STUDIA UBB. CHEMIA, LV, 3, 2010

PM6 MODELING OF ALPHA HELICAL POLYPEPTIDE STRUCTURES

ATTILA-ZSOLT KUN[a], ALEXANDRU LUPAN[a],
RADU SILAGHI-DUMITRESCU[a]

ABSTRACT. This paper presents a computational study of an alpha-Gly_{10} polypeptide (neutral and zwitterionic) using the PM6 semiempirical method and its variants implemented in Mopac2009 program. Statistical correlations with high level computational methods are discussed in all cases.

Keywords: *glycine, alpha-helix, PM6, semiempirical calculation, optimization*

INTRODUCTION

Computational study of proteins has been a challenging problem. This is due to their size and complex structure, based on weak interactions. High level computational methods such as *"ab initio"* and DFT are still computationally expensive although the computer evolution is tremendous. Less expensive approaches like molecular mechanics or the classical semiempirical methods, although easy in terms of computational effort, are not accurate enough in modeling large systems such as proteins [1].

The PM6 semiempirical method developed by Stewart [2] is a novel and more accurate among its class. It represents an update of the well-known NDDO method (core-core interactions and optimized parameters). Significant improvements of the predicted geometries as compared to the common AM1 and PM3 methods are also reported. In order to reduce the computational effort a new method of solving the self-consistent field equations has been proposed, namely Mozyme [3], which is a localized molecular orbital approach, much more efficient for large systems than the classical methods based on matrix algebra.

The Mopac2009 [4] implementation of the PM6 method and its augmented variants D2 and DH2 has been chosen here for studying a simple computationally-modeled Gly_{10} polypeptide. The different optimization algorithms available in the program were tested to see their performance in energetic and structural terms. All calculations were done using the Mozyme SCF method specially designed for modeling large structures.

---

[a] *Babeş-Bolyai University, Faculty of Chemistry and Chemical Engineering, 11, Arany Janos Str., RO-400028 Cluj-Napoca, Romania, rsilaghi@chem.ubbcluj.ro; zak@chem.ubbcluj.ro*

RESULTS AND DISCUSSIONS

The following tables show the different results obtained for the Gly_{10} model (cf. Methods section). As seen in Table 1, for the Gly_{10} neutral structure computed in vacuum the performance of the EF, EF-HESS1 and L-BFGS optimizers is roughly the same: the calculated heats of formation are roughly the same (within a difference of 1-2 kcal/mol), and the residual gradients of the optimized structures are low. The standard BFGS and DFP methods are giving poorer results in energy, with greater residual gradients. The change to precise optimization criteria has only little influence on the final energies.

Table 1. Heats of formation and residual gradients for the Gly_{10} neutral form computed in vacuum. See also the Methods section

Model Hamiltonian	Optimization method	Normal calculation		Precise calculation		Δ energy
		Energy	Gradient	Energy	Gradient	
PM6	DFP	-468.581	55.38226	-476.200	24.74704	7.61849
PM6	EF	-483.601	1.26385	-484.186	0.75005	0.58480
PM6	EF-HESS1	-483.329	1.40090	-484.177	0.90926	0.84806
PM6	BFGS	-474.173	11.15584	-479.449	2.85466	5.27662
PM6	L-BFGS	-481.251	1.59716	-483.699	1.42570	2.44824
PM6-D2	DFP	-495.630	11.49224	-499.523	2.64713	3.89263
PM6-D2	EF	-504.138	1.52330	-505.804	0.94282	1.70246
PM6-D2	EF-HESS1	-504.145	1.53369	-505.745	0.76028	1.60008
PM6-D2	BFGS	-498.226	6.71154	-499.428	2.85579	1.20209
PM6-D2	L-BFGS	-504.099	1.57670	-504.742	2.02749	0.64335
PM6-DH2	DFP	-497.731	10.27741	-503.475	4.02141	5.74371
PM6-DH2	EF	-508.041	3.05565	-507.973	4.31984	-0.06739
PM6-DH2	EF-HESS1	-508.358	1.63295	-507.911	6.62062	-0.44659
PM6-DH2	BFGS	-500.591	14.10671	-500.801	15.08551	0.20989
PM6-DH2	L-BFGS	-507.982	1.52834	-509.030	5.44155	1.04775

For the Gly_{10} neutral structure computed in vacuum the PM6 BFGS optimization method gives the correlation values closest to the M062X simulation with average bond differences of (0.3-0.4 A). PM6-D2 DFP lays also in the same area of precision. Both PM6-D2 BFGS variants have a correlation coefficient of less than 0.9 but globally all four BFGS methods are in top five in terms of these comparisons. All the PM6-DH2 methods fail in terms of correlation coefficients and the maximum and average differences are greater.

