Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.
This journal is © the Owner Societies 2023

Molecular Dynamics and Raman Optical Activity Spectra Reveal Nucleotide Conformation Ratios in Solution

Věra Schrenková, Mohammed Siddhique Para Kkadan, Jiří Kessler, Josef Kapitán, and Petr Bouř

Contents

Table S1. Examples of different parameters used in the GAFF, GAFF2, RNA.OL3 and DNA.Bsc1 force fields.[a]

	GAFF	GAFF2	RNA.OL3	DNA.Bsc1
Equilibrium Bond Lengths (Å)				
P-O	1.487	1.487	1.480	1.480
O3'-H	0.973	0.973	0.960	0.960
C4-C5	1.373	1.373	1.370	1.370
Bond Force Constants (kcal mol^{-1} Å^{-2})				
P-O	479.5	529.5	525.0	525.0
O3'-H	371.4	563.5	553.0	553.0
C4-C5	500.9	416.1	520.0	520.0
Equilibrium Bond Angles (degrees)				
O5'-C5'-C4'	108.0	108.0	109.5	109.5
C4'-C3'-O3'	110.2	110.2	109.5	109.5
O4'-C1'-N9	109.0	109.0	109.5	109.5
Bond Angle Force Constant (kcal mol^{-1} rad^{-2})				
O5'-C5'-C4'	68.9	85.3	50.0	50.0
C4'-C3'-O3'	67.5	84.6	50.0	50.0
O4'-C1'-N9	71.3	109.3	50.0	50.0
Dihedral Angles[b] B (kcal/mol) / φ (degree) / n				
P-O5'-C5'-C4'	0.383 / 0 / 3 3.950 / 180 / 1	0.383 / 0 / 3 3.950 / 180 / 1	0.383 / 0 / 3 3.950 / 180 / 1	0.383 / 0 / 3 3.950 / 180 / 1
O4'-C1'-C2'-O2'	0.144 / 0 / 3 1.175 / 0 / 2	1.010 / 0 / 3 0.000 / 0 / 2 0.020 / 180 / 1	0.144 / 0 / 3 1.175 / 0 / 2	0.144 / 0 / 3 1.175 / 0 / 2
O4'-C1'-N9-C8	-	-	0.00 / 0 / 2 2.50 / 0 / 1	0.44 / 210/ 3 1.73 / 4 / 2 0.81 / 88 / 1

[a] The general amber force field (GAFF)[1] was later updated to the second-generation GAFF (GAFF2),[2] e.g. using quantum mechanics (QM) calculations on more model compounds. RNA.OL3[3] was developed from the ff99 force field[4] highly exploiting QM calculations for sugars and phosphates, also with refinements of the glycosidic torsion, DNA.Bsc1[5] development similarly includes QM and solvation models. Technically, DNA.Bsc1 largely consists of RNA.OL3 completed by other parameters. In out calculations, partial atomic charges as obtained from a QM computations were the same for all the four force fields.

[b] Energy ~ $B[1+\cos(n\tau-\varphi)]$, where B is the barrier and φ is the phase.

1. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157-1174.
2. W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, 1995, **117**, 5179-5197.
3. M. Zgarbová, M. Otyepka, J. Šponer, A. Mládek, P. Banáš, T. E. Cheatham, III and P. Jurečka, *J. Chem. Theory Comput.*, 2011, **7**, 2886-2902.
4. A. Spasic, J. Serafini and D. H. Mathews, *J. Chem. Theory Comput.*, 2012, **8**, 2497-2505.
5. I. Ivani, P. D. Dans, A. Noy, A. Pérez, I. Faustino, A. Hospital, J. Walther, P. Andrio, R. Goñi, A. Balaceanu, et al., *Nat. Methods*, 2016, **13**, 55-58.

Table S2. Relative energies (*E*, kcal/mol), populations (*p*, %) and geometries of the lowest-energy minima (<2.5 kcal/mol) of the four nucleotides as obtained by WHAM.

Conformer		*E*	*p*	γ	χ	χ ,γ integration range
dTMP						
anti / g-	I	2.1	8	-57	-147	$\gamma \in$ (-120,80)
anti / t	II	0	92	177	-149	$\gamma \notin$ (-120,80)
rCMP						
anti / g$^-$	I	0	58	-57	179	$\chi \notin$ (-10,120), *anti*
syn / g$^-$	III	0.5	42	-57	63	$\chi \in$ (-10,120), *syn*
rGMP						
anti / g$^-$	I	0	68	-61	-177	$\gamma \in$ (-105,0), $\chi \notin$ (-90,90)
anti / t	II	0.2	24	-177	179	$\gamma \notin$ (-105,0), $\chi \notin$ (-90,90)
syn / g$^-$	III	1.3	6	-59	61	$\gamma \in$ (-105,0), $\chi \in$ (-90,90)
syn / t	IV	2.0	2	-177	63	$\gamma \notin$ (-105,0), $\chi \in$ (-90,90)
rAMP						
anti / t	I	0	79	-179	179	$\gamma \in$ (-105,0), $\chi \notin$ (-90,90)
anti / g$^-$	II	1.3	16	-65	179	$\gamma \notin$ (-105,0), $\chi \notin$ (-90,90)
syn / t	IV	2.0	5	179	67	$\chi \in$ (-90,90)

Table S3. Experimental and raw[a] computed frequencies (cm^{-1}), assignment of selected Raman bands.

