_{10} models. Labels are as in Table 1. The HF/3-21G**/vacuum geometry is not listed, as the helix was completely destroyed upon geometry optimization

Zwitterionic Gly_{10} [a]	Initial Vacuum NH···OC	Amber Gaussian Vacuum NH···OC	PM6 MOPAC Water NH···OC	HF/3-21G* Water NH···OC	M062X/6-31G** Water NH···OC	M062X/6-311+G** Water NH···OC
1–5	1.74	3.75	2.06	2.04	2.25	2.24
2–6	1.74	4.6	1.94	1.84	2.04	1.99
3–7	1.74	7.75	1.95	1.89	2.02	1.97
4–8	1.74	2.01	2	1.88	2.01	1.98
5–9	1.74	2.56	1.98	1.89	2.28	2.09
6–10	1.74	6.41	2.03	1.96	2.65	2.41
$(CO\cdots HN)_{avg}$	1.74	4.91	1.99	1.92	2.21	2.11
Helix length[b]	13.91	12.52	14.44	14.09	14.85	14.78

Table 3 CO···NH hydrogen bond lengths and helix lengths for the neutral Ile_{10} structure. Labels are as in Table 1

Neutral ILE_{10}	Initial Vacuum NH···OC	Amber Hyper Vacuum NH···OC	PM3 Hyper Vacuum NH···OC	PM6 Gaussian Vacuum NH···OC	PM6 MOPAC Water NH···OC	HF/3-21G* Vacuum NH···OC	M062X/6-31G** Vacuum NH···OC	M062X/6-31G** Water NH···OC	HF/3-21G* Water NH···OC
1–5	1.72	2.13	2.43	2.16	1.94	4.09	2.05	1.97	2.28
2–6	1.72	2.01	2.53	2.11	2.01	2.06	2.16	2.07	1.99
3–7	1.72	2.26	2.55	2.21	2.05	2.57	2.36	2.2	2.28
4–8	1.72	2.18	2.56	2.14	2.03	4.84	2.26	2.14	2.32
5–9	1.72	2.13	2.8	2.84	2.06	3.52	3.09	2.18	2.11
6–10	1.72	2.05	3	3.27	2.04	3.49	3.32	2.14	2.2
$(CO\cdots HN)_{avg}$	1.72	2.12	2.65	2.45	2.02	3.42	2.54	2.11	2.2
Helix length	13.86	14.43	14.65	15.05	14.26	15.89	14.96	14.42	14.77

hydrogen-bond length was only 0.07 Å longer than that for HF/3-21G* with water as solvent).

α-Gly_{50}

Figure S2 and Table S2 of the ESM illustrate results obtained for a neutral Gly_{50} α-helix, in order to gauge the extent to which the results obtained for the decaglycine model can be used to derive trends observable in larger models. The HF/3-21G* approach again yields results close to those expected for the canonical α-helix, with the MOPAC PM6 method performing slightly worse than for the decaglycine model (average hydrogen-bond length of 2.22 Å in Gly_{50} vs. 1.98 Å in Gly_{10}). The Gaussian09 implementation of the PM6 method reshapes the helix so that a hydrogen bond is formed in the optimized geometry between amino acids 1 and 6, as opposed to the hydrogen bond between amino acids 1 and 5 in the starting geometry (with this trend conserved throughout the helix: 2–7, 3–8, 4–9 instead of 2–6, 3–7, 4–8). Results for the UFF method are also listed in Table 1; in agreement with what was seen in the decaglycine model, hydrogen bonds are predicted to be too long (2.7 Å, almost 1 Å longer than the expected value), and the helix length is also distinctly different from the canonical value as well as the value computed with HF/3-21G*. Just as was seen for the decaglycine models, all of the methods indicated that the hydrogen bonds were asymmetric throughout the helix.

Model data for the zwitterionic Gly_{50} are shown in Table S3 and Fig. S3 of the ESM. The MOPAC PM6 method again performs remarkably well both in terms of hydrogen bond length and overall helix length. The Amber and UFF force fields each predict unreasonably long hydrogen bonds at the ends of the helix (2.6–6 Å); also, the helix lengths predicted by these force fields differ by more than 10 Å from the canonical value. The Gaussian09 implementation of the semiempirical method PM6 rearranges the α-helical structure in the same way as it did for the neutral model, so that 1–6 hydrogen bonds are preferentially formed over 1–5; furthermore, the N- and C- termini of the chain are eliminated from the molecule in the form of NH_3 and CO_2 during geometry optimization.

