

Symmetry-Constrained Monte Carlo for the crystal structure prediction of small organic molecules

Giovanni Macetti,^{*,†} Luca Sironi,[†] Margherita Vacchini,[†] Leonardo Lo Presti[†]

[†]Università degli Studi di Milano, Dipartimento di Chimica, Via C. Golgi 19, 20133 Milano, Italy

* giovanni.macetti@unimi.it

Supporting Information

Contents

S1 Symmetry-Constrained Monte Carlo	3
S1.1 Work flow	3
S1.2 Input parameters	5
S1.3 Input file	6
S1.3.1 Benzoic Acid - BZA	6
S1.3.2 Acrolein - ACR	8
S1.3.3 Isocaffeine - ISC	10
S1.4 Structures overlap algorithm	12
S2 Ranking results	14
S2.1 Benzoic Acid	14
S2.2 Acrolein	18
S2.3 Isocaffeine	22
S3 Structures Overlap	26

S1 Symmetry-Constrained Monte Carlo

S1.1 Work flow

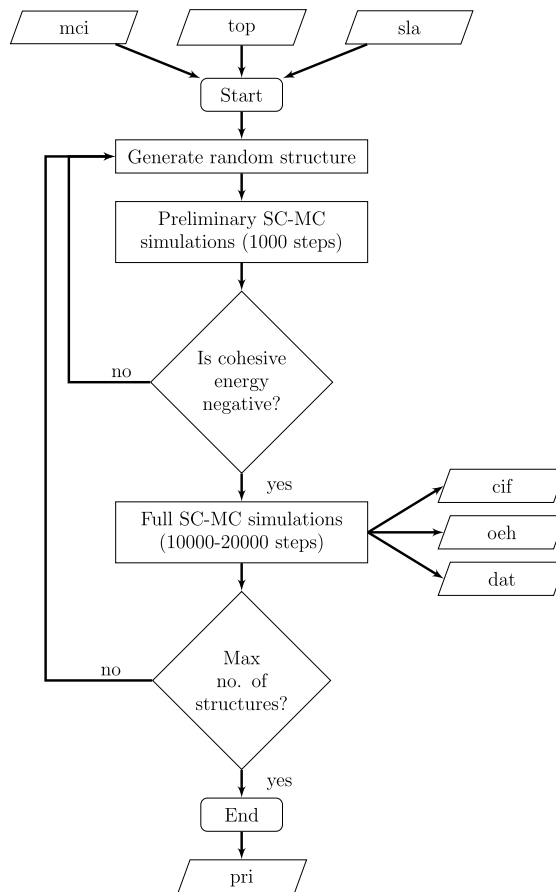


Figure S1: Flowchart of the Symmetry-Constrained Monte Carlo (SC-MC) algorithm.

The SC-MC algorithm is a fully automated procedure that generates a selected number of plausible structures across different space groups. Figure S1 shows a schematic flow chart that illustrates the overall algorithm. For those seeking a more rigorous and detailed description, the following list outlines the actual step-by-step workflow of the method:

1. Prepare the three required input files: (a) a topology file (*top*) and (b) a slave file (*sla*), as described in the *MiCMoS* manual for the Monte Carlo section¹; (c) an input file (*mci*). The input file contains all the information for the random search of crystal structures, including the space groups to be explored, the experimental conditions (temperature, pressure), and the algorithm parameters (cutoff, force field, number of steps, number of structures to generate, etc.). A more detailed description of the input file can be found in Sections S1.2 and S1.3.

2. The algorithm begins by generating random structures using lattice parameters in the range 5–25 Å for a , b , and c , and 90–120° for the α , β , and γ angles.
3. Following user's preference, one or more molecules are randomly positioned within the asymmetric unit by selecting a set of coordinates for the molecular centers of mass and three Euler angles for their rigid-body orientations.
4. A preliminary SC-MC simulation is performed. The degrees of freedom (DOFs) that can relax include symmetry-independent lattice parameters (as defined by the chosen Bravais lattice), the position and orientation of the molecule as a rigid body, and the internal DOFs specified in the topology and slave files. The full crystal packing is built on-the-fly at each step, updating all molecules consistently with the active symmetry operators. Typically, 1000 steps are sufficient at this stage.
5. After the preliminary SC-MC run, the cohesive energy of the generated structure is evaluated. If the energy is negative, the structure is considered promising and the algorithm proceeds to the next step. Otherwise, the structure is discarded and the cycle restarts from step 2.
6. For each accepted structure, the count of valid structures is incremented by one, and a longer SC-MC simulation is launched. At this stage, a total of 10,000 to 20,000 steps are typically performed.
7. Upon completion of the simulation, the minimum-energy structure along the trajectory is extracted. Its energy and lattice parameters are stored in memory and printed in the final ranking. Additionally, the minimum-energy geometry is saved in various formats for visualization (*cif*, universal file) and further processing (*oeh*, MiCMoS format). The full trajectory (*dat* file) is also saved for potential analysis.
8. If the target number of structures has been reached, the algorithm proceeds to the final step. Otherwise, the cycle restarts from step 2.
9. A final table of all predicted structures, sorted by space group and ranked by cohesive energy, is printed in the *pri* file.

The SC-MC manual, as well as the source code, can be freely downloaded from the MiCMoS website: https://sites.unimi.it/xtal_chem_group/index.php/research/33-sc-mc .

S1.2 Input parameters

Table S1: Input parameters for the Symmetry-Constrained Monte Carlo

Parameter	Benzoic Acid	Acrolein	Isocaffeine
Structure search			
Space Groups	6 [2,4,14,15,19,61]	6 [2,4,14,15,19,61]	6 [2,4,14,15,19,61]
No. of structures per run	200-500	200-500	200-500
No. of run (total)	10	10	10
No. of run with restrain [SG]	4	-	-
Lattice ratio (a,b,c)	[1,1,4]	-	-
Monte Carlo parameters			
Force field	LJC	LJC	LJC
Cutoff	20 Å	20 Å	20 Å
Temperature	300 K	133 K	300 K
Pressure	1 bar	1 bar	1 bar
Seeds generation			
Preliminary No. steps	1000	1000	1000
Production No. step	10000-20000	10000	10000

S1.3 Input file

S1.3.1 Benzoic Acid - BZA

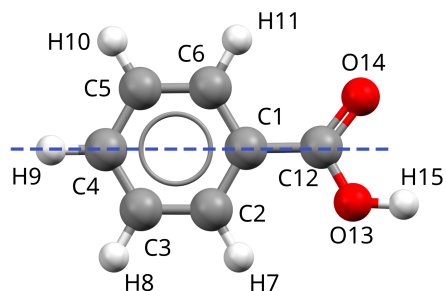


Figure S2: Atomic labels for the benzoic acid molecule. The dashed blue line represents the reference axis used for the rotational correlation analysis.

Benzoic acid topology (*top*) file:

```
#BENZAC02 'P 21/c' topology MC
11
 1 -0.00013 -0.00497 0.17693 12 0.1028
 2 0.00727 1.20076 -0.50719 12 -0.1934
 3 0.00790 1.19827 -1.89054 12 -0.0681
 4 -0.00200 0.02343 -2.58619 12 -0.1361
 5 -0.00762 -1.17770 -1.90681 12 -0.0449
 6 -0.00426 -1.19365 -0.52904 12 -0.2206
 7 0.01252 2.13453 0.03549 2 0.1289
 8 0.01632 2.13602 -2.42626 2 0.1074
 9 -0.00540 0.03424 -3.66606 2 0.1134
10 -0.01472 -2.10748 -2.45637 2 0.1055
11 -0.00491 -2.13536 -0.00030 2 0.1482
 4 nslav-u
12 -1 0 1 2 3 10 0.6016
13 -1 0 12 1 2 28 -0.5857
14 0 0 12 1 13 27 -0.5303
15 -1 0 13 12 14 6 0.4714
 0 ncore-v
 0 nslav-v
112.0 0.0 volu-u,volu-v
 0 nstr-u
 0 nstr-v
 0 nbend-u
 0 nbend-v
 4 ntors-u
```

```
2 1 12 13 14.00 -1.0 2.0 C- C- C- 0
1 2 6 12 100.00 -1.0 1.0 C- C- C- C
14 12 13 15 35.00 -1.0 2.0 C- C- O- H
12 1 13 14 100.00 -1.0 1.0 C- C- O- 0
0 ntors-v
2 nlist-u
7 15 11 15
0 nlist-v
0.000 0.000 0.000 0.000
0 nextr
```

Benzoic acid slave (*sla*) file:

```
4
1.484 120.41 180.0 0 0 0 12 0 0
1.259 118.76 0.9 0 0 3 13 0 0
1.267 0.00 0.0 0 0 0 14 0 0
1.000 122.00 -0.2 0 0 3 15 0 0
```

S1.3.2 Acrolein - ACR

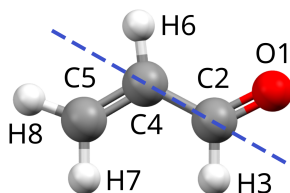


Figure S3: Atomic labels for the acrolein molecule. The dashed blue line represents the reference axis used for the rotational correlation analysis.