dTMP	ν_{exp}	ν_{sim}	assignment
Raman			
1	498	491	sugar and base deformation
2	751-791	765-798	base out of plane deformation
3	979	1006	*ν*(P=O)
4	1205	1212	*ν*(C–N), *ν*(C–C), *δ*(C'–H)
5	1241	1274	base def., *δ*(CH)
6	1377	1397	*δ*(CH)
7	1663	1699	*ν*(C=O), *ν*(C=N), *ν*(C=C)
ROA			
I	341	372	
II	751-793	762-799	
III	1093	1094	
IV	1208	1233	
V	1278	1307	
VI	1316	1354	
VII	1367	1395	
VIII	1448	1486	
IX	1657	1697	

rCMP	ν_{exp}	ν_{sim}	assignment
Raman			
1	599	603	sugar and base deformation
2	784	797	base out of plane deformation
3	873	888	sugar and base deformation
4	980	1006	*ν*(P=O)
5	1243	1233	base def., δ(CH)
6	1295	1325	*δ*(CH)
7	1530	1573	*ν*(C–C), *ν*(C–N)
8	1608	1639	*δ*(NH_2)
9	1656	1689	*ν*(C=O), *ν*(C=C), *δ*(NH_2)
ROA			
I	287	270	
II	720	727	
III	788	796	
IV	1113	1143	
V	1246	1229	
VI	1352	1365	
VII	1412	1445	
VIII	1532	1574	

Table S2 (cont.)

rGMP	ν_{exp}	ν_{sim}	assignment
Raman			
1	502	504	sugar and base deformation
2	673	691	sugar and base deformation
3	979	1007	*ν*(P=O)
4	1179	1212	sugar and base deformation
5	1324	1367	*δ*(CH)
6	1367	1407	*δ*(CH)
7	1417	1451	*ν*(C=N), *ν*(C=C), *δ*(CH)
8	1488	1536	*ν*(C=C), *ν*(C=N), *δ*(NH)
9	1577	1617	*ν*(C=O), *ν*(C=C), *ν*(C=N), *δ*(NH)
ROA			
I	368	387	
II	588	606	
III	868	886	
IV	1129	1086	
V	1179	1118	
VI	1215	1243	
VII	1320	1358	
VIII	1358	1394	
IX	1425	–	
X	–	1495	

rAMP	ν_{exp}	ν_{sim}	assignment
Raman			
1	729	749	base out of plane deformation
2	979	1007	*ν*(P=O)
3	1308	1341	base def., *δ*(CH)
4	1339	1389	*δ*(CH)
5	1379	1417	*ν*(C=C), *ν*(C=N), *δ*(CH)
6	1484	1525	*δ*(CH), *δ*(NH)
7	1510	1550	*ν*(C=C), *ν*(C=N), *δ*(CH)
8	1581	1629	*ν*(C=C), *ν*(C=N), *δ*(NH)
ROA			
I	244	264	
II	561	586	
III	886	881	
IV	1241	1242	
V	1311	1343	
VI	1350	1414	
VII	1510	1549	

[a] Only the P=O force constant was scaled (see ref. 54 in the main text), so that the PO stretching had similar systematic frequency error as other vibrations, and resultant spectrum could be more easily compared to the experiment, e.g. after one-step scaling of all frequencies.

ν … stretching, *δ* … bending.

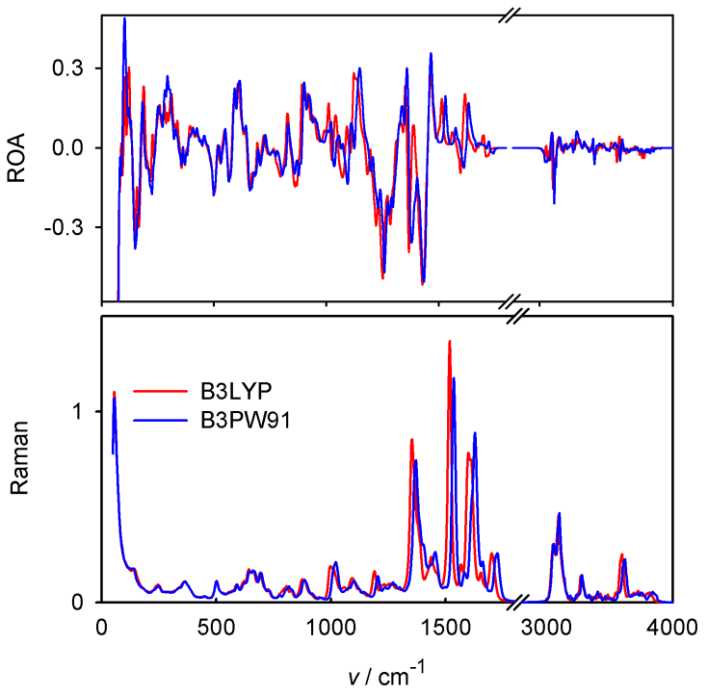

Figure S1. ROA and Raman spectra of a rGMP MD snapshot calculated with the B3LYP and B3PW91 functionals (for 6-31++G** and CPCM).

