

Supporting Information

1

S.1 Full Experimental Results

2

Table S 1: Sample weights and calculated yields for experimental work.

Sample No.	Dried Weight (g)	Calcined Mass (g)	Organics Mass (g)	Organics Content (%)	Yield (%)
1-1	0.49	0.019	0.474	96	13
1-2	9.01	0.526	8.479	94	78
1-3	0.00	0.000	0.004	100	0
1-4	1.07	0.565	0.503	47	84
1-5	0.80	0.015	0.788	98	9
1-6	10.54	0.438	10.099	96	57
1-7	0.02	0.008	0.012	61	5
1-8	1.26	0.712	0.547	43	94
1-9	1.40	0.047	1.352	97	29
1-10	10.91	0.489	10.420	96	82
1-11	0.15	0.065	0.090	58	40
1-12	1.26	0.543	0.712	57	81
1-13	1.47	0.048	1.422	97	30
1-14	13.23	0.573	12.655	96	80
1-15	0.14	0.063	0.078	56	39
1-16	0.96	0.440	0.521	54	58
2-1	1.20	0.52	0.68	56	87
2-2	0.57	0.28	0.29	50	47
2-3	0.26	0.10	0.16	63	16
2-4	1.30	0.55	0.75	58	91
2-5	0.47	0.23	0.24	50	39
2-6	0.21	0.10	0.11	51	17
3-1	1.14	0.48	0.66	59	80
3-2	1.20	0.47	0.73	61	78
3-3	1.14	0.44	0.69	61	74
3-4	1.19	0.39	0.80	67	65
3-5	1.16	0.49	0.67	58	81
4-1	0.73	0.39	0.34	47	65
4-2	0.84	0.48	0.36	42	80
4-3	0.78	0.42	0.36	46	69
4-4	0.49	0.25	0.25	50	41
4-5	0.87	0.47	0.40	46	78
4-6	0.88	0.51	0.37	42	86
4-7	0.64	0.34	0.30	47	57

Table S 2: Summary of the porosity data for all samples that were characterized by N₂ adsorption and XRD. The variables are fully described in Section 2.3. Comparison is made to available material properties for MCM-41 samples from literature (Samples A, B and C).

Sample No.	Primary Isotherm Type	S _{BET} (m ² g ⁻¹)	V _{pore} (cm ³ g ⁻¹)	d _{pore} (nm)	d ₁₀₀ (nm)	a ₀ (nm)	th _{wall} (nm)	I ₂₀₀ /I ₁₁₀
A [1, 2]	IV(a)	1040	0.79	4.00	3.98	4.60	0.60	0.62
B [3]	-	-	-	-	3.48	4.02	-	0.65
C [4]	IV(a)	1312	0.86	3.14	3.05	3.52	0.38	0.00
1-2	IV(a)	1120	0.906	2.48	4.03	4.65	2.17	0.00
1-4	IV(a)	1005	0.828	2.45	3.88	4.48	2.03	0.28
1-6	IV(a)	970	0.721	2.51	3.98	4.60	2.09	0.00
1-8	IV(a)	1030	0.839	2.30	3.81	4.40	2.10	0.27
1-10	IV(a)	1181	0.966	2.63	3.80	4.39	1.76	0.00
1-12	IV(b)	1259	0.834	2.12	3.50	4.04	1.92	0.06
1-14	IV(a)	1217	1.003	2.55	3.82	4.41	1.86	0.04
1-16	IV(b)	1210	0.850	2.14	3.68	4.25	2.11	0.08
2-1	IV(b)	1259	0.971	2.31	3.56	4.11	1.80	0.12
2-2	IV(b)	1165	0.841	2.15	3.40	3.93	1.78	0.46
2-3	IV(a)	1159	0.843	2.13	3.44	3.97	1.84	-
2-4	IV(a)	1193	0.857	2.22	3.52	4.06	1.84	0.11
2-5	IV(b)	1144	0.824	2.14	3.43	3.96	1.82	0.50
2-6	IV(a)	1194	0.898	2.25	3.47	4.01	1.76	0.07
3-1	IV(b)	1149	0.789	2.18	3.39	3.91	1.73	0.06
3-2	IV(b)	1203	0.835	2.17	3.38	3.90	1.73	0.00
3-3	IV(b)	1218	0.938	2.40	3.63	4.19	1.79	0.28
3-4	IV(a)	1273	0.969	2.35	-	-	-	0.00
3-5	IV(b)	1218	0.911	2.24	3.50	4.04	1.80	0.20
4-1	IV(b)	983	0.706	2.20	3.55	4.10	1.90	0.52
4-2	IV(b)	961	0.736	2.27	3.68	4.25	1.98	0.49
4-3	IV(b)	926	0.679	2.24	3.67	4.24	1.94	0.00
4-4	IV(b)	1186	0.830	2.09	3.44	3.97	1.88	0.18
4-5	IV(a)	783	0.608	2.80	-	-	-	0.00
4-6	IV(a)	878	0.686	2.74	-	-	-	0.00
4-7	IV(b)	1142	0.841	2.10	3.40	3.93	1.83	0.29