α-Ile10

Isoleucine-containing α-helices were also examined to check the extent to which side chains larger than that of glycine affect the performance of each computational method in describing

Table 4 CO···NH hydrogen bond lengths and helix lengths for the Ile_{10} zwitterion. Labels are as in Table 1

Zwitterionic Ile_{10}	Initial Spartan NH···OC	Amber Hyper Vacuum NH···OC	HF/3-21G* Gaussian Vacuum NH···OC	PM6 MOPAC Water NH···OC	HF/3-21G* Gaussian Water NH···OC	M062X/6-31G** Gaussian Water NH···OC
1–5	1.72	2.06	1.86	2.01	3.76	2.04
2–6	1.72	2.05	4.43	2.01	1.99	2.09
3–7	1.72	3.58	5.88	2.04	2.34	2.19
4–8	1.72	3.9	4.38	2.09	3.96	2.16
5–9	1.72	3.13	2	2.09	2.09	2.34
6–10	1.72	3.24	3.43	2.03	2.46	2.37
$(CO\cdots HN)_{avg}$	1.72	-	-	2.04	2.76	2.19
Helix length	13.86	14.34	-	14.4	15.62	14.56

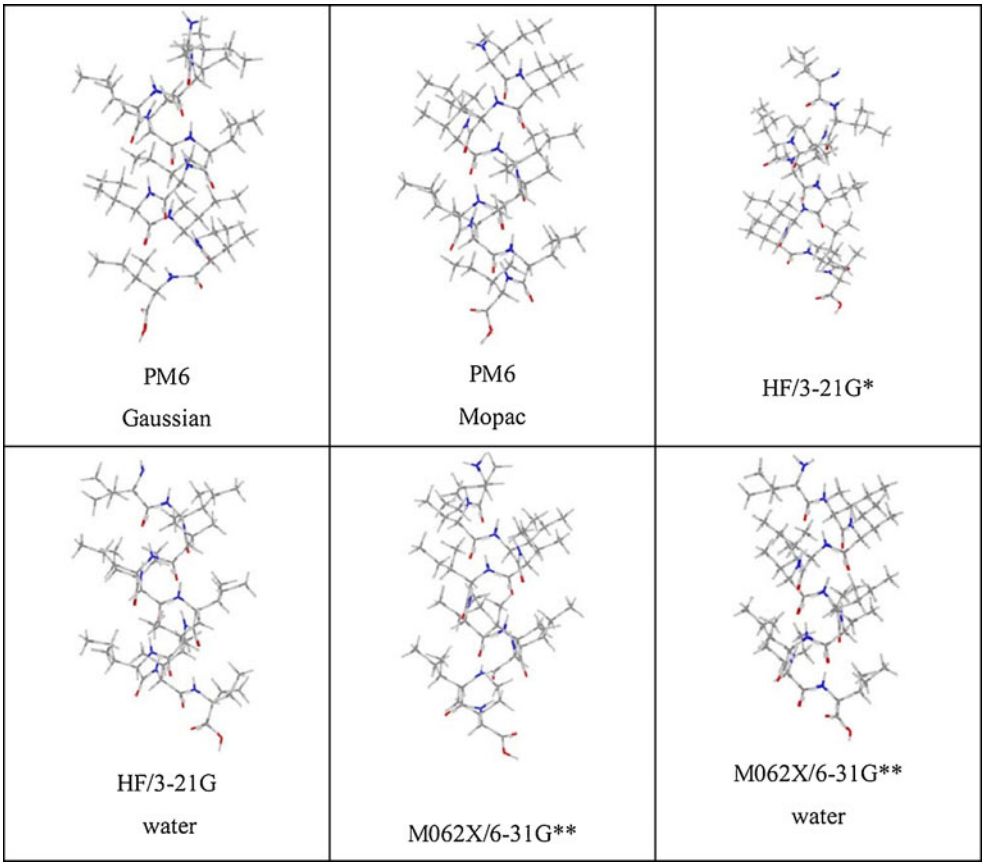

Fig. 3 Graphical representation of the optimized neutral Ile_{10} geometries obtained by three different methods

this element of secondary structure. Results for Ile_{10} (Tables 3 and 4 and Figs. 3 and 4) and Ile_{50} (Table S4) were generally similar to those obtained for the glycine models.