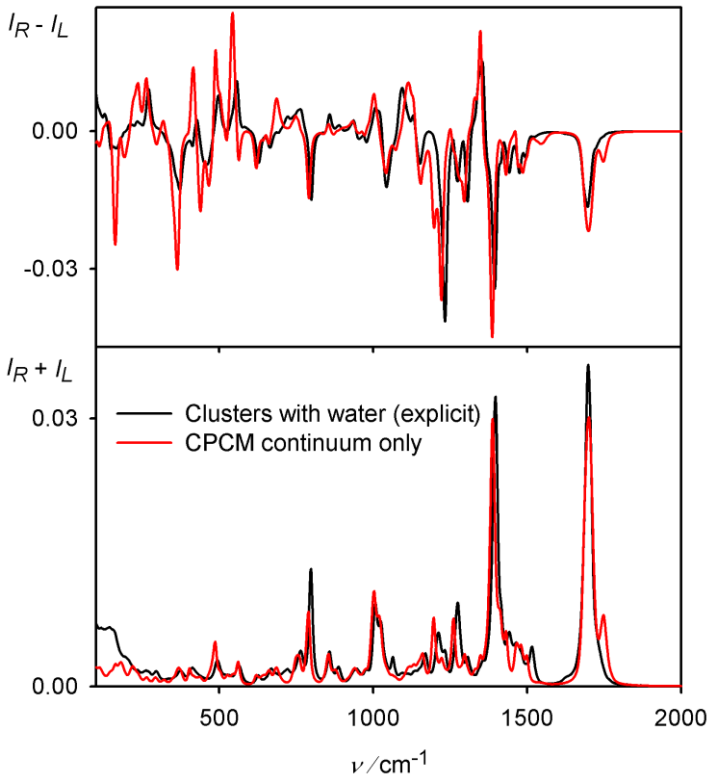

Figure S2. dTMP ROA and Raman spectra simulated from MD snapshots with explicit water molecules, and with CPCM only.

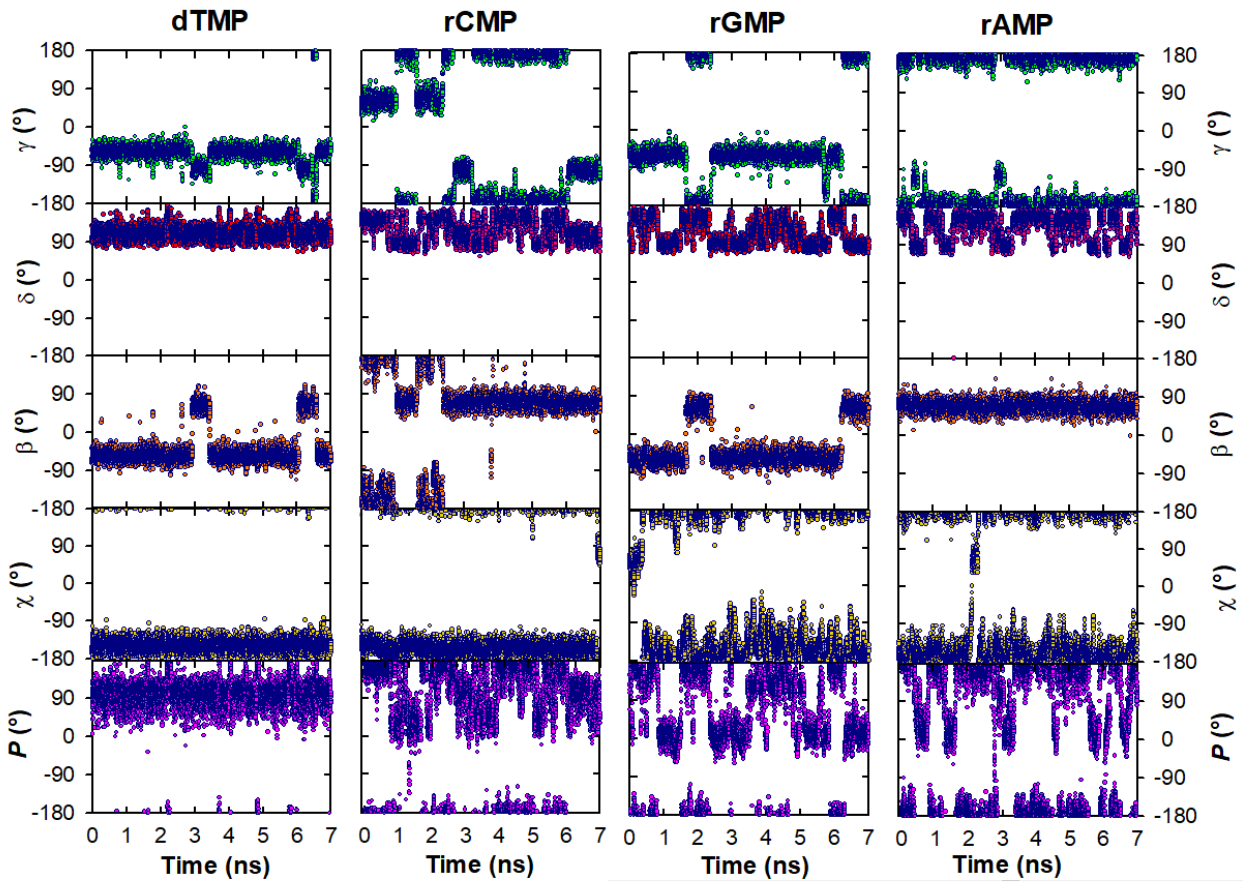

Figure S3. Time-evolution of selected geometrical parameters in the four nucleosides during MD, GAFF2 force field.