Acrolein topology (*top*) file:

```
#AXOSOW 'P b c a topology
4
1 -0.00010 0.10637 1.70341 27 -0.4791
2 0.00011 -0.35222 0.58610 10 0.5224
3 0.00035 -1.42822 0.49298 2 -0.0024
4 0.00009 0.42117 -0.64195 12 -0.1933
4 nslav-u
5 -1 0 4 2 1 12 -0.4022
6 0 0 4 2 5 2 0.1608
7 -1 0 5 4 2 2 0.1811
8 0 0 5 4 7 2 0.2127
0 ncore-v
0 nslav-v
59.7 0.0 volu-u,volu-v
0 nstr-u
0 nstr-v
0 nbend-u
0 nbend-v
5 ntors-u
1 2 4 5 11.00 -1.0 2.0 0- C- C- C
2 1 3 4 100.00 -1.0 1.0 C- O- H- C
2 4 5 7 50.00 -1.0 1.0 C- C- C- H
4 2 5 6 100.00 -1.0 1.0 C- C- C- H
5 4 7 8 100.00 -1.0 1.0 C- C- H- H
0 ntors-v
0 nlist-u
0 nlist-v
0.410 235.0 650.0 77000.0
0 nextr
```

Acrolein slave (*sla*) file:

```
4
1.31 121.85 180.00 0 0 3 5 0 0
1.08 0.00 0.00 0 0 0 6 0 0
1.08 120.0 0.00 0 0 0 7 0 0
1.08 0.0 0.00 0 0 0 8 0 0
```

S1.3.3 Isocaffeine - ISC

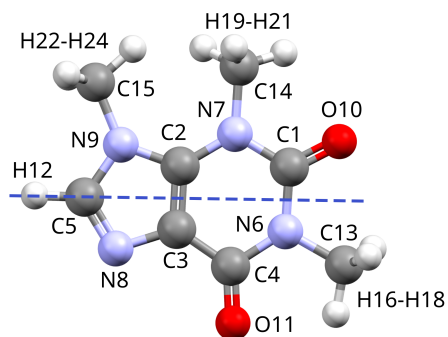


Figure S4: Atomic labels for the isocaffeine molecule. The dashed blue line represents the reference axis used for the rotational correlation analysis.

Isocaffeine topology (*top*) file:

```
#ISOCOFF 'P 22/n' topology
15
 1 0.01640 1.14390 -1.25543 10 0.4013
 2 0.00489 0.12111 0.87950 12 -0.2423
 3 0.00665 -1.13574 0.34551 12 0.1661
 4 -0.00447 -1.34098 -1.05625 10 0.4519
 5 0.00365 -1.39092 2.44154 12 0.0469
 6 -0.02176 -0.13895 -1.78628 17 -0.1296
 7 -0.00874 1.26716 0.12099 17 0.0722
 8 0.01094 -2.09102 1.34659 18 -0.5041
 9 0.01375 -0.02962 2.23506 17 0.3324
10 0.05381 2.11791 -1.98147 27 -0.4937
11 0.00399 -2.41632 -1.64207 27 -0.4839
12 -0.00916 -1.83687 3.42513 2 0.1465
13 -0.03142 -0.22639 -3.25552 13 -0.2609
14 -0.03289 2.61998 0.68034 13 -0.2340
15 -0.00999 0.94018 3.32454 13 -0.6563
 3 nslav-u
16 17 18 13 6 1 3 0.1296
19 20 21 14 7 2 3 0.1143
22 23 24 15 9 5 3 0.2186
 0 ncore-v
 0 nslav-v
167.2 0.0 volu-u,volu-v
 0 nstr-u
 0 nstr-v
```

```
0 nbend-u
0 nbend-v
3 ntors-u
1 6 13 16 2.00 -1.0 2.0 C- N- C- H
1 7 14 19 2.00 -1.0 2.0 C- N- C- H
2 9 15 22 2.00 -1.0 2.0 C- N- C- H
0 ntors-v
0 nlist-u
0 nlist-v
0.000 0.000 0.0 0.00000
0 nextr
```

Isocaffeine slave (*sla*) file:

```
3
1.08 109.5 60.0 0 0 3 16 17 18
1.08 109.5 60.0 0 0 3 19 20 21
1.08 109.5 60.0 0 0 3 22 23 24
```

S1.4 Structures overlap algorithm

The data produced by the SC-MC program at the end of the CSP search are raw, as there are inevitably duplicates among the printed structures. To reduce the total number of molecules, custom Fortran and Python utilities were developed to extract and compare the obtained structures. One of these utilities is a FORTRAN90 program called *compare_structures*, which is designed to cross-compare all the generated structures and produce a list of independent ones. The algorithm works by comparing two clusters of molecules, each built from a central molecule and consisting of 15 to 30 molecules, as specified by the user. The program attempts to optimize the overlap between the two sets of molecules by roto-translating them, and then evaluates the level of similarity by calculating the root-mean-square deviation (RMSD) of non-hydrogen atoms. If the RMSD is smaller than a given threshold, the higher-energy polymorph of the two is discarded, while the lower-energy one is saved back into the ranked list.

The following list outlines the actual step-by-step workflow of the method:

1. A pair of polymorphs/structures is selected for comparison. One is designated as the *reference* structure, while the other is referred to as the *target* structure (or structures (1) and (2), for brevity).
2. A cluster of N molecules from the *reference* structure is constructed, starting from the molecule(s) in the asymmetric unit and adding additional molecules in order of increasing center-of-mass (COM) distance. By default, N is set to 15, but the user can adjust this value as desired.
3. A cluster of N molecules is also built from the *target* to be compared. Since the two structures may differ significantly, a robust and flexible algorithm has been implemented to identify the most similar cluster. Several clusters of (2) are generated by considering different central molecules (i.e., each molecule within the primitive cell, one at a time). Internal symmetry can also be taken into account: for example, in benzoic acid, the torsion around the Ph-COOH bond has a periodicity of 180° if the acidic hydrogen is not considered. The generation of cluster (2) follows this workflow:
 - (a) A set of molecules from structure (2) is built, usually containing $2 \cdot N$ molecules (where N is the number of molecules in cluster (1)), starting from the central molecule. This forms a "super-cluster".
 - (b) Sub-clusters of increasing size are constructed for both structures, starting from their respective central molecules and adding one molecule at each step.
 - (c) The COMs of the two sub-clusters are calculated, and the systems are translated so that the COMs coincide with the origin of the reference frame.

- (d) The Kabsch rotation algorithm is used to find the optimal rotation matrix that minimizes the root-mean-square deviation (RMSD) of predicted equivalent atoms². The entire super-cluster (2) is then rotated using the calculated matrix.
 - (e) A one-to-one matching is defined between the molecules of cluster (1) and those of the super-cluster (2), in order of increasing COM distance from the central molecule. The best-overlapping molecules are flagged for inclusion in the sub-clusters at the next iteration (step b).
 - (f) The cycle restarts from step (b) until all N molecules in cluster (1) are matched with N molecules from the super-cluster (2). These selected molecules are saved as the final cluster (2).
4. Once the most similar cluster is built, the RMSD between the two clusters is calculated.
 5. The minimum RMSD value among all the possible cluster (2) configurations is stored in memory and printed at the end of the procedure.

S2 Ranking results

S2.1 Benzoic Acid

Table S2: Top 50 predicted polymorphs of benzoic acid from Symmetry-Constrained Monte Carlo (SC-MC) simulations. The acronyms and units of measurement (in square brackets) are as follow: (Poly.): Rank of the obtained polymorph; (SG): space group; (Run): No. of the run where this polymorph was found; (ID): Original ID No. of the polymorph; (Coh. E.): cohesive energy [$\text{kJ} \cdot \text{mol}^{-1}$]; (a, b, c, ϑ): lattice parameters a , b and c [\AA] and angles α , β and γ when $\neq 90^\circ$ [$^\circ$]; (Vol.): primitive cell volume [\AA^3]; (Dens.): primitive cell density [$\text{g} \cdot \text{cm}^{-3}$]; (Pack.): packing coefficient.