Table 2. CO...NH hydrogen bond lengths and helix lengths for the Gly_{10} neutral model form computed in vacuum. The order is from the N-terminal ending of the polypeptide chain. See also the Methods section for details

Neutral vacuum	1-5	2-6	3-7	4-8	5-9	6-10	CO…NH avg	Helix length	Correl	Max Dif	Avg Dif
Reference	2.51	2.13	2.15	2.23	2.68	3.08	2.46	14.90	-	-	-
PM6 DFP	1.85	2.01	1.95	1.90	1.92	2.17	1.97	14.10	0.55	0.91	0.50
PM6 DFP pr	1.97	1.99	2.10	1.99	2.06	2.30	2.07	14.33	0.74	0.78	0.40
PM6 EF	2.26	2.02	2.13	2.08	2.09	2.10	2.11	14.36	0.18	0.98	0.35
PM6 EF pr	1.97	1.99	2.10	1.99	2.06	2.30	2.07	14.33	0.74	0.78	0.40
PM6 EF HESS1	2.14	2.01	2.26	2.19	2.12	2.10	2.14	14.44	-0.22	0.98	0.36
PM6 EF HESS1 pr	2.27	2.04	2.07	2.05	2.07	2.21	2.12	13.77	0.61	0.87	0.35
PM6 BFGS	2.04	1.99	2.01	2.03	2.04	2.07	2.03	14.51	0.91	1.01	0.43
PM6 BFGS pr	2.13	2.06	2.10	2.09	2.20	2.37	2.16	14.79	0.97	0.71	0.31
PM6 L-BFGS	2.32	2.70	2.10	2.15	2.16	2.10	2.26	14.73	-0.44	0.98	0.40
PM6 L- BFGS pr	2.36	2.06	2.07	2.08	2.09	2.12	2.13	14.28	0.23	0.96	0.33
PM6 D2 DFP	1.90	1.97	2.01	1.97	2.02	2.17	2.01	14.24	0.70	0.91	0.46
PM6 D2 DFP pr	2.04	2.01	2.06	2.04	2.13	2.25	2.09	14.61	0.92	0.83	0.38
PM6 D2 EF	2.13	1.99	2.16	2.11	2.02	2.06	2.08	14.39	-0.20	1.02	0.39
PM6 D2 EF pr	2.28	1.99	2.04	2.03	1.96	2.72	2.17	12.94	0.79	0.72	0.29
PM6 D2 EF HESS1	2.12	2.00	2.17	2.10	2.00	2.09	2.08	14.26	-0.12	0.99	0.39
PM6 D2 EF pr HS1	2.36	1.99	2.01	2.37	1.95	2.74	2.24	12.97	0.65	0.73	0.27
PM6 D2 BFGS	1.98	1.98	2.07	2.03	2.11	2.42	2.10	14.41	0.84	0.66	0.37
PM6 D2 BFGS pr	2.04	2.00	2.05	2.01	2.10	2.41	2.10	14.55	0.89	0.67	0.36
PM6 D2 L-BFGS	2.17	1.99	2.18	2.10	2.02	2.06	2.09	14.32	-0.19	1.02	0.39
D2 L-BFGS pr	2.20	2.03	2.06	2.04	2.02	2.16	2.09	13.83	0.53	0.92	0.38
PM6 DH2 DFP	1.87	1.95	1.98	1.91	1.97	2.09	1.96	14.19	0.62	0.99	0.50
PM6 DH2 DFP pr	2.01	2.01	2.09	2.05	2.09	2.17	2.07	14.58	0.72	0.91	0.39
PM6 DH2 EF	2.09	2.02	2.33	2.18	2.00	2.03	2.11	14.53	-0.55	1.05	0.42
PM6 EF pr	2.11	2.02	2.32	2.22	2.00	2.04	2.12	14.55	-0.54	1.04	0.40
PM6 EF HESS1	2.11	1.99	2.29	2.18	2.03	2.03	2.11	14.47	-0.47	1.05	0.41
PM6 EF HESS1 pr	2.13	1.99	2.25	2.23	2.03	1.99	2.10	14.49	-0.56	1.09	0.39
PM6 DH2 BFGS	1.89	2.03	2.12	2.00	2.11	2.28	2.07	14.26	0.59	0.80	0.39
PM6 DH2 BFGS pr	1.90	2.04	2.12	2.00	2.12	2.32	2.08	14.27	0.63	0.76	0.38
PM6 DH2 L-BFGS	2.11	2.01	2.32	2.22	2.00	2.04	2.12	14.49	-0.52	1.04	0.40
PM6 DH2 L-BFGS pr	2.13	2.03	2.20	2.12	2.04	2.03	2.09	14.18	-0.52	1.05	0.39

On optimizing the Gly_{10} neutral structure in water (Table 3.) the clear winner is the L-BFGS method, producing the best energies and residual gradients in all of the cases. The DFP optimizer is generally still the worst, producing poor energies and great residual gradients, but for the PM6-D2 Hamiltonian equals the other methods (in this case all the energies lie in a 2-3 kcal/mol interval). The use of precise optimization criteria does not influence severely the results, except for the PM6-DH2 Hamiltonian, where the energies are rising and the residual gradients are greater in this case. This can be the effect of the known limitation of the gradient calculation in current implementation:

the optimized geometry is not the exact minimum of PM6-DH2 energy. This can be ignored for weak H-bonds, but for many strong H-bonds as in a peptide may be a serious problem.