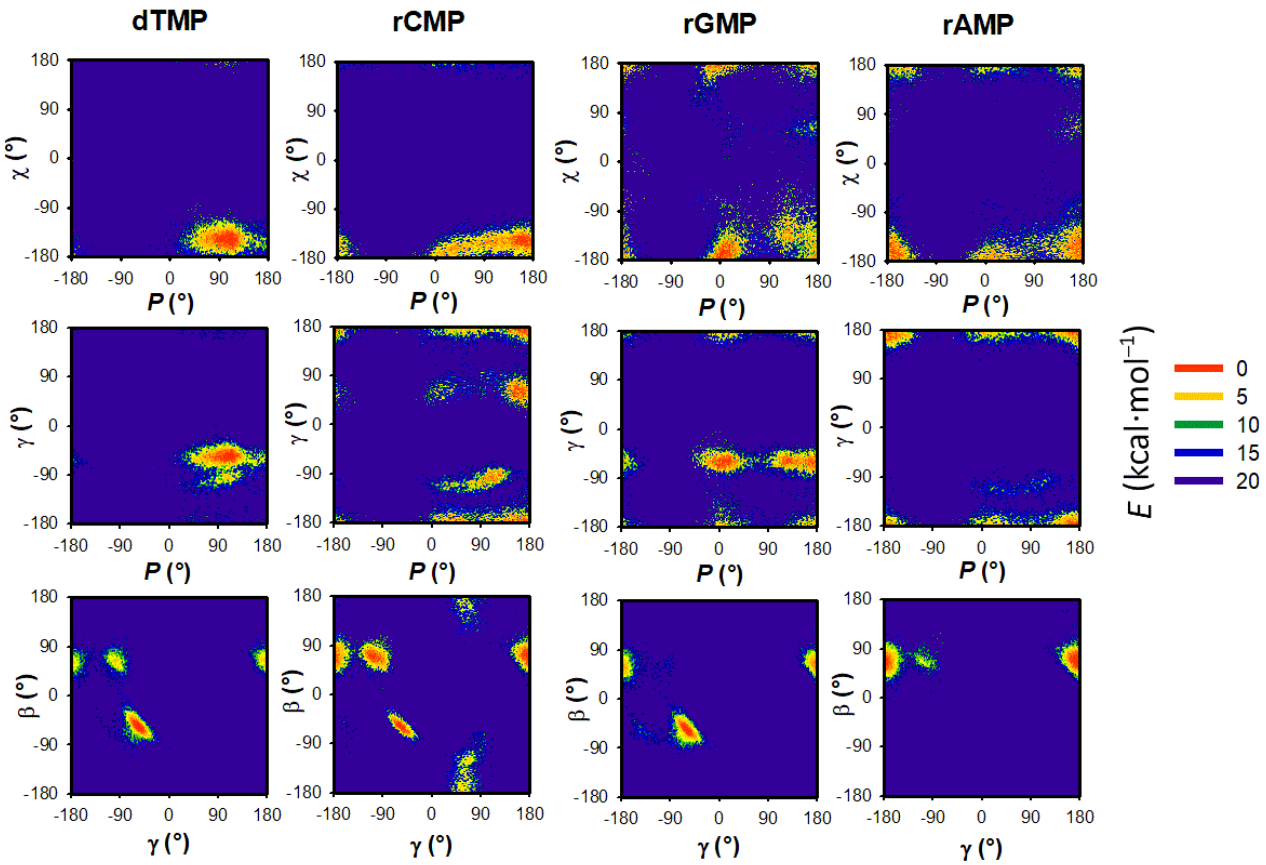

Figure S4. Dependence of the free energy on the backbone torsion angles (β, γ and δ), glycosidic bond torsion angle (χ) and pseudorotation angle (*P*), as obtained from free MD, GAFF2 force field.

