Chemistry–A European Journal

Supporting Information

Bringing Machine-Learning Enhanced Quantum Chemistry and Microwave Spectroscopy to Conformational Landscape Exploration: the Paradigmatic Case of 4-Fluoro-Threonine

V. Barone,* M. Fusè, R. Aguado, S. Potenti, I. León, E. R. Alonso, S. Mata, F. Lazzari, G. Mancini, L. Spada, A. Gualandi, P. G. Cozzi, C. Puzzarini,* and J. L. Alonso*

Contents

List of Figures

List of Tables

1 Computational Details

1.1 Exploration of the conformational PES

Evolutionary algorithms (EA) are particularly well suited for the exploration of rugged potential energy surfaces (PESs) like the conformational landscapes of flexible molecules [1]. Based on previous experience [2, 3], we have adopted a purposely tailored version of $(\lambda + \mu)$ Island Model (IM) [4], IMEA[2, 5]). A detailed description of the method and implementation details can be found in refs. [6, 2, 5]. The current version of the code is available under the GPL3 license at https://github.com/tuthmose/IM_EA, while the full set of parameters selected in the present context are given in Table 1.

The parameters of Table 1 were employed in four replicas, each consisting of about 4000 constrained geometry optimizations by a low-cost semi-empirical method (GFN2-XTB [7]). Then, a first reduction to about 1000 structures was obtained by applying a threshold of 1250 cm^{-1} (15 kJ mol^{-1}) with respect to the absolute energy minimum. These structures were then compared with each other by applying the Partition Around Medoids (PAM)[8] clustering method, using the root-mean-square deviations of heavy atom positions and rotational constants to build the feature space. The PAM algorithm optimizes the partition of objects in the data set into k distinct groups, with the best value of k being determined by applying suitable external validation criteria aimed to estimate the cluster compactness and separation [9]. Three criteria were selected here: the change in the slope of the Within Sum of Squares error (WSS) curve (the so-called elbow), together with the Silhouette (SI, to be maximized) and Davies Bouldin (DBI, to be minimized) indexes. Further comparison with the 21 structures directly derived from threonine and full geometry optimizations at the B3LYP/jun-cc-pVDZ level[10, 11, 12] (also including empirical dispersion contributions [13]; hereafter B3), leaves 7 additional candidates for the successive exploitation.

Parameter	Value
Population size	100
Number of generations (max)	50
Selection pressure	0.5
Selection method	Tournament
Tournament size	2
Elitism (last 10% generations)	T
Crossover method	SBX
Crossover probability	0.5
Mutation rate (parents)	0.3
Mutation rate (children)	0.5
Number of islands	4
Migration frequency	4
Migration size	0.05

Table 1: Run parameters and values for PES exploration.

1.2 Exploitation and property evaluation

Full geometry optimizations and Hessian evaluations at the rDSD level (double-hybrid revDSD-PBEP86-D3BJ functional with jun-cc-pvTZ[14, 15, 12]) provided the true energy minima shown in Table S2. Next, relaxed geometry scans were performed to detect possible relaxation paths between pairs of adjacent conformers. For purposes of illustration, some relaxed scans round the χ_3 dihedral angle are shown in Figure S1. In order to improve the accuracy of the electronic energies, the relative stability of the stationary points was further refined by single-point energy evaluations using the jun-Cheap-F12 composite scheme (junChS-F12) on top of rDSD geometries [16, 17]:

$$E(\text{junChS-F12}) = E(\text{CC-F12/TZ}) + \Delta E(\text{MP2-F12/CBS}) + \Delta E(\text{MP2-F12/CV}) \; .$$

The starting point of this composite scheme is the evaluation of the electronic energy using the CCSD(T)-F12b(3C/FIX) method [18, 19, 20] (hereafter simply referred to as CC-F12), within the frozen-core (fc) approximation, in conjunction with the jun-cc-pVTZ basis set (hereafter TZ). The second term on the right-hand side of the equation above incorporates the correction due to the extrapolation to the complete basis set (CBS) limit, which is carried out at the fc-MP2-F12 level employing the jun-cc-pVTZ and jun-cc-pVQZ (QZ) basis sets and the standard n^{-3} two-point formula[21]. The last term is the core-vale (CV) correlation contribution, which is incorporated as the difference between between all-electron (ae) and fc MP2-F12 calculations, both in conjunction with the cc-pCVTZ-F12 basis set.[22] The various contributions are collected in Table S3.

As mentioned in the main text, equilibrium rotational constants (A_e, B_e, C_e) were straightforwardly derived from the rDSD optimized geometries, whereas vibrational contributions (Δvib) leading to effective rotational constants (A_0, B_0, C_0) were computed in the framework of the generalized second-order perturbation theory (GVPT2)[24] at the B3 level. Improved equilibrium rotational constants were obtained by exploiting the linear regression approach (LRA)[25], which is detailed in the Methods section of the main text. The reader is also referred to ref. [26].

As detailed in the Methods section of the main text, thermodynamic functions at the harmonic level, quadrupole coupling constants and dipole moments were acomputed at the rDSD level (see Table S4), Anharmonic corrections to zero point energies (ZPEs) were also obitaned in the framework of GVPT2 at the B3 level, with contributions of low-frequency vibrations to ZPEs and entropies being treated as explained in the Methods section of the main text.

DFT and GVPT2 calculations were performed using the Gaussian16 suite of programs[27]. Geometry optimizations were always performed using tight convergence criteria (i.e., 1 × 10–5 hartree/bohr and 4 × 10–5 bohr on RMS force and displacements, respectively, with maximum values being 1.5 times larger). Cubic and semi-diagonal force constants were obtained from finite differences of analytical Hessians, and employing a stepsize of 0.01 amu$^{1/2}$ Å for the displacements along the mass-weighted normal coordinates. Explicitly correlated computations were carried out with the MOLPRO package [28]. Quasi-harmonic corrections to entropies and free energies [29] were computed employing the GoodVibes program [30].

Table 2: Relative electronic energies (ΔE_{el}; in cm^{-1}) and dihedral angles (ϕ' = LP-N-C^{α}-C', ψ = N-C-C-O(H), ω = C-C-O-H, χ_1 = N-C^{α}-C^{β}-O, χ_2 = C^{α}-C^{β}-O-H, χ_3 = C^{α}-C^{β}-C^{γ}-F; in degrees) of the low-energy conformers of 4FT (at the rDSD level). In the last column, the equivalent conformers of Thr are reported.

	ΔE_{el}	ϕ'	ψ	ω	χ_1	χ_2	χ_3	Thr equiv.[a]
I'gg$^-$g$^-$	0.0	103.7	-175.2	180.0	53.3	-42.5	-59.1	
I'gg$^-$g	271.3	88.6	-172.2	-178.5	43.7	-33.6	50.4	I'gg$^-$ (I'b)
I'gg$^-$t	892.9	103.4	175.5	179.6	55.1	-46.4	-173.3	
IIggg$^-$	-79.0	-35.5	24.2	-6.6	62.7	78.2	-50.6	
IIggg	315.1	-34.9	22.1	-6.4	53.4	79.2	54.9	IIgg (IIb)
IIggt	628.7	-33.0	21.1	-5.9	57.5	84.7	-168.1	
IIg$^-$tg$^-$	1862.8	41.4	-24.9	5.5	-67.2	167.9	-86.9	
IIg$^-$tg	1217.0	30.2	-15.6	3.6	-48.5	-175.6	60.3	
IIg$^-$tt	217.3	31.7	-16.4	3.3	-55.0	-173.8	175.2	IIg$^-$t (IIa)
I'g$^-$gg$^-$	439.4	162.9	142.5	175.0	-57.0	41.2	-75.4	
I'g$^-$gg	482.6	158.3	151.3	178.2	-49.4	36.1	51.5	I'g$^-$g (Ia)
I'g$^-$gt.	857.4	161.0	145.9	178.2	-52.64	41.69	-171.61	
III'ggg$^-$	453.4	-176.3	74.3	-176.3	56.2	66.5	-62.4	III'gg (IIIb)
III'ggg	1903.3	-161.0	67.4	-176.6	60.3	64.2	69.4	
III'ggt	816.8	-171.7	62.9	-177.9	60.7	72.0	-163.5	
IItg$^-$g$^-$	-195.5	-23.6	12.5	-2.8	-169.4	-49.3	-59.8	
IItg$^-$g	841.0	-32.6	15.7	-3.5	-179.9	-42.5	42.4	IItg$^-$ (IIc)
IItg-t	1092.8	-25.4	14.8	-3.0	-165.1	-55.5	-161.9	
III'g$^-$gg$^-$	840.8	168.8	-56.0	-178.9	-56.1	39.7	-71.0	
III'g$^-$gg	1006.9	163.9	-41.9	-178.5	-49.0	34.3	53.6	III'g$^-$g (IIIa)
III'g$^-$gt	1159.7	166.2	-45.0	-178.5	-52.6	40.7	-169.6	
Igtg$^-$	935.7	179.8	-179.9	-179.0	59.9	165.3	-66.5	
Igtg	relaxes to IIggg$^-$							Igt
Igtt	324.0	-173.4	179.5	-179.5	64.5	-174.7	-179.3	
IIgtg$^-$	relaxes to IIggg$^-$							
IIgtg	relaxes to IIggg							IIgt
IIgtt	433.9	-23.5	10.5	-2.7	46.9	178.9	176.8	
IIIgtg$^-$	1197.4	-178.9	11.9	178.4	57.2	149.1	-64.7	
IIIgtg	1915.0	-178.1	-6.1	-178.8	65.9	-93.4	83.8	
IIIgtt	667.7	-174.5	-1.2	-179.8	61.7	-176.4	179.8	IIIgt
IIg$^-$g$^-$g$^-$	2107.2	39.6	-25.8	6.0	-66.5	-64.8	-81.5	
IIg$^-$g$^-$g	290.6	29.2	-16.7	4.4	-47.2	-82.0	69.2	IIg$^-$g$^-$
IIg$^-$g$^-$t	relaxes to IIg$^-$tt							

[a] Symbols employed in ref. [23], when available, are given within parentheses.