Table 3. Heats of formation and residual gradients for the Gly_{10} neutral form computed in water. See also the Methods section

Model Hamiltonian	Optimization method	Normal calculation		Precise calculation		Δ energy
		Energy	Gradient	Energy	Gradient	
PM6	DFP	-535.427	50.61209	-535.734	61.14301	0.30768
PM6	EF	-540.011	6.54133	-540.014	13.14678	0.00287
PM6	EF-HESS1	-541.551	2.24855	-541.631	3.81095	0.08057
PM6	BFGS	-535.405	13.40674	-541.603	15.13316	6.19708
PM6	L-BFGS	-544.179	2.74499	-543.992	1.19717	-0.18779
PM6-D2	DFP	-559.642	8.42636	-560.376	8.87667	0.73325
PM6-D2	EF	-561.384	3.31497	-559.319	6.49577	-2.06447
PM6-D2	EF-HESS1	-560.526	5.35116	-561.698	3.35221	1.17194
PM6-D2	BFGS	-560.720	11.71167	-561.088	22.12513	0.36771
PM6-D2	L-BFGS	-563.167	2.08386	-562.861	1.25650	-0.30570
PM6-DH2	DFP	-565.760	14.20699	-565.226	27.70459	-0.53395
PM6-DH2	EF	-565.534	4.56001	-561.237	7.24253	-4.29728
PM6-DH2	EF-HESS1	-567.081	5.46998	-566.310	12.93433	-0.77113
PM6-DH2	BFGS	-558.271	32.45811	-565.573	32.06091	7.30236
PM6-DH2	L-BFGS	-574.409	1.51347	-570.567	2.23486	-3.84207

Table 4. CO...NH hydrogen bond lengths and helix lengths for the Gly_{10} neutral model computed in water. The order is from the N-terminal ending of the polypeptide chain. See also the Methods section for details.

Neutral water	1-5	2-6	3-7	4-8	5-9	6-10	CO…NH avg	Helix length	Correl	Max Dif	Avg Dif
Reference	2.12	1.97	2.02	1.99	2.11	2.26	2.08	14.61	-	-	-
PM6 DFP	1.82	1.91	1.95	1.92	1.87	1.96	1.91	14.10	0.03	0.30	0.17
PM6 DFP pr	1.84	2.00	2.03	1.96	1.90	2.00	1.96	14.13	-0.21	0.28	0.14
PM6 Ef	1.90	1.95	1.98	1.96	1.98	2.05	1.97	14.30	0.54	0.22	0.11
PM6 EF pr	1.89	1.94	1.98	1.96	1.98	2.04	1.97	14.29	0.49	0.23	0.11
PM6 EF HESS1	1.98	1.94	2.00	1.96	2.01	2.04	1.99	14.50	0.86	0.22	0.09
PM6 EF HESS1 pr	1.94	1.96	1.97	1.96	2.01	2.09	1.99	14.46	0.80	0.18	0.09
PM6 BFGS	1.90	1.86	1.89	1.93	1.91	1.98	1.91	14.22	0.78	0.28	0.17
PM6 BFGS pr	1.98	1.97	2.01	2.00	2.02	2.15	2.02	14.45	0.84	0.14	0.06
PM6 L-BFGS	2.09	1.98	1.98	2.04	2.00	2.07	2.03	14.71	0.64	0.19	0.07
PM6 L-BFGS pr	2.03	1.94	1.98	1.99	1.96	1.96	1.98	14.30	0.07	0.30	0.10
PM6 D2 DFP	1.82	1.90	1.97	1.92	1.89	2.00	1.92	14.10	0.25	0.30	0.16
PM6 D2 DFP pr	1.84	1.90	1.94	1.94	1.93	2.00	1.93	14.13	0.38	0.28	0.15
PM6 D2 EF	1.90	1.93	1.97	1.95	1.96	2.04	1.96	14.30	0.61	0.22	0.12

