Supplementary information for

On the Use of a Volume Constraint to Account for Thermal Expansion Effects on the Low-Frequency Vibrations of Molecular Crystals

Fergus J.M. Rogers^{*a}, Krunal Radhanpura^b, Joseph Horvat^c, and David Farrant^{*b}

^a Research School of Chemistry, Australian National University, Canberra, Australian Capital Territory 0200, Australia

^b Commonwealth Scientific and Industrial Research Organisation (CSIRO), Lindfield, NSW 2070, Australia

^c School of Physics and Institute for Superconducting and Electronic Materials, University of Wollongong, Wollongong, Australia, NSW, 2522

Table of Contents

Appendix S1 Comparison of the α and β Spectra of Para-Aminobenzoic acid (PABA)	S2
Appendix S2 Fitting of Experimental Modes 3 and 4 in L-Tartaric acid	S3
Appendix S3 Complete Tables and Charts of Errors	S4
Appendix S4 Average Simulated Spectrum – Scaled	S11
Appendix S5 Complete List of Raw Simulated Spectra	S12
Appendix S6 THz Vibrational Modes of Para-Aminobenzoic acid (PABA)	S20

Appendix S1 Comparison of the α and β Spectra of Para-Aminobenzoic acid (PABA)

Taking the crystal structure submitted by Gracin and Fischer (Fig. S1), the THz absorption spectrum of β -PABA was simulated using the PBE-D3(BJ)/6-311G(2d,2p) method for comparison with α -PABA. Irrespective of whether a volume constraint was applied, the spectral profile of β -PABA was clearly at odds with the experimental spectrum (Fig. S2).



Fig. S1 Crystal structure of β-PABA submitted by Gracin and Fischer.¹



Fig. S2 Absorption spectrum of β -PABA plotted assuming a FWHM of 6 cm⁻¹ compared with the roomtemperature THz-TDS spectrum (grey) of the commercial powder.

Appendix S2 Fitting of Experimental Modes 3 and 4 in L-Tartaric acid

Due to the distorted peak shapes of modes 3 and 4 in L-tartaric acid, we applied a Lorentzian fitting protocol to estimate the central frequencies of these absorption bands.^{2,3} In this case, a Lorentzian function was fitted to a selection of data points composing the experimental feature, using Mathematica's "FindFit" function.⁴

 $Lorentzian(v) = \frac{A}{1 + b(v - v_0)^2}$

Where A is the maximum absorption coefficient (set to the experimental value), v_0 is the central frequency to be determined, and b is a free parameter.



Fig. S3 Lorentzian fits of experimental absorption bands 3 and 4 in L-tartaric acid.

Appendix S3 Complete Tables and Charts of Errors

Table S1 Summary	of errors	(as compared	to experiment) for unconstrained	simulations of	I-tartaric acid
Tubic JI Junnin	01 01 01 01 5	las comparea	i to experiment	i or unconstruincu	Jiniulations of	