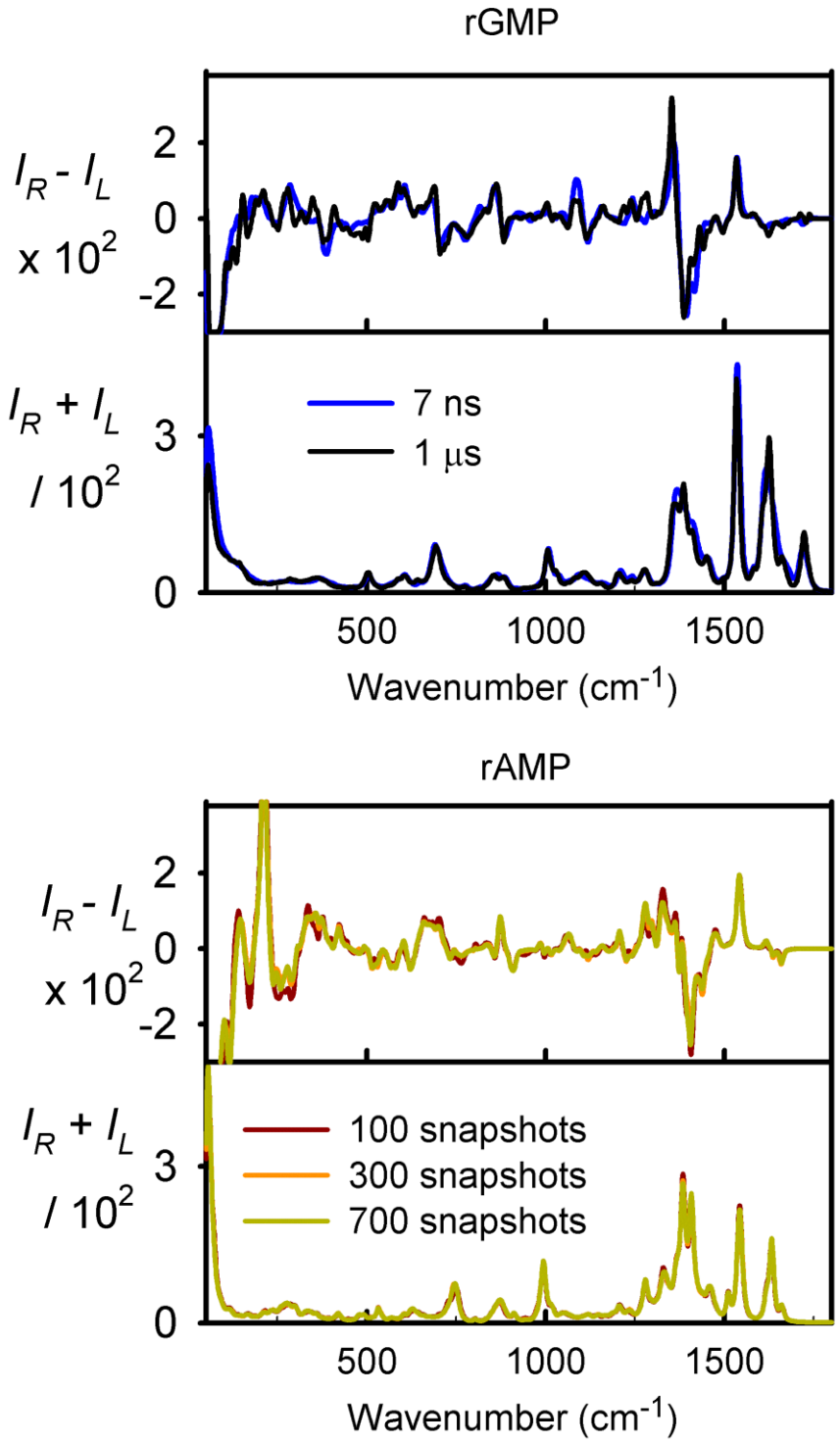

Figure S5. Convergence tests: rGMP spectra generated from 7 ns (700 snapshots, 10 ps intervals) and 1 μs (100 snapshots, 10 ns intervals), and rAMP spectra generated with 100, 300 and 700 snapshots (10 ps intervals), free MD runs.

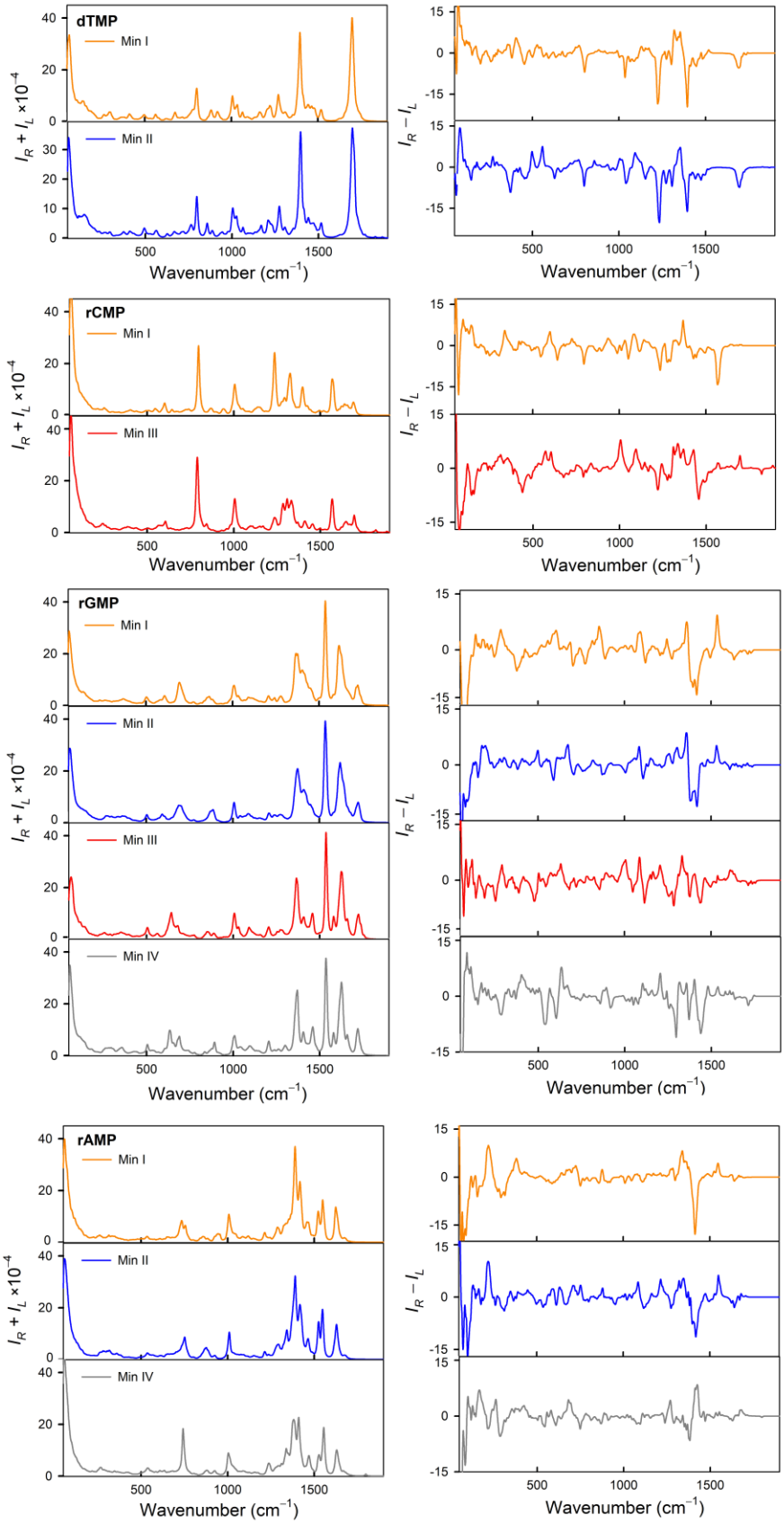

Figure S6. Raman and ROA spectra calculated for conformers corresponding to WHAM minima. Each spectrum is an average of at least 50 snapshots.