Neutral water	1-5	2-6	3-7	4-8	5-9	6-10	CO…NH avg	Helix length	Correl	Max Dif	Avg Dif
PM6 D2 EF pr	1.87	1.90	1.93	1.92	1.93	1.98	1.92	14.23	0.54	0.28	0.16
PM6 D2 EF HESS1	1.90	1.90	1.95	1.93	1.95	1.97	1.93	14.37	0.56	0.29	0.15
PM6 D2 EF HS1 pr	1.92	1.93	1.95	1.93	1.97	2.04	1.96	14.37	0.81	0.22	0.12
PM6 D2 BFGS	1.83	1.93	2.00	1.96	1.93	2.03	1.95	14.14	0.19	0.29	0.13
PM6 D2 BFGS pr	1.85	1.92	1.98	1.96	1.94	2.03	1.95	14.16	0.33	0.27	0.13
PM6 D2 L-BFGS	1.99	1.92	1.94	1.93	1.99	2.12	1.98	14.49	0.98	0.14	0.10
PM6 D2 L-BFGS pr	1.97	1.91	1.94	1.93	2.00	2.15	1.98	14.50	0.96	0.15	0.10
PM6 DH2 DFP	1.82	1.86	1.99	1.92	1.90	1.97	1.91	14.16	0.19	0.30	0.17
PM6 DH2 DFP pr	1.81	1.85	1.95	1.87	1.88	1.97	1.89	14.14	0.41	0.31	0.19
PM6 DH2 EF	1.86	1.88	1.91	1.90	1.90	1.93	1.90	14.27	0.41	0.33	0.18
PM6 DH2 EF pr	1.82	1.82	1.84	1.84	1.84	1.86	1.84	14.15	0.61	0.40	0.24
PM6 DH2 EF HESS1	1.94	1.89	1.98	1.93	1.94	1.97	1.94	14.51	0.55	0.29	0.14
PM6 DH2 EF pr HS1	1.88	1.86	1.95	1.90	1.91	1.93	1.91	14.40	0.36	0.33	0.17
PM6 DH2 BFGS	1.76	1.81	1.86	1.82	1.81	1.85	1.82	14.02	0.07	0.41	0.26
PM6 DH2 BFGS pr	1.84	1.80	1.98	1.93	1.88	1.96	1.90	14.12	0.30	0.30	0.18
PM6 DH2 L-BFGS	1.96	1.96	1.95	1.98	2.03	1.84	1.95	13.77	-0.61	0.42	0.13
PM6 DH2 L-BFGS pr	1.93	1.91	1.99	2.03	1.98	1.97	1.97	14.55	-0.06	0.29	0.12

When passing to the neutral Gly_{10} computed in water we notice that the correlations of the values as compared to the DFT computation are even better for the differences but worse for the correlation coefficient. The average of maximum difference for all methods is 0.27Å with a maximum of 0.42, and once again we notice the worst results for PM6-DH2 class. The two PM6-D2 L-BFGS variants stand the most precise of all, the best one with a correlation coefficient of 0.98, maximum difference of 0.14Å and average difference of 0.1Å.

Gly_{10} zwitterionic structure optimized in vacuum (Table 5.) is a highly unrealistic situation. Albeit the DFP optimizer still gives the highest energies and greatest residual gradients. The EF variants and the L-BFGS are producing better energies and gradients – but sometimes locating different minima on the potential energy surface, not the initial α-helix. As the geometries are changing, the difference between normal and precise calculations is great.

Table 5. Heats of formation and residual gradients for the Gly_{10} zwitterionic form computed in vacuum. See also the Methods section for details

Model Hamiltonian	Optimization method	Normal calculation		Precise calculation		Δ energy
		Energy	Gradient	Energy	Gradient	
PM6	BFGS	-328.555	16.25501	-478.472	2.10161	149.9168
PM6	DFP	-320.047	31.29806	-332.729	46.93261	12.68219
PM6	EF	-460.921	1.61288	-505.141	1.06650	44.22029
PM6	EF-HESS1	-471.336	1.71480	-460.808	1.34189	-10.52790
PM6	L-BFGS	-464.872	1.57470	-491.120	1.97629	26.24755
PM6-D2	BFGS	-371.332	22.11510	-493.162	5.88488	121.82980
PM6-D2	DFP	-362.805	12.16014	-463.076	12.07646	100.27150

Model Hamiltonian	Optimization method	Normal calculation		Precise calculation		Δ energy
		Energy	Gradient	Energy	Gradient	
PM6-D2	EF	-466.328	2.08065	-504.396	1.04559	38.06803
PM6-D2	EF-HESS1	-484.981	1.68821	-526.832	1.07549	41.85040
PM6-D2	L-BFGS	-455.669	1.61520	-507.447	1.20085	51.77824
PM6-DH2	BFGS	-382.392	9.40206	-383.229	5.57377	0.83670
PM6-DH2	DFP	-366.183	23.50733	-378.312	12.72887	12.12884
PM6-DH2	EF	-465.014	3.94958	-373.395	4.83304	-91.61840
PM6-DH2	EF-HESS1	-398.738	5.90495	-429.697	7.28087	30.95856
PM6-DH2	L-BFGS	-478.451	1.92075	-521.345	1.29656	42.89381

When performing the computations on the zwitterionic model in vacuum most of the methods fail so that table 6 presents only those methods that succeeded in maintaining the alpha helix. This is not surprising as for the reference density functional calculation in these conditions the alpha-helix is not stable. Even for the neutral form computed in vacuum the variation of helix length among all the options tested shows the imperfections of such approach. For all the options tested with a solvent model the geometry is more regular and the maximum deviation of helix length is less than 0.6Å.