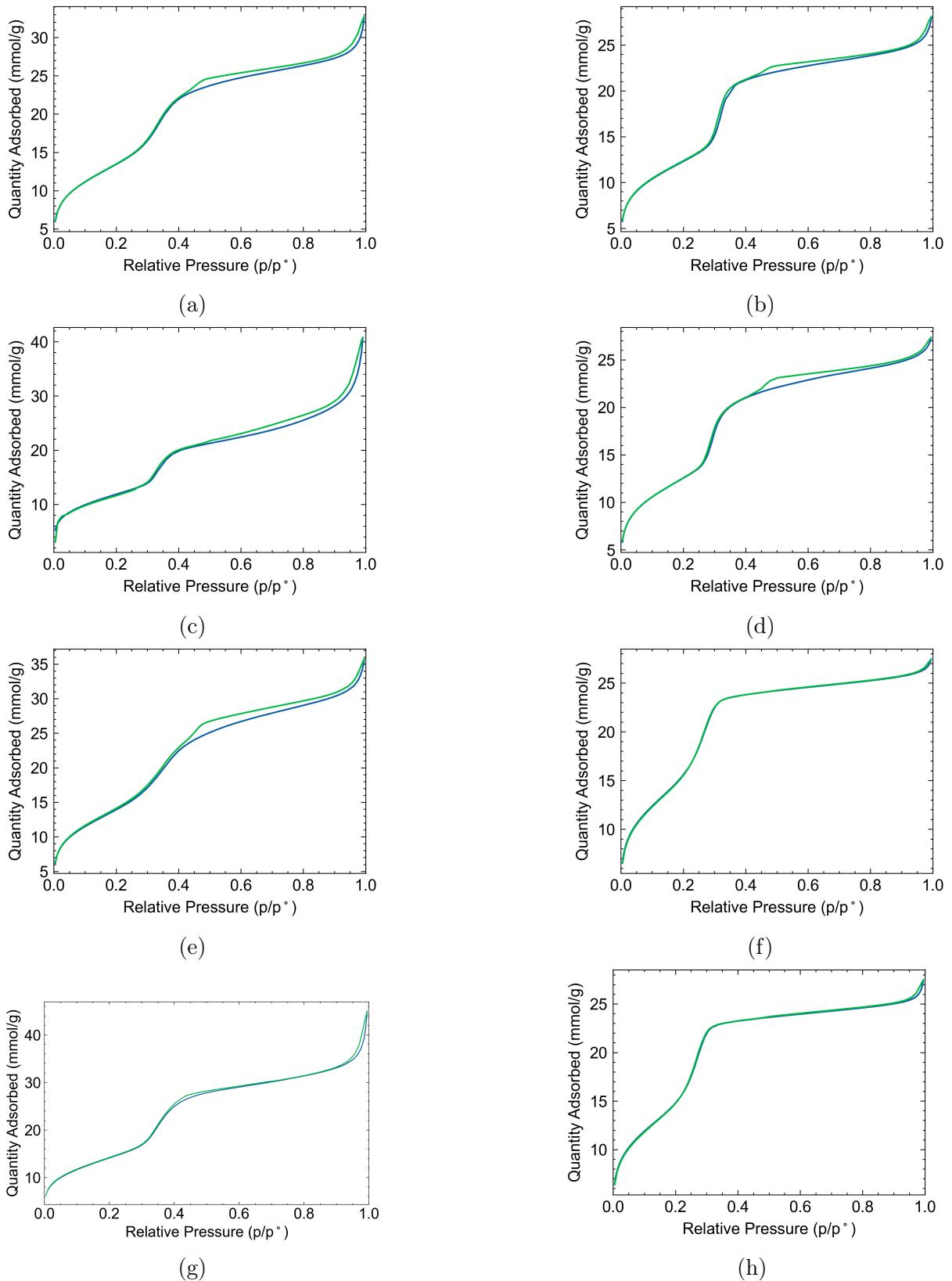


Figure S 1: N₂ adsorption isotherms for two-level four-factor screening samples. Letters a-h represent samples 1-2, 1-4, 1-6, 1-8, 1-10, 1-12, 1-14 and 1-16, sequentially. The blue line indicates adsorption while the green line indicates desorption.