Poly.	SG	Run	ID	Coh. E.	a	b	c	β	Vol	Dens.	Pack.
<i>exp</i>	14	-	-	-97.5	5.4496	5.1283	21.9500	97.37	613.95	1.321	0.730
<i>i</i>	19	1	108	-99.7407	6.0591	25.7447	3.8442	-	599.65	1.353	0.747
<i>ii</i>	19	1	41	-98.8714	6.1964	8.6358	11.4496	-	612.68	1.324	0.732
<i>iii</i>	4	1	121	-98.0081	12.8306	6.4911	3.8633	102.960	313.56	1.294	0.715
<i>iv</i>	14	2	139	-97.0100	3.9971	5.8274	26.2785	97.640	606.66	1.337	0.739
<i>v</i>	14	2	65	-96.5141	5.4799	5.1472	21.7574	91.982	613.32	1.323	0.731
<i>vi</i>	14	2	153	-96.4937	5.4894	5.2204	21.8051	90.144	624.87	1.298	0.717
<i>vii</i>	19	1	140	-96.2786	5.7928	9.6422	10.4532	-	583.87	1.389	0.768
<i>viii</i>	61	1	64	-96.0301	25.0649	6.4249	7.5482	-	1215.56	1.335	0.737
<i>ix</i>	15	1	136	-95.7315	13.5434	6.6577	14.2119	102.150	1252.75	1.295	0.716
<i>x</i>	19	1	182	-95.6870	12.5910	7.5359	6.4510	-	612.10	1.325	0.732
<i>xi</i>	61	1	43	-95.6465	6.5795	28.9838	6.6140	-	1261.28	1.286	0.711
<i>xii</i>	14	1	82	-95.3160	4.3977	6.0116	24.0187	97.020	630.23	1.287	0.711
<i>xiii</i>	19	1	155	-94.9472	6.2505	6.3057	16.5941	-	654.03	1.240	0.685
<i>xiv</i>	61	1	179	-94.7661	6.3775	12.0893	15.7262	-	1212.48	1.338	0.739
<i>xv</i>	14	2	173	-94.7098	5.5793	5.1662	22.6259	94.494	650.16	1.248	0.689
<i>xvi</i>	19	1	61	-94.4658	6.4239	12.2931	8.2145	-	648.69	1.250	0.691
<i>xvii</i>	15	1	20	-94.3465	14.9702	5.7163	15.0553	94.194	1284.88	1.263	0.698
<i>xviii</i>	61	1	149	-94.0167	11.5244	15.2082	7.0092	-	1228.47	1.321	0.730
<i>xix</i>	61	1	106	-94.0114	14.0600	6.4790	13.8809	-	1264.49	1.283	0.709
<i>xx</i>	19	1	105	-93.8794	15.8380	10.5564	3.9230	-	655.90	1.237	0.683
<i>xxi</i>	61	1	147	-93.6812	26.5404	3.7472	11.9880	-	1192.23	1.361	0.752
<i>xxii</i>	14	1	195	-93.6792	7.5522	6.3978	14.1699	103.088	666.87	1.216	0.672
<i>xxiii</i>	19	1	169	-93.6358	11.2384	9.0727	6.3734	-	649.84	1.248	0.690
<i>xxiv</i>	19	1	130	-93.3506	11.6206	6.1523	9.2117	-	658.58	1.232	0.680
<i>xxv</i>	61	1	57	-93.2769	14.3852	6.4424	14.0132	-	1298.68	1.249	0.690
<i>xxvi</i>	19	1	57	-93.2445	11.4525	6.8233	8.1573	-	637.45	1.273	0.703
<i>xxvii</i>	61	1	45	-93.1664	15.4551	6.3545	13.9805	-	1373.01	1.182	0.653
<i>xxviii</i>	19	1	168	-92.8906	13.1463	6.8546	6.9222	-	623.78	1.300	0.719
<i>xxix</i>	15	1	169	-92.8650	12.3870	7.1252	14.9528	90.611	1319.66	1.229	0.679
<i>xxx</i>	61	1	185	-92.8441	12.2429	15.0807	7.0902	-	1309.07	1.239	0.685
<i>xxxi</i>	14	2	171	-92.8231	5.8758	3.7900	27.3735	97.018	605.03	1.341	0.741
<i>xxxii</i>	19	1	175	-92.8183	12.1493	9.2330	6.0592	-	679.69	1.194	0.659

Poly.	SG	R.	ID	Coh. E.	a	b	c	β	Vol.	Dens.	Pack.
<i>xxiii</i>	19	1	103	-92.5466	5.8023	7.6850	15.2601	-	680.46	1.192	0.659
<i>xxxiv</i>	14	1	174	-92.1113	5.3737	15.7137	7.8485	99.140	654.32	1.240	0.685
<i>xxxv</i>	14	2	155	-92.0559	4.0560	24.1754	6.6034	90.939	647.41	1.253	0.692
<i>xxxvi</i>	4	1	21	-91.9179	11.9063	5.7339	4.8478	94.979	329.71	1.230	0.680
<i>xxxvii</i>	61	1	6	-91.7799	17.1077	6.6794	11.1931	-	1279.04	1.268	0.701
<i>xxxviii</i>	19	1	72	-91.7660	13.1467	12.7881	3.8673	-	650.18	1.248	0.689
<i>xxxix</i>	14	1	131	-91.5634	9.4878	6.4425	10.8370	100.695	650.91	1.246	0.689
<i>xl</i>	15	1	15	-91.4994	18.6813	5.3479	14.0632	100.085	1383.27	1.173	0.648
<i>xli</i>	19	1	111	-91.3957	14.4723	6.4824	7.0127	-	657.90	1.233	0.681
<i>xliv</i>	19	1	191	-91.3528	11.2870	6.5422	9.3300	-	688.95	1.177	0.650
<i>xliv</i>	15	1	1	-91.2703	15.2638	6.1406	13.7055	102.958	1251.88	1.296	0.716
<i>xliv</i>	19	1	152	-91.1691	6.0203	9.1110	11.6864	-	641.01	1.266	0.699
<i>xlvi</i>	15	1	81	-91.1159	13.3361	6.1009	16.6948	97.860	1345.58	1.206	0.666
<i>xlvi</i>	19	1	14	-91.1113	3.8166	18.6081	9.5899	-	681.06	1.191	0.658
<i>xlvi</i>	15	1	132	-90.7760	22.4621	7.6143	7.8310	91.870	1338.65	1.212	0.670
<i>xlvi</i>	19	1	64	-90.7716	13.7415	9.5570	5.1183	-	672.17	1.207	0.667
<i>xlvi</i>	19	1	79	-90.6600	8.7797	12.7957	6.0402	-	678.57	1.195	0.660
<i>l</i>	19	1	183	-90.6499	6.7686	5.5773	17.6044	-	664.57	1.221	0.674

Table S3: Lattice parameters obtained from Molecular Dynamics simulations for the benzoic acid polymorphs. Standard deviations are reported in parentheses. The experimental structure is highlighted in bold.

Rank(Poly.)	SG	$a / \text{\AA}$	$b / \text{\AA}$	$c / \text{\AA}$	$\beta / ^\circ$	$V / \text{\AA}^3$	Dens. / $g\text{cm}^{-3}$	Pack.
<i>i (i)</i>	19	5.569(16)	22.35(6)	4.956(10)	-	617(3)	1.315(5)	0.726
<i>ii (xiii)</i>	19	7.10(4)	6.064(17)	14.70(5)	-	632(5)	1.283(7)	0.709
<i>iii (viii)</i>	61	22.10(7)	6.767(11)	8.51(2)	-	1273(5)	1.275(4)	0.704
<i>iv (vii)</i>	19	5.852(17)	8.17(6)	13.34(12)	-	637(5)	1.273(5)	0.703
<i>v (ii)</i>	19	5.864(15)	8.17(6)	13.33(13)	-	639(5)	1.270(5)	0.701
<i>vi (xi)</i>	61	6.617(18)	27.09(9)	7.161(12)	-	1284(6)	1.264(4)	0.698
<i>vii (xii)</i>	14	4.309(11)	6.332(13)	23.91(6)	96.91(3)	648(3)	1.253(5)	0.692
<i>viii (ix)</i>	15	14.77(4)	6.78(2)	13.36(5)	104.13(7)	1298(7)	1.250(6)	0.690
<i>ix (x)</i>	19	13.58(7)	7.09(3)	6.825(11)	-	657(5)	1.234(7)	0.682
<i>x (iii)</i>	4	12.35(4)	5.89(2)	4.428(18)	100.85(7)	316(2)	1.282(6)	0.708
<i>xi (iv)</i>	14	5.22(2)	5.363(18)	22.41(7)	101.74(6)	614(4)	1.321(8)	0.730
<i>xii (xv)</i>	14	5.22(2)	5.608(14)	21.03(8)	94.526(17)	614(4)	1.321(10)	0.730
<i>xiii (v)</i>	14	5.34(2)	5.633(13)	20.56(8)	92.043(15)	618(4)	1.313(7)	0.725
<i>xiv (vi)</i>	14	5.17(2)	5.087(18)	23.65(10)	90.125(1)	622(3)	1.305(8)	0.721
<i>exp</i>	14	5.497(15)	5.42(2)	21.17(7)	97.64(4)	625(4)	1.297(6)	0.716
<i>xvi (xxiii)</i>	14	5.36(3)	5.86(4)	20.59(7)	103.45(9)	629(6)	1.289(9)	0.712
<i>xvii (xvii)</i>	15	13.34(6)	5.824(11)	16.96(4)	93.317(13)	1316(7)	1.233(8)	0.681

Table S4: Cohesive energy (E. Coh.) as well as rotational (Rot. Dif.) and translational (Trasl. Dif.) deviations from the initial ideal crystal structure, obtained from Molecular Dynamics simulations for the benzoic acid polymorphs. Standard deviations are reported in parentheses and limited to two significant digits to emphasize subtle differences. The experimental structure is highlighted in bold.