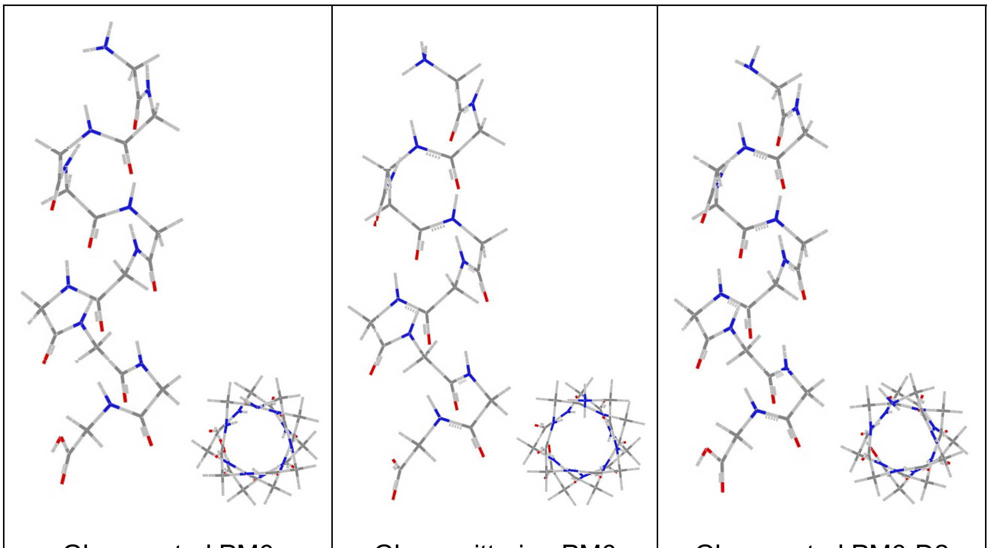
Table 6. CO...NH hydrogen bond lengths and helix lengths for the Gly_{10} zwitterion model computed in vacuum. The order is from the N-terminal ending of the polypeptide chain

Zwitterion vacuum	1-5	2-6	3-7	4-8	5-9	6-10	$(CO…NH)_{avg}$	Helix length
DFP	1.94	1.98	1.98	2.06	2.11	2.43	2.08	14.19
DFP pr	2.09	1.97	2.35	2.39	2.31	2.71	2.30	14.37
BFGS	2.02	1.96	2.14	2.12	2.23	2.75	2.20	14.32
D2 DFP	2.25	2.03	2.23	2.39	2.55	3.18	2.44	14.66
D2 BFGS	2.73	2.22	3.22	3.09	3.85	5.07	3.36	15.12
DH2 DFP	2.25	2.05	2.21	2.36	2.57	3.42	2.48	14.73
DH2 DFP pr	2.68	2.12	2.14	2.83	3.33	3.69	2.80	15.33
DH2 EF pr	2.55	2.15	2.18	2.59	3.22	3.65	2.72	15.32
DH2 BFGS	2.81	2.14	2.99	2.82	3.60	5.31	3.28	15.49
DH2 BFGS pr	2.76	2.21	3.59	2.69	3.92	4.70	3.31	15.55

As in case of optimizing Gly_{10} neutral structure in water, in case of the zwitterionic form in water (Table 7.), the L-BFGS optimizer produces the best results: lowest energies (in all cases) and smallest residual gradients (excepting one case). Surprisingly in this table, the standard EF method gives the worst results (excepting one case), and the DFP optimizer performs better, but still producing great residual gradients. The EF-HESS1 variant and the BFGS methods are producing intermediate results disputing the second place, with the latter giving generally greater residual gradients. The differences between normal and precise optimization criteria calculations are small - a few kcal/mol.

Table 7. Heats of formation and residual gradients for the Gly_{10} zwitterionic form computed in water. See also the Methods section

