Supplementary Information to:

Tailored anharmonic potential energy surfaces for infrared signatures

Janine Hellmers, Pascal Czember, and Carolin König

[1] *Institute of Physical Chemistry and Electrochemistry, Leibniz University Hannover*

Callinstraße 3A, D-30167 Hannover, Germany

E-mail:carolin.koenig@pci.uni-hannover.de

S1 Calculated potential energy surfaces

Molecule	Coordinates	Electronic-structure method	M	Coupling level
Catechol	FALCON (grow)	r2SCAN-3c	2	2 MC
			6	2MC
				3MC
				4MC
			12	2MC
				3MC
				4MC
		B2PLYP-D3BJ/aug-cc-pVTZ	6	2MC
	FALCON (partitioning)	r2SCAN-3c	36	2MC
				3MC (group)
			6	4MC
		B2PLYP-D3BJ/aug-cc-pVTZ	6	2MC
	Normal	r2SCAN-3c	36	2MC
				3MC (group)
			6	4MC
		B2PLYP-D3BJ/aug-cc-pVTZ	6	2MC
Uracil	FALCON (grow)	r2SCAN-3c	2	2MC
			9	2MC
				3MC
				4MC
	Normal	r2SCAN-3c	2	2MC
			9	2MC
				3MC
				4MC
			36	2MC
			36	3MC (group)
		B2PLYP-D3BJ/aug-cc-pVTZ	2	2MC
			6	2MC

Tab. SI: Overview of all calculated potential energy (PES) and dipole moment surfaces (DMS) for catechol and uracil using different underlying coordinates M. "Grow" denotes FALCON coordinates obtained through the growing scheme, while "partitioning" indicates FALCON coordinates representing the full vibrational space with vibrations localized to a predefined subset.

S2 Uracil

S2.1 Quasi-harmonic frequencies and mode visualization

Iteration	M	Initial atoms in growing group
2	2	4,8 + 11,7
5	9	3,4,7,8,11
6	12	2,3,4,7,8,11
7	15	1,2,3,4,7,8,11
9	21	1,2,3,4,6,7,8,11
10	24	1,2,3,4,5,6,7,8,9,11
12	30	all

Tab. SII: Initial atoms in the growing group for the respective iteration step of the FALCON growing scheme for uracil applying a Hessian-based coupling estimate (degeneracy value is 0.0001 a.u.). The initial group labeling refers to Fig. 3a in the main text (B2PLYP-D3/aug-cc-pVTZ). M refers to the number of resulting (quasi-)harmonic modes for the respective reduced-space.

S2.1.1 Normal modes of uracil

a Q0: 148   b Q1: 167   c Q2: 386   d Q3: 401   e Q4: 521   f Q5: 541

g Q6: 559   h Q7: 569   i Q8: 688   j Q9: 735   k Q10: 767   l Q11: 772

m Q12: 826   n Q13: 970   o Q14: 984   p Q15: 996   q Q16: 1092   r Q17: 1205
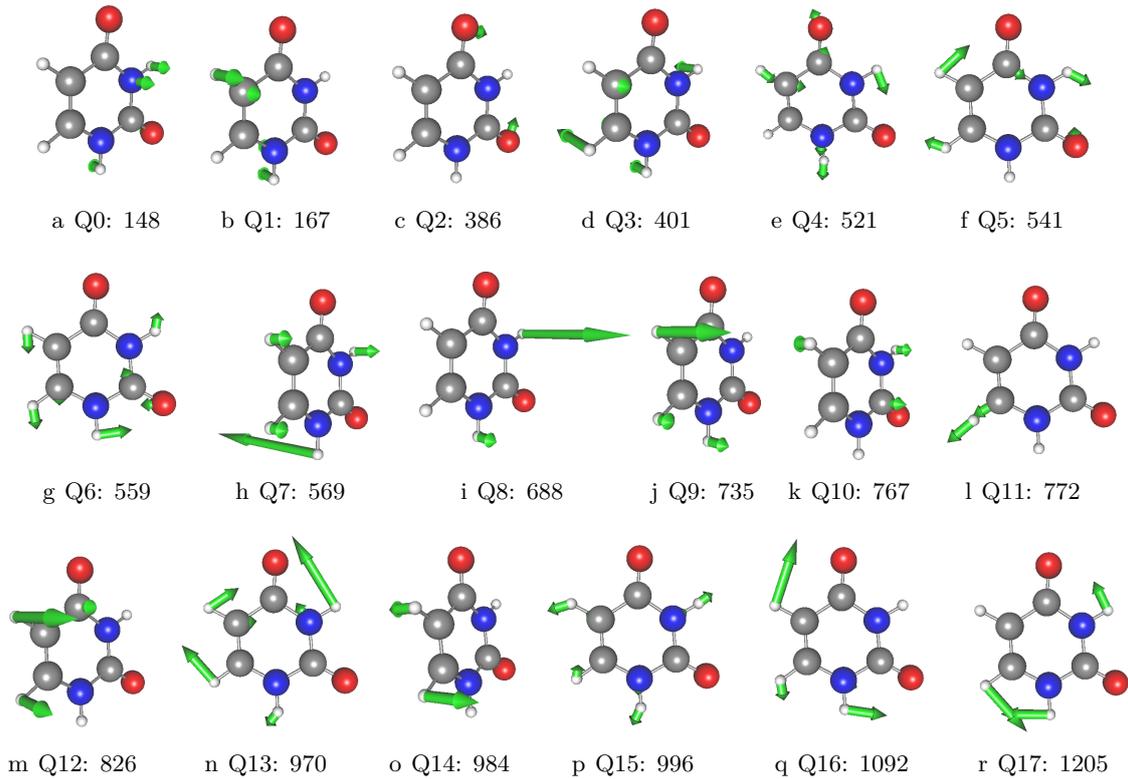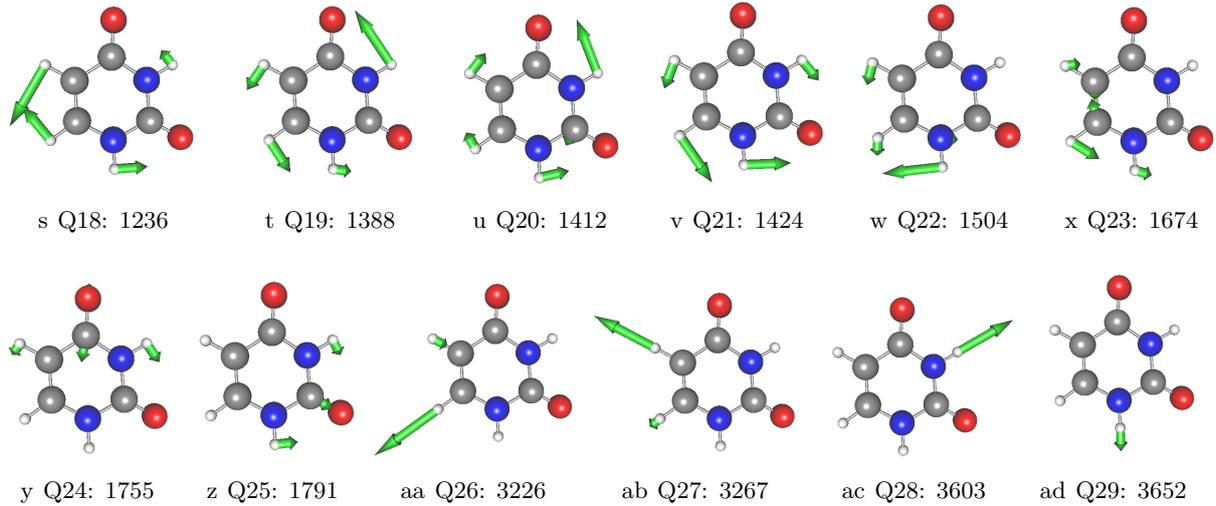
Fig. S1: Visualisation of all normal modes and their harmonic frequencies in cm^{-1} (B2PLYP-D3/aug-cc-pVTZ).