Rank(Poly.)	SG	E. Coh. / $kJ\ mol^{-1}$	Rot. Dif. / %	Trasl. Dif. / Å
<i>i (i)</i>	19	-101.6(2.4)	16.1(1.5)	20.6(2.4)
<i>ii (xiii)</i>	19	-99.8(2.2)	7.1(1.2)	6.4(2.1)
<i>iii (viii)</i>	61	-99.6(1.4)	29.7(2.7)	19.6(1.9)
<i>iv (vii)</i>	19	-98.8(2.7)	9.5(4.6)	16.5(7.6)
<i>v (ii)</i>	19	-98.7(2.8)	4.9(3.7)	6.0(5.0)
<i>vi (xi)</i>	61	-98.2(1.5)	1.56(0.48)	4.5(1.1)
<i>vii (xii)</i>	14	-97.9(1.3)	2.71(0.79)	1.02(0.38)
<i>viii (ix)</i>	15	-97.8(2.3)	2.46(0.85)	13.6(2.5)
<i>ix (x)</i>	19	-97.1(2.6)	4.7(1.4)	8.2(3.1)
<i>x (iii)</i>	4	-96.9(2.0)	2.61(0.79)	5.7(2.2)
<i>xi (iv)</i>	14	-96.3(2.0)	7.9(1.0)	31.3(5.2)
<i>xii (xv)</i>	14	-95.9(3.8)	1.84(0.51)	3.2(1.2)
<i>xiii (v)</i>	14	-95.7(2.2)	1.14(0.39)	2.37(0.59)
<i>xiv (vi)</i>	14	-95.1(1.8)	2.59(0.77)	2.23(0.95)
<i>exp</i>	14	-94.8(1.6)	1.11(0.34)	0.88(0.62)
<i>xvi (xxiii)</i>	14	-93.8(2.7)	1.43(0.51)	1.23(0.55)
<i>xvii (xvii)</i>	15	-91.3(2.2)	2.67(0.80)	5.2(1.0)

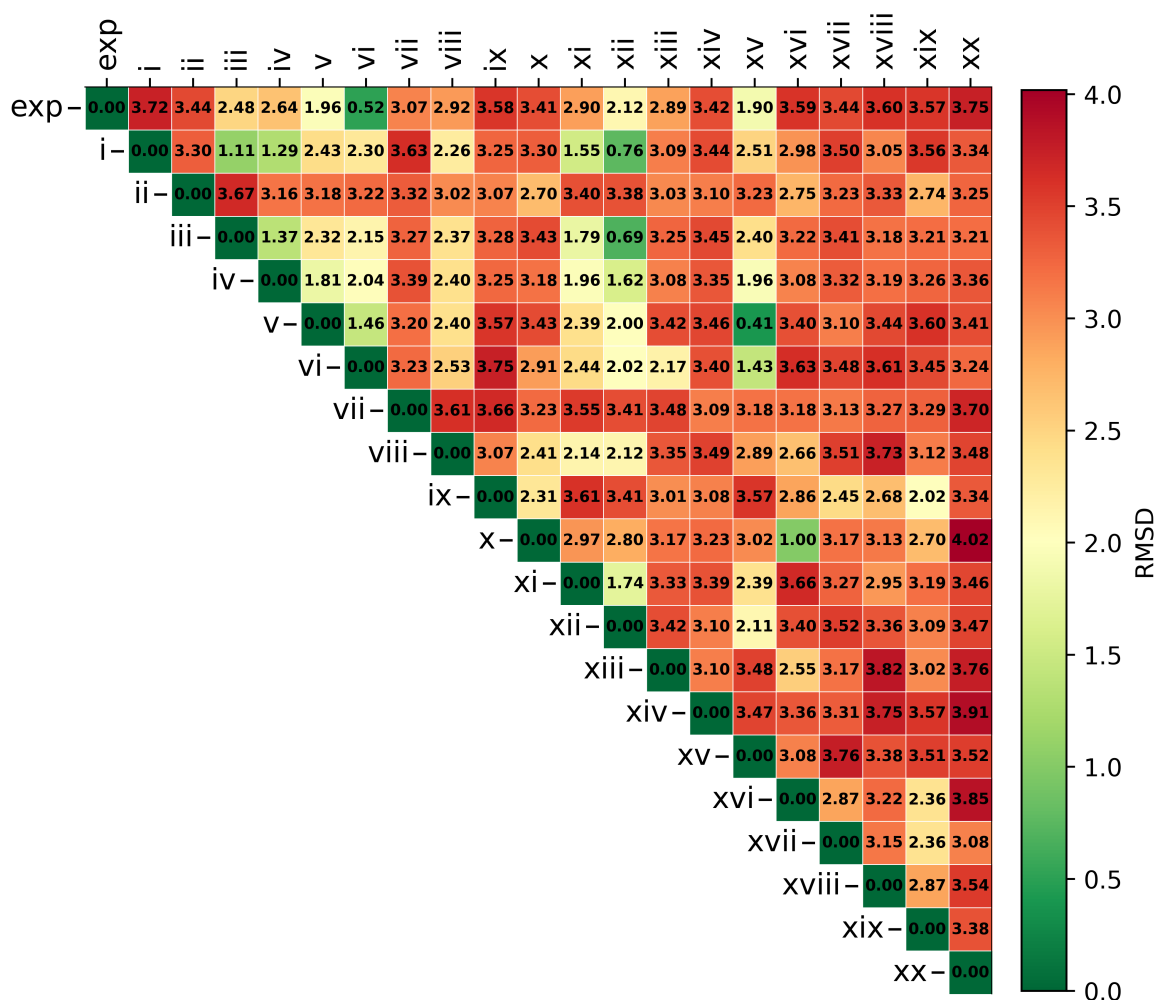


Figure S5: Root-mean-square deviation (RMSD) heatmap comparing the 15-molecule clusters of the top 20 ranked benzoic acid polymorphs obtained from the symmetry-constrained Monte Carlo (SC-MC) screening. Structure labels correspond to those reported in Table S2. RMSD values are displayed within the cells and visually enhanced using a color scale ranging from red (poor agreement) to green (excellent agreement).

S2.2 Acrolein

Table S5: Top 50 predicted polymorphs of acrolein from Symmetry-Constrained Monte Carlo (SC-MC) simulations. The acronyms and units of measurement (in square brackets) are as follow: (Poly.): Rank of the obtained polymorph; (SG): space group; (Run): No. of the run where this polymorph was found; (ID): Original ID No. of the polymorph; (Coh. E.): cohesive energy [$kJ \cdot mol^{-1}$]; (a, b, c, ϑ): lattice parameters a , b and c [\AA] and angles α , β and γ when $\neq 90^\circ$ [$^\circ$] in standard notation; (Vol.): primitive cell volume [\AA^3]; (Dens.): primitive cell density [$g \cdot cm^{-3}$]; (Pack.): packing coefficient.