Model Hamiltonian	Optimization method	Normal calculation		Precise calculation		
		Energy	Gradient	Energy	Gradient	Δ energy
PM6	BFGS	-554.153	6.84385	-551.934	30.19219	-2.21933
PM6	DFP	-547.085	39.02397	-551.908	15.07784	4.82284
PM6	EF	-539.728	11.73500	-539.541	11.36281	-0.18665
PM6	EF-HESS1	-551.387	11.60122	-552.919	10.75917	1.53230
PM6	L-BFGS	-557.027	2.26565	-557.941	2.01639	0.91368
PM6-D2	BFGS	-574.565	8.82374	-573.498	27.91870	-1.06748
PM6-D2	DFP	-571.552	12.20055	-566.640	27.55250	-4.91186
PM6-D2	EF	-561.899	15.35613	-561.669	11.35523	-0.22992
PM6-D2	EF-HESS1	-571.262	5.42663	-572.221	4.37181	0.95881
PM6-D2	L-BFGS	-578.515	2.10795	-577.025	3.96587	-1.48974
PM6-DH2	BFGS	-572.623	20.20676	-577.505	28.83444	4.88212
PM6-DH2	DFP	-575.463	16.35665	-574.727	29.94503	-0.73570
PM6-DH2	EF	-577.901	3.70453	-574.184	5.51010	-3.71653
PM6-DH2	EF-HESS1	-578.140	6.15222	-579.308	2.57518	1.16867
PM6-DH2	L-BFGS	-582.775	1.43065	-582.618	8.31848	-0.15647



Gly_{10}-neutral PM6 BFGS in vacuum

Gly_{10}-zwitterion PM6 BFGS in water

Gly_{10}-neutral PM6-D2 L-BFGS in water

Figure1. Optimized structures of the best correlated optimized structures.

Table 8. CO...NH hydrogen bond lengths and helix lengths for the Gly_{10} zwiterion model computed in water. The order is from the N-terminal ending of the polypeptide chain. See also the Methods section

Zwitterion water	1-5	2-6	3-7	4-8	5-9	6-10	CO…NH avg	Helix length	Correl	Max Dif	Avg Dif
Reference	2.25	2.04	2.02	2.01	2.28	2.65	2.21	14.85	-	-	-
PM6 DFP	1.87	1.87	1.99	1.89	1.97	1.98	1.93	14.15	0.41	0.67	0.28
PM6 DFP-pr	1.92	1.91	1.96	1.96	1.99	2.11	1.98	14.26	0.84	0.54	0.23
PM6 ef	1.84	1.83	1.84	1.85	1.85	1.89	1.85	14.12	0.87	0.76	0.36
PM6 EF pr	1.83	1.83	1.84	1.84	1.84	1.88	1.84	14.11	0.82	0.77	0.37
PM6 EF HESS1	1.97	1.91	1.98	1.98	1.95	2.08	1.98	14.42	0.77	0.57	0.23
PM6 EF pr HS1	1.99	1.92	1.96	1.97	1.98	2.06	1.98	14.40	0.91	0.59	0.23
PM6 BFGS	2.04	1.90	1.94	1.97	2.05	2.11	2.00	14.41	0.91	0.54	0.21
PM6 BFGS pr	1.94	1.91	1.98	1.97	1.99	2.08	1.98	14.28	0.80	0.57	0.23
PM6 L-BFGS	2.15	2.00	2.00	2.09	2.00	2.07	2.05	14.60	0.25	0.58	0.18
PM6 L-BFGS pr	2.12	1.92	2.05	2.04	1.95	3.04	2.19	12.06	0.87	0.39	0.17
PM6 D2 DFP	1.87	1.89	1.96	1.94	1.95	2.05	1.94	14.19	0.65	0.60	0.27
PM6 D2 DFP pr	1.84	1.89	1.94	1.90	1.86	1.94	1.90	14.09	0.13	0.71	0.31
PM6 D2 EF	1.83	1.83	1.84	1.85	1.84	1.89	1.85	14.11	0.77	0.76	0.36
PM6 D2 EF pr	1.83	1.82	1.84	1.84	1.84	1.88	1.84	14.11	0.82	0.77	0.37
PM6 D2 EF HESS1	1.95	1.90	1.97	1.96	1.92	2.08	1.96	14.40	0.75	0.57	0.25
PM6 D2 EF HS1 pr	1.95	1.93	1.91	1.92	1.96	2.10	1.96	14.36	0.97	0.55	0.25
PM6 D2 BFGS	2.03	1.86	1.90	1.96	2.03	2.07	1.98	14.35	0.83	0.58	0.23
PM6 D2 BFGS pr	1.95	1.89	1.96	1.97	2.00	2.09	1.98	14.29	0.85	0.56	0.23
PM6 D2 L-BFGS	2.02	1.95	2.01	2.03	1.90	1.96	1.98	14.20	-0.38	0.69	0.24
PM6 D2 L-BFGS pr	2.06	1.94	1.95	2.00	1.98	2.03	1.99	14.44	0.59	0.62	0.22
PM6 DH2 DFP	1.86	1.81	1.95	1.89	1.88	2.01	1.90	14.18	0.61	0.64	0.31
PM6 DH2 DFP pr	1.85	1.83	1.93	1.88	1.87	1.98	1.89	14.15	0.60	0.67	0.32
PM6 DH2 EF	1.95	1.90	1.92	1.92	1.93	2.03	1.94	14.40	0.94	0.62	0.27
PM6 DH2 EF pr	1.88	1.86	1.87	1.88	1.88	1.94	1.89	14.24	0.91	0.71	0.32
PM6 DH2 EF h1	1.98	1.90	1.94	1.92	1.94	2.02	1.95	14.50	0.89	0.63	0.26
PM6 DH2 EF pr h1	2.02	1.96	1.93	1.92	1.99	2.03	1.98	14.58	0.87	0.62	0.23
PM6 DH2 BFGS	1.92	1.83	1.83	1.88	1.86	1.98	1.88	14.20	0.87	0.67	0.33
PM6 DH2 BFGS pr	1.92	1.84	1.96	1.93	1.94	2.00	1.93	14.23	0.62	0.65	0.28
PM6 DH2 L-BFGS	2.08	1.90	1.97	2.05	2.02	1.98	2.00	14.66	0.11	0.67	0.22
PM6 DH2 L-BFGS pr	2.09	1.90	1.97	2.06	2.02	1.99	2.01	14.68	0.14	0.66	0.22