Theoretical Method	Volume (Å ³)	A (Å)	в (Å)	C (Å)	Beta (°)	Cell RMSD (%)	Frequency RMSD (THz)	Frequency SD	Scaling Factor
BLYP/6-311G(2d,2p)	-13.703	-0.007	-0.176	-0.085	1.526	2.69	0.107	0.053	0.99
PBE/pob-DZVP-rev2	-15.933	-0.055	-0.207	-0.075	1.069	3.05	0.124	0.065	0.97
PBE/6-311G(2d,2p)	-13.011	-0.057	-0.127	-0.096	0.813	2.41	0.126	0.050	1.01
B3LYP/6-311G(2d,2p)	-17.319	-0.046	-0.192	-0.120	1.518	3.28	0.128	0.055	0.97
PBE0/pob-DZVP-rev2	-20.305	-0.087	-0.244	-0.110	1.092	3.83	0.139	0.103	0.97
PBE0/6-311G(d,p)	-17.795	-0.090	-0.162	-0.137	1.065	3.29	0.157	0.060	0.97
PBE/6-311G(d,p)	-13.307	-0.058	-0.130	-0.096	0.925	2.47	0.158	0.082	0.97
PBE/pob-TZVP-rev2	-12.912	-0.113	-0.062	-0.126	0.336	2.37	0.161	0.056	0.95
PBE0/pob-TZVP-rev2	-17.259	-0.148	-0.099	-0.158	0.337	3.16	0.171	0.051	0.94
B3LYP/pob-DZVP-rev2	-20.398	-0.048	-0.310	-0.066	1.383	4.05	0.173	0.120	0.97
B3LYP/6-311G(d,p)	-17.679	-0.045	-0.195	-0.124	1.579	3.35	0.175	0.065	0.94
BLYP/6-311G(d,p)	-14.057	-0.007	-0.181	-0.085	1.618	2.76	0.180	0.108	0.94
BLYP/pob-DZVP-rev2	-16.801	-0.010	-0.273	-0.047	1.677	3.44	0.181	0.164	1.00
B3LYP/pob-TZVP-rev2	-17.703	-0.098	-0.148	-0.145	0.987	3.26	0.198	0.056	0.94
BLYP/pob-TZVP-rev2	-14.184	-0.056	-0.131	-0.113	1.111	2.64	0.204	0.078	0.94
PBE0/6-31G(d,p)	-24.488	-0.126	-0.242	-0.169	1.541	4.54	0.215	0.080	0.92
PBE/6-31G(d,p)	-21.296	-0.096	-0.223	-0.142	1.375	3.97	0.231	0.126	0.91
B3LYP/6-31G(d,p)	-24.610	-0.084	-0.275	-0.160	1.965	4.65	0.267	0.119	0.89
BLYP/6-31G(d,p)	-22.180	-0.050	-0.268	-0.137	2.005	4.25	0.283	0.158	0.89
Average	-17.628	-0.067	-0.192	-0.115	1.259	3.34	0.178	0.087	0.95

 Table S2
 Summary of errors (as compared to experiment) for constrained simulations of L-tartaric acid.

Theoretical Method	Volume (Å ³)	A (Å)	В (Å)	C (Å)	Beta (°)	Cell RMSD (%)	Frequency RMSD (THz)	Frequency SD	Scaling Factor
PBE/pob-TZVP-rev2	-0.417	-0.079	0.112	-0.070	-0.577	1.11	0.038	0.020	0.99
B3LYP/pob-TZVP-rev2	-0.328	-0.052	0.094	-0.066	-0.201	0.91	0.054	0.039	1.00
BLYP/pob-TZVP-rev2	-0.293	-0.014	0.057	-0.049	0.217	0.57	0.068	0.037	0.99
B3LYP/6-311G(2d,2p)	-0.335	0.010	0.021	-0.031	0.317	0.32	0.087	0.029	1.02
PBE0/6-31G(d,p)	-0.321	-0.042	0.062	-0.042	-0.309	0.62	0.089	0.059	1.00
PBE0/6-311G(d,p)	-0.316	-0.040	0.070	-0.050	-0.174	0.68	0.091	0.066	1.00
PBE0/pob-TZVP-rev2	-0.338	-0.109	0.139	-0.078	-0.930	1.40	0.099	0.079	1.01
PBE0/pob-DZVP-rev2	-0.323	-0.027	0.015	-0.008	-0.391	0.27	0.104	0.064	0.99
B3LYP/6-311G(d,p)	-0.322	0.011	0.024	-0.033	0.359	0.35	0.111	0.076	0.99
PBE/6-311G(2d,2p)	-0.359	-0.016	0.035	-0.034	-0.153	0.38	0.121	0.067	1.02
PBE/pob-DZVP-rev2	-0.382	-0.003	-0.018	0.011	-0.107	0.17	0.125	0.081	1.00
BLYP/6-311G(2d,2p)	-0.334	0.042	-0.013	-0.014	0.645	0.40	0.130	0.048	1.03
BLYP/pob-DZVP-rev2	-0.302	0.051	-0.073	0.040	0.533	0.72	0.133	0.103	1.02
B3LYP/pob-DZVP-rev2	-0.344	0.020	-0.032	0.014	0.229	0.31	0.137	0.115	1.01
B3LYP/6-31G(d,p)	-0.326	0.011	0.014	-0.026	0.251	0.26	0.139	0.097	0.98