Poly.	SG	Run	ID	Coh. E.	a	b	c	Angles	Vol	Dens.	Pack.
<i>exp</i>	61	-	-	-43.0	7.1340	9.6940	9.9300	-	686.73	1.085	0.696
<i>i</i>	61	1	458	-45.6397	8.3612	7.1898	10.4580	-	628.69	1.185	0.760
<i>ii</i>	61	1	37	-45.5760	6.4018	10.6134	9.1382	-	620.90	1.200	0.770
<i>iii</i>	61	1	11	-45.5726	6.6761	13.9311	6.9635	-	647.65	1.150	0.738
<i>iv</i>	61	1	333	-45.2700	9.5939	6.7998	9.7985	-	639.22	1.165	0.748
<i>v</i>	19	1	96	-45.2000	4.1343	6.5243	11.7462	-	316.83	1.175	0.754
<i>vi</i>	14	1	246	-45.1949	3.8385	6.7058	12.2016	95.119	312.81	1.190	0.764
<i>vii</i>	19	1	399	-45.1736	6.5351	4.1233	11.7581	-	316.83	1.175	0.754
<i>viii</i>	19	1	254	-45.1520	10.6806	4.1017	7.2359	-	317.00	1.175	0.754
<i>ix</i>	19	1	28	-45.1512	4.0693	10.6804	7.2755	-	316.21	1.178	0.756
<i>x</i>	61	1	334	-45.1172	6.7777	12.1636	7.8207	-	644.75	1.155	0.741
<i>xi</i>	14	1	298	-45.0589	4.6348	6.6887	10.3955	95.200	320.94	1.160	0.745
<i>xii</i>	61	1	61	-44.9546	7.7223	7.1832	11.4001	-	632.37	1.178	0.756
<i>xiii</i>	15	1	122	-44.9032	13.3780	3.8206	12.4892	101.846	624.75	1.192	0.765
<i>xiv</i>	14	1	304	-44.8382	5.3014	11.7287	5.2514	97.122	324.00	1.149	0.738
<i>xv</i>	4	1	64	-44.8147	5.4480	6.2758	4.5918	92.587	156.84	1.187	0.762
<i>xvi</i>	19	1	475	-44.7967	6.0143	4.8347	10.6922	-	310.90	1.198	0.769
<i>xvii</i>	19	1	447	-44.7361	4.9424	12.3407	5.1598	-	314.71	1.183	0.759
<i>xviii</i>	14	1	431	-44.7323	6.2239	7.1333	7.1536	101.382	311.35	1.196	0.768
<i>xix</i>	19	1	67	-44.6876	12.3654	5.0953	4.9478	-	311.74	1.195	0.767
<i>xx</i>	4	1	311	-44.6697	5.4163	6.5555	4.5267	89.705	160.73	1.159	0.744
<i>xxi</i>	19	1	296	-44.6562	7.2400	6.3766	6.7245	-	310.45	1.200	0.770
<i>xxii</i>	19	1	163	-44.6527	6.9044	7.2774	6.3002	-	316.56	1.176	0.755
<i>xxiii</i>	19	1	299	-44.6327	6.2027	10.6919	4.6753	-	310.06	1.201	0.771
<i>xxiv</i>	14	1	454	-44.3847	5.4468	4.7079	12.6698	103.124	316.41	1.177	0.755
<i>xxv</i>	14	1	400	-44.3693	6.3593	4.0608	12.3733	98.790	315.77	1.179	0.757
<i>xxvi</i>	61	1	160	-44.3334	7.2995	9.0822	9.1674	-	607.75	1.226	0.786
<i>xxvii</i>	61	1	72	-44.2794	5.5772	9.5975	11.8804	-	635.92	1.171	0.752
<i>xxviii</i>	19	1	18	-44.1705	7.0440	7.8625	5.8028	-	321.38	1.159	0.744
<i>xxix</i>	14	1	377	-44.1547	7.0495	6.9802	6.7188	91.572	330.49	1.127	0.723
<i>xxx</i>	19	1	42	-44.1116	6.4543	6.7556	7.0831	-	308.84	1.206	0.774
<i>xxxi</i>	19	1	142	-44.0402	6.5028	6.8590	7.1144	-	317.32	1.174	0.753
<i>xxxii</i>	19	1	185	-43.9861	7.0188	7.9111	5.8699	-	325.93	1.143	0.733
<i>xxxiii</i>	15	1	292	-43.9783	13.2557	5.2215	9.5773	99.054	654.63	1.138	0.730
<i>xxxiv</i>	61	1	79	-43.9452	9.7960	8.9944	7.5113	-	661.81	1.125	0.722
<i>xxxv</i>	14	1	399	-43.9181	6.8095	5.0907	9.5184	107.388	314.87	1.183	0.759

Poly.	SG	R.	ID	Coh. E.	a	b	c	Angles	Vol.	Dens.	Pack.
<i>xxxvi</i>	61	1	456	-43.9013	5.5643	9.5879	12.2409	-	653.05	1.141	0.732
<i>xxxvii</i>	4	1	180	-43.8773	6.1161	6.6815	3.8958	89.767	159.20	1.170	0.751
<i>xxxviii</i>	19	1	438	-43.8327	6.6224	4.0552	11.8274	-	317.62	1.173	0.752
<i>xxxix</i>	4	1	287	-43.8289	6.3369	6.8181	3.8887	103.786	163.17	1.141	0.732
<i>xl</i>	4	1	448	-43.8199	5.0327	5.8890	5.3653	96.253	158.07	1.178	0.756
<i>xli</i>	15	1	372	-43.8113	12.3734	6.8416	7.5114	90.183	635.87	1.171	0.752
<i>xlii</i>	19	1	197	-43.7896	6.7633	4.2378	11.7277	-	336.13	1.108	0.711
<i>xliii</i>	4	1	366	-43.7887	6.4335	6.7968	3.8655	105.689	162.73	1.144	0.734
<i>xliv</i>	14	1	439	-43.7671	6.7383	5.2967	9.4234	101.035	330.11	1.128	0.724
<i>xlv</i>	19	1	468	-43.7424	6.8238	6.7967	6.9865	-	324.02	1.149	0.738
<i>xlvi</i>	14	1	30	-43.7365	4.7414	12.8360	5.4448	86.795	330.86	1.126	0.722
<i>xlvii</i>	19	1	68	-43.7086	6.8706	6.9946	6.8900	-	331.12	1.125	0.722
<i>xlviii</i>	4	1	25	-43.6792	6.8442	6.1821	3.8405	91.822	162.42	1.146	0.736
<i>xl ix</i>	2	1	372	-43.6488	5.3492	5.8602	5.1066	90.943	159.00	1.171	0.752
								96.184			
								87.622			
<i>l</i>	19	1	12	-43.6458	11.7443	6.7237	4.2357	-	334.47	1.113	0.715

Table S6: Lattice parameters obtained from Molecular Dynamics simulations for the acrolein polymorphs. Standard deviations are reported in parentheses. The experimental structure is highlighted in bold.

Rank(Poly.)	SG	$a / \text{\AA}$	$b / \text{\AA}$	$c / \text{\AA}$	$\beta / ^\circ$	$V / \text{\AA}^3$	Dens. / gcm^{-3}	Pack.
<i>i (iv)</i>	61	9.590(8)	7.125(5)	9.797(8)	-	669.42 0.91	1.113(1)	0.714
<i>ii (iii)</i>	61	6.846(5)	13.742(11)	7.220(5)	-	679.24(87)	1.097(1)	0.704
<i>iii (xv)</i>	4	5.360(4)	6.632(4)	4.695(5)	92.489(4)	166.74(24)	1.117(1)	0.716
<i>iv (i)</i>	61	8.501(7)	7.582(8)	10.415(7)	-	671.3(1.0)	1.110(1)	0.712
v (exp)	61	7.120(6)	9.756(9)	9.659(9)	-	670.9(1.1)	1.110(2)	0.712
<i>vi (xx)</i>	4	5.341(7)	6.678(5)	4.650(6)	89.717(1)	165.85(33)	1.123(1)	0.720
<i>vii (ix)</i>	19	4.251(3)	10.884(9)	7.317(6)	-	338.54(46)	1.100(1)	0.705
<i>viii (x)</i>	61	6.971(5)	12.304(12)	7.980(6)	-	684.45(98)	1.088(1)	0.698
<i>ix (viii)</i>	19	10.845(8)	4.258(3)	7.376(6)	-	340.61(44)	1.093(1)	0.701
<i>x (xiii)</i>	15	13.776(11)	4.060(3)	12.476(8)	102.217(14)	681.99(86)	1.092(1)	0.700
<i>xi (xiv)</i>	14	5.324(4)	12.341(12)	5.197(4)	97.228(8)	338.75(49)	1.099(1)	0.705
<i>xii (ii)</i>	61	6.703(7)	10.556(8)	9.567(9)	-	676.9(1.1)	1.100(1)	0.705
<i>xiii (xi)</i>	14	4.844(5)	6.675(6)	10.438(8)	95.413(6)	335.99(53)	1.108(1)	0.711
<i>xiv (vii)</i>	19	6.687(5)	4.261(5)	11.896(14)	-	338.96(62)	1.099(1)	0.705
<i>xv (vi)</i>	14	4.244(9)	6.810(5)	11.877(16)	95.817(19)	341.50(89)	1.091(1)	0.700
<i>xvi (xvi)</i>	19	6.279(6)	5.006(4)	10.777(9)	-	338.75(51)	1.099(1)	0.705
<i>xvii (xii)</i>	61	7.849(38)	8.049(39)	10.851(23)	-	685.5(4.9)	1.087(2)	0.697
<i>xviii (xviii)</i>	14	6.405(6)	7.255(6)	7.477(6)	101.206(14)	340.82(51)	1.093(2)	0.701
<i>xix (xvii)</i>	19	4.961(6)	12.851(11)	5.312(6)	-	338.66(63)	1.100(1)	0.705
<i>xx (xix)</i>	19	12.848(12)	5.312(6)	4.970(6)	-	339.20(64)	1.098(2)	0.704
<i>xxi (v)</i>	19	5.379(4)	6.358(5)	10.127(8)	-	346.34(46)	1.075(1)	0.689

Table S7: Cohesive energy (E. Coh.) as well as rotational (Rot. Dif.) and traslational (Trasl. Dif.) deviations from the initial ideal crystal structure, obtained from Molecular Dynamics simulations for the acrolein polymorphs. Standard deviations are reported in parentheses and limited to two significant digits to emphasize subtle differences. The experimental structure is highlighted in bold.