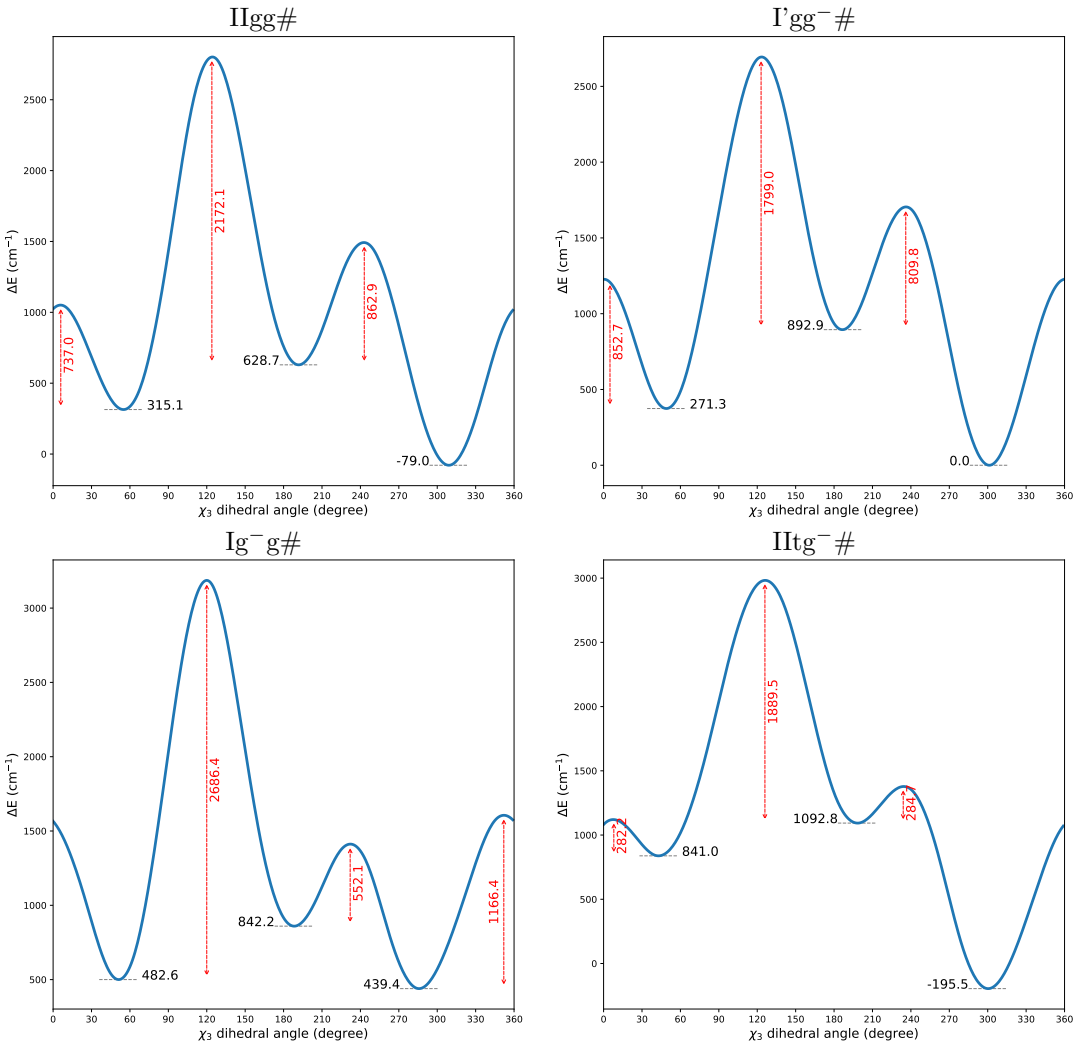

Figure 1: Potential energy profiles (rDSD level; in cm^{-1}) along the χ_3 dihedral angle. In all plots, the absolute energy minimum corresponds to the g^- orientation ($\chi_3 \approx 300°$).

Table 3: Relative energy of the different conformers (in cm^{-1}; 1 kJ/mol = 83.59 cm^{-1}): different contributions to the junChS-F12 approach (evaluated on top of rDSD geometries).

	CC-F12/TZ	MP2-F12/TZ	MP2-F12/QZ	CBS	CBS+CV	B3
I'gg$^-$g$^-$	0.0	0.0	0.0	0.0	0.0	0.0
I'gg$^-$g	262.0	253.2	260.2	274.0	276.4	125.9
I'gg$^-$t	920.0	940.2	943.8	926.3	929.5	924.4
IIggg$^-$	-50.8	-19.1	-9.0	-33.2	-34.9	-87.8
IIggg	363.1	396.6	410.5	387.2	391.2	101.1
IIggt	696.0	709.5	724.5	721.9	721.4	580.2
IIg$^-$tt	261.5	214.3	221.6	274.1	274.0	75.3
I'g$^-$gg$^-$	503.8	525.9	535.1	519.6	530.2	269.6
I'g$^-$gg	544.8	537.7	544.9	557.3	566.5	322.2
I'g$^-$gt	947.0	942.9	946.9	953.8	965.3	672.1
III'ggg$^-$	490.7	595.1	604.5	506.9	516.8	407.3
III'ggt	862.8	965.1	978.4	885.8	899.5	792.9
IItg$^-$g$^-$	-170.1	-127.7	-111.4	-141.8	-137.7	-521.9
IItg$^-$g	857.7	917.6	941.4	898.8	910.5	638.7
IIttg$^-$t	1164.0	1217.0	1230.7	1187.7	1194.7	849.9
III'g$^-$gg$^-$	922.7	945.3	951.3	933.0	944.9	701.2
III'g$^-$gg	1079.9	1057.4	1064.3	1091.8	1101.2	851.9
III'g$^-$gt	1259.0	1248.3	1252.1	1265.6	1275.8	979.8
Igtg$^-$	986.7	1045.6	1048.0	990.9	1001.8	932.3
Igtt	372.5	415.2	424.0	387.8	395.0	263.5
IIgtt	495.1	472.0	482.1	512.6	508.2	283.2
IIIgtt	732.8	760.7	769.5	748.0	756.8	562.1
IIg$^-$g$^-$g	317.4	283.5	293.3	334.3	338.3	128.0
Mean(AE)	12.5	27.5	31.6	4.7	0.0	203.4
Max(AE)	32.4	69.6	78.7	9.5	0.0	384.2

Table 4: Relative electronic energies (ΔE_{el}), harmonic enthalpies at 0 K and harmonic Gibbs free energies at 298 K (ΔG°_{H}) in cm^{-1}; ^{14}N nuclear quadrupole coupling constants (χ, in MHz) and dipole moment components (μ, in debyes) of the low-energy conformers of 4FT evaluated at the rDSD level.

	ΔE_{el}	$\Delta H^{\circ}_{0,H}$	ΔG°_{H}	χ_{aa}	χ_{bb}	χ_{cc}	μ_a	μ_b	μ_c
I'gg$^-$g$^-$	0.0	0.0	0.0	0.4662	1.1833	-1.6541	-0.1602	0.2344	1.3179
I'gg$^-$g	271.3	280.9	348.7	1.2779	2.7202	-3.9981	-0.6290	-1.1466	-2.8058
I'gg$^-$t	892.9	812.4	798.0	-0.6618	3.0106	-2.3488	-1.6864	1.1363	-3.2721
IIggg$^-$	-79.0	22.2	75.1	-3.0318	0.7818	2.2500	-2.6337	-1.5756	1.7777
IIggg	315.1	369.4	390.0	-2.7725	0.8710	1.9016	2.6175	0.7640	-1.0749
IIggt	628.7	637.0	649.9	-4.2115	2.3251	1.8864	-0.8759	-4.2112	-0.9919
IIg$^-$tg$^-$	1862.8	1851.4	1909.2	0.5631	2.2698	-2.8329	-3.4832	4.3103	-0.5055
IIg$^-$tg	1217.0	1209.2	1215.5	-1.6575	2.5453	-0.8878	-1.8832	3.7462	-1.1583
IIg$^-$tt	217.3	286.2	349.4	-0.2645	2.3471	-2.0826	-0.7267	3.8673	0.7290
I'g$^-$gg$^-$	439.4	420.6	429.5	-3.8457	1.9995	1.8463	-1.0357	1.1808	-1.7257
I'g$^-$gg	482.6	487.2	438.9	-2.2969	0.4853	1.8116	-2.4520	0.6923	-0.2626
I'g$^-$gt	857.4	811.7	737.4	-1.9945	0.3152	1.6792	3.7741	-0.3299	1.7041
III'ggg$^-$	453.4	456.0	431.3	-3.9309	1.3010	2.6299	-2.6935	0.5673	0.3533
III'ggg	1903.3	1840.4	1782.4	-2.0127	0.1703	-2.0732	3.0036	-0.2622	-1.8232
III'ggt	816.8	781.7	739.0	-2.2829	-0.0221	2.3050	3.8992	-1.5437	-0.7216
IItg$^-$g$^-$	-195.5	-3.7	167.5	-3.2770	2.0751	1.2018	1.3244	2.3436	-2.5310
IItg$^-$g	841.0	970.1	1132.7	-3.6887	2.7363	0.9524	-1.4392	4.5419	-1.0467
IItg$^-$t	1092.8	1158.8	1217.9	-4.3149	2.7555	1.5593	1.3416	4.9369	-2.8428
III'g$^-$gg$^-$	840.8	800.0	775.6	-3.8937	2.1201	1.7736	-0.8033	2.1493	1.1058
III'g$^-$gg	1006.9	966.8	676.0	-2.5623	0.8245	1.7378	-3.4122	1.6786	1.5630
III'g$^-$gt	1159.7	1090.8	943.5	-1.8751	0.1252	1.7499	-4.1257	1.0693	0.7086
Igtg$^-$	935.7	775.8	658.2	-4.1820	2.6896	1.4924	0.2134	-1.3008	1.7329
Igtt	324.0	218.0	167.0	-2.5149	1.6381	0.8768	0.3612	0.8121	-1.5642
IIgtt	433.9	444.4	484.6	-4.2430	2.5572	1.6858	2.8759	-4.0390	1.8598
IIIgtg$^-$	1197.4	1018.1	655.4	-4.1072	2.5093	1.5979	-1.1352	0.5334	1.4631
IIIgtg	1915.0	1827.7	1661.4	-3.6795	2.7658	0.9137	-0.3859	1.0418	-0.7404
IIIgtt	667.7	588.1	510.7	-2.6103	1.6969	0.9134	-1.3952	-1.4899	0.3445
IIg$^-$g$^-$g$^-$	2107.2	2097.9	2168.4	0.3174	2.3357	-2.6531	-2.8566	4.2137	1.3311
IIg$^-$g$^-$g	290.6	431.3	445.6	-2.0827	2.4738	-0.3911	1.6343	-3.3607	0.6551

	ZPE_A - ZPE_H	$T\Delta S_{QH}$ - $T\Delta S_H$
I'gg$^-$g$^-$	0.0	0.0
I'gg$^-$g	-0.4	12.5
I'gg$^-$t	8.0	4.6
IIggg$^-$	2.4	5.7
IIggg	7.5	-8.3
IIggt	-9.6	-3.1
Iig$^-$tg$^-$	18.4	26.3
Iig$^-$tg	-34.6	7.7
IIg$^-$tt	5.1	26.8
I'g$^-$gg$^-$	10.7	4.2
I'g$^-$gg	-2.2	-34.0
I'g$^-$gt.	11.6	-32.9
III'ggg$^-$	11.9	-7.0
III'ggg	-3.4	-26.6
III'ggt	9.4	-17.3
IItg$^-$g$^-$	7.9	52.9
IItg$^-$g	4.0	13.8
IItg-t	-14.6	24.4
III'g$^-$gg$^-$	2.5	-14.9
III'g$^-$gg	-0.6	-74.6
III'g$^-$gt	10.9	-160.4
Igtg$^-$	40.1	-14.7
Igtt	18.8	-41.7
IIgtt	-17.4	21.5
IIIgtg$^-$	8.3	-154.1
IIIgtg	10.0	-75.9
IIIgtt	12.6	-23.7
IIg$^-$g$^-$g$^-$	-7.6	34.9
IIg$^-$g$^-$g	-13.0	30.3

Table 5: Anharmonic corrections to ZPEs and quasi-Harmonic corrections to entropies ($T\Delta S$).