Table 3 illustrates that for the neutral Ile_{10} model, the Amber force field performs as well as it did for Gly_{10}. The relative performance levels of the semiempirical

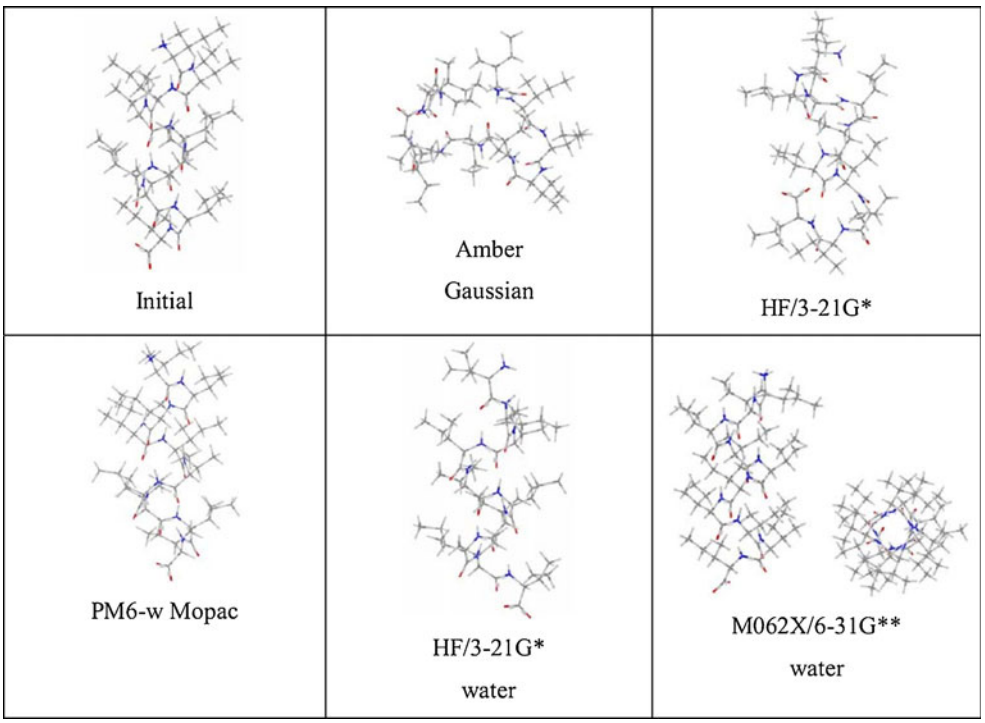

Fig. 4 Graphical representation of the optimized zwitterionic Ile_{10} geometries obtained by various methods

Table 5 Alpha helix CO···NH bond lengths for the calmodulin fragment

Long loop of 3CLN	PDB NH···OC	Amber Gaussian Vacuum NH···OC	PM6 MOPAC Water NH···OC	HF/3-21G* Gaussian Vacuum NH···OC
Phe65–Leu69	1.79	1.8	1.96	2.29
Pro66–Thr70	2.19	2.01	2.02	2.01
Glu67–Met71	2.06	1.82	1.95	2.34
Phe68–Met72	1.79	1.92	1.79	1.89
Leu69–Ala73	2.04	1.99	2	1.92
Thr70–Arg74	2.74	1.72	1.99	2.23
Met71–Lys75	2.21	2.72	1.87	1.92
Met72–Met76	2.32	5.44	1.98	2.09
Ala73–Lys77	2.6	6.5	1.96	1.98
Arg74–Asp78	2.67	3.26	1.88	1.84
Lys75–Thr79	2.94	4.02	1.82	3.37
Met76–Asp80	2.89	1.86	1.92	1.77
Lys77–Ser81	2.67	2.02	1.99	1.94
Asp78–Glu82	2.31	6.29	1.92	4.1
Thr79–Glu83	2.84	3.26	1.88	2.64
Asp80–Glu84	3	2.39	2.01	1.83
Ser81–Ile85	2	2	2.03	1.96
Glu82–Arg86	2.04	2.71	1.93	1.96
Glu83–Glu87	2.45	1.98	1.86	1.73
Glu84–Ala88	2.12	3.41	1.93	1.93
Ile85–Phe89	2.31	3.67	1.92	2.58
Arg86–Arg90	2.71	3.11	1.88	1.66
Glu87–Val91	2.27	-	2.19	1.85
Ala88–Phe92	1.89	-	2.5	3.87
$(CO\cdots HN)_{avg}$	2.09	2.18	1.97	1.95
RMSD	-	4.98	1.23	3.74