Rank(Poly.)	SG	E. Coh. / $kJ\ mol^{-1}$	Rot. Dif. / %	Trasl. Dif. / Å
<i>i</i> (iv)	61	-40.80(1)	1.31(37)	0.40(13)
<i>ii</i> (iii)	61	-40.80(1)	1.11(37)	0.39(14)
<i>iii</i> (xv)	4	-40.60(1)	1.42(49)	0.70(21)
<i>iv</i> (i)	61	-40.60(1)	0.87(29)	0.59(20)
<i>v</i> (exp)	61	-40.60(1)	1.47(30)	0.21(6)
<i>vi</i> (xx)	4	-40.40(1)	0.90(32)	0.27(14)
<i>vii</i> (ix)	19	-40.20(1)	1.15(30)	0.44(12)
<i>viii</i> (x)	61	-40.10(1)	1.14(36)	0.30(10)
<i>ix</i> (viii)	19	-40.10(1)	1.28(81)	0.40(15)
<i>x</i> (xiii)	15	-40.00(1)	1.50(54)	0.80(21)
<i>xi</i> (xiv)	14	-40.00(1)	1.35(48)	0.61(22)
<i>xii</i> (ii)	61	-39.90(1)	1.47(57)	0.84(22)
<i>xiii</i> (xi)	14	-39.90(1)	1.25(75)	0.39(17)
<i>xiv</i> (vii)	19	-39.80(1)	1.12(35)	0.38(14)
<i>xv</i> (vi)	14	-39.60(1)	2.4(1.2)	2.02(78)
<i>xvi</i> (xvi)	19	-39.30(1)	1.12(0.58)	0.60(17)
<i>xvii</i> (xii)	61	-39.20(2)	8.6(15.1)	2.8(1.5)
<i>xviii</i> (xviii)	14	-39.10(1)	1.18(0.45)	0.59(0.19)
<i>xix</i> (xvii)	19	-39.10(1)	2.7(1.2)	0.61(0.19)
<i>xx</i> (xix)	19	-38.90(1)	3.5(3.7)	0.72(0.22)
<i>xxi</i> (v)	19	-38.70(1)	18.7(2.3)	16.84(0.89)

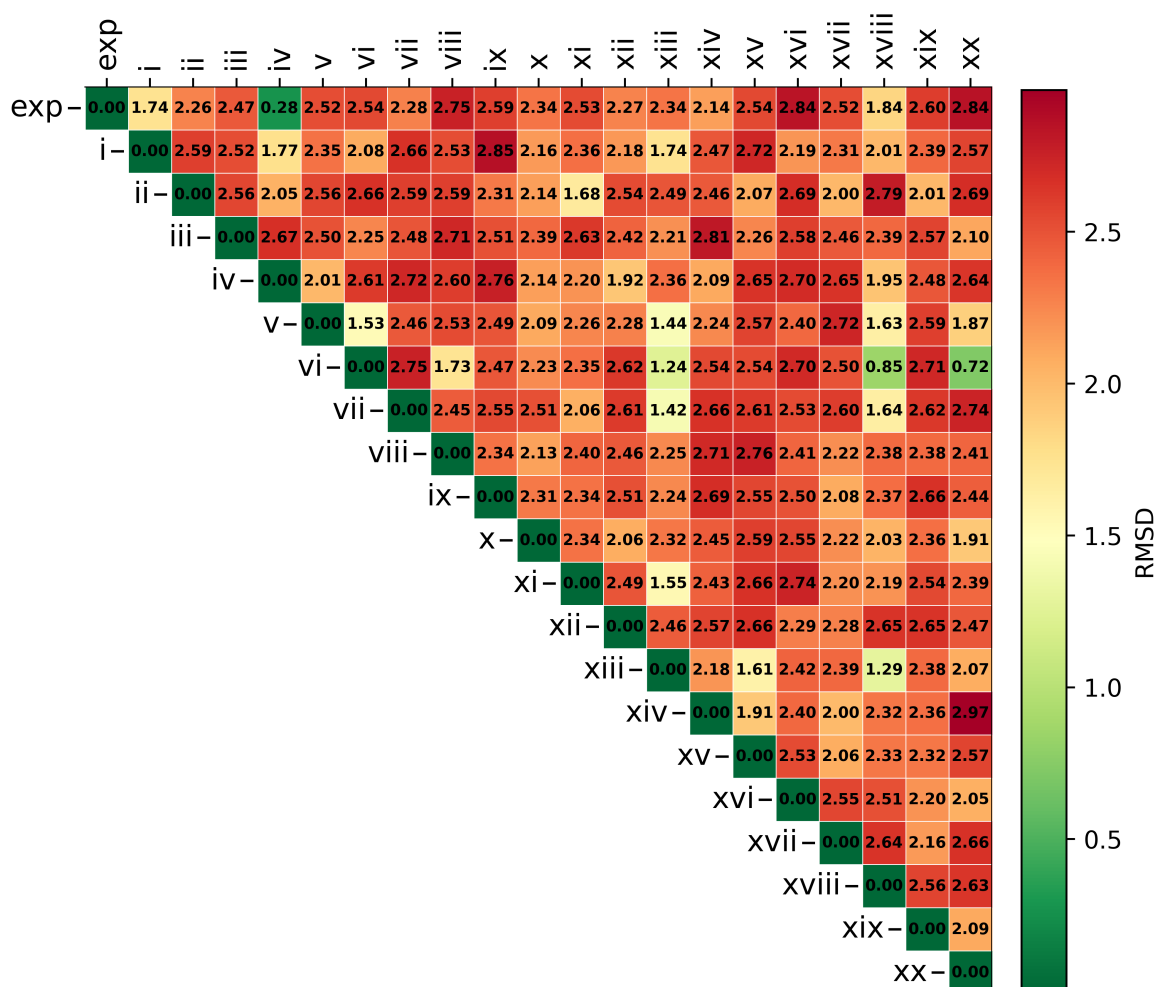


Figure S6: Root-mean-square deviation (RMSD) heatmap comparing the 15-molecule clusters of the top 20 ranked acrolein polymorphs obtained from the symmetry-constrained Monte Carlo (SC-MC) screening. Structure labels correspond to those reported in Table S5. RMSD values are displayed within the cells and visually enhanced using a color scale ranging from red (poor agreement) to green (excellent agreement).

S2.3 Isocaffeine

Table S8: Top 50 predicted polymorphs of isocaffeine from Symmetry-Constrained Monte Carlo (SC-MC) simulations. The acronyms and units of measurement (in square brackets) are as follow: (Poly.): Rank of the obtained polymorph; (SG): space group; (Run): No. of the run where this polymorph was found; (ID): Original ID No. of the polymorph; (Coh. E.): cohesive energy [$kJ \cdot mol^{-1}$]; (a, b, c, ϑ): lattice parameters a , b and c [\AA] and angles α , β and γ when $\neq 90^\circ$ [$^\circ$]; (Vol.): primitive cell volume [\AA^3]; (Dens.): primitive cell density [$g \cdot cm^{-3}$]; (Pack.): packing coefficient.

Poly.	SG	Run	ID	Coh. E.	a	b	c	Angles	Vol	Dens.	Pack.
<i>exp</i>	14	-	-	-141.7	7.7170	7.9150	13.6460	92.86	832.46	1.550	0.804
<i>i</i>	14	1	479	-151.0469	7.0595	13.6079	8.7548	104.355	814.77	1.583	0.821
<i>ii</i>	14	4	128	-150.7582	8.2113	14.8327	7.0143	92.135	853.71	1.511	0.783
<i>iii</i>	14	1	18	-150.2648	6.2804	11.7950	11.3180	94.573	835.73	1.544	0.800
<i>iv</i>	61	1	160	-148.7332	14.7314	6.9129	16.9040	-	1721.45	1.499	0.777
<i>v</i>	14	5	163	-148.4522	7.8901	7.8600	13.5321	91.948	838.72	1.538	0.797
<i>vi</i>	14	1	364	-147.9714	9.1340	8.8322	10.5183	90.102	848.55	1.520	0.788
<i>vii</i>	14	2	132	-147.0226	12.4140	8.6186	7.8869	92.669	842.92	1.530	0.794
<i>viii</i>	14	2	112	-145.9286	8.6520	14.7455	6.8893	103.789	853.59	1.511	0.784
<i>ix</i>	14	5	453	-144.5755	3.9403	12.0266	18.0073	101.590	835.94	1.543	0.800
<i>x</i>	19	1	283	-144.4613	18.0763	3.8288	12.5617	-	869.41	1.484	0.769
<i>xi</i>	19	1	99	-144.3102	11.2066	4.8533	15.4850	-	842.22	1.532	0.794
<i>xii</i>	19	1	491	-143.2426	18.1472	5.3198	8.8428	-	853.68	1.511	0.784
<i>xiii</i>	15	1	98	-143.1219	11.1776	8.8313	16.8771	92.338	1664.60	1.550	0.804
<i>xiv</i>	14	1	394	-142.9234	4.4838	14.2840	13.3119	93.486	851.00	1.516	0.786
<i>xv</i>	14	3	69	-142.9118	5.6870	8.8407	16.8033	90.195	844.82	1.527	0.792
<i>xvi</i>	14	1	28	-142.7275	7.3297	15.3855	7.5957	98.326	847.54	1.522	0.789
<i>xvii</i>	15	1	137	-142.6145	15.7805	8.1769	13.8611	102.805	1744.10	1.479	0.767
<i>xviii</i>	19	1	23	-142.5806	14.3159	8.7079	7.0277	-	876.08	1.472	0.763
<i>xix</i>	61	1	347	-142.5269	15.7497	7.6641	14.3790	-	1735.65	1.486	0.771
<i>xx</i>	61	1	283	-142.3596	12.9504	14.9313	8.7937	-	1700.40	1.517	0.787
<i>xxi</i>	19	1	443	-142.2061	12.6975	17.0555	3.9860	-	863.21	1.494	0.775
<i>xxii</i>	19	1	59	-141.8930	8.8028	14.0187	6.9513	-	857.83	1.504	0.780
<i>xxiii</i>	14	3	167	-141.6704	7.7703	8.5140	13.4909	89.602	892.49	1.445	0.750
<i>xxiv</i>	61	1	457	-141.4394	15.0690	8.4863	13.4460	-	1719.49	1.500	0.778
<i>xxv</i>	61	1	108	-141.2738	17.0613	14.9110	6.8515	-	1743.03	1.480	0.767
<i>xxvi</i>	14	1	498	-141.2454	13.7451	7.4983	8.5216	101.173	861.64	1.497	0.776
<i>xxvii</i>	61	1	264	-140.7063	14.2957	8.5948	13.8429	-	1700.85	1.517	0.787
<i>xxviii</i>	61	1	141	-140.6968	7.0609	16.0372	15.2722	-	1729.37	1.492	0.774
<i>xxix</i>	14	1	308	-140.6404	13.4520	8.7447	7.5195	104.553	856.16	1.507	0.781
<i>xxx</i>	61	1	95	-140.3259	15.8678	15.9862	6.7728	-	1718.04	1.502	0.779
<i>xxxi</i>	14	1	113	-140.1153	7.8504	16.1026	6.8201	91.486	861.86	1.497	0.776
<i>xxxii</i>	61	1	451	-139.9633	29.8032	8.4146	6.9430	-	1741.19	1.482	0.768
<i>xxxiii</i>	14	3	11	-139.8638	6.9052	8.3314	15.6522	103.392	875.99	1.472	0.764
<i>xxxiv</i>	14	5	87	-139.7410	13.2462	8.3257	8.0739	93.768	888.50	1.452	0.753
<i>xxxv</i>	14	2	199	-139.5385	15.1189	8.5405	6.9167	100.632	877.76	1.470	0.762