S2.2 Anharmonic spectra peaks

			VCC[1]		VCC[2]		VCC[3]		VCC[4]	
			C4=O8	C11=O7	C4=O8	C11=O7	C4=O8	C11=O7	C4=O8	C11=O7
FC	2 M	2MC	–	–	1708	1754	–	–	–	–
	9 M	2 MC	–	–	–	–	–	–	1716	1749
		3 MC	–	–	–	–	–	–	1709	1747
		4 MC	–	–	–	–	–	–	1709	1747
NC	2 M	2MC			1758	1802	–	–	–	–
	9 M	2 MC	1749	1788	1746	1785	1746	1785	1746	1785
		3 MC	1748	1788	1747	1786	1747	1786	1747	1786
		4 MC	1748	1788	1748	1788	1747	1786	1747	1786
	30 M	LL1	–	–	1732	1765	–	–	–	–
		LL2	–	–	1743	1777	–	–	–	–
		ML1	–	–	1738	1759	–	–	–	–
		ML2	–	–	1733	1754	–	–	–	–
Experiment[1]			1728.2–1733.2 (1731)	1761.4–1763.7 (1763)						

Tab. SIII: C4=O8 and C11=O7 peak positions of VCC[1]–VCC[4] damped linear response calculations for several potentials and dipole moment surfaces. For information on the nomenclature see main text. The entries for the 2MC coupled surfaces of two coordinates for FALCON (FC) and normal coordinates (NC), respectively, at the VCC[2] column refer to FVCI calculations.

4

S2.3 VCC chain length convergence check

a VCC4 (9 M 2MC)   b VCC4 (9 M 3MC)   c VCC4 (9 M 4MC)
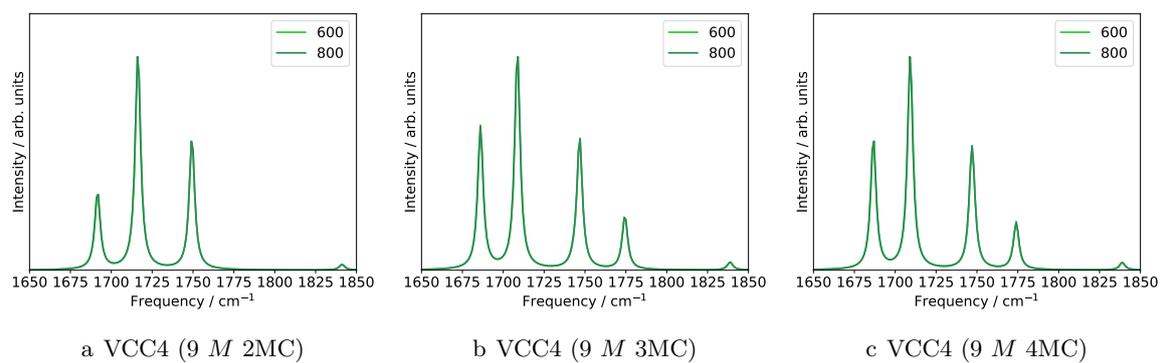
Fig. S2: Chain length check of the nine FALCON mode reduced space of uracil for VCC[4] damped linear response calculations on 2MC, 3MC, and 4MC potentials (r2SCAN-3c).

a VCC3 (9 M 2MC)   b VCC4 (9 M 2MC)   c VCC3 (9 M 3MC)

d VCC4 (9 M 3MC)   e VCC3 (9 M 4MC)   f VCC4 (9 M 4MC)

g LL1

h LL2

i ML1

j ML2
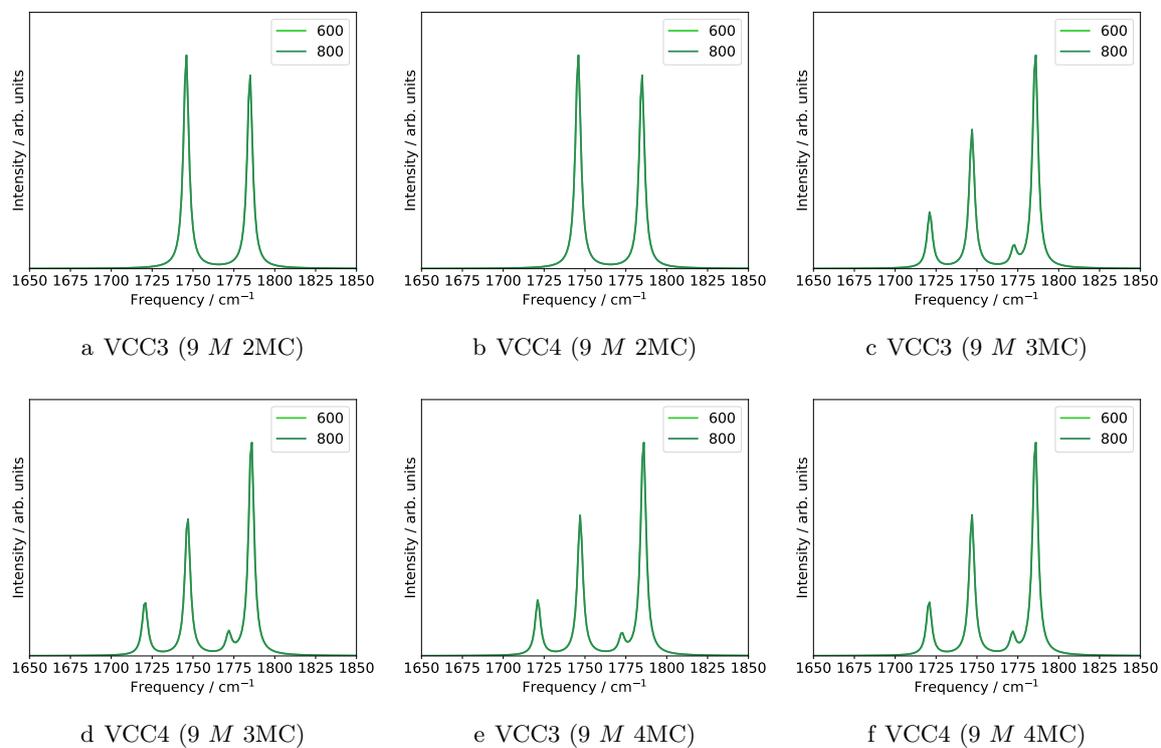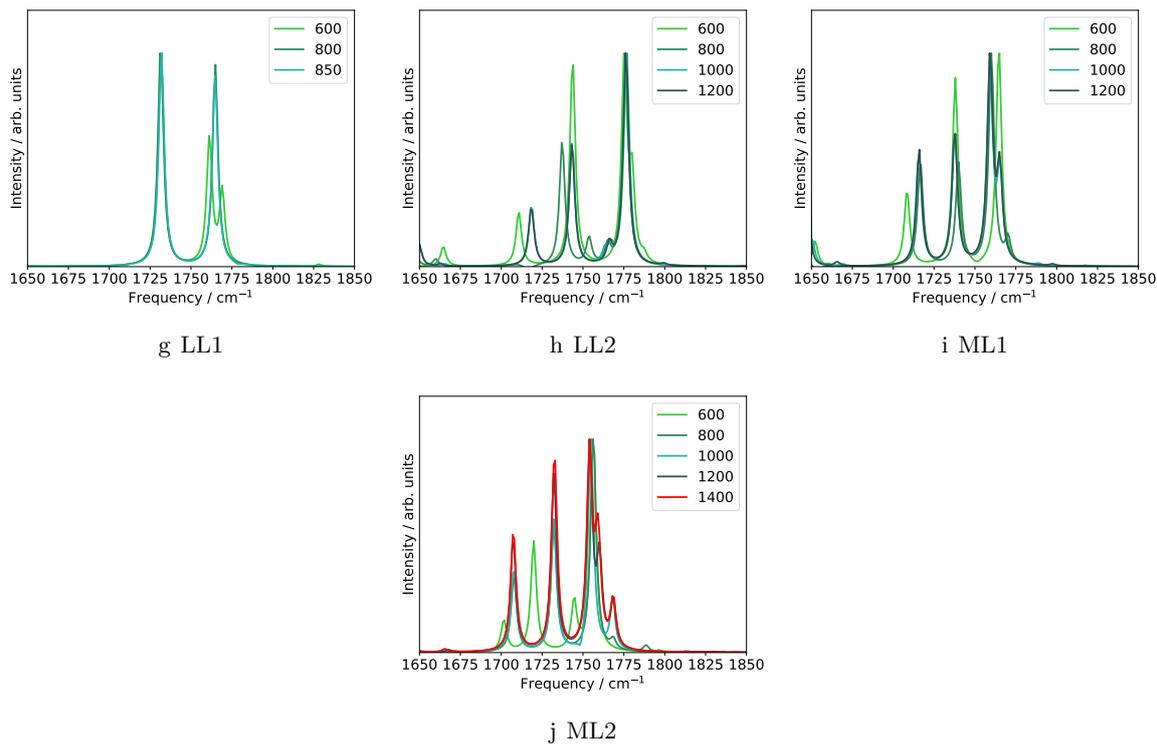
Fig. S3: Chain length check of the nine normal mode reduced spaces of uracil for VCC[4] damped linear response calculations on 2MC, 3MC, and 4MC potentials (r2SCAN-3c) and VCC2 chain length checks for all ML property surfaces generated for uracil. Details on LL1, LL2, ML1 and ML2 can be found in the main text.