Table 6: Equilibrium rotational constants at the rDSD and LRA-corrected (rDSD-LRA) levels together with vibrational corrections (B3 level). All values are in MHz.

	rDSD			rDSD-LRA			B3		
	A_e	B_e	C_e	A_e	B_e	C_e	$\Delta_{vib}A$	$\Delta_{vib}B$	$\Delta_{vib}C$
I'gg$^-$g$^-$	2496.98	1033.63	994.31	2510.02	1038.00	998.61	21.30	9.26	8.49
I'gg$^-$g	2827.90	1047.60	949.60	2842.41	1052.09	953.71	26.14	11.01	7.62
I'gg$^-$	2952.24	953.45	858.37	2967.3	957.5	862.1	38.15	6.82	4.91
IIggg$^-$	2610.52	1075.83	946.97	2624.08	1080.44	951.05	24.57	10.96	9.84
IIggg	2900.13	1076.35	917.92	2914.89	1080.90	921.86	28.23	11.93	6.41
IIggt	3074.08	949.47	836.89	3089.52	953.64	840.56	37.65	10.08	4.77
IIg$^-$tg$^-$	1844.47	1545.34	994.63	1853.38	1552.21	999.02	13.06	16.35	10.00
IIg$^-$tg	2228.98	1290.23	934.85	2239.56	1295.96	939.05	7.64	17.62	10.37
IIg$^-$tt	2371.71	1142.32	849.47	2383.00	1147.48	853.30	21.43	9.74	7.78
I'g$^-$gg$^-$	1848.28	1552.55	998.28	1856.95	1559.51	1002.63	17.36	10.50	9.01
I'g$^-$gg	2073.53	1365.12	963.42	2083.54	1371.05	967.70	15.71	14.01	9.43
I'g$^-$gt	2143.46	1163.17	869.78	2153.86	1168.37	873.64	18.98	8.23	6.15
III'ggg$^-$	2521.17	1056.02	931.28	2534.37	1060.41	935.36	20.06	11.20	7.84
III'ggg	3087.04	1035.87	890.39	3102.86	1040.27	894.26	37.30	6.55	8.57
III'ggt	3131.68	928.23	829.23	3147.51	932.28	832.87	20.22	6.46	8.75
IItg$^-$g$^-$	1908.72	1435.22	1122.13	1918.17	1441.03	1127.26	13.45	14.03	11.64
IItg$^-$g	2335.72	1102.40	978.35	2346.66	1107.24	982.81	31.41	7.46	4.80
IItg$^-$t	2008.92	1437.36	1285.64	2018.56	1443.34	1291.19	21.08	13.43	5.25
III'g$^-$gg$^-$	1880.05	1514.30	1009.85	1888.90	1521.08	1014.25	14.63	10.30	7.02
III'g$^-$gg	2087.58	1323.94	972.47	2097.54	1329.78	976.80	23.05	12.67	5.97
III'g$^-$gt	2132.66	1160.08	882.56	2143.05	1165.31	886.47	22.54	6.74	3.47
Igtg$^-$	2487.60	1024.57	986.46	2500.55	1028.90	990.70	21.76	8.67	7.39
Igtt	2973.24	959.57	862.64	2988.52	963.75	866.41	37.22	6.64	5.97
IIgtt	3044.11	936.34	844.49	3059.34	940.47	848.22	29.84	7.31	6.33
IIIgtg$^-$	2521.58	1005.55	981.27	2534.73	1009.91	985.46	20.06	11.20	7.84
IIIgtg	3014.46	1015.58	926.20	3029.83	1019.95	930.17	33.74	25.60	2.01
IIIgtt	3022.11	945.71	861.19	3037.39	949.82	864.93	28.86	7.82	8.55
IIg$^-$g$^-$g$^-$	1855.11	1536.94	1004.05	1863.97	1543.81	1008.45	13.63	15.67	8.81
IIg$^-$g$^-$g	2308.27	1245.77	923.41	2319.25	1251.30	927.56	20.10	11.65	9.04

1.3 Geometries

rDSD Geometries

I'gg$^-$g$^-$

17

F	-0.031109	-0.029662	0.027017
C	0.001703	-0.018491	1.422293
C	1.443570	0.000542	1.905845
C	2.218273	-1.226169	1.392713
C	3.688926	-1.101674	1.743354
H	-0.521045	0.872445	1.770701
H	-0.515137	-0.911364	1.781095
H	1.930580	0.913314	1.562300
O	1.450009	0.016333	3.319278
H	1.288473	-0.903744	3.581287
H	2.153932	-1.240701	0.299197
N	1.678125	-2.426815	2.023801
H	2.440470	-3.035683	2.300845
H	1.075788	-2.936814	1.392122
O	4.325741	-1.930522	2.348475
O	4.221726	0.043229	1.276670
H	5.153307	0.044226	1.541585

I'gg$^-$g

17

F	0.021651	-0.053687	0.005668
C	0.008664	-0.004236	1.402343
C	1.417309	0.014994	1.967280
C	2.289975	1.098548	1.289814
C	3.480874	1.469301	2.156117
H	-0.533413	0.901345	1.688416
H	-0.520846	-0.888704	1.758163
H	1.316326	0.247037	3.030713
O	2.056251	-1.235392	1.859613
H	2.517191	-1.203444	1.003627
H	1.693482	2.017161	1.197235
N	2.790458	0.586022	0.023675
H	3.676222	1.017722	-0.208279
H	2.121861	0.725661	-0.720931
O	4.624251	1.518106	1.773806
O	3.109617	1.794643	3.411683
H	3.919604	2.038187	3.883454

I'gg$^-$t

17

F	3.041212	-0.609708	1.126516
C	2.187135	-0.027977	0.203030
C	0.743942	-0.296670	0.592789
C	-0.202064	0.202944	-0.517288
C	-1.624944	-0.236920	-0.219707
H	2.392562	1.045112	0.186561
H	2.410610	-0.464743	-0.775692
H	0.616641	-1.370940	0.732488
O	0.401600	0.318184	1.815524
H	0.290634	1.258614	1.610974
H	0.075099	-0.274078	-1.466179
N	-0.159305	1.661385	-0.568362

H	-1.099402	2.024582	-0.683113
H	0.407389	1.992724	-1.336982
O	-2.574059	0.506110	-0.154518
O	-1.714591	-1.570308	-0.063305
H	-2.643302	-1.767111	0.128547

IIggg$^-$

17

F	0.002690	-0.019821	0.015013
C	0.009268	-0.010659	1.408066
C	1.442354	-0.005313	1.906781
C	2.287838	-1.103001	1.254903
C	3.696902	-1.068364	1.862025
H	-0.520507	0.878505	1.748915
H	-0.516231	-0.901781	1.756870
H	1.907749	0.958798	1.687049
O	1.344180	-0.212065	3.310589
H	2.173741	0.088835	3.699859
H	2.406712	-0.867005	0.193435
N	1.759039	-2.462479	1.360153
H	1.057668	-2.637836	0.652039
H	1.329314	-2.598176	2.270442
O	4.206635	-0.059951	2.290053
O	4.316087	-2.251087	1.866859
H	3.640528	-2.891527	1.559259

IIggg

17

F	0.004650	-0.047839	0.013817
C	0.010137	0.000783	1.407082
C	1.412172	0.003128	1.972307
C	2.307035	1.119039	1.420650
C	3.591362	1.188069	2.267162
H	-0.514354	0.912175	1.704459
H	-0.530595	-0.875105	1.767122
H	1.306443	0.165752	3.050479
O	2.009002	-1.264080	1.717507
H	2.696030	-1.398790	2.380003
H	1.807534	2.083860	1.561713
N	2.688392	0.993665	0.012942
H	1.946055	1.329988	-0.586569
H	2.801491	0.006401	-0.202254
O	3.621837	0.919524	3.444765
O	4.663256	1.584483	1.581024
H	4.353212	1.632712	0.648438

IIggt

17

F	0.006150	0.020613	0.023295
C	0.007182	-0.000619	1.408192
C	1.428182	-0.000446	1.924635
C	1.479830	-0.318068	3.422809
C	2.952062	-0.403205	3.861157
H	-0.532636	0.883973	1.753442
H	-0.517124	-0.905600	1.728456
H	1.997385	-0.762446	1.386711
O	1.973104	1.293413	1.694074

H	2.929541	1.201020	1.623383
H	1.059584	-1.313772	3.600076
N	0.782572	0.640417	4.283685
H	0.871001	1.573636	3.891786
H	-0.202063	0.426053	4.369926
O	3.827193	-0.819627	3.142056
O	3.170470	0.008225	5.111917
H	2.307837	0.363699	5.415133

IIg$^-$tg$^-$

17

F	-0.943156	-1.990235	-0.653023
C	-1.519970	-1.051561	0.188748
C	-1.039179	0.339710	-0.213113
C	0.278826	0.757751	0.440556
C	1.350342	-0.320555	0.245483
H	-2.604235	-1.111791	0.065100
H	-1.249700	-1.301132	1.215704
H	-0.912156	0.349081	-1.302869
O	-1.978616	1.335042	0.193247
H	-2.777325	1.237501	-0.334904
H	0.098411	0.833932	1.517619
N	0.721301	2.008534	-0.180493
H	1.379267	2.492221	0.420938
H	-0.077600	2.622044	-0.298561
O	1.467235	-1.266715	0.980056
O	2.140421	-0.115978	-0.818499
H	1.880170	0.757234	-1.174326

IIg$^-$tg

17

F	-1.792748	-0.767570	1.647215
C	-1.537654	-0.889279	0.289450
C	-0.758900	0.314870	-0.208820
C	0.596747	0.473412	0.498641
C	1.520112	-0.662322	0.032383
H	-0.975211	-1.808127	0.126238
H	-2.502867	-0.945893	-0.223365
H	-0.584004	0.174994	-1.285933
O	-1.466285	1.529643	0.017084
H	-2.351443	1.447394	-0.351076
H	0.435420	0.356866	1.572777
N	1.192963	1.762128	0.131879
H	1.782994	2.103820	0.881462
H	0.458437	2.446559	-0.007655
O	1.404053	-1.804045	0.404087
O	2.440611	-0.275424	-0.857111
H	2.311119	0.696335	-0.939861

IIg$^-$tt

17

F	0.014699	-0.008377	0.021165
C	0.016166	-0.010671	1.421260
C	1.467629	-0.006545	1.859132
C	1.632615	0.109671	3.376786
C	1.047313	-1.141063	4.045799
H	-0.494010	0.896127	1.751042

H	-0.519260	-0.892995	1.763759
H	1.946726	-0.933329	1.515830
O	2.140198	1.117010	1.302320
H	1.963587	1.113414	0.354693
H	1.047979	0.968951	3.720332
N	3.060450	0.211717	3.689459
H	3.200625	0.732274	4.547508
H	3.532470	0.720065	2.949622
O	-0.138977	-1.321290	4.169578
O	1.968496	-2.020830	4.447954
H	2.829675	-1.591800	4.246113

