Supporting Information

Resorcinol Crystallization from the Melt: A New Ambient Phase and New "Riddles"

Qiang Zhu,[1] Alexander G. Shtukenberg,[2] Damien J. Carter,[3] Tang-Qing Yu,[4] Jingxiang Yang,[2] Ming Chen,[4] Paolo Raiteri,[3] Artem R. Oganov,[1] Boaz Pokroy,[5] Iryna Polishchuk,[5] Peter J. Bygrave,[6] Graeme M. Day,[6] Andrew L. Rohl,[3]* Mark E. Tuckerman[4,7]* and Bart Kahr [2]*

[1]Department of Geosciences, Stony Brook University, Stony Brook, NY, 11794, USA

[2]Department of Chemistry and Molecular Design Institute, New York University, New York City, NY, 10003, USA

[3]Curtin Institute for Computation, Nanochemistry Research Institute and Department of Chemistry, Curtin University, P.O. Box U1987, Perth, 6845, Western Australia, Australia

[4]Department of Chemistry and Courant Institute, New York University, New York City, NY, 10003, USA

[5]Department of Materials Science and Engineering and the Russell Berrie Nanotechnology Institute, Technion Israel Institute of Technology, Haifa 32000, Israel

[6]School of Chemistry, University of Southampton, Highfield, Southampton, SO17 1BJ, UK

[7]New York University-East China Normal University Center for Computational Chemistry at NYU Shanghai, 3663 Zhongshan Road North, Shanghai 200062, China

*a.rohl@curtin.edu.au, mark.tuckerman@nyu.edu, bk66@nyu.edu

Force Field Generation:

We initially attempted to predict the structure of the new polymorph using the molecular dynamics approach of Buch *et al*.[1] Here, NVT molecular dynamics (MD) simulations of a small cell at a high temperature are employed to explore phase space. Periodic snapshots are taken from the MD trajectory and optimized at constant pressure. The resorcinol specific force field of Chatchawalsaisin et al.[2] was used although a number of changes

were made. Firstly we utilized RESP-A1A partial atomic charges calculated using the RESP ESP charge Derive (RED) Server web service[3], as these charges underpin the AMBER force field from which the bonded terms are taken. These charges are tabulated in Table S1.

Table S1 Derived partial atomic charges for resorcinol. Atom numbering follows Ref[2]

Atom	Atom type	Charge
1	HC	0.1528
2	C	-0.2033
3 / 7	C	0.2035
4 / 6	C	-0.1471
5	C	-0.2620
8 / 9	O	-0.5294
10 / 11	HO	0.4009
12 / 13	HC	0.1366
14	HC	0.1835

Lennard-Jones 9-6 potentials were fitted to the Buckingham potentials utilised by Chatchawalsaisin et al. as the latter tend to minus infinity at short distances, which is problematical with the high temperature MD. The C-C-C-H and C-C-C-O out of plane potentials were fitted to a $k.d^2$ potential, where d is the distance of the middle atom to the plane of the other three again for robustness; the fitted value of k for both was 14.7 kJ/mol/Å^2. Finally we refitted the C-C-O-H torsional term to the quantum mechanical data presented by Chatchawalsaisin et al. to include the 1-4 interactions yielding a force constant of 4.8 kJ/mol.

Twelve starting structures for the MD were generated by randomly inserting four resorcinol molecules into a unit cell with the same dimensions as the α form. 4 ns of MD with a timestep of 0.5 fs was run on each of these at 400 K, 450 K, 500 K, 550 K and 600 K with the trajectory written every 5 ps using the program GULP[4]. GULP was also used to optimize all of the configurations within the trajectory file at constant pressure. Although both the α and β forms were found with this method, a number of new structures were found that were lower in potential energy than these two. These were then optimized using density functional theory (DFT) calculations using the Quantum ESPRESSO code [5] using the projector-augmented wave (PAW) method [6]. The details are found in the main manuscript. All of these structures were found to be higher in energy than the α and β forms. A sample of 42 of the generated structures were taken (including α and β) and optimized using the DFT procedure outlined above and then the Lennard-Jones potentials were fitted to match the DFT derived lattice parameters and energies. The final values for the Lennard-Jones parameters are given in table S2.