methods are also the same: the only reasonable result (monotonous hydrogen-bond lengths that are well below 2.5 Å) is provided by MOPAC PM6. On the other hand, the HF method, which described Gly10 reasonably well, fails here, yielding hydrogen-bond lengths as long as 4.8 Å. The M06 functional appears to be more reliable in this respect, as the average hydrogen-bond length is ~2.5 Å in both Gly_{10} and Ile_{10}.

Table 4 illustrates geometry optimization for the Ile_{10} zwitterion. The Amber force field fails to provide reasonable hydrogen-bonding geometries, while the MOPAC PM6 once again provides reasonable results. Unexpectedly, the HF method fails not only in vacuum but also in solvent: in the latter case, hydrogen bonds as long as 4 Å are predicted, as opposed to the reasonable geometry predicted for Gly_{10} (see Table 2). By contrast, the M06 functional with solvation provides a distinctly more reasonable geometry and is, in this respect, the one method—alongside Mopac PM6—that provides reasonable descriptions for all of the models examined thus far.

α-Ile-50

Table S4 and Figs. S4 and S5 of the ESM show data for neutral and zwitterionic α-Ile_{50} structures. With the Amber force field, the two ends of the helix are distorted (with hydrogen-bond lengths as long as 5.4 Å), but the rest of the chain retains reasonable hydrogen-bond lengths for an α-helix. For the neutral structure, the UFF force field yields hydrogen bonds of >3 Å (data not shown).

Calmodulin model

Calmodulin shows α-helical regions, and thus represents a reasonable test case to use to check the extent to which the results obtained for models such as Gly10/50 and Ile10/50 can be extrapolated to experimentally known peptide/protein structures. The central part of calmodulin features a long α-helix (28 residues, overall charge −2). This portion of the protein was extracted from the X-ray diffraction structure,