I'g$^-$gg$^-$

17

F	0.027513	-0.003030	0.002548
C	0.014549	-0.008229	1.394498
C	1.440772	-0.008308	1.911951
C	2.159755	-1.364963	1.791027
C	2.262949	-1.794561	0.337876
H	-0.493016	0.894173	1.735876
H	-0.535631	-0.891313	1.726105
H	2.014366	0.734152	1.338803
O	1.367440	0.346049	3.276893
H	2.229860	0.104932	3.646786
H	1.577564	-2.101867	2.347296
N	3.466447	-1.209745	2.426595
H	4.100992	-0.763350	1.769983
H	3.865248	-2.113820	2.649990
O	3.116136	-1.404582	-0.422874
O	1.293762	-2.664130	-0.006371
H	1.382724	-2.822659	-0.958388

I'g$^-$gg

17

F	-0.050322	0.056541	0.065484
C	0.009218	0.025321	1.452472
C	1.445605	-0.010026	1.927121
C	2.327833	1.099712	1.306478
C	1.879525	2.461262	1.808672
H	-0.507482	0.910563	1.828133
H	-0.503173	-0.876862	1.789948
H	1.430740	0.136714	3.019862
O	1.990582	-1.267136	1.603572
H	2.951762	-1.138654	1.624017
H	2.210942	1.049896	0.224002
N	3.712840	0.809710	1.663322
H	3.907973	1.189926	2.585081
H	4.350332	1.258854	1.017651
O	2.308944	2.990896	2.807201
O	0.923588	3.012804	1.034938
H	0.683284	3.856818	1.447185

I'g$^-$gt

17

F	-2.297758	-1.952821	-0.545778
C	-1.321622	-1.331622	0.219429
C	-1.135486	0.092666	-0.253023

C	0.077833	0.786442	0.411931
C	1.374482	0.113366	-0.003513
H	-1.652820	-1.339590	1.260358
H	-0.402785	-1.910994	0.116213
H	-0.967763	0.082853	-1.340260
O	-2.301145	0.815879	0.067045
H	-2.051385	1.748846	-0.012503
H	-0.038852	0.717035	1.495036
N	0.035809	2.190697	0.015511
H	0.453229	2.288425	-0.905754
H	0.581427	2.760760	0.650139
O	1.976338	0.367483	-1.019279
O	1.781357	-0.821515	0.882301
H	2.591867	-1.214160	0.522877

III'ggg$^-$

17

F	-0.005197	-0.021731	0.020356
C	0.005233	-0.008394	1.413703
C	1.445414	-0.004275	1.890282
C	2.190166	-1.268116	1.454462
C	3.656263	-1.158691	1.846722
H	-0.499454	0.896631	1.752438
H	-0.519642	-0.894773	1.767599
H	1.953485	0.864432	1.455155
O	1.388230	0.102359	3.304582
H	2.298044	0.237564	3.600032
H	2.132768	-1.355892	0.367134
N	1.566180	-2.448684	2.047903
H	1.574350	-2.370200	3.059204
H	2.060588	-3.292267	1.784915
O	4.113003	-0.405130	2.681032
O	4.419637	-2.055397	1.189958
H	5.321135	-1.963950	1.534044

III'ggg

17

F	-2.695350	-0.665242	0.217067
C	-2.124611	0.069532	-0.805462
C	-0.739188	0.573954	-0.450206
C	0.328797	-0.521857	-0.353212
C	1.694778	0.124845	-0.148518
H	-2.080160	-0.561762	-1.697060
H	-2.773864	0.927394	-0.991989
H	-0.445435	1.232363	-1.280898
O	-0.844026	1.312631	0.753208
H	0.003360	1.763080	0.865756
H	0.369046	-1.056612	-1.308657
N	0.025219	-1.478323	0.705669
H	-0.278604	-0.982529	1.536594
H	0.836379	-2.038707	0.936108
O	1.906277	1.280811	0.151659
O	2.690669	-0.772314	-0.301962
H	3.519435	-0.310063	-0.105243

III'ggt

17

F	-3.157828	0.258572	-0.516035
C	-2.065062	-0.559614	-0.254827
C	-0.814027	0.286864	-0.155424
C	0.439370	-0.582818	-0.278709
C	1.685700	0.294938	-0.256432
H	-2.249282	-1.085163	0.681135
H	-1.983035	-1.278665	-1.073028
H	-0.824930	1.011203	-0.977448
O	-0.844377	0.955699	1.096681
H	-0.166321	1.641389	1.058265
H	0.428817	-1.094574	-1.246603
N	0.441855	-1.594087	0.775561
H	0.364311	-1.134303	1.677194
H	1.302928	-2.127159	0.764683
O	1.718333	1.486554	-0.041928
O	2.800925	-0.430585	-0.485308
H	3.551481	0.179189	-0.423273

IItg$^-$g$^-$

17

F	1.046351	1.593984	1.913193
C	1.522118	0.395396	1.359743
C	1.340862	0.399872	-0.153916
C	-0.122721	0.550036	-0.612524
C	-0.990608	-0.575197	-0.031831
H	2.581535	0.312236	1.599631
H	0.971929	-0.425017	1.823914
H	1.904284	1.239283	-0.570898
O	1.916238	-0.782510	-0.671408
H	1.285309	-1.495347	-0.477434
H	-0.120564	0.379997	-1.693137
N	-0.753831	1.844636	-0.344829
H	-0.363132	2.272520	0.486559
H	-0.628637	2.484000	-1.117733
O	-0.598508	-1.719159	0.066078
O	-2.208985	-0.197044	0.336860
H	-2.223317	0.773806	0.172540

IItg$^-$g

17

F	0.575953	0.030118	2.063887
C	1.579259	0.634651	1.305940
C	1.363458	0.384002	-0.181577
C	-0.116033	0.534800	-0.596153
C	-0.982417	-0.649820	-0.133684
H	1.560922	1.702296	1.541810
H	2.537492	0.211370	1.607082
H	1.932076	1.162065	-0.704873
O	1.895236	-0.848642	-0.607534
H	1.216357	-1.523177	-0.430055
H	-0.140713	0.477982	-1.688407
N	-0.757509	1.789164	-0.199007
H	-0.494688	2.061483	0.740837
H	-0.521754	2.547219	-0.825449
O	-0.596058	-1.797228	-0.136253
O	-2.226569	-0.320236	0.205283
H	-2.266688	0.655446	0.122961

IItg$^-$t

17

F	2.780256	0.648757	1.712521
C	1.462204	0.395976	1.366543
C	1.326600	0.381709	-0.150256
C	-0.137456	0.554507	-0.603427
C	-1.030882	-0.537388	0.001502
H	0.851692	1.181848	1.824774
H	1.178089	-0.569783	1.787816
H	1.897675	1.227403	-0.543766
O	1.895711	-0.777293	-0.717691
H	1.316409	-1.513994	-0.471377
H	-0.144541	0.365681	-1.680353
N	-0.733896	1.872111	-0.367565
H	-0.363807	2.308939	0.468724
H	-0.554490	2.499842	-1.140136
O	-0.653500	-1.677044	0.158725
O	-2.260219	-0.135183	0.319842
H	-2.271520	0.825403	0.118935

III'g$^-$gg$^-$

17

F	0.119580	-2.366637	0.708702
C	1.233708	-1.763998	0.131750
C	1.168483	-0.266064	0.360103
C	0.075147	0.429169	-0.465179
C	-1.317898	-0.087893	-0.150310
H	2.128519	-2.169467	0.604600
H	1.237363	-2.003311	-0.933730
H	0.975319	-0.083109	1.426040
O	2.428476	0.245246	-0.024253
H	2.290982	1.195464	-0.152103
H	0.259337	0.216632	-1.519317
N	0.225399	1.872488	-0.253737
H	-0.281839	2.394793	-0.958029
H	-0.154540	2.134427	0.650149
O	-2.035509	-0.672529	-0.922804
O	-1.685714	0.222079	1.115794
H	-2.569171	-0.152133	1.247460

III'g$^-$gg

17

F	1.363222	-2.024921	-1.238157
C	1.214910	-1.783773	0.120013
C	1.144593	-0.296764	0.393792
C	0.057572	0.412270	-0.444704
C	-1.316339	-0.107396	-0.058888
H	0.302804	-2.286325	0.450553
H	2.080445	-2.206266	0.632744
H	0.896403	-0.174893	1.460750
O	2.400354	0.273575	0.109729
H	2.225306	1.213035	-0.053252
H	0.219823	0.149963	-1.489500
N	0.237089	1.854046	-0.284089
H	-0.224528	2.362066	-1.027991
H	-0.161965	2.168143	0.593711
O	-1.940539	-0.951608	-0.651999
O	-1.769713	0.471952	1.082464
H	-2.631796	0.072052	1.269960

III'g$^-$gt

17

F	2.036091	-2.399390	0.949605
C	1.182788	-1.772149	0.052933
C	1.156269	-0.285510	0.329337
C	0.037088	0.427579	-0.461998
C	-1.330047	-0.100970	-0.065173
H	0.192999	-2.218513	0.169041
H	1.557659	-1.956309	-0.956121
H	0.975498	-0.133037	1.403352
O	2.400498	0.258207	-0.049221
H	2.253039	1.212545	-0.125458
H	0.168963	0.199960	-1.520986
N	0.214585	1.865716	-0.269901
H	-0.290942	2.392411	-0.971324
H	-0.140099	2.146832	0.637900
O	-1.986546	-0.887529	-0.703367
O	-1.734678	0.397967	1.127111
H	-2.595524	-0.002654	1.319408

Igtg$^-$

17

F	2.096190	0.176544	-1.956437
C	2.054566	0.303756	-0.570153
C	0.755388	-0.298416	-0.059540
C	-0.461630	0.502625	-0.524545
C	-1.744195	-0.213525	-0.107983
H	2.910405	-0.238280	-0.159252
H	2.115333	1.361005	-0.317246
H	0.663613	-1.322683	-0.430938
O	0.737786	-0.266637	1.366356
H	1.338611	-0.938171	1.701869
H	-0.440377	0.503413	-1.618879
N	-0.381875	1.867051	-0.043410
H	-0.461924	1.864844	0.968160
H	-1.172798	2.398310	-0.387315
O	-2.639481	0.278582	0.531771
O	-1.785382	-1.484982	-0.572750
H	-2.630769	-1.855725	-0.279662

Igtt

17

F	0.006629	0.005336	0.019626
C	0.009671	0.007802	1.419834
C	1.454623	-0.004762	1.872409
C	1.565468	-0.021762	3.396377
C	3.036232	-0.093419	3.806456
H	-0.505208	0.905710	1.755006
H	-0.521076	-0.884817	1.757267
H	1.941008	-0.899428	1.470214
O	2.132877	1.164062	1.428629
H	2.000956	1.235349	0.476960
H	1.103241	-0.950046	3.750016
N	0.861494	1.111016	3.965967
H	1.277847	1.962544	3.601878
H	1.011068	1.135853	4.968194
O	3.585529	0.679715	4.549755
O	3.655313	-1.167053	3.264879
H	4.571924	-1.143277	3.576761