S2.3.1 Excitation levels in VCC damped response spectra of the nine normal mode space

Before performing calculations on the full vibrational space, we investigated the performance of lower excitation levels on the nine-mode reduced space in normal coordinates. VCC[1] to VCC[4] calculations for different potential MC orders are shown in Fig. S4.
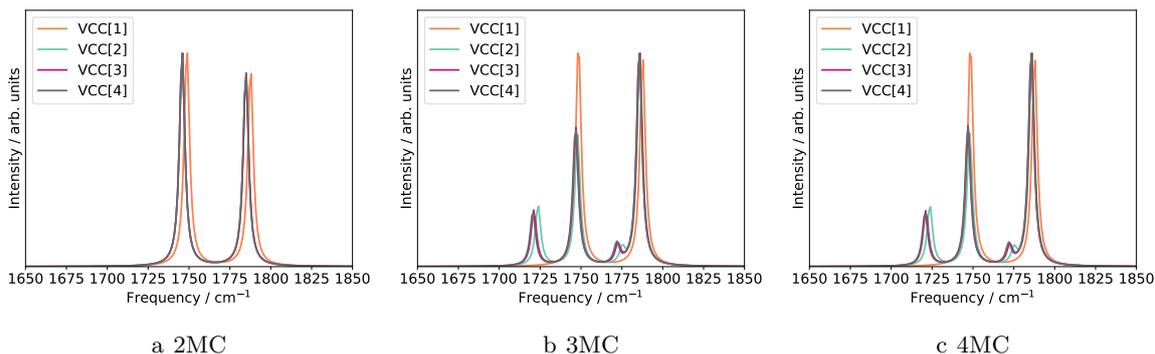
Fig. S4: VCC[1]–VCC[4] damped linear response spectra of the nine-mode (normal) reduced space of uracil for a 2MC (a), 3MC (b), and 4MC (c) coupled potential. All peak position values for the calculations can be found in Tab. SIII.

For the 2MC potential, all excitation levels of the vibrational wave function exhibit coinciding results. The VCC[1] calculation deviates slightly from the peak positions by 3 cm^{-1} to the VCC[2] calculation, which is identical to the VCC[3] and VCC[4] calculations. For all excitation levels, we obtain matching intensity heights. For the 3MC potential, the VCC[2], VCC[3], and VCC[4] spectra agree well, whereas the VCC[1] calculation leads to non-convergent intensities of the peak around 1750 cm^{-1} and the peak around 1725 cm^{-1} is completely missing. These observations apply to the 4MC potential as well. These calculations suggest that the coupling order of the potential exerts a higher influence on the prediction of the infrared spectrum than higher excitations (beyond VCC[2]) in the correlated wave function. Nevertheless, a description based on the correlated wave function remains indispensable, as VCC[1] fails to accurately capture the correct intensities.

S3 Catechol

S3.1 Quasi-harmonic frequencies and FALCON generated subspaces

The atoms of the FALCON growing scheme spanning each vibrational subspace are listed in Tab. SIV.

Iteration	Number of modes	Atom numbers in growing group
2	2	1, 3 + 2, 4
3	6	1, 3, 5 + 2, 4, 6
6	12	1, 2, 3, 4, 5, 6
9	21	1, 2, 3, 4, 5, 6, 7, 8, 9
11	27	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11
14	36	all

Tab. SIV: Initial atoms in the growing group for the respective iteration step of the FALCON growing scheme for catechol applying a Hessian-based coupling estimate with a degeneracy value of 0.0001 a.u. The atom number labeling refers to Fig. 3b in the main text (B2PLYP-D3/aug-cc-pVTZ). M refers to the number of (quasi) harmonic modes.

The first created space is represented by the two OH stretching vibrations and is indicated by blue circles in Fig. 3b in the main text. In the next iteration step, the two carbon atoms C5 and C6 of the benzene ring are fused resulting in six FALCON coordinates (red circles in Fig. 3b in the main text). These coordinates are visualized in Fig. S5.

a 1263   b 1308   c 1318   d 1379   e 3784   f 3839
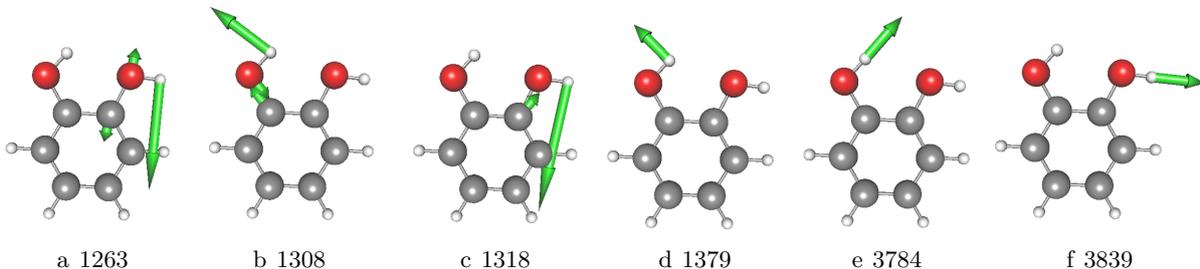
Fig. S5: Visualisation of local FALCON coordinates and their quasi-harmonic frequencies in cm^{-1} resulting from the six-mode reduced-space obtained by the FALCON algorithm (B2PLYP-D3/aug-cc-pVTZ).

Note, that these intra-vibrations are localized to their subsets, and no relative motions are included in the tailored subspace. In the sixth step, the two subspaces representing six DOFs are merged into one vibrational subspace (black circle in Fig. 3b in the main text) describing twelve vibrational DOFs. A visualization of these coordinates can be found in Fig. S6.

a Q6: 253 b Q7: 339 c Q8: 447 d Q9: 688 e Q10: 947 f Q11: 1144

g Q12: 1276 h Q13: 1320 i Q14: 1456 j Q15: 1534 k Q16: 3784 l Q17: 3839
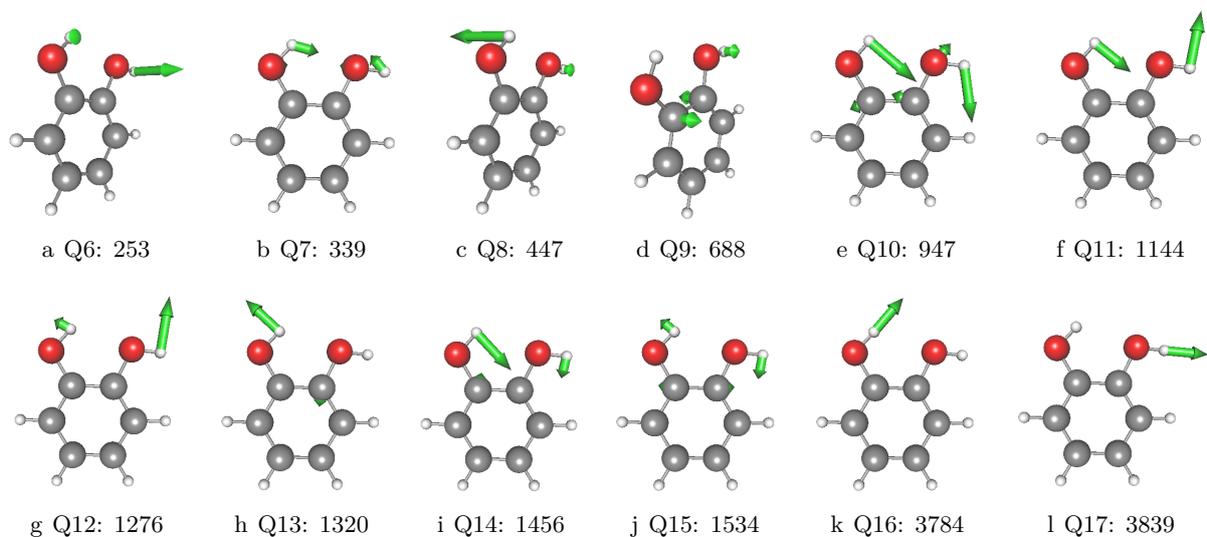
Fig. S6: Visualisation of local FALCON coordinates and their quasi-harmonic frequencies in cm^{-1} resulting from the twelve-mode reduced-space obtained by the FALCON algorithm (B2PLYP-D3/aug-cc-pVTZ).