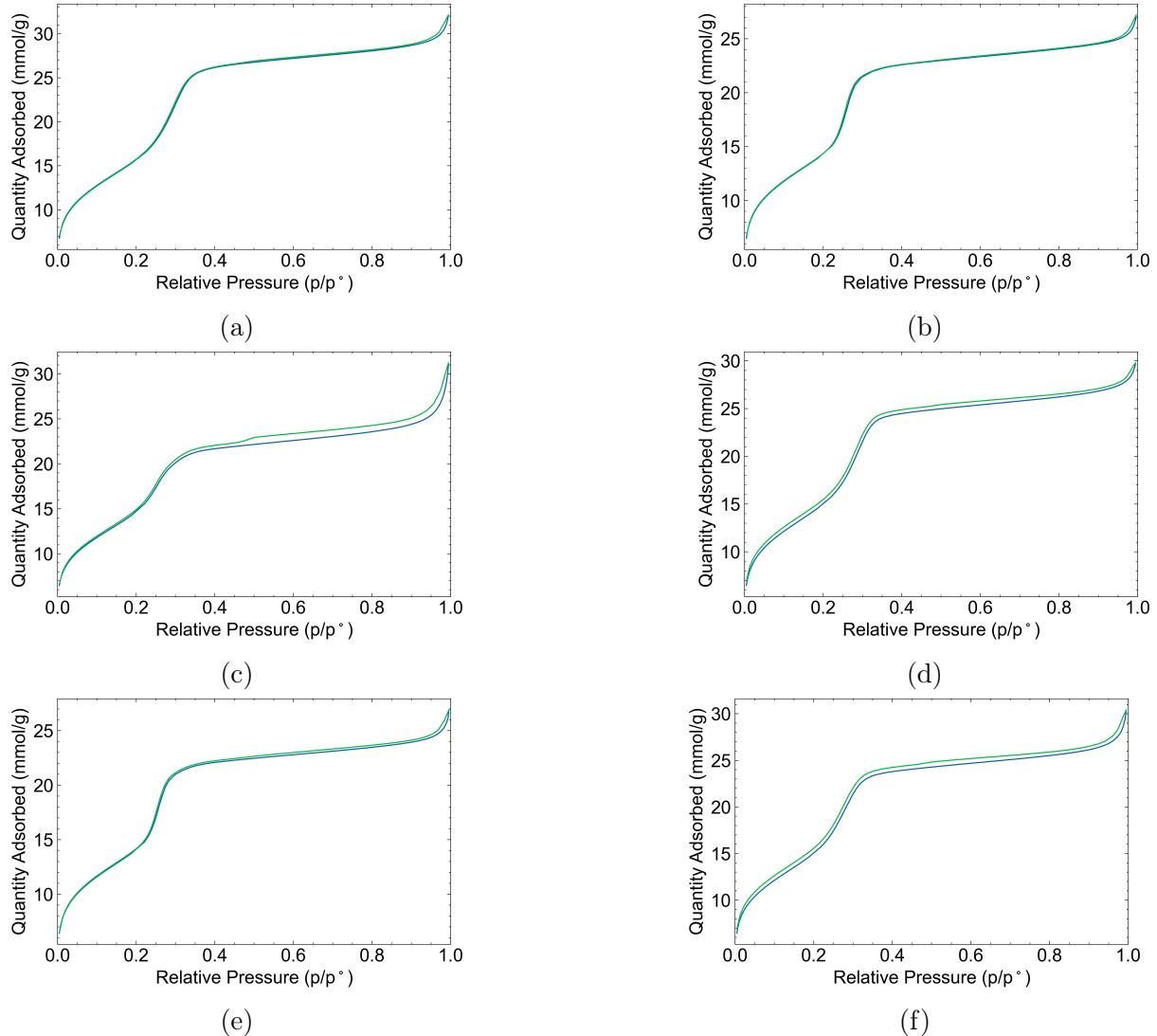


Figure S 2: N₂ adsorption isotherms for component ratio investigation samples. Letters a to f represent samples 2-1 to 2-6, sequentially. The blue line indicates adsorption while the green line indicates desorption.