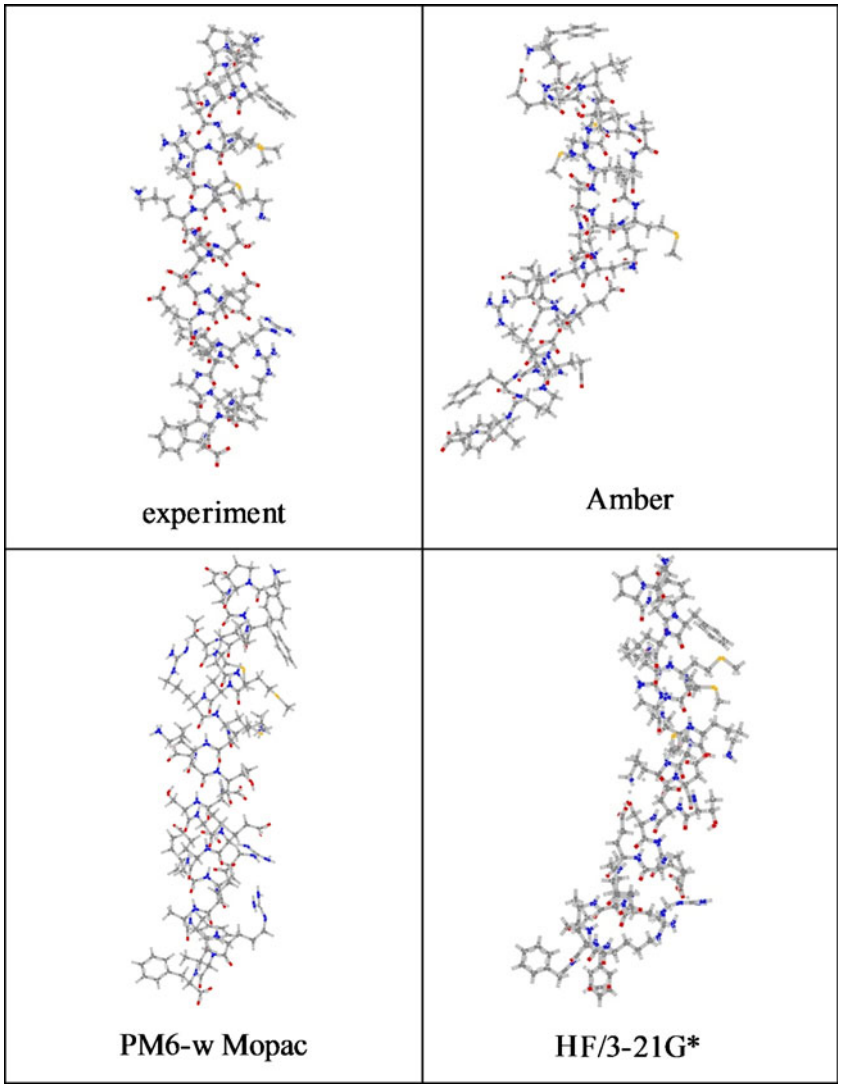

Fig. 5 Graphical representation of the optimized 3CLN long-loop geometries obtained by various methods

pdb code 3CLN [48], and used in computations as detailed in Table 5. The N- and C-termini were modeled as ammonium and carboxylate groups, respectively, and an overall charge of −2 was assigned (as determined by the number of Asp/Glu and Lys/Arg amino acids, whose side chains were modeled as charged −1 and +1, respectively). Three methods with reasonable computational costs were employed in this case. As shown in Table 5 and Fig. 5, the Amber method does not model the helical structure properly, predicting a well-defined bending point midway through the structure (cf. Fig. 5) and several unacceptably long (>3 Å) hydrogen-bonding distances (amino acids 74–85). A bent helix is also predicted by the HF approach, which unfortunately also eliminates the C-terminal carboxylate as CO_2 upon geometry optimization. The MOPAC PM6 method is the only method that completely preserves the α-helical structure; however, in this case, the tendency to give a canonical structure is exaggerated: even though there are hydrogen-bond distances as long as 3 Å in the experimental structure, the PM6 method predicts that all of the hydrogen bonds in the structure are in the ~1.9–2.0 Å range.

1ALG

The PDB structure with a code of 1ALG is a protein composed of 24 residues with a global α-helix structure (although some of the terminal amino acids are not part of the helix) [49]. Computations were performed on the complete structure of this protein without any truncation; data are shown in Table 6 and Fig. 6. Interestingly, methods that gave less impressive results for the smaller/simpler models discussed above (e.g., the Gaussian implementation of PM6, or HF/3-21G* in vacuum) are found to reproduce the regularity of the helical structure of the complete, experimentally known, protein 1ALG reasonably well. In fact, even where the experimental structure shows deviations from an α-helix, such as for the