Within the 9th iteration step, the atoms C7, C8, and H9 are further merged into the growing group, with 21 vibrational coordinates spanning the vibrational space. In the 11th iteration, the atoms H11 and C11 are further merged into the growing group. With the last iteration step, all DOFs are included again resulting in 36 normal modes.

S3.2 Anharmonic spectra peaks

	Level		VCC[1]		VCC[2]		VCC[3]		VCC[4]	
			OH_δ	OH_α	OH_δ	OH_α	OH_δ	OH_α	OH_δ	OH_α
2 FC	LL	2MC			3618	3677	–	–	–	–
6 FC	LL	2MC	3598	3656	3593	3661	3593	3652	3593	3655
		3MC	3595	3695	3606	3668	3604	3652 (3674)	3604	3652 (3672)
		4MC	3595	3653	3606	3668	3604	3657 (3674)	3604	3652 (3672)
	HL	2MC	–	–	–	–	–	–	3590	3652
	ML	4MC	–	–	–	–	–	–	3604	3665
12 FC	LL	2MC	–	–	–	–	3607	3738	–	–
	LL	3MC	–	–	–	–	?	?	–	–
	LL	4MC	–	–	–	–	3606	3675	–	–
8 FC	LL	2MC	–	–	–	–	–	–	3581	3643
(from 12 FC)	LL	3MC	–	–	–	–	–	–	3613	3671
	LL	4MC	–	–	–	–	–	–	3602	3664
36 FC	HL	2MC	–	–	3602	3664	–	–	–	–
	LL	4MC	–	–	3613	3683	–	–	–	–
	ML	4MC	–	–	3616	3688	–	–	–	–
36 NC	HL	2MC	–	–	3600	3655	–	–	–	–
	LL	4MC	–	–	3632	3703	–	–	–	–
	ML	4MC	–	–	3632	3709	–	–	–	–
Experiment[2]			3603	3663						

Tab. SV: Bound OH (OH_δ) and free OH (OH_α) peak positions of catechol for VCC[1]–VCC[4] damped linear response calculations and 2MC, 3MC, and 4MC potentials for different vibrational coordinates. 2, 6, and 12 FALCON (FC) indicating the reduced spaces generated with the growing scheme. The 8 FC are FALCON coordinates selected from the twelve-mode FALCON space, whereas 36 FC and 36 normal coordinates (NC) represent the full vibrational space of catechol. For more information on the 36 FC space, we refer to the main text.

S3.3   VCC chain length convergence check
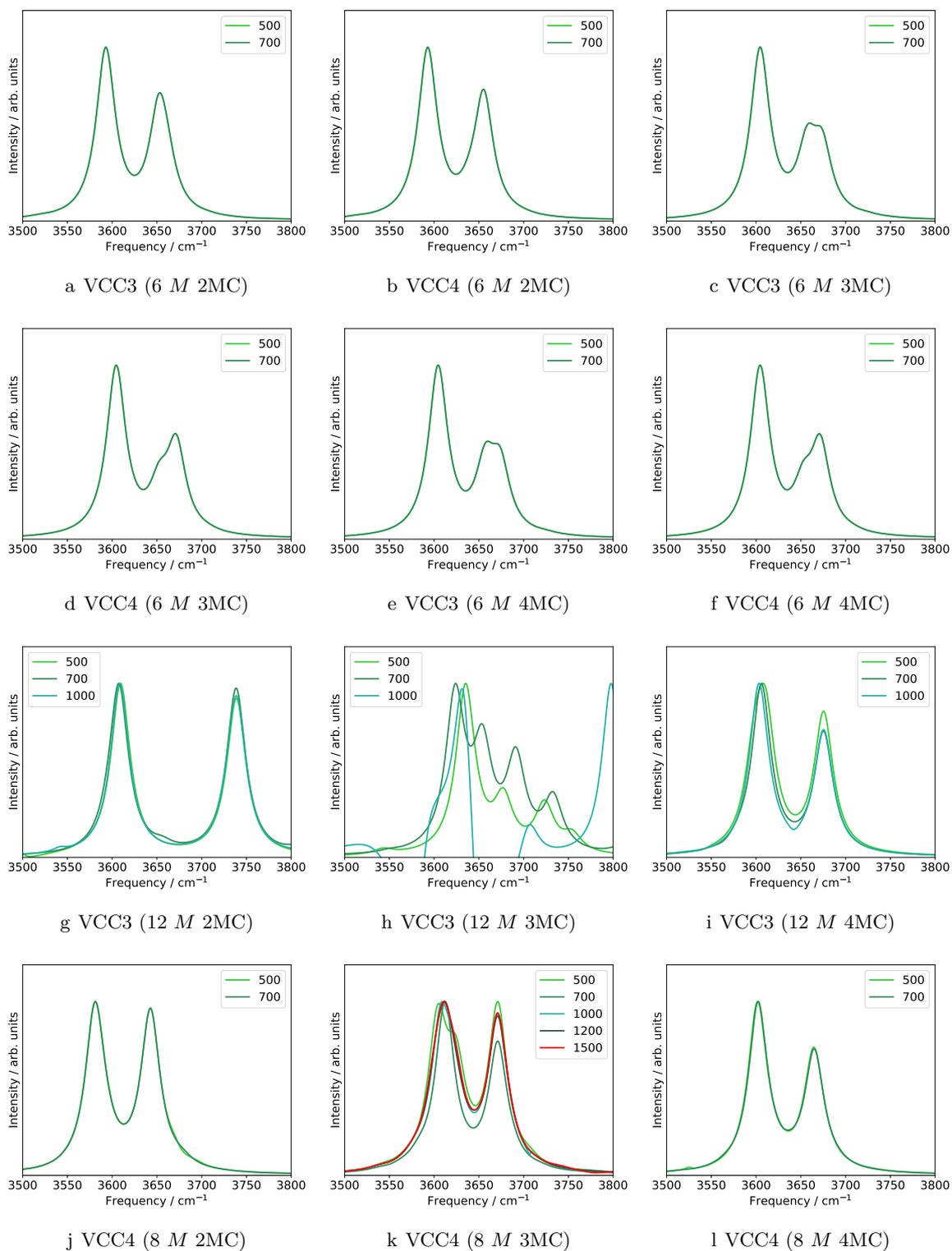
Fig. S7: Chain length convergence for the six-, twelve-, and eight-FALCON mode reduced space of catechol for VCC[3] and VCC[4] damped response calculations on 2MC, 3MC, and 4MC potentials (r2SCAN-3c). The eight-mode space consist of the modes from the twelve-mode space excluding modes Q6, Q7, Q8, Q9 (Fig. S6 a-d).
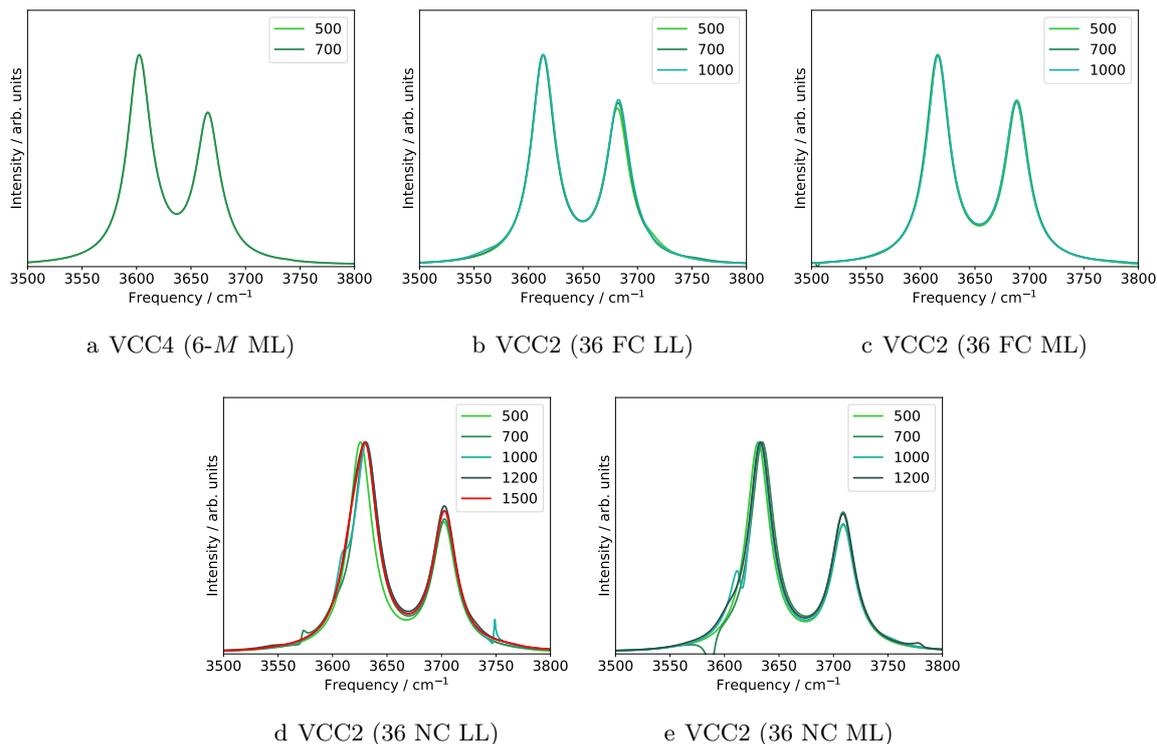
Fig. S8: Chain length check for different multilevel (ML) spectra of catechol.