Table S2 Lennard Jones parameters for resorcinol

		A (Å^9 kJ/mol)	B (Å^6 kJ/mol)
C	C	151284.7324	3171.00
C	O	90317.4601	2463.00
O	O	26991.9910	0
C	HC	19734.7941	828.75
O	HC	18661.8185	660.10
C	HO	7648.3229	319.00
HC	HC	1375.2233	0

O	HO	553.2396	0
HO	HC	1656.0560	0
HO	HO	169.8626	0

An appreciation for the quality of this force field can be gained from comparing the experimentally observed lattice parameters to the calculated ones as in Table S3

Table S3 Experimental and calculated lattice parameters for resorcinol

	A (Å)	**b (Å)**	**c (Å)**
α-resorcinol			
Expt.[7]	10.53	9.53	5.60
Potentials	10.5169 **(-0.12%)**	9.5624 **(0.34%)**	5.4844 **(-2.06%)**
DFT (XDM)	10.4135 **(-1.11%)**	9.3666 **(-1.71%)**	5.5957 **(0.00%)**
β-resorcinol			
Expt [8]	7.934	12.606	5.511
Potentials	8.0096 **(0.95%)**	12.4968 **(-0.87%)**	5.3671 **(-2.61%)**
DFT (XDM)	7.8240 **(-1.39%)**	12.5264 **(-0.63%)**	5.4111 **(-1.81%)**

Crystallographic Information File (CIF) for *ε* resorcinol polymorph from USPEX

```
data_qiang042615_fixcell_after
_audit_creation_date              2015-10-19
_audit_creation_method            'Materials Studio'
_symmetry_space_group_name_H-M    'P 21 21 21'
_symmetry_Int_Tables_number       19
_symmetry_cell_setting            orthorhombic
loop_
_symmetry_equiv_pos_as_xyz
  x,y,z
  -x+1/2,-y,z+1/2
  -x,y+1/2,-z+1/2
  x+1/2,-y+1/2,-z
_cell_length_a                    17.7750
_cell_length_b                    10.7150
_cell_length_c                    5.7330
_cell_angle_alpha                 90.0000
_cell_angle_beta                  90.0000
_cell_angle_gamma                 90.0000
loop_
_atom_site_label
_atom_site_type_symbol
_atom_site_fract_x
_atom_site_fract_y
_atom_site_fract_z
_atom_site_U_iso_or_equiv
_atom_site_adp_type
_atom_site_occupancy
C1     C     0.99075   1.02220   1.05173   1.00000
C2     C     0.92999   1.01733   0.89739   1.00000
C3     C     1.05164   0.94074   1.02881   1.00000
C4     C     0.93016   0.92860   0.71892   1.00000
C5     C     1.05074   0.85368   0.84815   1.00000
C6     C     0.99057   0.84622   0.69269   1.00000
O7     O     0.99311   1.10871   1.23030   1.00000
O8     O     0.86902   0.92616   0.57175   1.00000
H9     H     0.94180   1.14088   1.26307   1.00000
H10    H     0.88234   1.08007   0.91794   1.00000
H11    H     1.09835   0.94540   1.15109   1.00000
H12    H     1.09786   0.78964   0.82939   1.00000
H13    H     0.99035   0.77711   0.55417   1.00000
H14    H     0.87406   0.85776   0.45444   1.00000
C15    C     0.21507   0.71478   1.13417   1.00000
C16    C     0.26734   0.69523   0.95586   1.00000
C17    C     0.22411   0.81258   1.29348   1.00000
C18    C     0.33001   0.77377   0.93991   1.00000
C19    C     0.28691   0.88982   1.27354   1.00000
C20    C     0.34040   0.87155   1.09858   1.00000
O21    O     0.15353   0.63833   1.15996   1.00000
O22    O     0.38058   0.75030   0.76458   1.00000
H23    H     0.15450   0.56342   1.05616   1.00000
H24    H     0.26057   0.61816   0.83376   1.00000
H25    H     0.18320   0.82607   1.43246   1.00000
H26    H     0.29441   0.96537   1.39837   1.00000
H27    H     0.38990   0.93111   1.08740   1.00000
H28    H     0.42546   0.80554   0.77330   1.00000
#END
```