Poly.	SG	R.	ID	Coh. E.	a	b	c	ϑ	Vol.	Dens.	Pack.
<i>xxxvi</i>	4	1	82	-139.2661	4.1117	15.3259	6.9010	94.262	433.66	1.487	0.771
<i>xxxvii</i>	14	5	309	-139.0579	9.7752	11.5117	8.0525	102.079	886.08	1.456	0.755
<i>xxxviii</i>	14	1	327	-138.9518	5.0154	11.1504	15.6839	94.618	874.26	1.475	0.765
<i>xxxix</i>	14	5	389	-138.8866	14.1450	7.6211	7.9038	91.009	851.90	1.514	0.785
<i>xl</i>	14	4	460	-138.4945	6.7072	16.1262	8.6987	101.014	923.53	1.397	0.724
<i>xli</i>	14	2	113	-138.4851	15.1961	4.8708	11.3935	98.014	835.08	1.545	0.801
<i>xlII</i>	14	1	96	-138.3429	8.1337	15.6718	7.0438	89.980	897.86	1.437	0.745
<i>xlIII</i>	19	1	202	-138.3306	12.1090	4.9697	14.0174	-	843.54	1.529	0.793
<i>xlIV</i>	14	3	135	-138.3229	7.4831	8.0114	15.0815	105.374	871.78	1.480	0.767
<i>xlV</i>	19	1	96	-138.0724	14.2064	8.2525	7.6102	-	892.20	1.446	0.750
<i>xlVI</i>	19	1	140	-137.8220	13.2384	7.9435	8.4873	-	892.51	1.445	0.750
<i>xlVII</i>	61	1	230	-137.7259	7.6675	25.0458	8.9221	-	1713.39	1.506	0.781
<i>xlVIII</i>	4	1	147	-137.7235	4.0858	15.4740	7.0985	92.062	448.51	1.438	0.746
<i>xlIX</i>	14	3	92	-137.6861	7.1238	8.3930	14.9555	103.671	868.86	1.485	0.770
<i>l</i>	14	5	363	-137.5558	4.8567	15.0694	12.0813	97.733	876.16	1.472	0.763

Table S9: Lattice parameters obtained from Molecular Dynamics simulations for the isocaffeine polymorphs. Standard deviations are reported in parentheses. The experimental structure is highlighted in bold.

Rank(Poly.)	SG	$a / \text{\AA}$	$b / \text{\AA}$	$c / \text{\AA}$	$\beta / ^\circ$	$V / \text{\AA}^3$	Dens. / gcm^{-3}	Pack.
<i>i (exp)</i>	14	8.165(13)	8.041(14)	13.203(25)	93.128(9)	866(3)	1.490(4)	0.773
<i>ii (v)</i>	14	7.939(8)	8.074(16)	13.550(25)	91.957(4)	868(3)	1.486(5)	0.770
<i>iii (vi)</i>	14	8.859(25)	8.825(13)	11.481(25)	90.091(1)	898(3)	1.437(5)	0.745
<i>iv (vii)</i>	14	10.949(4)	8.539(14)	9.497(31)	91.955(13)	887(5)	1.454(4)	0.754
<i>v (ii)</i>	14	8.449(15)	14.354(26)	7.330(13)	92.102(4)	888(3)	1.452(5)	0.753
<i>vi (xv)</i>	14	5.324(27)	8.925(15)	18.93(11)	90.162(2)	899(7)	1.434(4)	0.744
<i>vii (xiv)</i>	14	6.473(12)	11.215(15)	12.333(18)	95.436(15)	891(2)	1.447(3)	0.750
<i>viii (xii)</i>	19	20.275(67)	5.130(14)	8.812(26)	-	917(5)	1.407(4)	0.730
<i>ix (ix)</i>	14	4.468(11)	11.305(26)	17.967(23)	103.198(33)	884(3)	1.460(4)	0.757
<i>x (iii)</i>	14	6.406(13)	11.309(13)	12.362(26)	94.270(15)	893(3)	1.444(3)	0.749
<i>xi (iv)</i>	61	14.429(21)	7.348(11)	16.941(24)	-	1796(5)	1.436(3)	0.745
<i>xii (i)</i>	14	7.886(11)	13.246(28)	8.801(14)	105.992(34)	884(3)	1.460(5)	0.757
<i>xiii (xvi)</i>	14	7.619(17)	15.243(24)	7.917(12)	98.304(27)	910(3)	1.418(3)	0.735
<i>xiv (xviii)</i>	19	13.216(3)	8.352(16)	8.138(17)	-	898(3)	1.436(4)	0.745
<i>xv (viii)</i>	14	8.791(18)	14.451(39)	7.296(14)	103.220(36)	902(4)	1.430(4)	0.741
<i>xvi (xix)</i>	61	15.695(31)	7.716(19)	15.306(4)	-	1854(8)	1.392(3)	0.722
<i>xvii (xiii)</i>	15	11.028(26)	8.938(13)	18.401(53)	92.12(1)	1813(7)	1.423(3)	0.738
<i>xviii (xx)</i>	61	12.900(24)	14.698(29)	9.393(17)	-	1781(6)	1.449(4)	0.751
<i>xix (xvii)</i>	15	15.449(34)	8.648(19)	14.517(29)	101.958(39)	1897(7)	1.360(3)	0.705
<i>xx (xi)</i>	19	11.697(27)	4.992(11)	15.602(33)	-	911(3)	1.416(4)	0.734
<i>xxi (x)</i>	19	15.402(36)	5.090(11)	11.664(28)	-	914(4)	1.411(3)	0.732

Table S10: Cohesive energy (E. Coh.) as well as rotational (Rot. Dif.) and traslational (Trasl. Dif.) deviations from the initial ideal crystal structure, obtained from Molecular Dynamics simulations for the isocaffeine polymorphs. Standard deviations are reported in parentheses and limited to two significant digits to emphasize subtle differences. The experimental structure is highlighted in bold.