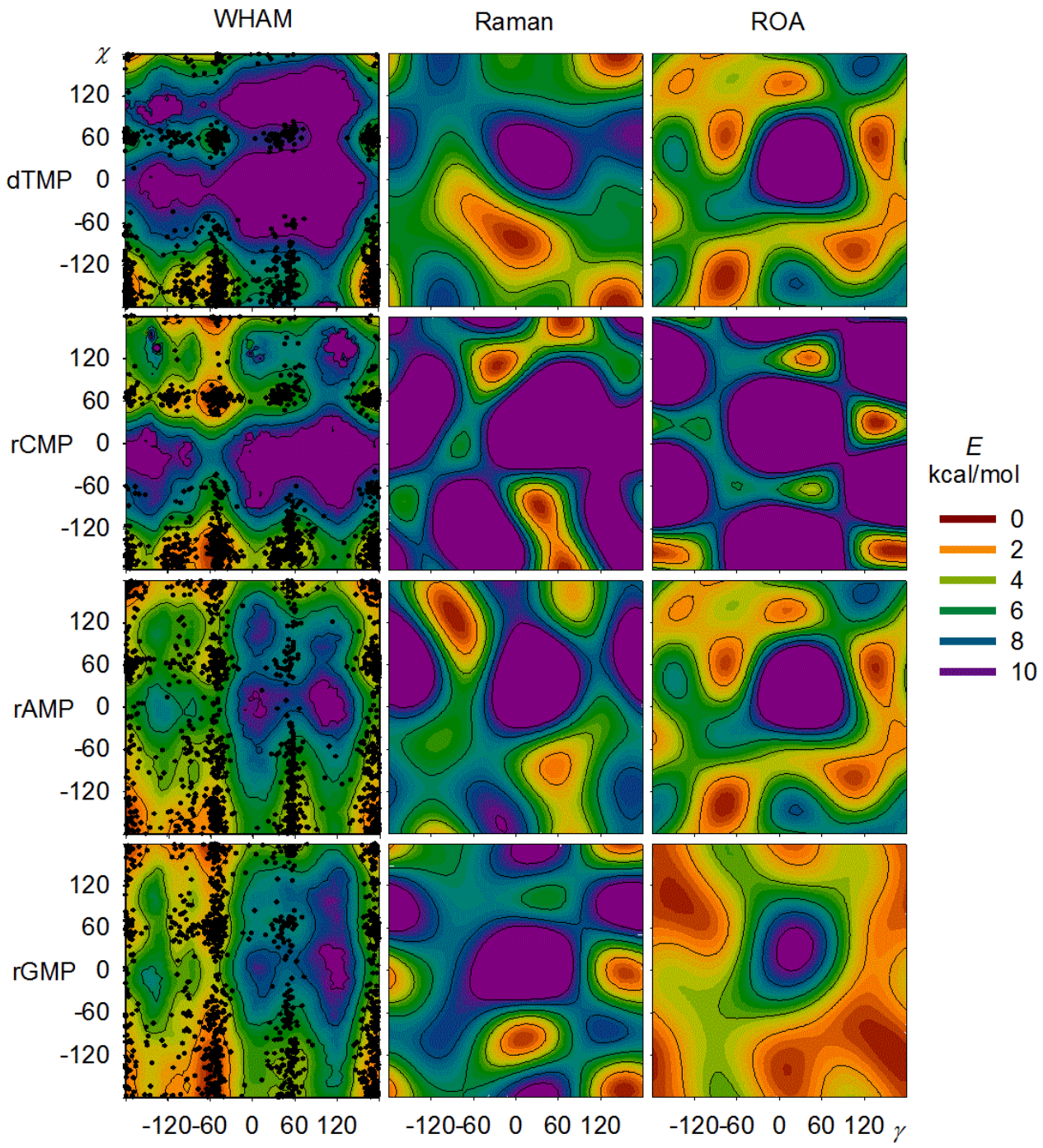

Figure S7. Dependence of the free energy on the (γ,χ) angles obtained by MD-WHAM, and arbitrary maps obtained by decomposition of experimental Raman and ROA spectra. The snapshots used in the decomposition are indicated by the black dots.