Table 6 Alpha helix CO···NH hydrogen-bond lengths for the 1ALG structure

	1ALG Charge −1	PDB NH···OC	Amber Gaussian Vacuum NH···OC	PM6 Gaussian Vacuum NH···OC	PM6 MOPAC Water NH···OC	HF/3-21G* Vacuum NH···OC	HF/321G* Water NH···OC
1	Met8–Gly4	2.58	-	2.28	1.84	1.88	1.93
2	Leu9–Cys5	2.29	2.95	2	1.97	2.13	2.13
3	Glu10–Asp6	1.94	1.82	2	1.95	2.02	1.83
4	Gly11–Glu7	2.14	2	2.27	2.21	1.94	1.93
5	Phe12–Met8	2.33	1.8	2.02	1.82	1.85	1.9
6	Ala13–Leu9	2.02	1.88	2.25	1.88	1.9	1.84
7	Val14–Gln10	1.94	1.84	2.03	2.02	1.91	1.81
8	Ala15–Gly11	2.14	1.93	2.01	2	1.94	2.17
9	Val16–Phe12	1.83	1.85	2	2.01	1.89	1.97
10	Lys17–Ala13	2.25	2.05	2.21	1.93	2.34	1.95
11	Met18–Val14	1.91	3.78	1.85	1.8	1.85	1.97
12	Gly19–Ala15	1.74	2.04	1.89	2.14	1.74	1.91
	$(CO\cdots HN)_{avg}$	2.09	2.17	2.07	1.96	1.95	1.95
	RMSD	-	2.74	2.38	1.18	1.59	1.30

first two hydrogen bonds (which are longer) and the last hydrogen bond (which is distinctly shorter), the computational methods fail to predict these deviations.

Summary

Several computational methods for performing the geometric optimization of simple α-helical protein models were tested. Methods that performed well with larger protein structures appeared to be unable to describe simpler α-helices properly. One of the implementations of the semiempirical method PM6, alongside density functional calculations including solvation (with the M06 functional performing better than the others), appear to provide reasonable results for all of test cases studied here, although the hydrogen-bonding distances obtained using these methods were typically within 0.2–0.6 Å of those expected for the canonical structure.

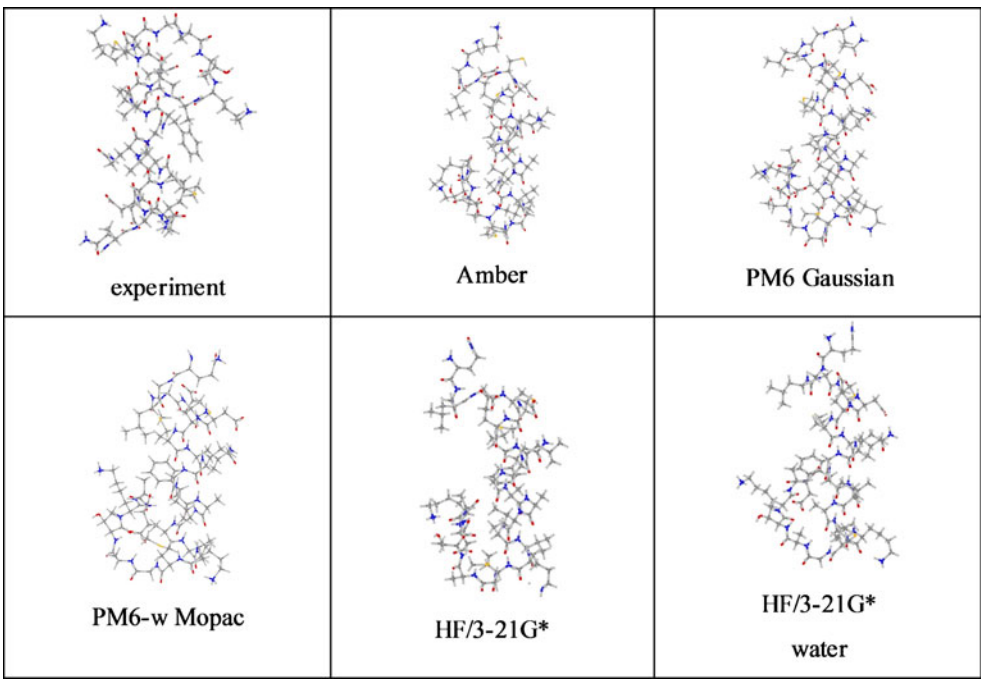

Fig. 6 Graphical representation of the optimized 1ALG geometries obtained by various methods

Acknowledgments Funding from the Romanian Ministry of Education and Research (grants PN II 312/2008 and Parteneriate 72168/2008-FLUORODENT) is gratefully acknowledged.