IIgtt

17

F	0.004184	0.003437	0.022724
C	0.008655	0.006930	1.417322
C	1.446988	0.001354	1.890092
C	1.545324	0.083817	3.413320
C	2.953303	-0.361875	3.858592
H	-0.517098	0.908143	1.740466
H	-0.522274	-0.885662	1.756119
H	1.932793	-0.923651	1.566183
O	2.141945	1.135861	1.387472
H	2.110756	1.100837	0.425695
H	0.865493	-0.653491	3.850662
N	1.239086	1.406078	3.962604
H	1.433315	2.118685	3.266478
H	0.273301	1.488971	4.248331
O	3.488862	-1.347020	3.418218
O	3.493961	0.412046	4.801554
H	2.837032	1.127157	4.945381

IIIgtg$^-$

17

F	2.094076	0.183512	-1.993021
C	2.079718	0.318989	-0.606305
C	0.803616	-0.300057	-0.061721
C	-0.441926	0.446460	-0.534899
C	-1.664671	-0.394773	-0.163872
H	2.952645	-0.203089	-0.208301
H	2.127084	1.379543	-0.362878
H	0.725791	-1.338164	-0.392063
O	0.818767	-0.214226	1.365036
H	1.236574	-1.003449	1.719054
H	-0.414403	0.456192	-1.629479
N	-0.416516	1.815368	-0.054497
H	-0.459412	1.821724	0.958387
H	-1.224555	2.324152	-0.389336
O	-1.748666	-1.587419	-0.337112
O	-2.674224	0.339326	0.351989
H	-3.395842	-0.283922	0.524606

IIIgtg

17

F	2.476723	1.236372	0.344578
C	2.094829	0.215506	-0.525446
C	0.759583	-0.371854	-0.082800
C	-0.460212	0.417780	-0.584015
C	-1.699327	-0.409983	-0.238425
H	2.037821	0.633759	-1.531072
H	2.874930	-0.545626	-0.475270
H	0.680330	-1.378235	-0.501040
O	0.724898	-0.509531	1.329397
H	1.237355	0.219378	1.698889
H	-0.413585	0.413978	-1.678016
N	-0.429271	1.790310	-0.120902
H	-0.490993	1.812824	0.890804
H	-1.236327	2.296924	-0.463891
O	-1.897854	-1.523952	-0.661584
O	-2.559727	0.229258	0.580700
H	-3.286173	-0.391185	0.743088

IIIgtt

17

F	0.008735	0.005199	0.019334
C	0.010615	0.004688	1.419011
C	1.454032	-0.000715	1.874797
C	1.566149	0.002086	3.398801
C	3.041791	-0.208297	3.749019
H	-0.509453	0.899520	1.755391
H	-0.514850	-0.892206	1.753568
H	1.948296	-0.899804	1.492413
O	2.130791	1.165670	1.419914
H	2.022468	1.211886	0.463870
H	1.057410	-0.894227	3.768091
N	0.913279	1.177484	3.945281
H	1.355462	2.008488	3.567747
H	1.029842	1.213035	4.950316
O	3.683845	-1.166323	3.392008
O	3.550632	0.773649	4.521074
H	4.478408	0.540041	4.674954

IIg$^-$g$^-$g$^-$

17

F	-0.609027	-1.760453	-1.501080
C	-1.410902	-0.993890	-0.670040
C	-0.796099	0.393277	-0.528167
C	0.338279	0.488456	0.502713
C	1.353984	-0.646672	0.332166
H	-2.392138	-0.896726	-1.137299
H	-1.493911	-1.512723	0.288383
H	-0.405826	0.681362	-1.507565
O	-1.792772	1.365105	-0.221025
H	-2.187112	1.142420	0.630061
H	-0.102453	0.347745	1.497655
N	1.002619	1.779480	0.321494
H	1.491362	2.061580	1.163641
H	0.311046	2.493388	0.119146
O	1.214170	-1.726318	0.845944
O	2.402838	-0.329373	-0.438808
H	2.305219	0.623092	-0.637232

IIg$^-$g$^-$g

17

F	-0.001857	-0.006064	0.028559
C	0.014684	0.013217	1.429957
C	1.464993	0.010037	1.872537
C	2.188687	1.336035	1.540836
C	1.634120	2.432653	2.459534
H	-0.513473	0.904802	1.761749
H	-0.484878	-0.894510	1.769598
H	1.504828	-0.151832	2.954001
O	2.149013	-1.089750	1.288417
H	1.866643	-1.141159	0.366617
H	1.950697	1.616004	0.510537
N	3.621088	1.167462	1.791061
H	4.168841	1.764525	1.183191
H	3.884088	0.207089	1.597359
O	0.556019	2.949703	2.295720
O	2.435076	2.728730	3.488380
H	3.241567	2.186338	3.339099

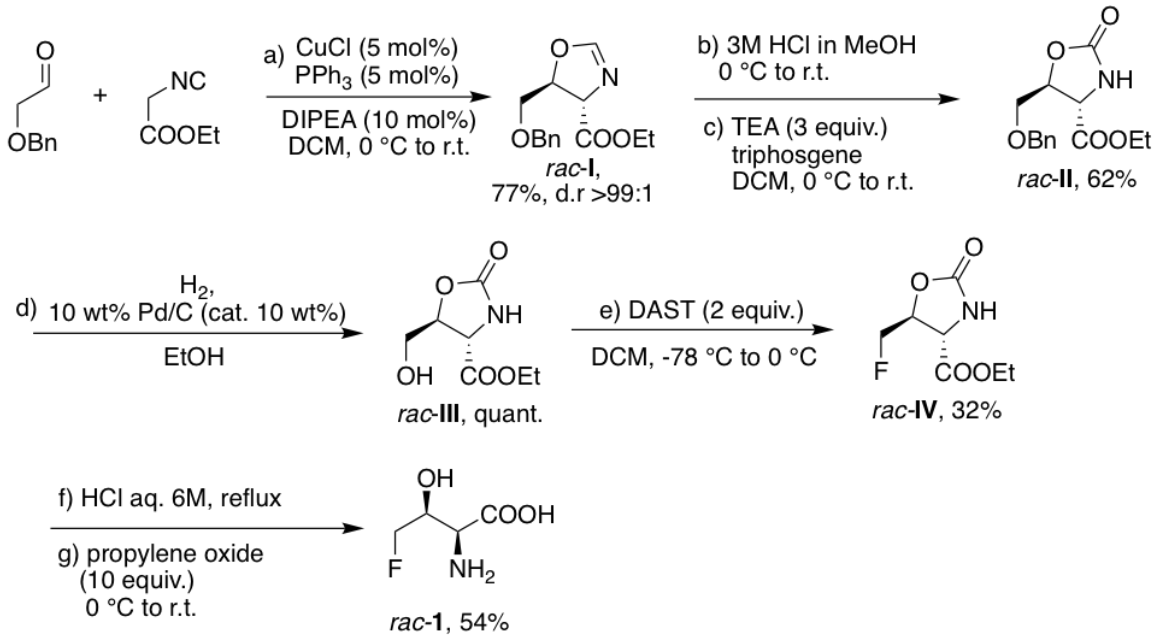

Figure 2: Diastereoisomeric synthesis of 4FT

2 Experimental details

2.1 Synthesis and characterization

An effective racemic, but diastereomeric synthesis of 4FT was recently proposed by some of the authors [3], and involved a fluorination step (see Figure 2) which was performed with DAST (diethylaminosulphur trifluoride) in order to reach the maximum possible yield. To obtain the required amount of 4FT, it was necessary to perform the reaction in a limited scale (200-300 mg of starting material) and to repeat it five times in order to avoid extensive decomposition upon DAST addition (experienced on larger scale). The produced 4FT sample was repeatedly washed with methanol and accurately dried under vacuum, thus leading to a spectroscopically pure compound as white powder.

White powder. Exp. m.p. 178–179°C (with decomposition, loss of CO_2), lit. m.p. 182–183°C.[31] ^{1}H NMR (400 MHz, D_2O) δ: 4.72 (1H, ddd, J = 46.6, 10.8, 3.7 Hz), 4.60 (1H, ddd, J = 47.0, 10.8, 3.1 Hz), 4.35 (1H, dddd, J = 24.9, 4.7, 3.7, 3.1 Hz), 3.87 (1H, d, J = 4.7 Hz). ^{19}F NMR (376.5 Hz, D_2O) δ: -232 (dt, J = 46.8, 24.9 Hz). ^{13}C NMR (101 MHz, D_2O) δ: 172.5, 85.6 (d, J = 167.3 Hz), 68.4 (d, J = 19.1 Hz), 56.8 (d, J = 4.4 Hz). ESI-MS m/z: 138.2 [M + H]+ (100); 92.2 [M – COOH]+ (15). Spectroscopic data agree with those reported in the literature.

2.2 Rotational spectroscopy

Amino acids, including their fluorinated derivatives, are molecules characterized by high melting points and decomposition at high temperatures. These characteristics make it impossible to transfer them into the gas phase using conventional heating methods. In order to lead 4FT molecules to the gas phase, we have employed the same experimental approach that has been successfully used for

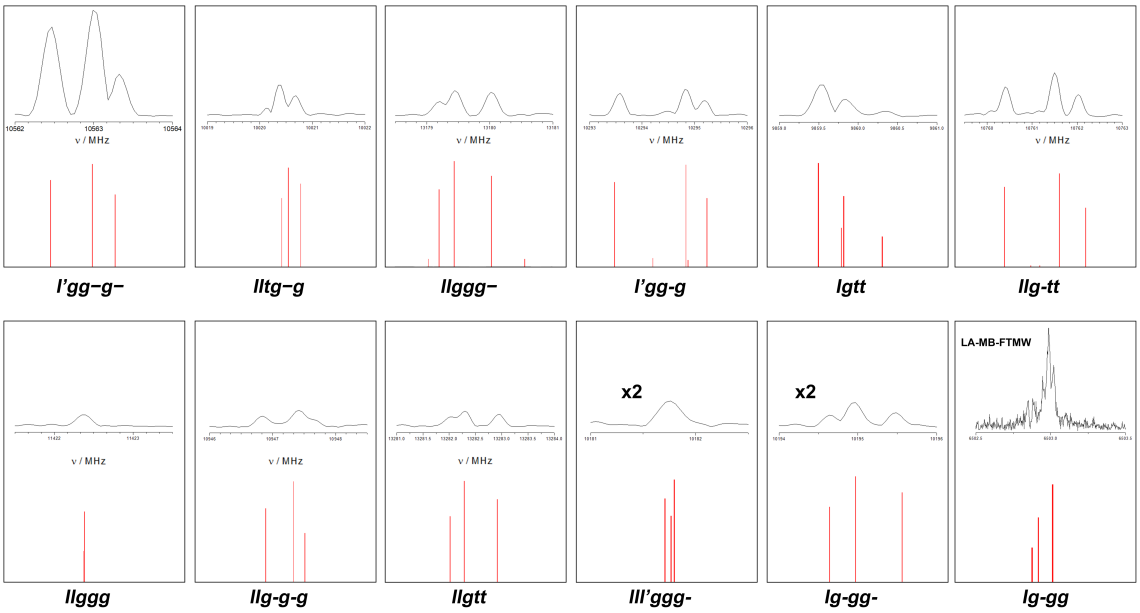

Figure 3: Examples of rotational transitions belonging to the twelve conformers of 4FT. For each transition, the experimental spectrum (top) and its predicted hyperfine pattern (bottom) are shown. Except for the Ig^{-}gg conformer, the experimental spectra have been obtained using the LA-CP-FTMW spectrometer.

the neutral forms of many amino acids [32, 33, 34, 35, 36], which is based on laser ablation (LA) technique. This makes it possible to characterize any solid molecule in the isolated conditions of a supersonic expansion [37].