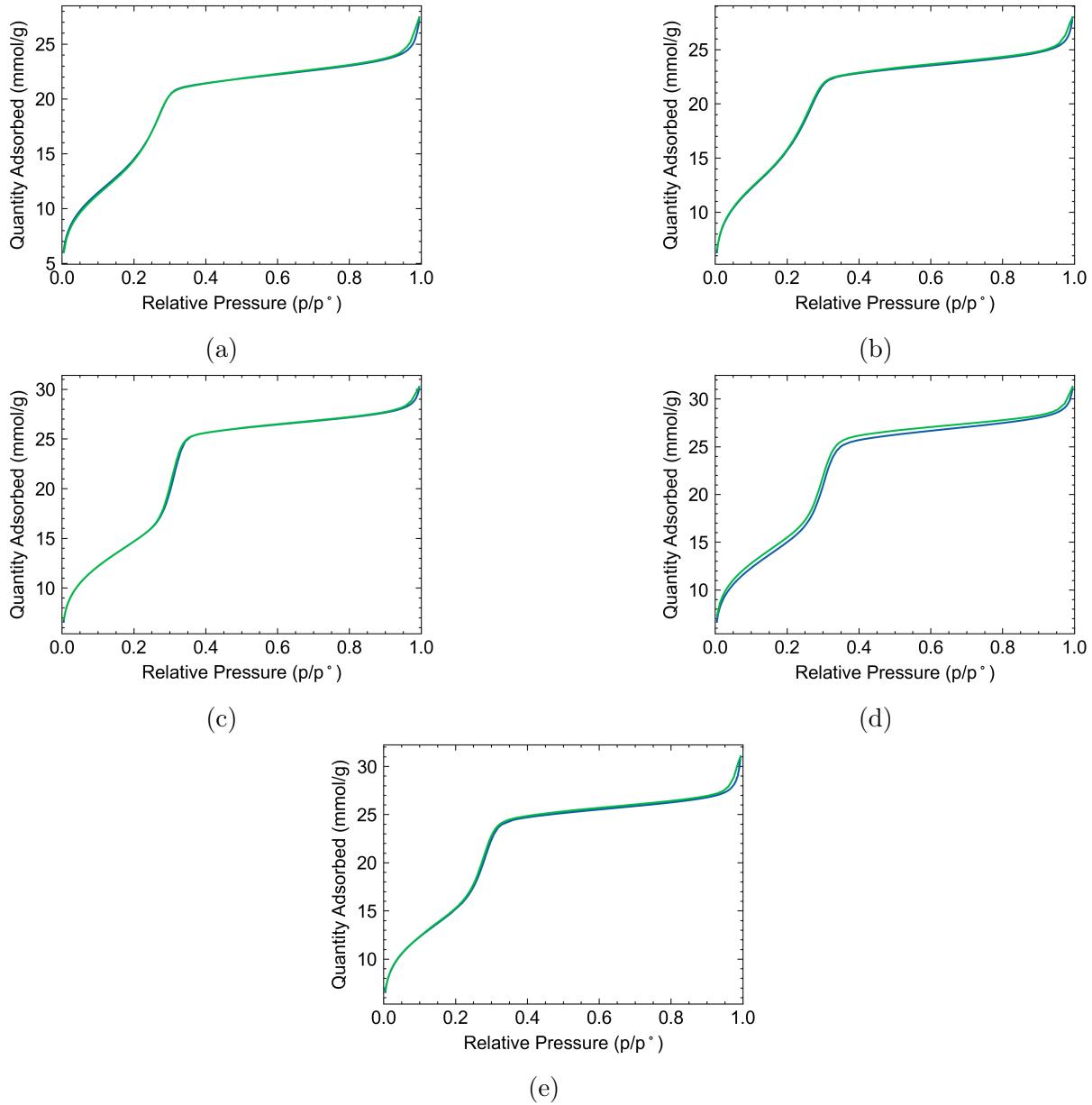


Figure S 3: N₂ adsorption isotherms for additive investigation samples with different additives: (a) L-arginine, (b) ammonia, (c) PEHA, (d) propylamine, (e) no additive.