Crystallographic Information File (CIF) for $P2_1$ resorcinol polymorph from USPEX

```
##################################################
#                                                #
#Generated by USPEX: Crystal Structure Prediction#
#                                                #
##################################################
data_P21 from VASP: -89.6286 eV/molecule

_symmetry_space_group_name_H-M 'P 1 21 1'
_symmetry_Int_Tables_number 4

_cell_length_a       9.38000
_cell_length_b       5.46600
_cell_length_c      10.59100
_cell_angle_alpha   90.00000
_cell_angle_beta    89.15300
_cell_angle_gamma   90.00000

loop_
_atom_site_label
_atom_site_type_symbol
_atom_site_fract_x
_atom_site_fract_y
_atom_site_fract_z
_atom_site_occupancy
C1  C    0.07459  -0.44955  -0.26409   1.00000
C2  C    0.03928   0.33016  -0.32291   1.00000
C3  C    0.17960  -0.29582  -0.31471   1.00000
C4  C    0.10964   0.26848  -0.43576   1.00000
C5  C    0.24774  -0.36149  -0.42792   1.00000
C6  C    0.21375   0.41976  -0.48940   1.00000
C7  C   -0.25447   0.12615  -0.08570   1.00000
C8  C   -0.28429  -0.09218  -0.02171   1.00000
C9  C   -0.32930   0.19175  -0.19350   1.00000
C10 C  -0.39141  -0.24384  -0.06699   1.00000
C11 C  -0.43737   0.03777  -0.23580   1.00000
C12 C  -0.46889  -0.18220  -0.17411   1.00000
O1  O    0.00676  -0.37209  -0.15504   1.00000
O2  O    0.34709  -0.20829  -0.48263   1.00000
O3  O   -0.15243   0.28565  -0.04406   1.00000
O4  O    0.48262   0.10644  -0.33804   1.00000
H1  H   -0.06552  -0.49432  -0.12455   1.00000
H2  H   -0.04176   0.20946  -0.28177   1.00000
H3  H    0.20427  -0.12301  -0.26810   1.00000
H4  H    0.08193   0.09764  -0.48274   1.00000
H5  H    0.26756   0.37000   0.42202   1.00000
H6  H    0.38056  -0.07841  -0.42307   1.00000
H7  H   -0.10292   0.22350   0.03300   1.00000
H8  H   -0.22416  -0.14219   0.06199   1.00000
H9  H   -0.30652   0.36396  -0.24211   1.00000
H10 H  -0.41477  -0.41479  -0.01782   1.00000
H11 H   0.44728  -0.30257  -0.20871   1.00000
H12 H  -0.45995   0.21192  -0.39688   1.00000
#END
```