Table 8 summarizes the hydrogen bond lengths for the zwitterion species computed in water. We already notice that the precision as compared to the reference is better, the maximum hydrogen bond differences do not exceed 0.8Å with an average of just 0.6 while the average lies at 0.2 (never less than 0.21 or more than 0.37). The correlation coefficient would then be the only mean of discerning among methods although just correlating the computed values may not be a perfect criterion.

CONCLUSIONS

A complete study of PM6 semiempirical optimization of alpha-Gly_{10} has been achieved in this paper. All the optimization methods available in Mopac2009 have been tested in all the possible situations (neutral and zwitterionic model in vacuum and in water). Most of the methods fail to model amfionic form in vacuum while the neutral form results are still bad. For the calculations in water the neutral model gives correlation values better than the amfionic form. Classical PM6 approach results are comparable with the PM6-D2 class while PM6-DH2 methods are not performing well in this case. The PM6-D2 Hamiltonian and CPCM continuous solvent model (water) using L-BFGS optimization method with standard convergence criteria gives the best correlation as compared to a high level DFT optimization of the same structure.

METHODS - COMPUTATIONAL OPTIONS

The alpha-helical model of Gly_{10} was built with the Spartan [5] program using the standard torsion angles. The model was either neutral (i.e. with the N- and C- terminal endings capped with hydrogen atoms to obtain $–NH_2$ or -COOH) or maintained as zwitterion ($-COO^-$ and $–NH_3^+$ respectively). The characteristic of an alpha-helix is the regularity imposed by the hydrogen bonds formed between CO and NH groups of different peptide bonds. Tables throughout manuscript show hydrogen bond lengths ordered starting from the N-terminal end. Geometry optimizations were performed either in vacuum or with a continuum solvent model (water). A high level computation (M062X/6-31G** [6] performed with Gaussian 09 [7]) is taken as reference geometry for statistical comparisons.

Tables depict for each optimization method the energy and the gradient for the normal and the precise calculation and the energy difference between these two is presented in the last column. All the CO...NH bond leghts are also presented as tables followed by detailed statistical analysis including:

- the average value of the six CO..NH bonds determining the alpha-helix.

- the length of the helix (measured each time between the C atom neighbouring the N-terminal end and the C atom neighbour to the C-terminal end).

- the correlation coefficient between the computed values and the reference (dft calculation) values computed by the statistical formula

$$Correl(x_{ref}, x_{calc}) = \frac{\sum (x_{ref} - \bar{x}_{ref})(x_{calc} - \bar{x}_{calc})}{\sqrt{\sum (x_{ref} - \bar{x}_{ref})^2 \sum (x_{calc} - \bar{x}_{calc})^2}}$$

- the maximum difference, i.e. the value representing the maximal difference between each computed alpha-helical H bond and the corresponding H bond of the reference structure.

- the average difference, i.e. the difference between the average alpha-helical H bond values of each optimization and the average of the corresponding refference structure.

Mopac2009 has implemented many methods, optimization algorithms and optional facilities designed to deal with a variety of systems.

Beside the basic PM6 method the Mopac2009 program has also two improved versions of this model Hamiltonian: PM6-D2 (augmented for dispersion interactions) and PM6-DH2 (augmented for dispersion and H-bond interactions) [8,9] which are parameterized to reproduce interaction energies for geometries obtained from high-level quantum mechanical calculations.

Five different geometry optimization algorithm variants present in Mopac2009 were tested. All geometry optimizations where done in Cartesian coordinates, as the system involves several big rings formed by hydrogen bonds, as using internal coordinates is often difficult in such cases. All energies (heat of formation, kcal/mol) and gradients (kcal/mol.Å) were recalculated in final single-point runs to prevent their alteration due to degradation of the wave function during optimization.

The default geometry optimizer in Mopac2009 is Baker's eigenvector following method [10]. Its performance was tested using the default starting Hess matrix (EF) and a calculated Hessian for the starting geometry (EF-HESS1). The full Hess matrix is constructed using single-sided derivatives, thus the calculation may be computationally prohibitive.