To the purpose of exploiting the LA technique, 4FT was finely powdered, mixed with a small amount of copolymeric binder, and finally pressed to form a cylindrical rod. This rod was kept in the dark and under vacuum conditions to favor its compaction and to avoid decomposition due to light or moisture. After some time, the sample rod was placed in the ablation nozzle of a newly constructed Laser-ablation Chirped-pulse Fourier-Transform Microwave (LA-CP-FTMW) spectrometer based on previous designs [38, 39] and designed to maximize the performance in the 6 to 14 GHz range. Then, a picosecond Nd:YAG laser (355 nm of wavelength, 20 mJ per pulse, 25 ps pulse width) was used to vaporize the sample, transferring the molecules of 4FT into the gas phase. The laser ablation products were supersonically expanded into a high-vacuum chamber utilizing a flow of carrier gas (Ne, 10 bar). Chirped pulses of 4 μs in the 3–7 GHz range were generated digitally using a 24 GS/s arbitrary waveform generator. The pulse frequency was doubled (6-14 GHz) using an active multiplier and amplified to about 300 W peak power using a traveling wave tube amplifier. Broadband microwave horn antennas then transmitted the resultant pulses into the vacuum chamber, interacting with the molecular supersonic expansion macroscopically polarizing 4FT molecules. The free induction decays (FIDs) were collected using a second horn antenna, amplified, and coherently averaged using a 50 GS/s digital oscilloscope. At a repetition rate of 2 Hz, up to 90 000 FIDs were averaged. The frequency-domain spectrum in the 6–14 GHz frequency range was obtained by taking a fast Fourier transform (FFT) following a Kaiser–Bessel window to improve the baseline resolution.

As stated in the main text, this LA-CP-FTMW spectrometer allowed the characterization of 11 conformers of 4FT. However, line intensities of the eleventh conformer close to the sensibility limit of this technique forced us to move to our LA-MB-FTMW spectrometer [40] to continue the search for additional conformers. This spectrometer is based on a Fabry-Perot resonator and is dedicated to maximizing performances at low-frequency ranges (2 to 8 GHz). Hence, the sample was placed on the ablation nozzle of the LA-MB-FTMW spectrometer. Again, a Nd:YAG picosecond laser working on the third harmonic was used as a vaporization tool. The ablated sample was seeded in a flow of Ne (10 bar of stagnation pressure) and supersonically expanded between the two aluminum mirrors of the Fabry–Pérot resonator. Then, a microwave pulse was applied to produce a macroscopic polarization of the species in the jet. Once the excitation ceases, 4 FIDs per gas pulse were captured in the time domain and coherently averaged. After a Fourier transformation, the frequency-domain spectrum was obtained. This technique makes measured transitions appear as Doppler doublets because the supersonic jet travels parallel to the resonator axis. Rest frequencies are therefore calculated as the arithmetic mean of the Doppler doublets and obtained with an accuracy better than 3 kHz.

The assignment of the 12 conformers of 4FT is described in detail in the Methods section of the main text. As mentioned there, for all conformers, the fit of the experimental lines exploits the asymmetric-top rigid-rotor Watson Hamiltonian combined, if required, with a term accounting for the nuclear quadrupole coupling.[39] All measured transitions and hyperfine components are collected in Tables S7 to S18 of the SI.

Rotational transitions of the observed 4FT conformers The assigned rotational transitions are classified according to the selection rules for the principal quantum number J (ΔJ = -1, 0, + 1 leads to P-, Q- and R-branches, respectively) and non-zero dipole moment component (a-, b- and c-type refers to μ_a, μ_b and μ_c dipole moment components, respectively). For asymmetric rotors like 4FT, rotational energy levels are identified by two further quantum numbers in addition to J, namely K_{-1} and K_{+1}: $J_{K_{-1},K_{+1}}$. Transitions are usually denoted by reporting the two energy levels connected, with the final energy level given as first.

Table S7. Measured frequencies for the hyperfine components and rotational transitions belonging to the conformer 1 (I'gg-g-) of 4-F-Thr obtained using LA-MB-FTMW spectroscopy. An uncertainty of 50 and 100 KHz was given for fitting the hyperfine-resolved and non-resolved transitions, respectively.

J'	K'_{-1}	K'_{+1}	F'	J''	K''_{-1}	K''_{+1}	F''	ν_{obs} / MHz	$\nu_{obs} - \nu_{calc}$ / MHz
3	1	2	3	2	0	2	2	7677.611	-0.008
3	1	2	4	2	0	2	3	7678.171	-0.008
2	2	0	1	1	1	0	0	8453.794	-0.002
2	2	0	2	1	1	0	1	8454.947	0.003
2	2	1	2	1	1	1	1	8494.334	0.006
2	2	1	3	1	1	1	2	8494.855	0.004
6	0	6	7	5	1	4	6	10328.004	0.007
3	2	1	3	2	1	1	2	10444.412	-0.003
3	2	2	3	2	1	2	2	10562.455	0.002
3	2	2	3	2	1	2	3	10563.007	0.000
3	2	2	2	2	1	2	1	10563.321	0.006
4	2	2	5	3	1	2	4	12417.928	-0.004
4	2	2	4	3	1	2	3	12418.366	-0.002
4	2	3	3	3	1	3	2	12651.724	-0.004
4	0	4		3	1	2		6499.565	-0.076
5	0	5		4	1	3		8429.391	-0.004
5	1	4		4	0	4		11923.323	0.099
3	3	0		2	2	0		13447.710	-0.017
3	3	1		2	2	1		13448.561	-0.011

Table S8. Measured frequencies for the hyperfine components of rotational transitions

J'	K'_{-1}	K'_{+1}	F'	J''	K''_{-1}	K''_{+1}	F''	ν_{obs} / MHz	$\nu_{obs} - \nu_{calc}$ / MHz
2	2	0	3	1	1	0	2	6950.953	0.017
2	2	0	2	1	1	0	1	6951.484	0.030
2	2	0	2	1	1	0	2	6951.789	0.025
2	2	1	3	1	1	1	2	7152.312	0.022
2	2	1	2	1	1	1	1	7152.676	0.011
2	2	1	2	1	1	1	2	7153.254	-0.016
3	1	3	3	2	0	2	2	7402.676	0.031
3	1	3	4	2	0	2	3	7402.415	0.039
3	1	3	2	2	0	2	1	7402.126	0.014
3	2	2	3	2	1	1	2	9080.112	0.032
3	2	2	2	2	1	1	1	9079.087	0.016
3	2	2	4	2	1	1	3	9079.441	0.009
4	1	3	4	3	2	2	4	9497.319	0.023
4	1	3	5	3	2	2	4	9498.176	0.039
3	2	1	3	2	1	1	2	9570.574	0.017
3	2	1	4	2	1	1	3	9570.219	0.021
3	2	1	2	2	1	1	1	9569.960	0.021
3	2	2	2	2	1	2	1	10020.127	-0.016
3	2	2	4	2	1	2	3	10020.376	0.049
3	2	2	3	2	1	2	2	10020.672	0.012
4	2	3	4	3	2	2	3	10127.173	0.044
4	2	3	5	3	2	2	4	10127.567	0.062
3	3	1	3	2	2	0	3	10757.823	-0.016
3	3	1	3	2	2	0	2	10757.002	-0.009
3	3	1	4	2	2	0	3	10756.610	0.002
3	3	0	2	2	2	0	2	10782.485	-0.012
3	3	0	4	2	2	0	3	10783.737	-0.007
3	3	0	3	2	2	0	2	10784.108	-0.004
4	2	2	4	3	2	1	3	10839.267	-0.002
4	2	2	5	3	2	1	4	10839.684	-0.019
4	2	3	3	3	1	2	2	11161.339	-0.010
4	2	3	5	3	1	2	4	11161.576	0.026
4	2	3	4	3	1	2	3	11162.043	0.017
5	1	4	5	4	2	3	4	12214.323	-0.012
5	1	4	6	4	2	3	5	12214.665	-0.053
4	3	2	5	3	2	1	4	13098.287	-0.045
4	3	2	4	3	2	1	3	13098.880	-0.024
5	2	4	6	4	1	3	5	13154.627	-0.054
5	2	4	5	4	1	3	4	13154.938	-0.049
5	3	2	5	4	3	1	4	13371.026	-0.053
5	3	2	6	4	3	1	5	13371.649	-0.017
4	3	2	5	3	2	2	4	13589.152	0.053
4	3	2	4	3	2	2	3	13589.321	-0.059

Table S9. Measured frequencies for the hyperfine components and rotational transitions belonging to the conformer 3 (IIggg−) of 4-F-Thr obtained using LA-MB-FTMW spectroscopy. An uncertainty of 50 KHz was given for fitting the hyperfine-resolved transitions, while uncertainties between 50 and 200 KHz were given for fitting non-resolved transitions.

J'	K'_{-1}	K'_{+1}	F'	J''	K''_{-1}	K''_{+1}	F''	ν_{obs} / MHz	$\nu_{obs} - \nu_{calc}$ / MHz
3	1	2	2	2	0	2	1	7986.550	-0.005
4	2	3	4	3	2	2	3	8058.625	0.003
4	3	2	5	3	3	1	4	8075.991	0.001
4	3	1	3	3	3	0	2	8077.603	0.006
4	3	1	5	3	3	0	4	8077.275	-0.003
2	2	1	3	1	1	0	2	8743.931	0.003
2	2	1	3	1	1	1	2	8859.491	0.002
2	2	0	3	1	1	1	2	8865.784	-0.003
3	2	2	2	2	1	2	1	10990.709	0.005
3	2	2	4	2	1	2	3	10991.098	0.006
3	2	1	2	2	1	2	1	11022.060	-0.003
4	2	3	3	3	1	3	2	13179.216	0.003
4	2	3	5	3	1	3	4	13179.445	-0.006
4	2	3	4	3	1	3	3	13180.026	-0.002
3	0	3		2	0	2		6022.810	0.036
4	0	4		3	1	3		6731.283	0.002
3	1	3		2	0	2		7293.457	0.073
4	1	4		3	1	3		7820.564	0.015
4	0	4		3	0	3		8001.905	0.014
4	1	3		3	1	2		8281.829	0.014
5	0	5		4	1	4		8869.784	0.032
4	1	4		3	0	3		9091.191	0.031
5	1	5		4	1	4		9765.328	0.067
5	0	5		4	0	4		9958.974	-0.047
5	2	4		4	2	3		10065.800	0.039
5	2	3		4	2	2		10186.496	0.044
5	1	4		4	1	3		10339.005	0.040
3	2	2		2	1	1		10644.295	-0.037
3	2	1		2	1	1		10675.640	-0.076
5	1	5		4	0	4		10854.540	0.010
6	0	6		5	1	5		10997.709	0.011
6	1	6		5	1	5		11704.084	-0.009
6	0	6		5	0	5		11893.199	-0.008
6	2	5		5	2	4		12067.395	-0.020
6	4	2		5	4	1		12117.270	-0.017
6	2	4		5	2	3		12269.644	-0.058
6	1	5		5	1	4		12386.008	-0.033
6	1	6		5	0	5		12599.588	-0.014
7	0	7		6	1	6		13100.763	-0.045
4	2	2		3	1	3		13272.703	0.045
7	1	7		6	1	6		13636.786	-0.035
7	0	7		6	0	6		13807.164	-0.039

Table S10. Measured frequencies for the hyperfine components of rotational transitions belonging to the conformer 4 (I'gg−g) of 4-F-Thr obtained using LA-CP-FTMW spectroscopy. An uncertainty of 50 KHz was given for measured transitions.