S3.3.1 Excitation levels in VCC damped response spectra

For the sake of computational feasibility, we examined the performance of lower excitation levels in the vibrational coupled cluster wave functions (VCC[1]–VCC[3]) for 2MC, 3MC, and 4MC PES of the six-mode reduced space in comparison to the VCC[4] reference calculation. The results are shown in Fig. S9.
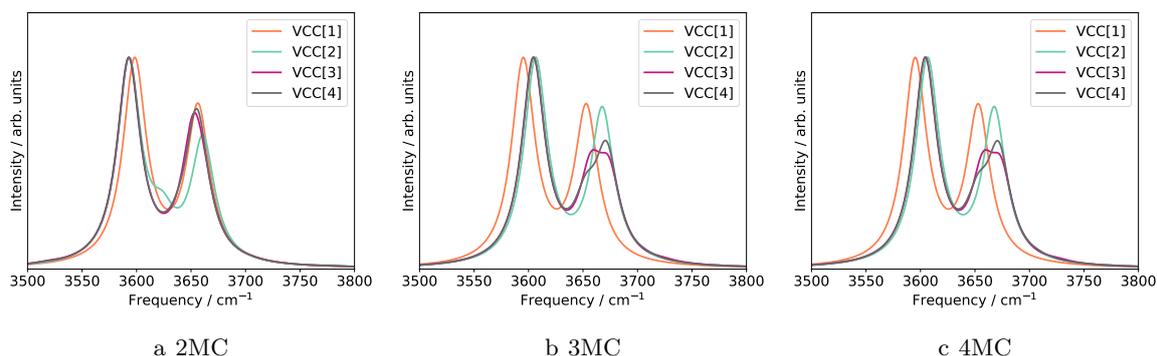
Fig. S9: VCC[1]-VCC[4] damped linear response spectra of the six-mode (FALCON) reduced space of catechol for a 2MC (a), 3MC (b) and 4MC (c) coupled potential. All peak position values for the calculations can be found in Tab. SV.

For the 2MC PES VCC[1]–VCC[4] calculations exhibit overall similar results. VCC[3] and

12

VCC[4] spectra are in close agreement, whereas the VCC[1] spectra differ marginally from the VCC[4] reference, and VCC[2] predicts a lower intensity of the free OH stretch in comparison to VCC[4]. The VCC[1]–VCC[4] calculations on the 3MC and 4MC PES are concurring and therefore only the results of 3MC are discussed here. The VCC[3] calculated spectrum exhibits a very similar shape to the VCC[4] reference, and also covers the resonance of the free fundamental OH stretching vibration, indicating a right-sided shoulder, whereas the VCC[4] reference indicates a left-sided shoulder. For the bound fundamental OH stretch, the VCC[2] spectrum agrees with the VCC[4] reference, whereas the intensity of the free OH stretch is overestimated. These results suggest, that a triple-excitation level can reproduce the VCC[4] spectrum, while the VCC[2] predicts the same peak position, but overestimate the intensity of the free OH stretching peak.

S3.3.2 Analysis of the twelve-mode space property surfaces

VCC[3] calculations on the twelve-mode FALCON space with a 3MC potential exhibit numerical instabilities as shown in Fig. S7h and were recovered with the inclusion of 4MC terms. After a visual analysis of all 2MC potential cuts of this PES, we identified potential cuts that showed sharp declines at already small displacements of the modes. These were found for the mode pairs consisting of a low-frequency mode that describes the out-of-plane wagging motions of the OH groups in conjunction with the respective OH stretch. The respective potential cuts and mode pairs $Q6/Q17$ and $Q8/Q16$ are shown in Fig. S10.
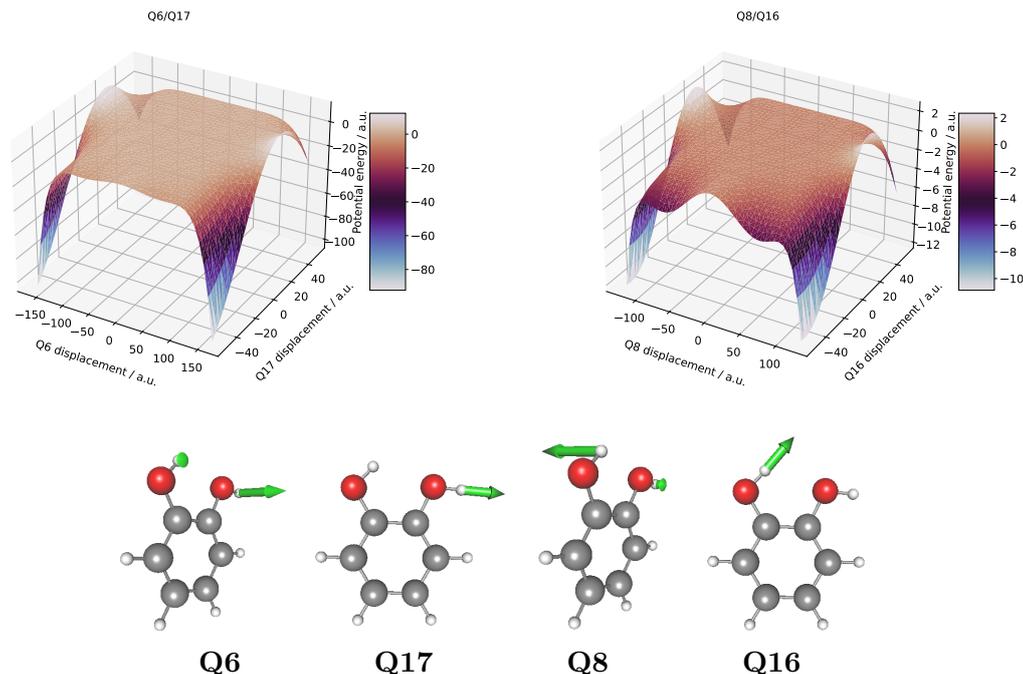
Fig. S10: Upper panel: Two-dimensional potential cut of the anharmonic potential of the Q6 mode (out-of-plane wagging) and Q17 mode (stretching) of the free OH group (left) and two-dimensional potential cut of the anharmonic potential of the Q8 (out-of-plane wagging) and Q16 (stretching) of the bound OH group (right).
Lower panel: Visualization of the respective FALCON modes.