BLYP/6-311G(d,p)	-0.332	0.043	-0.015	-0.012	0.700	0.43	0.157	0.095	0.99
PBE/6-311G(d,p)	-0.338	-0.016	0.036	-0.032	-0.032	0.37	0.163	0.133	0.98
PBE/6-31G(d,p)	-0.361	-0.017	0.032	-0.031	-0.211	0.36	0.167	0.128	0.97
BLYP/6-31G(d,p)	-0.317	0.044	-0.020	-0.012	0.487	0.38	0.187	0.124	0.97
Average	-0.336	-0.010	0.028	-0.028	0.034	0.53	0.116	0.077	1.00



Fig. S4 Bar chart showing the frequency RMSD for tartaric acid for all tested methods. Black bars are using the volume-constrained unit cell and white bars are using the fully-optimised unit cell. Standard deviations are given above each bar.

Table S3 Summary of electronic energies computed for tartaric acid using standard GGA functionals. ΔE is the difference in energy between the constrained cell and the unconstrained cell.

Method	Unconstrained Cell Energy (a.u.)	Constrained Cell Energy (a.u.)	ΔE (a.u.)
BLYP/6-311G(2d,2p)	-1215.183211	-1215.182053	0.001158
BLYP/6-311G(d,p)	-1215.147903	-1215.146636	0.001268
BLYP/6-31G(d,p)	-1214.768720	-1214.764226	0.004494
BLYP/pob-DZVP-rev2	-1213.864547	-1213.862429	0.002118
BLYP/pob-TZVP-rev2	-1215.133985	-1215.132780	0.001205
PBE/6-311G(2d,2p)	-1214.092516	-1214.091647	0.000870
PBE/6-311G(d,p)	-1214.056991	-1214.056027	0.000964
PBE/6-31G(d,p)	-1213.709632	-1213.706175	0.003457
PBE/pob-DZVP-rev2	-1212.798897	-1212.797275	0.001622
PBE/pob-TZVP-rev2	-1214.051266	-1214.050377	0.000889
Average Per Atom	-37.946274	-37.946218	0.000056

 Table S4
 Summary of electronic energies computed for tartaric acid using hybrid GGA functionals.

Method	Unconstrained Cell Energy (a.u.)	Constrained Cell Energy (a.u.)	ΔE (a.u.)
B3LYP/6-311G(2d,2p)	-1214.824736	-1214.822594	0.002142
B3LYP/6-311G(d,p)	-1214.787937	-1214.785525	0.002411
B3LYP/6-31G(d,p)	-1214.435165	-1214.429683	0.005482
B3LYP/pob-DZVP-rev2	-1213.501789	-1213.498535	0.003254
B3LYP/pob-TZVP-rev2	-1214.769841	-1214.767569	0.002273
PBE0/6-311G(d,p)	-1214.037419	-1214.035401	0.002017
PBE0/6-31G(d,p)	-1213.713849	-1213.709188	0.004661
PBE0/pob-DZVP-rev2	-1212.764849	-1212.761859	0.002990
PBE0/pob-TZVP-rev2	-1214.025341	-1214.023312	0.002029
Average per atom	-37.940489	-37.940395	0.000095