Crystallographic Information File (CIF) for ε resorcinol polymorph from refinement of synchrotron data at 200K

```
##############################################################################
# Gamma resorcinol after refinement as rigid body, T=200K
##############################################################################

data_findsym-STRUC-95
_symmetry_cell_setting          orthorhombic
_symmetry_space_group_name_H-M   'P 21 21 21'
_symmetry_Int_Tables_number      19
loop_
_symmetry_equiv_pos_site_id
_symmetry_equiv_pos_as_xyz
1 x,y,z
2 1/2-x,-y,1/2+z
3 -x,1/2+y,1/2-z
4 1/2+x,1/2-y,-z
_cell_length_a                 17.8998
_cell_length_b                 10.5680
_cell_length_c                 5.7221
_cell_angle_alpha              90.00000
_cell_angle_beta               90.00000
_cell_angle_gamma              90.00000
_cell_volume                   1082.42
loop_
_atom_site_label
_atom_site_type_symbol
_atom_site_fract_x
_atom_site_fract_y
_atom_site_fract_z
_atom_site_U_iso_or_equiv
_atom_site_adp_type
_atom_site_occupancy
C1    C    0.98823  1.02315  1.06017  0.00000  Uiso   1.00
C2    C    0.92802  1.01777  0.90509  0.00000  Uiso   1.00
C3    C    1.04917  0.94163  1.03662  0.00000  Uiso   1.00
C4    C    0.92881  0.92843  0.72523  0.00000  Uiso   1.00
C5    C    1.04890  0.85398  0.85458  0.00000  Uiso   1.00
C6    C    0.98930  0.84598  0.69835  0.00000  Uiso   1.00
O7    O    0.98995  1.11028  1.24012  0.00000  Uiso   1.00
O8    O    0.86821  0.92548  0.57736  0.00000  Uiso   1.00
H9    H    0.93880  1.14195  1.27301  0.00000  Uiso   1.00
H10   H    0.88033  1.08053  0.92612  0.00000  Uiso   1.00
H11   H    1.09545  0.94669  1.15948  0.00000  Uiso   1.00
H12   H    1.09607  0.78989  0.83531  0.00000  Uiso   1.00
H13   H    0.98956  0.77639  0.55874  0.00000  Uiso   1.00
H14   H    0.87369  0.85662  0.45905  0.00000  Uiso   1.00
C15   C    0.21748  0.71685  1.12949  0.00000  Uiso   1.00
C16   C    0.26938  0.69844  0.95031  0.00000  Uiso   1.00
C17   C    0.22651  0.81466  1.29190  0.00000  Uiso   1.00
C18   C    0.33166  0.77810  0.93660  0.00000  Uiso   1.00
C19   C    0.28891  0.89304  1.27416  0.00000  Uiso   1.00
C20   C    0.34202  0.87589  1.09837  0.00000  Uiso   1.00
O21   O    0.15633  0.63922  1.15312  0.00000  Uiso   1.00
O22   O    0.38187  0.75569  0.76030  0.00000  Uiso   1.00
H23   H    0.15726  0.56414  1.04699  0.00000  Uiso   1.00
H24   H    0.26262  0.62133  0.82578  0.00000  Uiso   1.00
H25   H    0.18589  0.82724  1.43151  0.00000  Uiso   1.00
H26   H    0.29640  0.96858  1.40138  0.00000  Uiso   1.00
H27   H    0.39121  0.93629  1.08889  0.00000  Uiso   1.00
H28   H    0.42646  0.81155  0.77062  0.00000  Uiso   1.00
#END
```

Crystal Structure Prediction study using GLEE quasi-random search

Overall Scheme

Rigid molecule structure generation with one molecule in the asymmetric unit (Z'=1) was performed using the GLEE program, starting from 11 different molecular conformations. Structure generation and geometry optimisation was conducted until 2,000 valid lattice energy-minimised structures were generated with Z'=1 in each of the space groups; 1, 2, 4, 5, 7, 9, 14, 15, 18, 19, 29, 33, 60, 61, 76, 78, 92, 96 and 144. From these results the planar starting conformations C6O2H6_OPLS_84, C6O2H6_OPLS_85, C6O2H6_OPLS_88 generated the lowest energy crystal structures and the 6 possible combinations of these was used to conduct structure prediction with two molecules in the asymmetric unit (Z'=2). The naming convention of C6O2H6_84x85 was used to signify combining conformer 84 and 85. Due to the increased search space, 10,000 valid energy-minimised structures were generated in each of the common Z'=2 space groups; 1, 2, 4, 5, 7, 9, 14, 15, 19, 29, 33 and 61. The resulting structures were clustered in each space group to remove duplicate structures. COMPACK was used for the structure comparison algorithm with a 25 molecule comparison and standard tolerances. We also used these clustering results to ensure good convergence was found for our search.