This conspicuous behavior suggests that these vibrational motions may not ideally be suited, offering a possible explanation for the unstable potentials, albeit not exhaustively analyzed.

Recently, Yang and co-authors[3] have shown that the coupling of low-frequency and OH-stretching vibrations in conjunction with strong anharmonic coupling can lead to significant unphysical behavior of VSCF energies, evident in truncated n-mode expansions, especially at the second order (2MC). Using normal modes from a trophineH+ molecule as an example, they observed this behavior originating from potential surface regions below the global minimum. While including selected 3MC terms mitigated this issue, such terms in the twelve-mode FALCON space of catechol failed to address the problem.

To address, if the unstable description of the twelve FALCON coordinates originates from the low-frequency coordinates, we performed VCC damped linear response calculations on the 2MC, 3MC, and 4MC potentials excluding these low-frequency modes (c.f. Fig. S6 Q6, Q7, Q8, Q9) resulting in a total vibrational space of eight modes. To ensure convergence of the VCC wave function, we performed VCC[4] calculations. The resulting spectra for each coupling potential, excluding low-frequency modes, are depicted in Fig. S11.
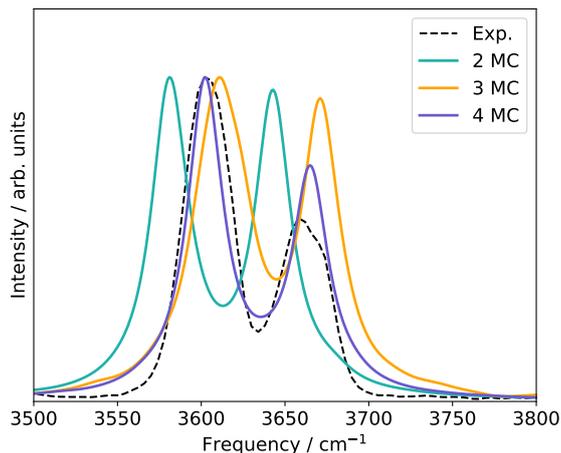
Fig. S11: VCC[4] damped linear response spectra on 2MC, 3MC, and 4MC potentials (r2SCAN-3c) of the twelve-mode space excluding the four modes that are lowest in frequency (cf. Fig. S6 a-d).

The spectrum on the 2MC potential is red-shifted compared to the experiment, and the predicted intensity ratio remains incorrect. However, the distance between the fundamental OH peaks corresponds to the experimental values, unlike the 2MC spectra with twelve FALCON coordinates (c.f. Fig. 9 in the main text). The calculation of the 3MC potential spectrum tends to be computationally more stable as no convergence issues in the chain length of the band-Lanczos algorithm occur. The spectrum based on the 3MC PES shows a slight blue shift towards the experimental values, approximately 10 cm^{-1} for the bound and 8 cm^{-1} for the free OH-stretch, respectively. Nonetheless, the intensity ratio does not agree with the experiment. The peak positions of the two fundamental stretches for the 4MC potential deviate only slightly, by approximately 1 cm^{-1} to the experimental spectrum: This spectrum closely resembles the experimental spectrum.

This indicates, that excluding low-frequency modes stabilizes the calculations. However, the inclusion of 4MC coupling terms is still necessary to obtain good agreement to experiment. From these findings, we conclude, that the six-mode FALCON space leads in the present case to a numerical more stable description than in the twelve-mode case. This is why we focused on the six-mode FALCON description in the main text.

S3.4 Full space mode visualization and 2MC potential cuts

S3.4.1 36 FALCON modes
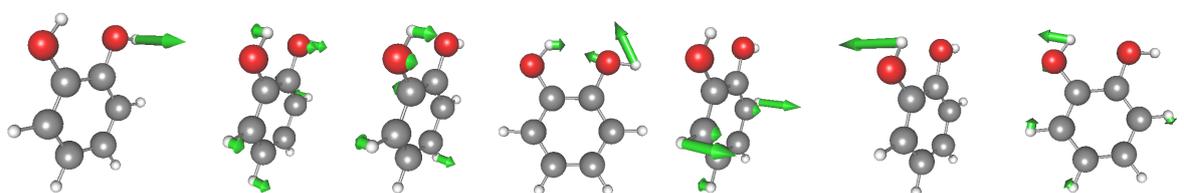
a Q6: 259 (IC)   b Q7: 293 (IC)   c Q8: 321 (IC)   d Q9: 400 (IC)   e Q10: 439   f Q11: 441 (IC)   g Q12: 518 (IC)
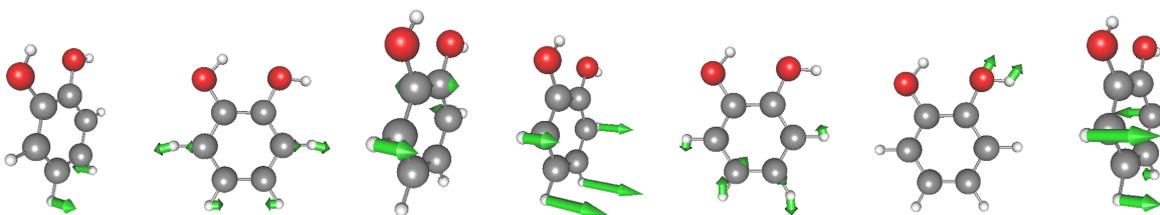
h Q13: 547   i Q14: 670   j Q15: 713   k Q16: 751   l Q17: 754   m Q18: 847   n Q19: 849
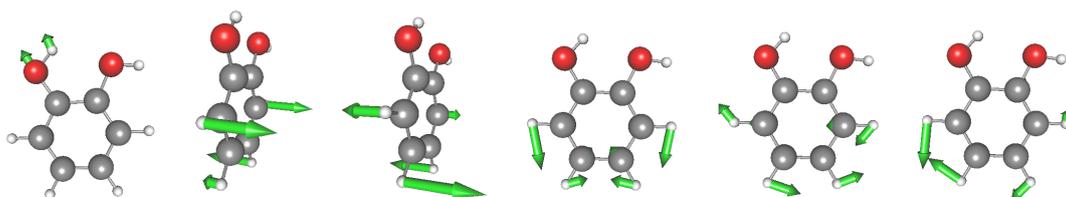
o Q20: 897   p Q21: 930   q Q22: 966   r Q23: 1052   s Q24: 1105   t Q25: 1167
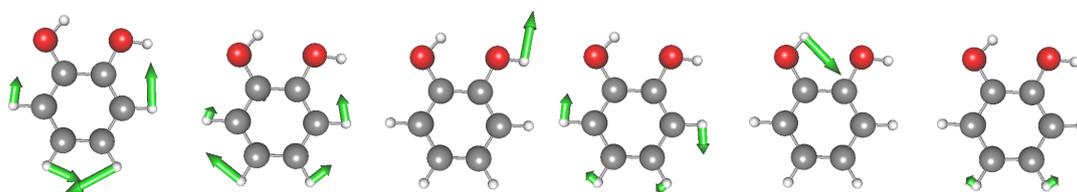
u Q26: 1175   v Q27: 1200   w Q28: 1244 (IC)   x Q29: 1294   y Q30: 1328 (IC)   z Q31: 1372
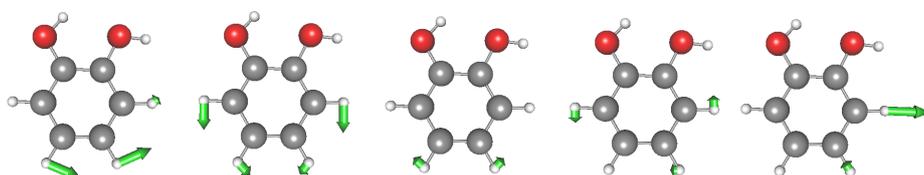
aa Q32: 1485   ab Q33: 1523   ac Q34: 1638   ad Q35: 1645   ae Q36: 3175