References

1. Noodleman L, Lovell T, Han WG, Li J, Himo F (2004) Quantum chemical studies of intermediates and reaction pathways in selected enzymes and catalytic synthetic systems. Chem Rev 104:459–508
2. Siegbahn PE, Blomberg MRA (2000) Transition-metal systems in biochemistry studied by high-accuracy quantum chemical methods. Chem Rev 100:421–437
3. Harris DL (2001) High-valent intermediates of heme proteins and model compounds. Curr Opin Chem Biol 5:724–735
4. Friesner RA, Dunietz BD (2001) Large-scale ab initio quantum chemical calculations on biological systems. Acc Chem Res 34:351–358
5. Gooding SR, Winn PJ, Jones GA, Ferenczy GG, Frusher MJ, Reynolds CA (2006) Classical polarization in hybrid QM/MM methods. J Phys Chem A 110:6487–6497
6. Warshel A, Parson WW (2001) Dynamics of biochemical and biophysical reactions: insight from computer simulations. Q Rev Biophys 34:563–679
7. Warshel A (2002) Molecular dynamics simulations of biological reactions. Acc Chem Res 35:385–395
8. Rosta E, Klahn M, Warshel A (2006) Towards accurate ab initio QM/MM calculations of free-energy profiles of enzymatic reactions. J Phys Chem B 110:2934–2941
9. Ryde U (2003) Combined quantum and molecular mechanics calculations on metalloproteins. Curr Opin Chem Biol 7:136–142
10. Silaghi-Dumitrescu R (2004) On the performance of the PM3 semiempirical method with heme complexes relevant to dioxygen and peroxide activation. Rev Chim 55:304–307
11. Tejero I, Gonzalez-Lafont A, Lluch JM (2007) A PM3/d specific reaction parameterization for iron atom in the hydrogen abstraction catalyzed by soybean lipoxygenase-1. J Comput Chem 28:997–1005
12. Mcnamara JP, Sundararajan M, Hillier IH, Ge J, Campbell A, Morgado C (2006) Can the semiempirical PM3 scheme describe iron-containing bioinorganic molecules? J Comput Chem 27:1307–1323
13. Mcnamara JP, Sundararajan M, Hillier IH (2005) Development of parameter sets for semi-empirical MO calculations of transition metal systems: iron parameters for iron–sulfur proteins. J Mol Graph Model 24:128–137
14. Csontos J, Palermo NY, Murphy RF, Lovas S (2008) Calculation of weakly polar interaction energies in polypeptides using density functional and local Moller–Plesset perturbation theory. J Comput Chem 29:1344–1352
15. Sulpizi M, Raugei S, Vandevondele J, Carloni P, Sprik M (2007) Calculation of redox properties: understanding short- and long-range effects in rubredoxin. J Phys Chem B 111:3969–3976
16. Jensen KP (2006) Iron–sulfur clusters: why iron? J Inorg Biochem 100:1436–1439
17. Sundararajan M, Hillier IH, Burton NA (2006) Structure and redox properties of the protein, rubredoxin, and its ligand and metal mutants studied by electronic structure calculation. J Phys Chem A 110:785–790
18. Machonkin TE, Westler WM, Markley JL (2005) Paramagnetic NMR spectroscopy and density functional calculations in the analysis of the geometric and electronic structures of iron–sulfur proteins. Inorg Chem 44:779–797
19. Vrajmasu VV, Munck E, Bominaar EL (2004) Theoretical analysis of the three-dimensional structure of tetrathiolato iron complexes. Inorg Chem 43:4867–4879
20. Kennepohl P, Solomon EI (2003) Electronic structure contributions to electron-transfer reactivity in iron–sulfur active sites: 3. Kinetics of electron transfer. Inorg Chem 42:696–708
21. Vrajmasu VV, Bominaar EL, Meyer J, Munck E (2002) Mossbauer study of reduced rubredoxin as purified and in whole cells. Structural correlation analysis of spin Hamiltonian parameters. Inorg Chem 41:6358–6371
22. Glaser T, Rose K, Shadle SE, Hedman B, Hodgson KO, Solomon EI (2001) S K-edge X-ray absorption studies of tetranuclear iron–sulfur clusters: mu-sulfide bonding and its contribution to electron delocalization. J Am Chem Soc 123:442–454
23. Sigfridsson E, Olsson MH, Ryde U (2001) Inner-sphere reorganization energy of iron–sulfur clusters studied with theoretical methods. Inorg Chem 40:2509–2519
24. Kloiber K, Weiskirchen R, Krautler B, Bister K, Konrat R (1999) Mutational analysis and NMR spectroscopy of quail cysteine and glycine-rich protein CRP2 reveal an intrinsic segmental flexibility of LIM domains. J Mol Biol 292:893–908
25. Nayal M, Di Cera E (1996) Valence screening of water in protein crystals reveals potential Na^+ binding sites. J Mol Biol 256:228–234
26. Rossi M, Blum V, Kupser P, Helden V, Bierau F, Pagel K, Meijer G, Scheffler M (2010) Secondary structure of Ac-Ala_n-$LysH^+$ polyalanine peptides (n=5, 10, 15) in vacuo: helical or not? J Phys Chem Lett 1:3465–3470
27. Hua S, Xu L, Li S (2011) Cooperativity in long α- and 3_{10}-helical polyalanines: both electrostatic and van der Waals interactions are essential. J Phys Chem B 115:11462–11469
28. Tkatchenko A, Rossi M, Blum V, Ireta J, Scheffler M (2011) Unraveling the stability of polypeptide helices: critical role of van der Waals interactions. Phys Rev Lett 106:118102
29. Wavefunction, Inc. (1998)Spartan 5.0, Wavefunction, Inc., Irvine
30. Frisch MJ, Trucks GW, H. B. Schlegel HB, Scuseria GE, Robb, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery, Jr JA, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell K, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas O, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) Gaussian 09, revision A.02. Gaussian Inc., Wallingford
31. Hypercube, Inc. (2007) HyperChem molecular modelling system, release 8.0 for Windows. Hypercube, Inc.
32. Stewart JJP (2009) MOPAC2009, version 10.153L. Stewart Computational Chemistry, Colorado Springs
33. Eckert F, Klamt A (2002) COSMO solvation model. AICHE J 48:369–385
34. Kudin KN, Scuseria GE, Cancès E (2002) A black-box self-consistent field convergence algorithm: one step closer. J Chem Phys 116:8255–8261
35. Kun ZA, Lupan A, Silaghi-Dumitrescu R (2010) PM6 modeling of alpha helical polypeptide structures. Studia Universitatis Babes-Bolyai Chemia 55:31–36
36. Korth M, Pitonak M, Rezac J, Hobza P (2010) A transferable H-bonding correction for semiempirical quantum-chemical methods. J Chem Theor Comput 6:344–352
37. Klamt A, Schüürmann G (1993) COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. J Chem Soc Perkin Trans 2:799–805
38. Baker J (1986) An algorithm for the location of transition states. J Comput Chem 7:385–395