From these rigid-molecule structures, those with a total energy within 20 kJ/mol of the global minimum were selected for further geometry optimisation in the CrystalOptimizer code. This step allows the molecular geometry to relax due to packing forces. This step allowed dihedrals angles within the hydroxyl groups to relax within each crystal structures.

Finally these flexible-molecule optimised structure had a single rigid-molecule optimisation performed that incorporate polarisation effects. The molecular energy and density is computed within a PCM environment, and the DMACRYS code was employed to optimise the crystal structure from the resulting polarised multipole representation.

Clustering was then performed across the entire set of structures to remove duplicates. These final structures form our crystal structure prediction landscape.

Rigid-Molecule Workflow

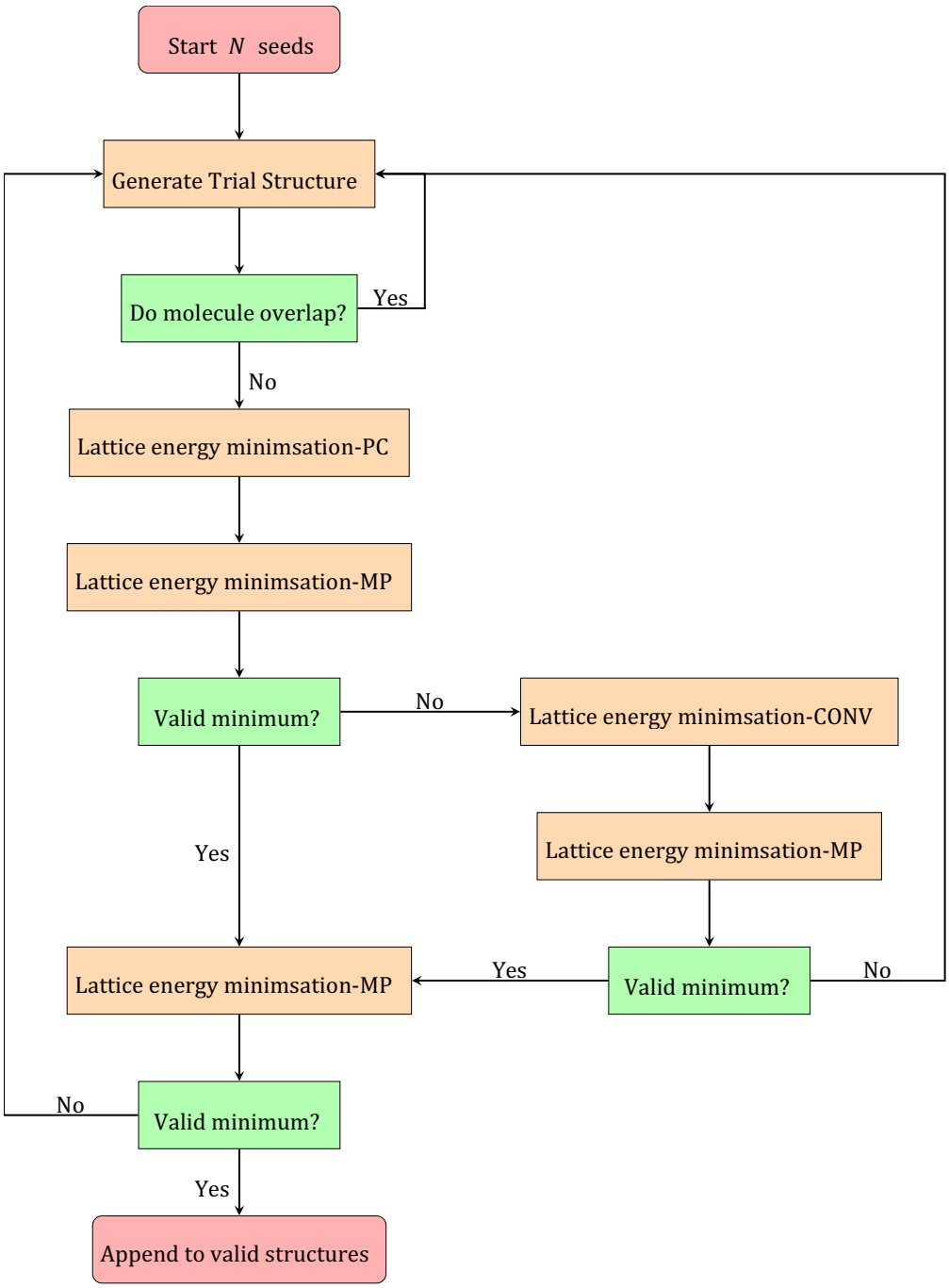

Figure S1: Flow diagram of GLEE CSP procedure. Conducted for each conformation for every space group

Structure Generation Settings

2000 valid structure were made in space groups for $Z' = 1$ search; space groups 1, 2, 4, 5, 7, 9, 14, 15, 18, 19, 29, 33, 60, 61, 76, 78, 92, 96 and 144 were used. 10000 valid

structures were made in space groups with 2 independent molecules in the asymmetric unit; space groups 1, 2, 4, 5, 7, 9, 14, 15, 19, 29, 33 and 61 were used. The SAT-Expansion technique was used.

DMACRYS: Rigid-Molecule Lattice Energy Minimisation

Options	Value
Basis set	6-311G**
Functional	B3LYP
Dispersion correction	GD3BJ
Point charge determination	CHELPG
Max PC lattice energy iterations	80
Multipole Method	DMA
Multipole Level	0-4
Max multipoles lattice energy iterations	1000
Buckingham potential	W99 with revised H···X terms for 6-311G**
Van der Waals cutoff	15 Angstrom
External iteration PCM	False

Table S4: Options used during rigid molecule search and optimisation

For every structure a minimisation using point charges (-PC) is performed with a maximum of 80 iterations. Every structures is then optimised using multipoles (-MP). If this fails then a constant volume optimisation (-CONV) is performed followed by a standard multipole optimisation (-MP). If the structure is valid from either the first multipole or constant volume branch, a secondary minimisation (-MP) is performed to further remove any problem structures.