Theoretical Method	Volume (Å ³)	A (Å)	В (Å)	C (Å)	Beta (°)	Cell RMSD (%)	Frequency RMSD (THz)	Frequency SD	Scaling Factor
BLYP/6-311G(2d,2p)	-29.817	0.035	-0.068	-0.167	0.808	2.35	0.090	0.062	0.95
PBE/6-311G(2d,2p)	-31.263	-0.060	-0.171	-0.128	-0.218	2.22	0.106	0.033	0.93
BLYP/6-311G(d,p)	-23.073	0.007	-0.132	-0.143	-0.833	1.93	0.122	0.018	0.91
BLYP/pob-DZVP-rev2	-28.324	0.043	-0.079	-0.158	0.862	2.24	0.122	0.061	0.94
PBE/6-311G(d,p)	-33.641	-0.059	-0.159	-0.144	-0.143	2.40	0.149	0.098	0.92
PBE/pob-DZVP-rev2	-26.287	-0.057	-0.178	-0.121	-1.066	2.00	0.162	0.018	0.89
BLYP/pob-TZVP-rev2	-51.235	-0.086	-0.248	-0.207	0.307	3.60	0.190	0.036	0.87
PBE/6-31G(d,p)	-33.838	-0.015	-0.135	-0.175	-0.107	2.56	0.203	0.051	0.87
BLYP/6-31G(d,p)	-50.623	0.008	-0.139	-0.244	1.541	3.76	0.235	0.033	0.85
PBE/pob-TZVP-rev2	-39.232	-0.098	-0.211	-0.161	-0.713	2.82	0.281	0.060	0.83
Average	-34.733	-0.028	-0.152	-0.165	0.044	2.59	0.166	0.047	0.90

Table S5 Summary of errors (as compared to experiment) for unconstrained simulations of α -lactose monohydrate.

Table S6 Summary of errors (as compared to experiment) for constrained simulations of α -lactose

monohydrate.

Theoretical Method	Volume (Å ³)	A (Å)	В (Å)	C (Å)	Beta (°)	Cell RMSD (%)	Frequency RMSD (THz)	Frequency SD	Scaling Factor
BLYP/6-311G(2d,2p)	-0.264	0.058	0.126	-0.060	0.151	0.70	0.063	0.024	0.96
PBE/6-311G(2d,2p)	-0.408	-0.015	0.074	-0.019	-0.306	0.28	0.070	0.018	0.95
BLYP/6-311G(d,p)	-0.266	0.058	0.143	-0.060	0.259	0.72	0.095	0.017	0.93
PBE/pob-DZVP-rev2	-0.209	0.044	0.075	-0.067	-0.751	0.75	0.099	0.045	0.94
BLYP/pob-DZVP-rev2	-0.337	-0.012	0.052	-0.034	-0.922	0.50	0.113	0.033	0.93
BLYP/pob-TZVP-rev2	-0.196	-0.011	0.097	-0.022	-0.180	0.30	0.117	0.036	0.91
PBE/6-311G(d,p)	-0.220	0.031	0.169	-0.069	-0.383	0.76	0.140	0.072	0.92
PBE/pob-TZVP-rev2	-0.241	0.044	0.152	-0.047	0.466	0.62	0.142	0.077	0.90
BLYP/6-31G(d,p)	-0.313	-0.024	0.096	-0.009	0.008	0.25	0.174	0.091	0.88
PBE/6-31G(d,p)	-0.232	-0.034	0.133	-0.033	-0.758	0.55	0.176	0.050	0.89
Average	-0.269	0.014	0.112	-0.042	-0.242	0.54	0.119	0.046	0.92



Fig. S5 Bar chart showing the frequency RMSD for lactose for all tested methods. Black bars are using the volume-constrained unit cell and white bars are using the fully-optimised unit cell. Standard deviations are given above each bar.

Table S7 Summary of electronic energies computed for lactose.

Method	Unconstrained Cell Energy (a.u.)	Constrained Cell Energy (a.u.)	ΔE (a.u.)
BLYP/6-311G(2d,2p)	-2749.387446	-2749.385492	0.001954
BLYP/6-311G(d,p)	-2749.299525	-2749.297252	0.002273
BLYP/6-31G(d,p)	-2748.472026	-2748.463575	0.008450
BLYP/pob-DZVP-rev2	-2746.476750	-2746.475425	0.001325
BLYP/pob-TZVP-rev2	-2749.259549	-2749.256513	0.003036
PBE/6-311G(2d,2p)	-2746.965367	-2746.962967	0.002400
PBE/6-311G(d,p)	-2746.880095	-2746.877145	0.002950
PBE/6-31G(d,p)	-2746.122400	-2746.113983	0.008418
PBE/pob-DZVP-rev2	-2744.114060	-2744.112264	0.001796
PBE/pob-TZVP-rev2	-2746.855980	-2746.851714	0.004266
Average per atom	-28.618576	-28.618538	0.000038