Rank(Poly.)	SG	E. Coh. / $kJ\ mol^{-1}$	Rot. Dif. / %	Trasl. Dif. / \AA
<i>i (exp)</i>	14	-138.40(4)	0.51(16)	0.88(35)
<i>ii (v)</i>	14	-138.00(6)	0.40(17)	0.21(13)
<i>iii (vi)</i>	14	-133.80(7)	1.73(46)	1.57(52)
<i>iv (vii)</i>	14	-133.80(6)	0.79(28)	11.0(2.7)
<i>v (ii)</i>	14	-133.70(5)	0.68(26)	0.75(29)
<i>vi (xv)</i>	14	-133.60(6)	1.34(67)	8.3(4.7)
<i>vii (xiv)</i>	14	-133.20(4)	6.13(46)	36.5(2.9)
<i>viii (xii)</i>	19	-133.20(6)	2.26(71)	4.1(1.8)
<i>ix (ix)</i>	14	-133.00(5)	0.77(18)	128.5(1.7)
<i>x (iii)</i>	14	-132.20(5)	0.38(10)	1.95(48)
<i>xi (iv)</i>	61	-132.10(5)	0.67(28)	0.75(21)
<i>xii (i)</i>	14	-131.50(6)	1.39(48)	2.34(54)
<i>xiii (xvi)</i>	14	-130.80(6)	0.70(28)	1.53(30)
<i>xiv (xviii)</i>	19	-130.40(6)	11.8(1.8)	8.1(1.5)
<i>xv (viii)</i>	14	-130.30(5)	3.3(1.4)	1.58(71)
<i>xvi (xix)</i>	61	-129.70(6)	0.89(36)	1.02(59)
<i>xvii (xiii)</i>	15	-128.00(6)	0.79(27)	2.30(95)
<i>xviii (xx)</i>	61	-127.90(4)	0.46(16)	0.92(31)
<i>xix (xvii)</i>	15	-126.30(5)	1.19(37)	1.06(31)
<i>xx (xi)</i>	19	-126.30(4)	0.88(30)	0.64(24)
<i>xxi (x)</i>	19	-125.50(4)	6.57(94)	23.3(2.6)

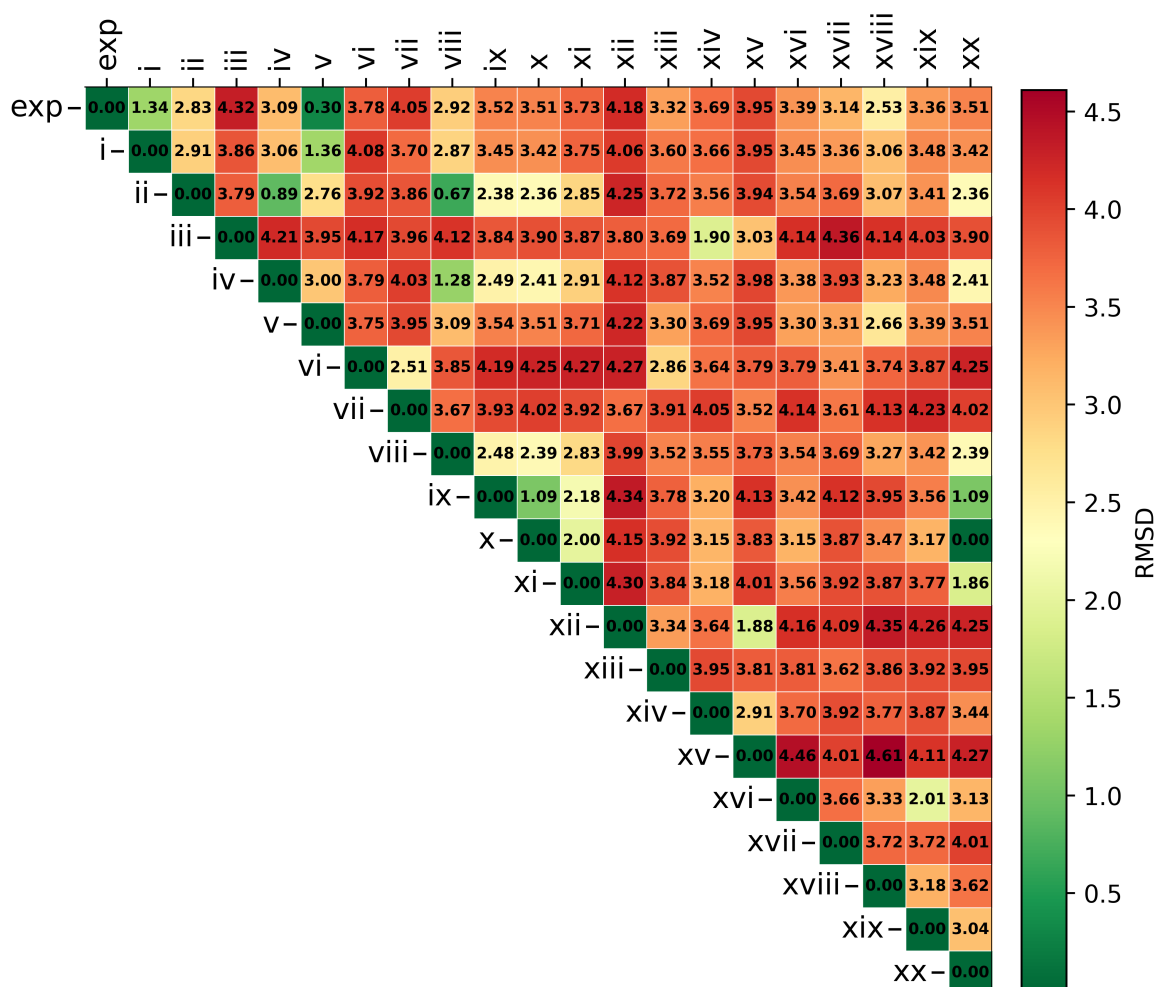


Figure S7: Root-mean-square deviation (RMSD) heatmap comparing the 15-molecule clusters of the top 20 ranked isocaffeine polymorphs obtained from the symmetry-constrained Monte Carlo (SC-MC) screening. Structure labels correspond to those reported in Table S8. RMSD values are displayed within the cells and visually enhanced using a color scale ranging from red (poor agreement) to green (excellent agreement).

S3 Structures Overlap

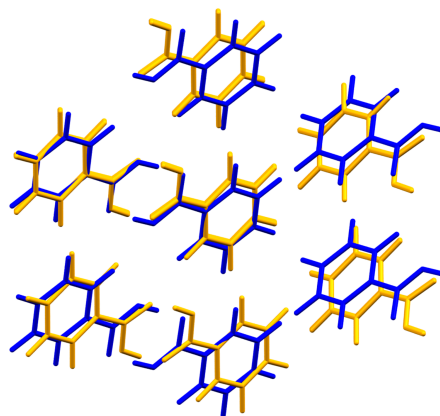


Figure S8: Overlap between the 15-molecules clusters of the best prediction (yellow) and experimental (BENZAC02, blue) structures of benzoic acid.

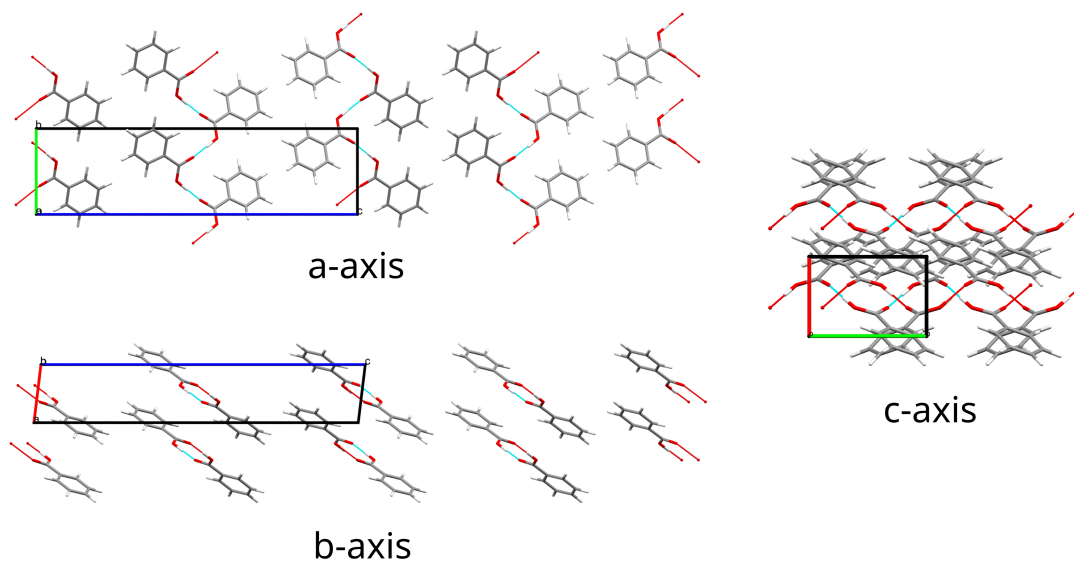


Figure S9: Averaged crystal structure of catemeric polymorph *BNZ-xii* (Benzoic acid), after 200 ps of equilibration through Molecular Dynamics, along the *a* axis (top left panel), *b* axis (bottom left panel) and *c* axis (right panel). The last 40 ps of the trajectory were used for the averaging.

References

- [1] Lo Presti, L. MiCMoS - Milano Chemistry Molecular Simulation Description and User Manual - Release 2.3, July 30, 2024. Università degli Studi di Milano: Milano, 2024; https://sites.unimi.it/xtal_chem_group/images/MiCMoS_package/v2.3_t/MiCMoS_v2.3_User_Manual.pdf.
- [2] Kabsch, W. A discussion of the solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A* **1978**, *34*, 827–828.