Starting conformers for the rigid-molecule crystal structure search

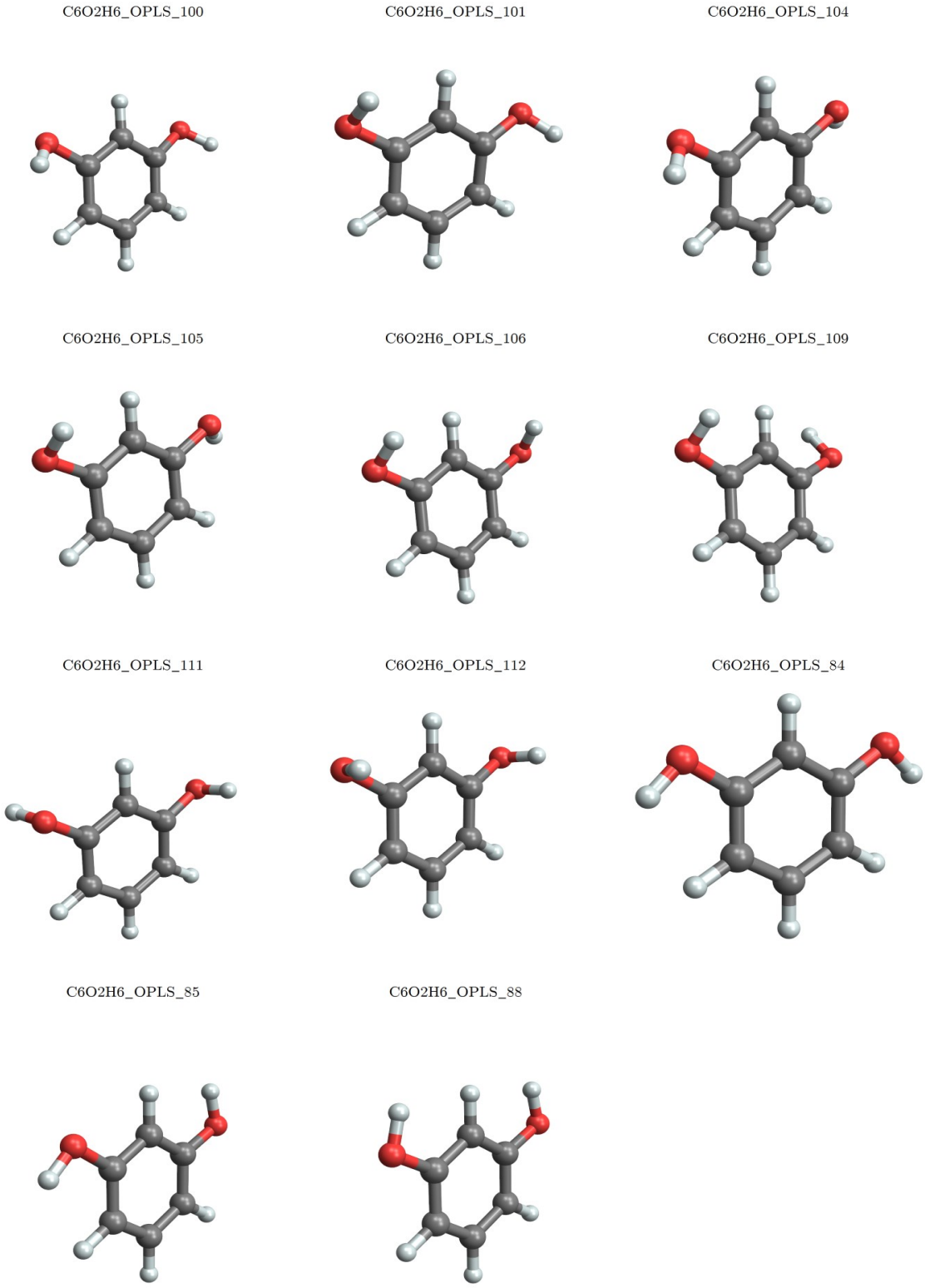

Figure S2: Conformers used in rigid search

Clustering using COMPACK

Standard COMPACK settings were used with a RMSD25 and incorporating hydrogen positions into the comparison.

CrystalOptimizer: Flexible-Molecule Total Energy Minimisation

Here we select which degrees-of-freedom are optimised directly in response to packing forces. The remaining are optimised in response to this new conformation, but in the gas phase. The flexible DOFs: O1-C5-C4-C3, O2-C3-C2-C1, H5-O1-C5-C4, H6-O2-C3-C2, H5-O1-C5, H6-O2-C3.

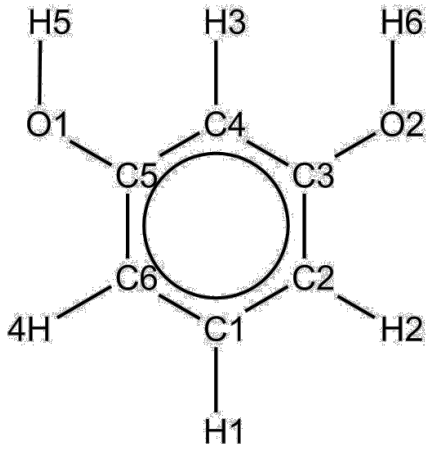

Figure S3: resorcinol molecule with labels

DMACRYS: Final PCM Rigid-Molecule Lattice Energy Minimisation

Options	Value
Basis set	6-311G**
Functional	B3LYP
Dispersion Correction	GD3BJ
Buckingham potential	W99 with revised H···X terms for 6-311G**
Van der Waals cutoff	25 Angstrom
External iteration PCM	True
Dielectric constant	$3.0\varepsilon_0$