Theoretical Method	Volume (Å ³)	A (Å)	В (Å)	C (Å)	Beta (°)	Cell RMSD (%)	Frequency RMSD (THz)	Frequency SD	Scaling Factor
PBE/6-311G(2d,2p)	-76.468	0.151	-0.138	-0.506	2.219	3.46	0.173	0.090	0.92
PBE/6-311G(d,p)	-87.588	0.174	-0.179	-0.498	1.906	3.94	0.192	0.088	0.91
BLYP/pob-DZVP-rev2	-107.284	0.022	-0.181	-0.623	2.165	4.56	0.236	0.101	0.88
PBE/pob-DZVP-rev2	-110.966	0.128	-0.237	-0.494	2.502	4.95	0.239	0.136	0.88
PBE/6-31G(d,p)	-157.002	-0.021	-0.297	-0.772	1.404	6.62	0.242	0.081	0.88
BLYP/6-311G(d,p)	-117.121	0.147	-0.261	-0.490	2.212	5.24	0.246	0.151	0.88
BLYP/6-311G(2d,2p)	-115.185	0.057	-0.226	-0.550	2.245	4.99	0.257	0.106	0.87
BLYP/6-31G(d,p)	-135.408	0.052	-0.274	-0.576	3.174	5.93	0.286	0.121	0.87
BLYP/pob-TZVP-rev2	-142.046	-0.015	-0.281	-0.620	1.893	6.05	0.299	0.115	0.85
PBE/pob-TZVP-rev2	-190.129	-0.053	-0.382	-0.821	1.421	8.08	0.310	0.127	0.86
Average	-123.920	0.064	-0.246	-0.595	2.114	5.38	0.248	0.112	0.88

Table S8 Summary of errors (as compared to experiment) for unconstrained simulations of α -paraaminobenzoic acid.

Table S9 Summary of errors (as compared to experiment) for constrained simulations of α -para-

aminobenzoic acid.

Theoretical Method	Volume (Å ³)	A (Å)	В (Å)	C (Å)	Beta (°)	Cell RMSD (%)	Frequency RMSD (THz)	Frequency SD	Scaling Factor
BLYP/pob-DZVP-rev2	-0.186	0.402	-0.014	-0.291	1.551	1.41	0.051	0.048	1.01
PBE/pob-TZVP-rev2	-0.634	0.787	-0.202	0.198	-1.894	3.19	0.073	0.039	1.04
PBE/6-311G(d,p)	-0.130	0.458	-0.024	-0.297	1.484	1.52	0.086	0.079	1.03
PBE/pob-DZVP-rev2	-0.156	0.536	-0.054	-0.206	2.279	1.87	0.098	0.078	1.04
BLYP/pob-TZVP-rev2	-0.267	0.525	-0.068	-0.202	-0.608	1.59	0.098	0.073	1.03
BLYP/6-311G(d,p)	-0.239	0.620	-0.082	-0.159	1.951	2.04	0.106	0.087	1.04
PBE/6-31G(d,p)	-0.127	0.399	0.001	-0.350	1.889	1.56	0.118	0.111	1.00
PBE/6-311G(2d,2p)	-0.232	0.580	-0.065	-0.198	2.160	1.95	0.122	0.116	1.03
BLYP/6-31G(d,p)	-0.243	0.568	-0.050	-0.248	2.443	1.98	0.125	0.117	0.99
BLYP/6-311G(2d,2p)	-0.208	0.359	0.023	-0.409	2.179	1.69	0.129	0.100	0.96
Average	-0.242	0.523	-0.054	-0.216	1.343	1.88	0.101	0.085	1.02



Fig. S6 Bar chart showing the frequency RMSD for PABA for all tested methods. Black bars are using the volume-constrained unit cell and white bars are using the fully-optimised unit cell. Standard deviations are given above each bar.

Table S10 Summary of electronic energies computed for PABA.