The Davidon–Fletcher–Powell (DFP) algorithm [11,12] was the first quasi-Newton generalized secant method implemented in Mopac. This method is superseded now by the Broyden-Fletcher-Goldfarb-Shanno (BFGS) procedure [13-16].

An advanced variant of the BFGS optimizer is the "Limited memory BFGS" function minimizer, which does not store the inverse Hessian, but instead calculates it as needed [17-19]. Because of this, the L-BFGS method uses very little storage, and is therefore suitable for optimizing very large systems. It is not as efficient as the other optimizers, but for large systems, it is often the only method that can be used. The Mopac2009 implementation (L-BFGS) is really the L-BFGS-B method, a variant which can handle box-constraints on the variables [20].

Different sets of calculations with normal and increased (PRECISE) convergence criteria were performed to estimate the usefulness of this computationally more demanding variant.

The structures were modeled in vacuum and in water, using the conductor-like screening model [21] as implemented in the program, using the dielectric constant of water 78.39 and solvent radius of 1.3 Å.

ACKNOWLEDGEMENTS

This work is supported by the Romanian Ministry for Education and Research (grant PCCE 312/2008).

The theoretical calculations are carried out using the resources of the Center for Molecular Modelling and Computational Quantum Chemistry acquired through a "Capacitati" program (Romanian Ministry for Education and Research, 130/14.09.2007).

REFERENCES

1. J.J.P. Stewart, *J. Mol. Model.*, **2009**, *15*, 765.
2. J.J.P. Stewart, *J. Mol. Model.*, **2007**, *13*, 1173.
3. J.J.P. Stewart, *Int. J. Quant. Chem.*, **1996**, *58*, 133.
4. Mopac2009, Version 10.153L, J. J. P. Stewart, Stewart Computational Chemistry, www.openmopac.net.
5. Spartan `06 Wavefunction Inc. 18401 Von Karman Avenue, Suite 370 Irvine, CA 92612 www.wavefun.com.
6. Y. Zhao, D.G. Truhlar, *Theor. Chem. Acc.*, **2008**, *120*, 215.
7. Gaussian 09, Revision A.02, Frisch, M.J.,Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Montgomery, J., J.A., Vreven, T., Kudin, K.N., Burant, J.C., Millam, J.M., Iyengar, S.S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G.A., 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, J.E., Hratchian, H.P., Cross, J.B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R.E., Yazyev, O., Austin, A.J., Cammi, R., Pomelli, C., Ochterski, J.W., Ayala, P.Y., Morokuma, K., Voth, G.A., Salvador, P., Dannenberg, J.J., Zakrzewski, V.G., Dapprich, S., Daniels, A.D., Strain, M.C., Farkas, O., Malick, D.K., Rabuck, A.D., Raghavachari, K., Foresman, J.B., Ortiz, J.V., Cui, Q., Baboul, A.G., Clifford, S., Cioslowski, J., Stefanov, B.B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R.L., Fox, D.J.,Keith, T., Al-Laham, M.A., Peng, C.Y., Nanayakkara, A., Challacombe, M., Gill, P.M.W., Johnson, B., Chen, W., Wong, M.W., Gonzalez, C., Pople, J.A., Gaussian, Inc., Wallingford CT, **2009**, www.gaussian.com.
8. M. Korth, M. Pitonák, J. Rezác, P. Hobza, *J. Chem. Theory Comp.*, **2010**, *6,* 344.
9. J. Rezác, J. Fanfrlik, D. Salahub, P. Hobza, *J. Chem. Theory Comp.*, **2009**, *5*, 1749.
10. J. Baker, *J. Comp. Chem.*, **1986**, *7*, 385.
11. W.C. Davidon, *SIAM J. Optimiz.*, **1991**, *1*, 1.
12. R. Fletcher, “Practical methods of optimization” (2nd ed.), John Wiley & Sons, New York, **1987**.
13. C.G. Broyden, *J. Inst. Math*. Appl., **1970**, *6*, 222.
14. R. Fletcher, *Comp. J.*, **1970**, *13*, 317
15. D. Goldfarb, *Math. Comput.*, **1970**, *24*, 23.
16. D.F. Shanno, *Math. Comput.*, **1970**, *24*, 647.
17. H. Matthies, G. Strang, *Int. J. Num. Meth. Eng*., **1979**, *14*, 1613.
18. J. Nocedal, *Math. Comput.*, **1980**, *35*, 773.
19. R.H. Byrd, J. Nocedal, R.B. Schnabel, *Math. Prog.*, **1994**, *63*, *4*, 129.
20. C. Zhu, R.H. Byrd, J. Nocedal, *ACM Trans. Math. Soft.*, **1997**, *23*, 550.
21. A. Klamt, G. Schüümann, *J. Chem. Soc. Perkin Trans. 2*, **1993**, 799.