J'	K'_{-1}	K'_{+1}	F'	J''	K''_{-1}	K''_{+1}	F''	ν_{obs} / MHz	$\nu_{obs} - \nu_{calc}$ / MHz
3	1	2	3	2	0	2	2	8098.624	0.070
3	1	2	3	2	0	2	3	8098.985	0.003
3	1	2	4	2	0	2	3	8099.820	-0.002
3	1	2	2	2	0	2	1	8100.312	-0.043
2	2	0	1	1	1	0	0	9417.789	0.070
2	2	0	2	1	1	0	2	9418.904	0.044
2	2	0	3	1	1	0	2	9419.281	-0.007
2	2	0	2	1	1	0	1	9420.030	-0.036
2	2	1	2	1	1	1	1	9511.200	0.051
2	2	1	1	1	1	1	1	9511.706	-0.020
2	2	1	3	1	1	1	2	9512.343	-0.036
4	1	3	4	3	0	3	3	10293.571	0.060
4	1	3	5	3	0	3	4	10294.833	0.000
4	1	3	3	3	0	3	2	10295.183	-0.039
3	2	1	2	2	1	1	1	11331.947	0.021
3	2	1	4	2	1	1	3	11332.385	0.000
3	2	1	3	2	1	1	2	11333.124	-0.032
3	2	2	3	2	1	2	2	11602.248	0.051
3	2	2	3	2	1	2	3	11603.455	-0.034
3	2	2	2	2	1	2	1	11604.125	-0.082
4	2	2	5	3	1	2	4	13219.378	0.011
4	2	2	4	3	1	2	3	13220.097	-0.007
4	2	3	4	3	1	3	3	13740.989	0.038
4	2	3	3	3	1	3	4	13742.187	-0.002
4	2	3	3	3	1	3	2	13742.644	-0.054

Table S11. Measured frequencies for the hyperfine components of rotational transitions belonging to the conformer 5 (Igtt) of 4-F-Thr obtained using LA-CP-FTMW spectroscopy. An uncertainty of 50 KHz was given for measured transitions.

J'	K'_{-1}	K'_{+1}	F'	J''	K''_{-1}	K''_{+1}	F''	ν_{obs} / MHz	$\nu_{obs} - \nu_{calc}$ / MHz
4	2	3	5	3	1	3	4	13744.474	0.060
4	2	2	4	3	1	2	3	13210.795	-0.006
4	2	2	5	3	1	2	4	13210.365	-0.017
4	2	2	3	3	1	2	2	13210.162	-0.037
4	2	3	5	3	1	2	4	13158.509	-0.017
4	2	3	4	3	1	2	3	13158.972	-0.013
5	1	4	6	4	0	4	5	11918.416	-0.076
3	2	1	4	2	1	2	3	11795.168	-0.034
3	2	2	4	2	1	2	3	11777.825	-0.021
3	2	2	3	2	1	2	2	11778.051	-0.039
5	1	5	6	4	0	4	5	10455.896	0.038
2	2	0	2	1	1	1	1	9863.302	0.025
2	2	0	3	1	1	1	2	9863.024	0.021
2	2	1	2	1	1	1	2	9860.343	0.000
2	2	1	2	1	1	1	1	9859.838	0.030
2	2	1	3	1	1	1	2	9859.532	0.008
2	2	1	1	1	1	1	1	9858.551	0.016
2	2	1	2	1	1	0	1	9762.438	0.020
2	2	1	3	1	1	0	2	9761.820	-0.007

Table S12. Measured frequencies for the hyperfine components of rotational transitions belonging to the conformer 6 (IIg−tt) of 4-F-Thr obtained using LA-CP-FTMW spectroscopy. An uncertainty of 50 KHz was given for measured transitions.

J'	K'_{-1}	K'_{+1}	F'	J''	K''_{-1}	K''_{+1}	F''	ν_{obs} / MHz	$\nu_{obs} - \nu_{calc}$ / MHz
6	2	5	6	5	1	4	5	13899.454	-0.012
6	2	5	6	5	1	4	5	13899.454	-0.012
6	2	5	7	5	1	4	6	13898.760	-0.032
4	2	2	3	3	1	3	2	13587.314	0.000
4	2	2	5	3	1	3	4	13586.896	0.007
4	2	2	4	3	1	3	3	13585.524	0.010
3	3	0	4	2	2	1	3	12885.485	-0.035
3	3	1	4	2	2	0	3	12834.321	0.031
7	1	7	6	6	0	6	5	12618.455	0.005
5	2	4	5	4	1	3	4	12599.586	0.035
5	2	4	6	4	1	3	5	12598.853	0.040
5	2	4	4	4	1	3	3	12598.597	-0.031
7	0	7	8	6	1	6	7	12422.246	-0.026
4	2	3	4	3	1	2	3	11196.035	0.012
4	2	3	5	3	1	2	4	11195.259	0.000
4	2	3	3	3	1	2	2	11194.998	0.005
6	1	6	7	5	0	5	6	11017.518	0.019
3	2	1	3	2	1	2	2	10760.411	0.011
3	2	1	4	2	1	2	3	10761.499	0.004
3	2	1	2	2	1	2	1	10762.025	-0.001
6	0	6	7	5	1	5	6	10638.864	-0.016
3	2	2	2	2	1	1	1	9654.178	0.011
3	2	2	3	2	1	1	3	9654.600	0.010
3	2	2	3	2	1	1	2	9655.359	0.007
5	1	5	5	4	0	4	4	9471.210	0.024
5	1	5	6	4	0	4	5	9470.843	0.016
5	0	5	6	4	1	4	5	8786.142	-0.071
2	2	0	2	1	1	1	1	8302.100	-0.024
2	2	0	3	1	1	1	2	8302.929	-0.016
2	2	0	1	1	1	1	0	8304.056	-0.016
4	1	4	4	3	0	3	3	7972.241	0.015
4	1	4	5	3	0	3	4	7971.691	-0.016
2	2	1	2	1	1	0	1	7963.140	-0.024
2	2	1	3	1	1	0	2	7962.441	-0.009
2	2	1	1	1	1	0	0	7961.423	0.001
4	0	4	5	3	1	3	4	6838.783	0.017
4	0	4	4	3	1	3	3	6838.588	0.029
3	1	3	3	2	0	2	2	6475.736	0.037
3	1	3	4	2	0	2	3	6475.044	0.012

Table S13. Measured frequencies for the rotational transitions belonging to the conformer 7 (IIggg) of 4-F-Thr obtained using LA-CP-FTMW spectroscopy. An uncertainty of 50 KHz was given for measured transitions.

J'	K'_{-1}	K'_{+1}	J''	K''_{-1}	K''_{+1}	ν_{obs} / MHz	$\nu_{obs} - \nu_{calc}$ / MHz
4	1	4	3	1	3	7640.428	-0.006
4	0	4	3	0	3	7872.674	0.064
4	2	2	3	2	1	8045.343	0.059
4	1	3	3	1	2	8247.760	-0.033
5	0	5	4	0	4	9778.749	0.017
5	2	4	4	2	3	9932.794	0.041
5	2	3	4	2	2	10107.145	0.051
5	1	4	4	1	3	10290.171	0.060
6	1	6	5	1	5	11422.374	0.004
6	2	5	5	2	4	11902.536	-0.023
6	2	4	5	2	3	12192.292	-0.086
6	1	5	5	1	4	12317.369	-0.040
7	1	7	6	1	6	13300.719	-0.010
7	0	7	6	0	6	13502.546	-0.035

Table S14. Measured frequencies for the hyperfine components and rotational transitions belonging to the conformer 8 (IIg−g−g) of 4-F-Thr obtained using LA-CP-FTMW spectroscopy. An uncertainty of 50 KHz was given for fitting the hyperfine-resolved transitions, while uncertainties between 50 and 300 KHz were given for fitting non-resolved transitions.

J'	K'_{-1}	K'_{+1}	F'	J''	K''_{-1}	K''_{+1}	F''	ν_{obs} / MHz	$\nu_{obs} - \nu_{calc}$ / MHz
4	1	4		3	1	3		7904.041	0.061
4	1	3		3	1	2		9144.622	-0.017
5	1	5		4	0	4		10101.340	0.104
4	2	3		3	1	2		11359.976	-0.164
5	2	3		4	2	2		11533.894	-0.011
6	0	6		5	1	5		11612.126	-0.026
6	1	6		5	0	5		11832.387	0.001
3	3	1		2	2	0		12592.403	-0.080
3	3	0		2	2	1		12662.651	-0.016
6	2	5		5	2	4		12704.242	0.096
7	0	7		6	1	6		13511.496	-0.050
7	1	7		6	0	6		13613.457	-0.081
3	1	3	4	2	0	2	3	6776.985	0.024
3	1	3	3	2	0	2	2	6777.521	-0.008
3	2	2	2	2	1	1	1	9684.189	0.040
3	2	2	4	2	1	1	3	9684.635	0.030
3	2	2	3	2	1	1	2	9685.461	0.036
3	2	1	3	2	1	2	2	10946.838	0.000
3	2	1	4	2	1	2	3	10947.416	-0.006
5	2	4	6	4	1	3	5	12892.026	0.021
5	2	4	5	4	1	3	4	12892.688	0.041

Table S15. Measured frequencies for the hyperfine components and rotational transitions belonging to the conformer 9 (IIgtt) of 4-F-Thr obtained using LA-CP-FTMW spectroscopy. An uncertainty of 50 and 100 KHz was given for fitting the hyperfine-resolved and non-resolved transitions, respectively.