af Q37: 3199   ag Q38: 3212   ah Q39: 3222   ai Q40: 3783   aj Q41: 3839
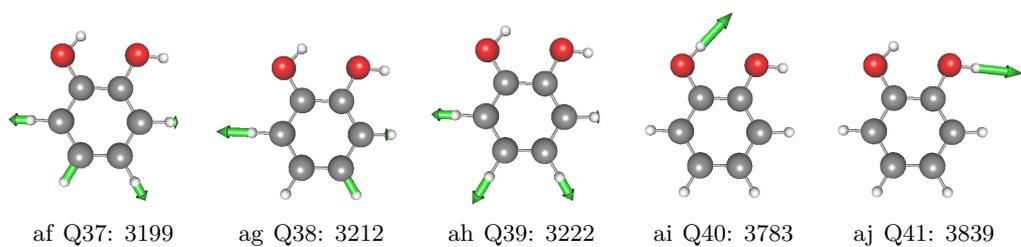
Fig. S12: Visualisation of the modes with their quasi-harmonic frequencies in cm^{-1} resulting from the generation of FALCON coordinates in the full vibrational space in which the vibrations are localized to the two OH-groups and the benzene ring (B2PLYP-D3/aug-cc-pVTZ). IC refers to inter-group FALCON modes.

S3.4.2   36 normal modes

a Q6: 190   b Q7: 247   c Q8: 296   d Q9: 310   e Q10: 440   f Q11: 447   g Q12: 461

h Q13: 559   i Q14: 564   j Q15: 590   k Q16: 724   l Q17: 757   m Q18: 782   n Q19: 852

o Q20: 866   p Q21: 932   q Q22: 967   r Q23: 1054   s Q24: 1111

t Q25: 1173   u Q26: 1178   v Q27: 1212   w Q28: 1273   x Q29: 1309   y Q30: 1369

z Q31: 1410   aa Q32: 1509   ab Q33: 1549   ac Q34: 1652   ad Q35: 1659   ae Q36: 3175

af Q37: 3199   ag Q38: 3212   ah Q39: 3222   ai Q40: 3784   aj Q41: 3839
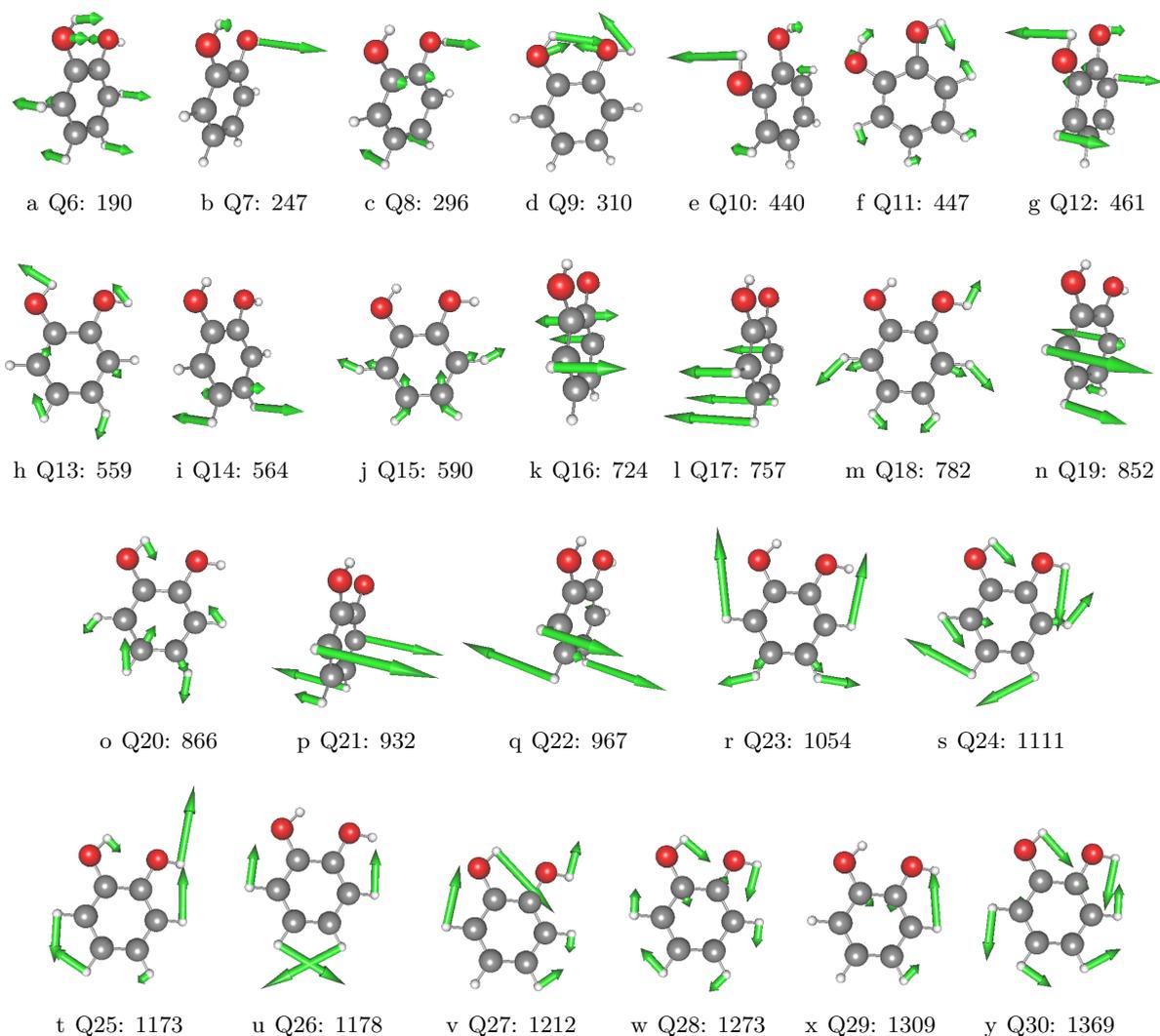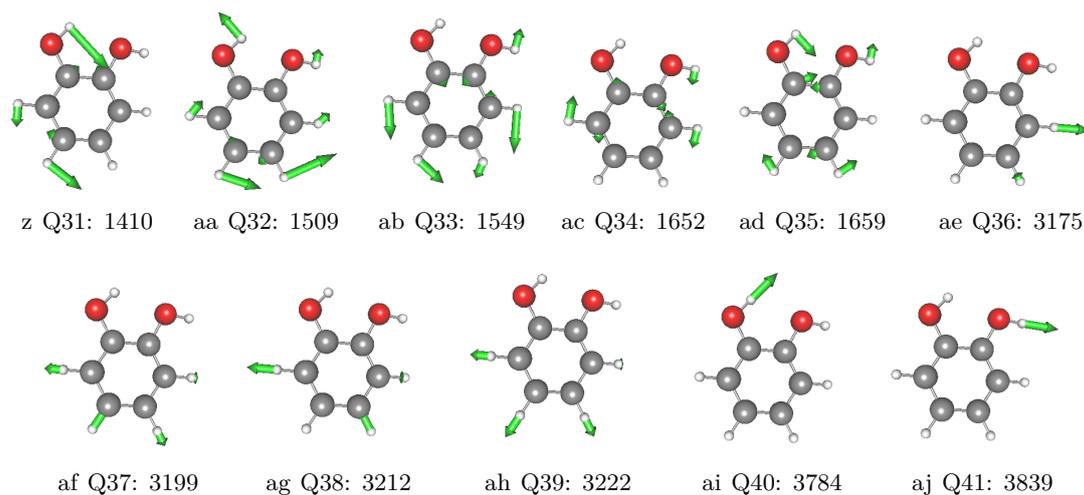
Fig. S13: Visualisation of all normal modes with their harmonic frequencies in cm^{-1} (B2PLYP-D3/aug-cc-pVTZ).