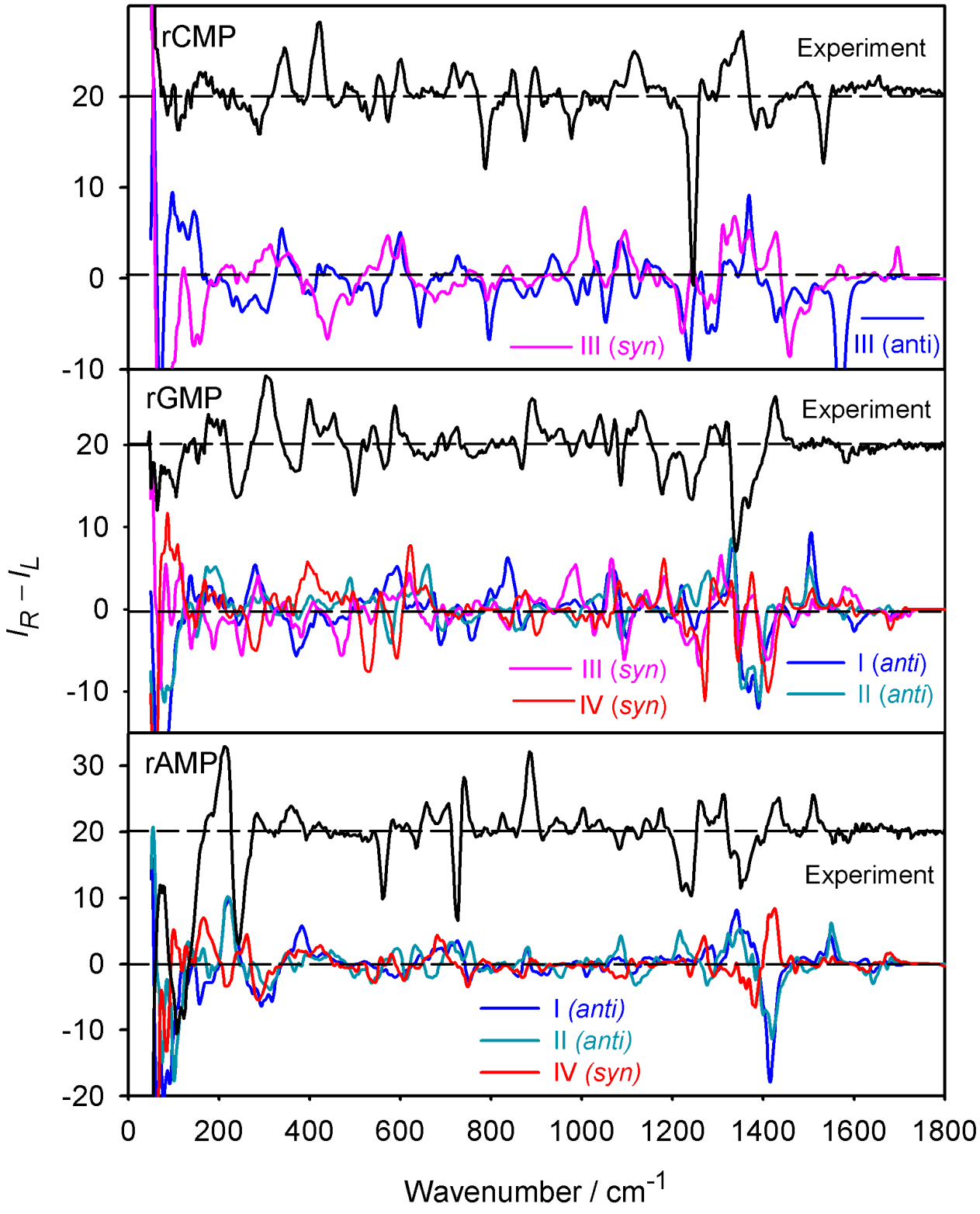

Figure S8. Calculated ROA spectra of the nucleotide conformers with *syn* or *anti* conformation of the base (MD average, frequency multiplied by 0.98), and the experiment. For dTMP the *syn* conformation was not present in MD. The experimental spectra are arbitrarily shifted and scaled in the intensity axis.