J'	K'_{-1}	K'_{+1}	F'	J''	K''_{-1}	K''_{+1}	F''	ν_{obs} / MHz	$\nu_{obs} - \nu_{calc}$ / MHz
5	1	4		4	1	3		9110.363	-0.018
6	1	5		5	1	4		10923.505	0.019
7	1	7		6	1	6		12097.471	0.003
7	0	7		6	0	6		12297.882	0.007
7	2	6		6	2	5		12430.126	0.013
7	2	5		6	2	4		12584.971	0.004
7	1	6		6	1	5		12731.172	0.020
8	1	8		7	1	7		13813.966	-0.025
3	1	3	2	2	0	2	1	7210.081	0.028
3	1	3	4	2	0	2	3	7210.504	0.009
3	1	3	3	2	0	2	2	7210.955	-0.010
4	1	4	4	3	0	3	3	8812.722	0.026
4	1	4	5	3	0	3	4	8812.323	-0.026
2	2	1	3	1	1	0	2	9953.651	0.019
2	2	1	2	1	1	0	1	9954.530	0.018
2	2	0	2	1	1	1	1	10048.745	0.001
2	2	0	3	1	1	1	2	10048.196	0.012
5	1	5	4	4	0	4	5	10379.577	0.049
5	1	5	6	4	0	4	5	10379.254	-0.044
3	2	2	3	2	1	1	2	11641.923	0.004
3	2	2	4	2	1	1	3	11641.079	0.007
3	2	2	2	2	1	1	1	11640.600	-0.002
3	2	1	4	2	1	2	3	11930.502	0.018
3	2	1	3	2	1	2	2	11930.951	-0.024
4	2	3	4	3	1	2	3	13282.942	-0.021
4	2	3	5	3	1	2	4	13282.289	-0.021
4	2	3	3	3	1	2	2	13282.017	-0.011
4	2	2	4	3	1	3	3	13875.776	0.016
4	2	2	5	3	1	3	4	13875.466	-0.027

Table S16. Measured frequencies for the rotational transitions belonging to the conformer 10 (III'ggg−) of 4-F-Thr obtained using LA-CP-FTMW spectroscopy. An uncertainty of 50 KHz was given for measured transitions.

J'	K'_{-1}	K'_{+1}	J''	K''_{-1}	K''_{+1}	ν_{obs} / MHz	$\nu_{obs} - \nu_{calc}$ / MHz
4	1	4	3	1	3	7680.373	-0.016
4	0	4	3	0	3	7864.296	0.031
5	2	3	4	2	2	10033.894	0.055
5	1	4	4	1	3	10181.753	0.006
6	2	5	5	2	4	11867.957	0.055
6	3	4	5	3	3	11930.951	-0.094
6	2	4	5	2	3	12090.012	-0.009
6	1	5	5	1	4	12194.784	-0.012
7	1	7	6	1	6	13386.931	0.003
7	0	7	6	0	6	13550.585	-0.005

Table S17. Measured frequencies for the hyperfine components of rotational transitions belonging to the conformer 11 (Ig−gg−) of 4-F-Thr obtained using LA-CP-FTMW spectroscopy. An uncertainty of 50 KHz was given for measured transitions.

J'	K'_{-1}	K'_{+1}	F'	J''	K''_{-1}	K''_{+1}	F''	ν_{obs} / MHz	$\nu_{obs} - \nu_{calc}$ / MHz
2	2	1	3	1	1	1	2	7091.659	-0.006
3	2	1	4	2	1	1	3	9820.165	0.019
3	2	1	3	2	1	1	2	9820.410	-0.024
3	1	2	4	2	0	2	3	9879.472	0.022
3	2	2	4	2	1	2	3	10194.951	0.008
3	2	2	2	2	1	2	1	10194.652	-0.009
3	2	2	3	2	1	2	2	10195.475	0.027
3	3	0	4	2	2	0	3	10510.572	0.043
3	3	0	3	2	2	0	2	10511.104	0.043
3	3	1	4	2	2	1	3	10687.392	0.011
3	3	1	3	2	2	1	2	10687.616	-0.073
4	2	2	5	3	1	2	4	13041.648	0.012
4	3	1	5	3	2	1	4	13242.505	-0.022
4	3	1	4	3	2	1	3	13242.890	-0.018
4	2	3	4	3	1	3	3	13455.355	-0.003
4	2	3	5	3	1	3	4	13454.971	-0.066

Table S18. Measured frequencies for the hyperfine components of rotational transitions belonging to the conformer 12 (Ig-gg) of 4-F-Thr obtained using LA-MB-FTMW spectroscopy. An uncertainty of 5 KHz was given for measured transitions.

J'	K'_{-1}	K'_{+1}	F'	J''	K''_{-1}	K''_{+1}	F''	ν_{obs} / MHz	$\nu_{obs} - \nu_{calc}$ / MHz
2	0	2	1	1	0	1	1	4501.6375	0.0003
2	0	2	3	1	0	1	2	4500.5375	0.0055
2	0	2	2	1	0	1	1	4500.4975	0.0045
3	0	3	2	2	0	2	1	6502.8693	0.0054
3	0	3	3	2	0	2	2	6502.9793	-0.0013
3	0	3	4	2	0	2	3	6502.9993	-0.0068
3	1	3	4	2	1	2	3	6269.8550	-0.0026
3	1	3	3	2	1	2	2	6269.6499	0.0007
3	1	2	3	2	1	1	2	7441.6030	-0.0026
3	1	2	4	2	1	1	3	7441.8030	0.0003

References

[1] J. H. Holland, *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*, Complex adaptive systems, MIT Press, 1st MIT press ed edition **1992**.

[2] G. Mancini, M. Fusè, F. Lazzari, B. Chandramouli, V. Barone, *J. Chem. Phys.* **2020**, *153*, 124110.

[3] S. Potenti, L. Spada, M. Fusè, G. Mancini, A. Gualandi, C. Leonardi, P. G. Cozzi, C. Puzzarini, V. Barone, *ACS Omega* **2021**, *6*, 13170.

[4] D. Whitley, *Statistics and Computing* **1994**, *4*, 65.

[5] G. Mancini, M. Fusè, F. Lazzari, V. Barone, *Digit. Disc.* **2022**, *1*, 10539.

[6] B. Chandramouli, S. Del Galdo, M. Fusè, V. Barone, G. Mancini, *Phys. Chem. Chem. Phys.* **2019**, *21*, 19921.

[7] C. Bannwarth, S. Ehlert, S. Grimme, *J. Chem. Theory Comput.* **2019**, *15*, 1652.

[8] L. Kaufmann, P. Rousseeuw, *Clustering by Means of Medoids*, Reports of the Faculty of Mathematics and Informatics, Delft University of Technology **1987**.

[9] J. Han, M. Kamber, *Data mining: concepts and techniques*, Elsevier, 3rd ed edition **2011**.

[10] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098.

[11] T. H. Dunning, *J. Chem. Phys.* **1989**, *90*, 1007.

[12] E. Papajak, J. Zheng, X. Xu, H. R. Leverentz, D. G. Truhlar, *J. Chem. Theory Comput.* **2011**, *7*, 3027.

[13] S. Grimme, S. Ehrlich, L. Goerigk, *J. Comput. Chem.* **2011**, *32*, 1456.

[14] G. Santra, N. Sylvetsky, J. M. L. Martin, *J. Phys. Chem. A* **2019**, *123*, 5129.

[15] T. H. Dunning, K. A. Peterson, A. K. Wilson, *J. Chem. Phys.* **2001**, *114*, 9244.

[16] J. Lupi, C. Puzzarini, C. Cavallotti, V. Barone, *J. Chem. Theory Comput.* **2020**, *16*, 5090.

[17] J. Lupi, S. Alessandrini, V. Barone, C. Puzzarini, *J. Chem. Theory Comput.* **2021**, *17*, 6974.

[18] T. B. Adler, G. Knizia, H.-J. Werner, *J. Chem. Phys.* **2007**, *127*, 221106.

[19] G. Knizia, T. B. Adler, H.-J. Werner, *J. Chem. Phys.* **2009**, *130*, 054104.

[20] H.-J. Werner, G. Knizia, F. R. Manby, *Mol. Phys.* **2011**, *109*, 407.

[21] T. Helgaker, W. Klopper, H. Koch, J. Noga, *J. Chem. Phys.* **1997**, *106*, 9639.

[22] J. G. Hill, S. Mazumder, K. A. Peterson, *J. Chem. Phys.* **2010**, *132*, 054108.

[23] J. L. Alonso, M. Pérez, M. E. Sanz, S. Blanco, *Phys. Chem. Chem. Phys.* **2009**, *11*, 617.

[24] J. Bloino, M. Biczysko, V. Barone, *J. Chem. Theory Comput.* **2012**, *8*, 1015.

[25] E. Penocchio, M. Piccardo, V. Barone, *J. Chem. Theory Comput.* **2015**, *11*, 4689.

[26] G. Ceselin, V. Barone, N. Tasinato, *J. Chem. Theory Comput.* **2021**, *17*, 7290.

[27] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, , D. J. F. and, Gaussian 16 Revision C.01, gaussian Inc. Wallingford CT 2016.

[28] H.-J. Werner, P. J. Knowles, F. R. Manby, J. A. Black, K. Doll, A. Heßelmann, D. Kats, A. Köhn, T. Korona, D. A. Kreplin, Q. Ma, T. F. Miller, A. Mitrushchenkov, K. A. Peterson, I. Polyak, G. Rauhut, M. Sibaev, *J. Chem. Phys.* **2020**, *152*, 144107.

[29] S. Grimme, *Chem. - A Eur. J.* **2012**, *18*, 9955.

[30] G. Luchini, J. Alegre-Requena, I. Funes-Ardoiz, R. Paton, *F1000Research* **2020**, *9*.

[31] C. Scolastico, E. Conca, L. Prati, G. Guanti, L. Banfi, A. Berti, P. Farine, U. Valcavi, *Synthesis* **1985**, *1985*.

[32] M. E. Sanz, J. C. López, J. L. Alonso, *Phys. Chem. Chem. Phys.* **2010**, *12*, 3573.

[33] J. L. Alonso, J. C. López, Microwave spectroscopy of biomolecular building blocks, in *Gas-Phase IR Spectroscopy and Structure of Biological Molecules*, pages 335–401, Springer **2015**.

[34] J. L. Alonso, I. Peña, J. C. López, E. R. Alonso, V. Vaquero, *Chemistry* **2019**, *25*, 2288.

[35] I. León, E. R. Alonso, S. Mata, J. L. Alonso, *Phys. Chem. Chem. Phys.* **2020**, *22*, 13867.

[36] C. Cabezas, I. León, E. R. Alonso, J. L. Alonso, The Shape of Dipeptides: Insights from Rotational Spectroscopy, in T. L. Thygesen (Editor), *Dipeptides and Tripeptides: Advances in Applications and Research*, Nova Science **2020**.

[37] A. I. Jimenéz, V. Vaquero, C. Cabezas, J. C. López, C. Cativela, J. L. Alonso, *J. Am. Chem. Soc.* **2011**, *133*, 10621.

[38] S. Mata, I. Peña, C. Cabezas, J. C. López, J. L. Alonso, *J. Mol. Spectrosc.* **2012**, *280*, 91.

[39] I. León, E. R. Alonso, S. Mata, C. Cabezas, J. L. Alonso, *Angew. Chem. Int. Ed.* **2019**, *58*, 16002.

[40] C. Bermúdez, S. Mata, C. Cabezas, J. Alonso, *Angew. Chem. Int. Ed.* **2014**, *53*, 11015.