Method	Unconstrained Cell Energy	Constrained Cell Energy	ΔE (aul)
BLYP/6-311G(2d.2p)	-3810.300363	-3810.288689	0.011674
BLYP/6-311G(d,p)	-3810.177050	-3810.163470	0.013580
BLYP/6-31G(d,p)	-3809.116965	-3809.094133	0.022832
BLYP/pob-DZVP-rev2	-3806.156467	-3806.105024	0.051443
BLYP/pob-TZVP-rev2	-3810.075177	-3810.055405	0.019772
PBE/6-311G(2d,2p)	-3806.794769	-3806.790654	0.004115
PBE/6-311G(d,p)	-3806.677672	-3806.671541	0.006131
PBE/6-31G(d,p)	-3805.724152	-3805.713973	0.010179
PBE/pob-DZVP-rev2	-3802.772577	-3802.742800	0.029778
PBE/pob-TZVP-rev2	-3806.588583	-3806.576971	0.011613
Average per atom	-27.995870	-27.995737	0.000133

Appendix S4 Average Simulated Spectrum – Scaled



Fig. S7 Room-temperature THz-TDS absorption spectrum (grey) compared with the average simulated spectrum produced using a volume constrained cell (red) and a fully optimised cell (blue). These were plotted according to equation 1 assuming a FWHM of 5, 2.5, and 6 cm⁻¹, for tartaric acid, lactose, and PABA, respectively, and frequency scaled to minimise the RMSD.











Fig. S8 Simulated spectra of L-tartaric acid plotted assuming a FWHM of 5 cm⁻¹. Selected modes are numbered and marked by the red dots.



S16



Fig. S9 Simulated spectra of α -Lactose monohydrate plotted assuming a FWHM of 2.5 cm⁻¹.



S18



Fig. S10 Simulated spectra of α -para-aminobenzoic acid plotted assuming a FWHM of 6 cm⁻¹.

Appendix S6 THz Vibrational Modes of Para-Aminobenzoic acid (PABA)

Computed using the PBE/6-311G(2d,2p) method using the volume constrained cell, unless otherwise indicated, and visualised in PyMOL using the PyVibMS plugin.⁵



Fig. S11 Mode 1 assigned to the first experimental absorption of PABA. Top: viewed down b axis, middle:

down c axis, bottom: down a axis.



Fig. S12 Mode 2 (strong) assigned to the second experimental absorption band of PABA. Top: viewed down b axis, middle: down c axis, bottom: down a axis.



Fig. S13 Mode 3 (weak) assigned to the second experimental absorption band of PABA. Top: viewed down b axis, middle: down c axis, bottom: down a axis.



Fig. S14 Mode 6 assigned to the third experimental absorption of PABA. Top: viewed down b axis, middle: down c axis, bottom: down a axis.



Fig. S15 Mode 9 assigned to the fourth experimental absorption of PABA using a volume-constrained unit cell. Top: viewed down b axis, middle: down c axis, bottom: down a axis.



Fig. S16 Mode 9 assigned to the fourth experimental absorption of PABA using the fully optimised unit cell. Top: viewed down b axis, middle: down c axis, bottom: down a axis.



Fig. S17 Mode 13 assigned to the fifth experimental absorption of PABA. Top: viewed down b axis, middle:

down c axis, bottom: down a axis.

References

- 1 S. Gracin and A. Fischer, *Acta Crystallogr. Sect. Sect. E: Struct. Rep. Online*, 2005, **61**, 1242-1244.
- 2 E. R. Brown, J. E. Bjarnason, A. M. Fedor and T. M. Korter, *Appl. Phys. Lett.*, 2007, **90**, 061908.
- 3 Z. Li, Z. Zhang, X. Zhao, H. Su, H. Zhang and J. Lan, J. Infrared Millim. Terahertz Waves, 2013, 34, 617-626.
- 4 I. Wolfram Research, *Mathematica*, Champaign, Illinois, 12.0 edn., 2019.
- 5 Y. Tao, W. Zou, S. Nanayakkara and E. Kraka, *J. Mol. Model.*, 2020, **26**, 290.