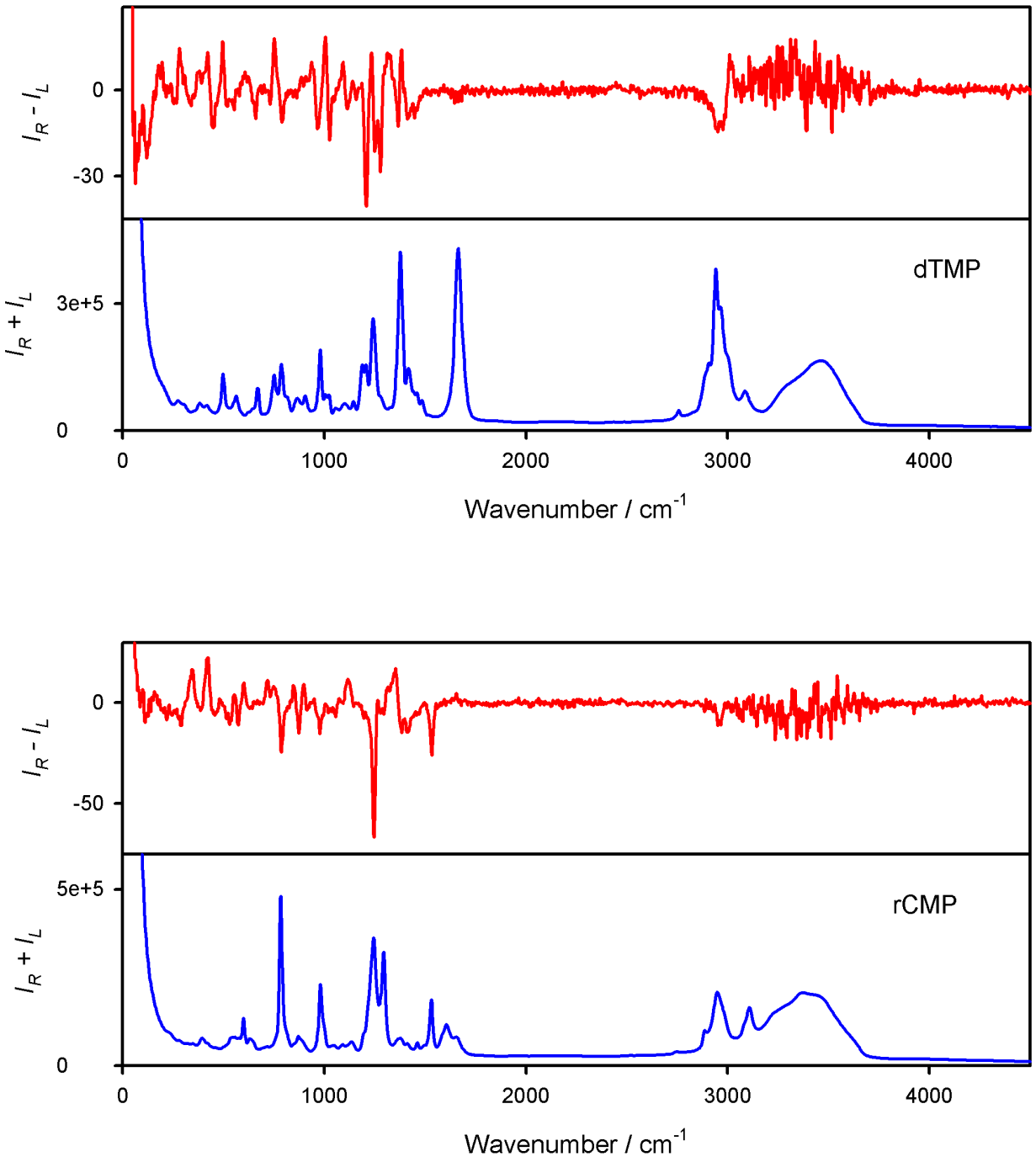

Figure S9. Experimental Raman and ROA spectra in the full range of wavenumbers.

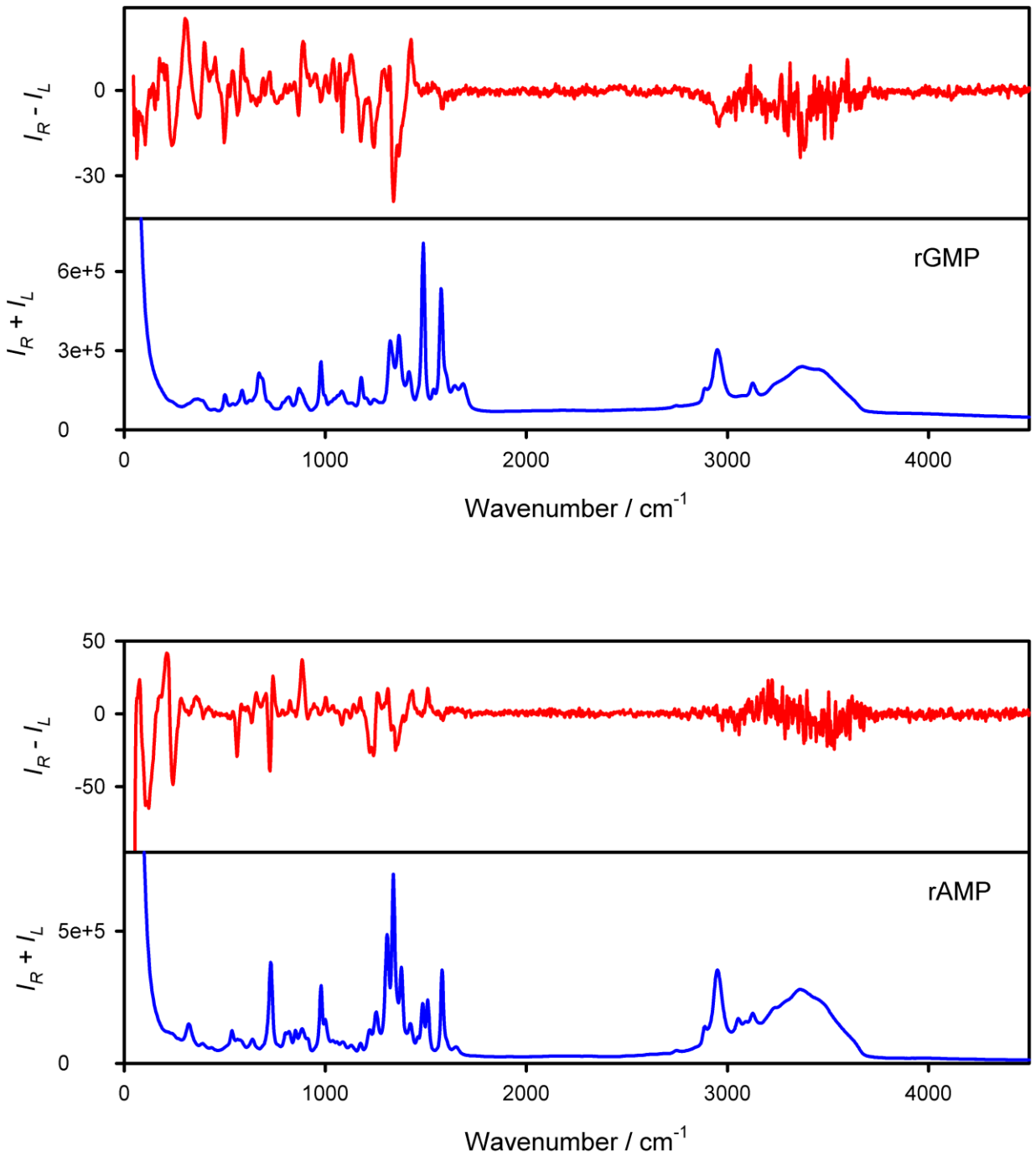

Figure S9 (cont.). Experimental Raman and ROA spectra in the full range of wavenumbers.