39. Davidon WC (1991) Variable metric method for minimization. SIAM J Optim 1:1–17
40. Fletcher R (1987) Practical methods of optimization, 2nd edn. Wiley, New York
41. Broyden CG (1970) The convergence of a class of double-rank minimization algorithms. II. The new algorithm. J Inst Math Appl 6:222–231
42. Fletcher R (1970) A new approach to variable metric algorithms. Comput J 13:317–322
43. Goldfarb D (1970) A family of variable-metric methods derived by variational means. Math Comput 24:23–26
44. Shanno DF (1970) Conditioning of quasi-Newton methods for function minimization. Math Comput 24:647–656
45. Matthies H, Strang G (1979) The solution of nonlinear finite element equations. Int J Num Meth Eng 14:1613–1626
46. Nocedal J (1980) Updating quasi-Newton matrices with limited storage. Math Comput 35:773–782
47. Byrd RH, Nocedal J, Schnabel RB (1994) Representations of quasi-Newton matrices and their use in limited memory methods. Math Prog 63:129–156
48. Babu YS, Bugg CE, Cook WJ (1988) Structure of calmodulin refined at 2.2 Å resolution. J Mol Biol 204:191–204
49. Nordhoff A, Tziatzios C, Van Den Broek JA, Schott MK, Kalbitzer HR, Becker K, Schubert D, Schirmer RH (1997) Solution structure of an HGR inhibitor. Eur J Biochem 245:273–282