S3.4.3   2MC potential of Q6/Q41 and Q11/Q40 in the 36 FALCON mode space
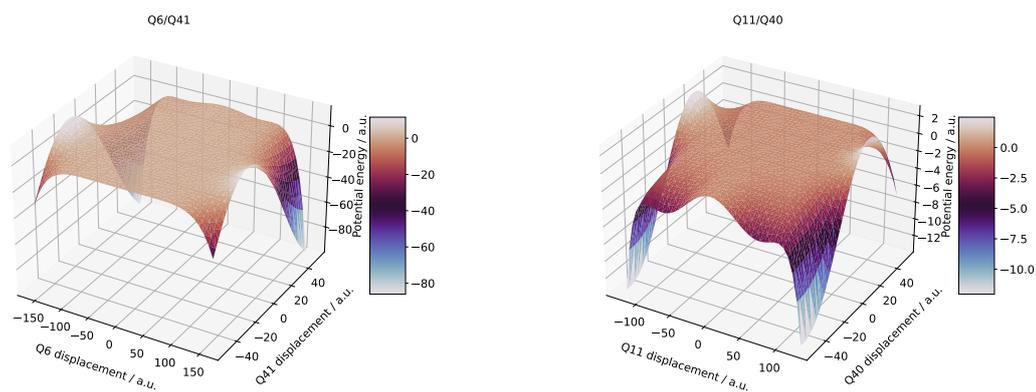
Fig. S14: 36 FALCON space: Two-dimensional potential cut of the anharmonic potential of the Q6 mode (relative out-of-plane wagging) and Q41 mode (stretching) of the free OH group (left) and two-dimensional potential cut of the anharmonic potential of the Q11 (relative out-of-plane wagging) and Q40 (stretching) of the bound OH group.

18

S3.4.4 2MC potential of Q7/Q41 and Q10/Q40 in the 36 normal mode space
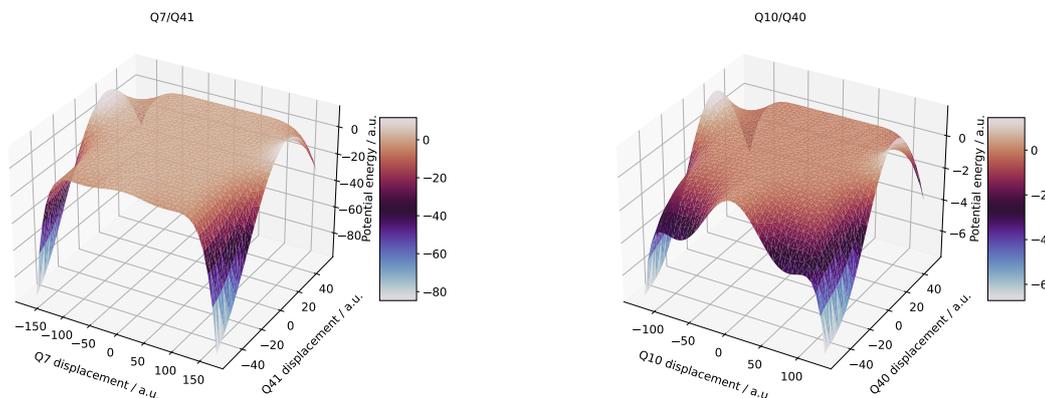
Fig. S15: 36 normal space: Two-dimensional potential cut of the anharmonic potential of the Q7 mode (relative out-of-plane wagging) and Q41 mode (stretching) of the free OH group (left) and two-dimensional potential cut of the anharmonic potential of the Q10 (relative out-of-plane wagging) and Q40 (stretching) of the bound OH group.

S3.5 Data to multilevel surface generation on 36 FALCON and 36 normal mode space

S3.5.1 Coupling Estimates

Tab. SVI: Largest coupling strength in a.u. obtained from a VSCF calculation on a 2MC potential on all 36 normal modes of catechol (r2SCAN-3c).

Mode 1	Mode 2	Coupling
7	41	5.8214
10	40	4.9144
12	40	3.1777
25	41	2.8142
27	40	2.5359
8	41	1.5025
31	40	1.0013
30	41	0.8014

Tab. SVII: Largest coupling strength in a.u. obtained from a VSCF calculation on a 2MC potential on all 36 FALCON modes of catechol (r2SCAN-3c).

Mode 1	Mode 2	Coupling
6	41	0.0234
11	40	0.0213
28	41	0.0114
30	40	0.0106

36 FALCON space				36 NC space			
3613 cm^{-1}		3683 cm^{-1}		3632 cm^{-1}		3703 cm^{-1}	
Coefficient	Mode	Coefficient	Mode	Coefficient	Mode	Coefficient	Mode
0.872	Q40$_1$	0.762	Q41$_1$	0.746	Q40$_1$	0.653	Q41$_1$
0.038	Q11$_2$	0.046	Q6$_2$	0.028	Q7$_2$Q24$_1$	0.051	Q7$_2$
0.017	Q30$_2$	0.015	Q28$_2$	0.021	Q10$_2$	0.019	Q27$_1$Q28$_1$
0.007	Q11$_4$	0.014	Q6$_2$Q28$_2$	0.011	Q10$_1$Q12$_1$	0.018	Q6$_1$Q14$_1$
0.005	Q30$_3$	0.007	Q30$_3$	0.010	Q27$_2$Q28$_1$	0.017	Q7$_2$Q24$_2$

Tab. SVIII: Coefficients to the fundamental vibrational excitations of the two OH stretching vibrations obtained in the VCC[2] response framework on the LL PES for the vibrational spaces described in FALCON and normal coordinates (NC) for catechol. For further information see main text.

S4   Estimated Number of Single Points

For a better comparison of the efficiency of the protocol, we estimated the number of single points for a PES generation on the full vibrational space for 4MC coupling terms. The total number of single-point N_{points} calculations required for the computation of the PES is given by

$$N_{\text{points}} = \sum_{k=1}^{n} \binom{M}{k} g_k^k.$$ (1)

Here, M refers to the total number of vibrational coordinates, the PES is a function thereof and g_k refers to the number of single points (SPs) per direction in the generation of the direct grid mode-coupling level k.[4] As the ADGA procedure places the grid points efficiently according to the needs of the vibrational wave function, the number of g_k per grid for the respective coupling level varies for each molecule and each mode-coupling level. However, to get a rough estimation of g_k for each mode-coupling level for uracil and catechol, respectively, g_k have been estimated from performed calculations with information on the total number of SP calculations N_k for the respective mode-coupling level with

$$g_k^k = \frac{N_k}{\binom{M}{k}}.$$ (2)

For the uracil molecule, the required SP calculations for the generation of the $9M$ 4MC in normal coordinates were taken. For the catechol molecule, the required SP calculations for the generation of the $12M$ 4MC in FALCON coordinates were taken. The resulting estimation for

g_k and the taken number of SP N_k are shown in Table SIX.

Molecule	PES	k	N_k	g_k^k	g_k
Uracil	9M 4MC	1	329	37	37
		2	5660	156	12
		3	43589	519	8
		4	2512	20	2

Tab. SIX: Number of SP N_k per mode-coupling level k and estimated SP calculations per grid g_k for the mode coupling level k for an ADGA PES generation on 9M and 4MC coupling terms of uracil and 12M and 4MC coupling terms of catechol.

Tab. SX displayed the estimated total number of SP required for a PES generation of catechol (36 M) and uracil (30 M) on the full vibrational space for 4MC coupling terms by applying eq. 2.

Molecule	N_{points}
Uracil	2 722 650

Tab. SX: Estimated total number of SPs required for a PES generation on the full vibrational space and 4MC coupling terms. This number is estimated by applying eq. 1.

References

[1] A. Y. Ivanov, A. M. Plokhotnichenko, E. D. Radchenko, G. G. Sheina, Y. P. Blagoi, *J. Mol. Struct. THEOCHEM* **1995**, *372*, 91–100.

[2] NIST, **2023**, `https://webbook.nist.gov/cgi/cbook.cgi?ID=C120809&Units=SI&Mask=80#IR-Spec`.

[3] E. L. Yang, J. J. Talbot, R. J. Spencer, R. P. Steele, *J. Chem. Phys.* **2023**, *159*.

[4] O. Christiansen, *Phys. Chem. Chem. Phys.* **2012**, *14*, 6672–6687.