Table S5: Options used during PCM section

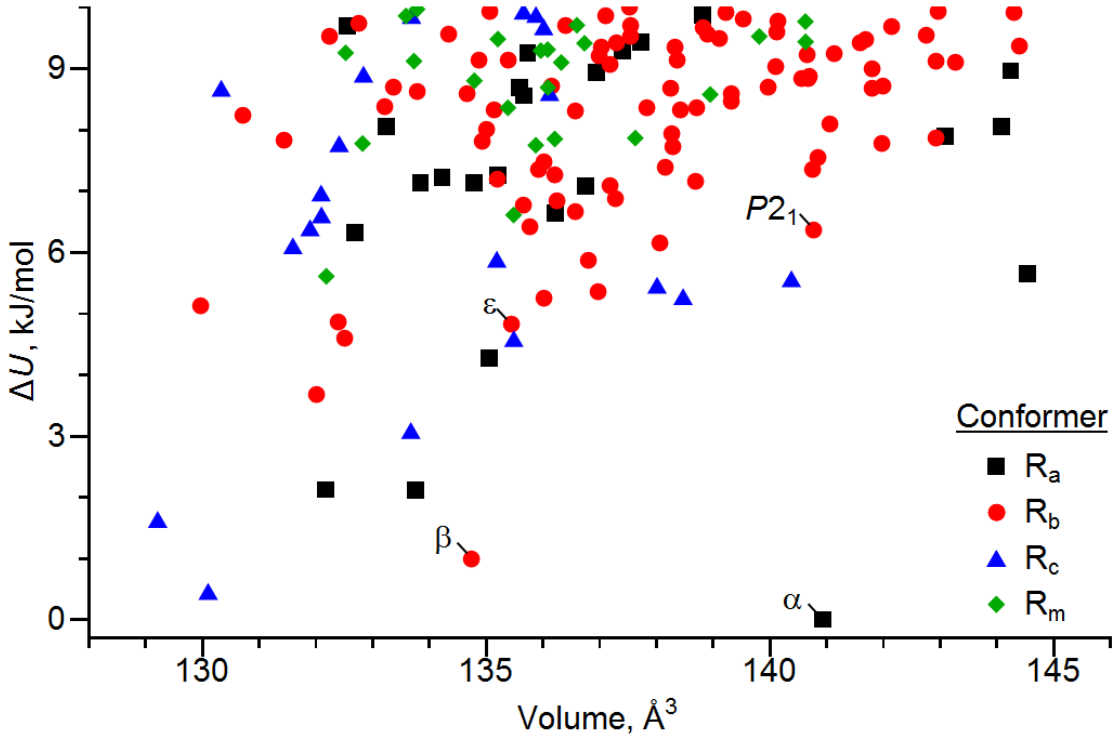

Figure S4: Final CSP Landscape from GLEE crystal structure prediction. All crystal structures within 10 kJ/mol of the global minimum from the GLEE search are submitted in CIF format and are also available from DOI: 10.5258/SOTON/XXXXXX

References

1. V. Bush, R. Martoňák, M. Parrinello, *J. Chem. Phys.,* **2005**,*123*, 051108.
2. J. Chatchawalsaisin, J. Kendrick, S. C. Tuble, J. Anwar, *CrystEngComm*, **2008**, *10*, 437.
3. E. Vanquelef, S. Simon, G. Marquant, E. Garcia, G. Klimerak, J. C. Delepine, P. Cieplak and F.-Y. Dupradeau, R.E.D. Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragment, Nucl. Acids Res. (Web server issue) **2011**, W511-W517,http://q4md-forcefieldtools.org/REDS/
4. J. D. Gale and A. L. Rohl, *Mol. Simul.,* **2003**, *29*, 291.
5. P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, C. A. Dal,.; S. de, Gironcoli; S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, R. M. Wentzcovitch, *J. Phys: Condensed Matter*, **2009**, *21*, 395502.
6. P. E. Blöchl, *Phys. Rev. B,* **1994**, *50*, 17953.
7. G. E. Bacon, R. J. Jude, *Z. Kristallogr*. **1973**, *138*, 19.
8. G. E. Bacon, E. J. Lisher, *Acta Crystallogr. Sect. B*., **1980**, *B36*, 1908.