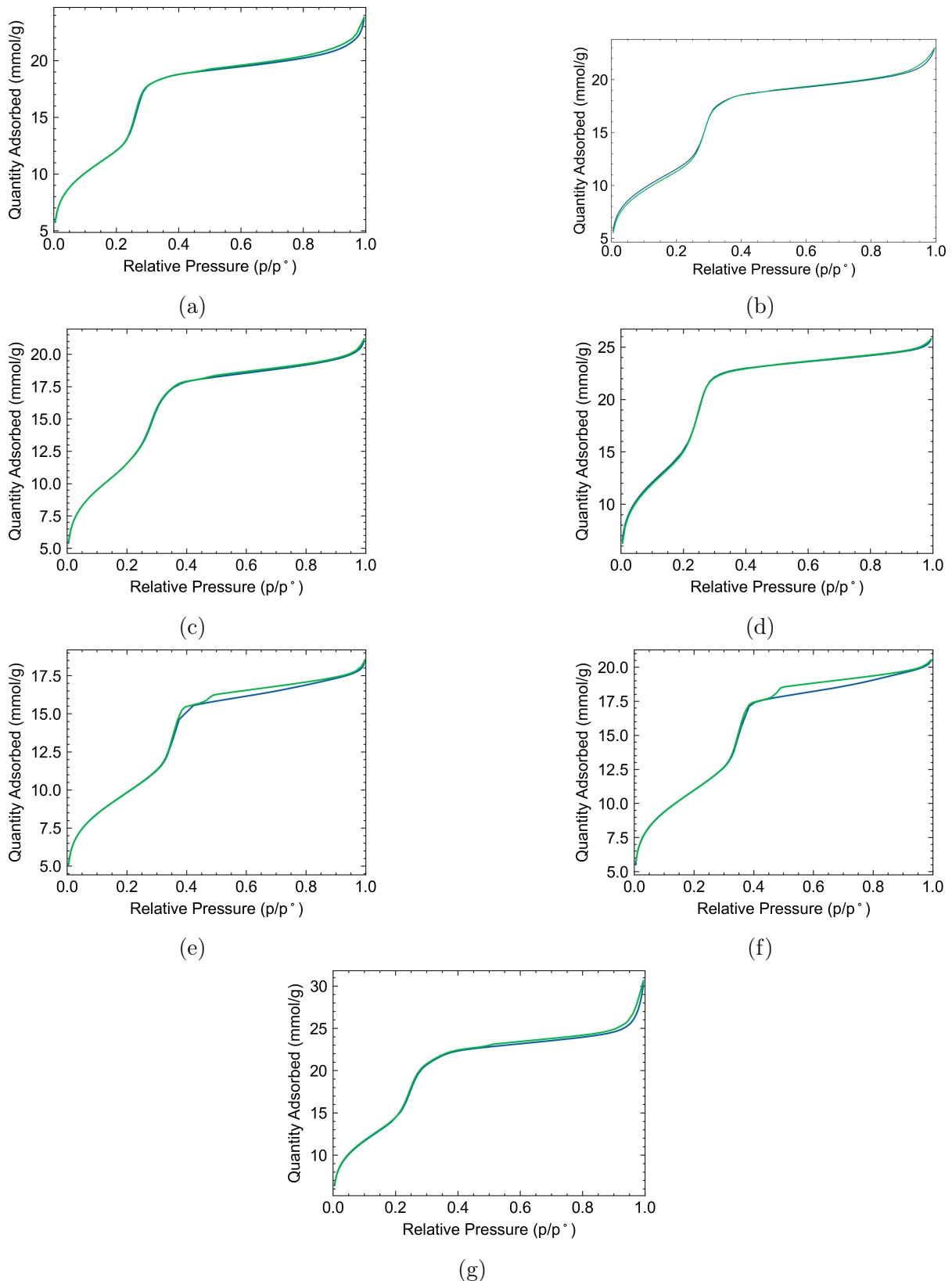


Figure S 4: N_2 adsorption isotherms for ordered mesoporous silica samples using PEHA as an additive in synthesis. Letters a to g represent samples 4-1 to 4-7, sequentially.

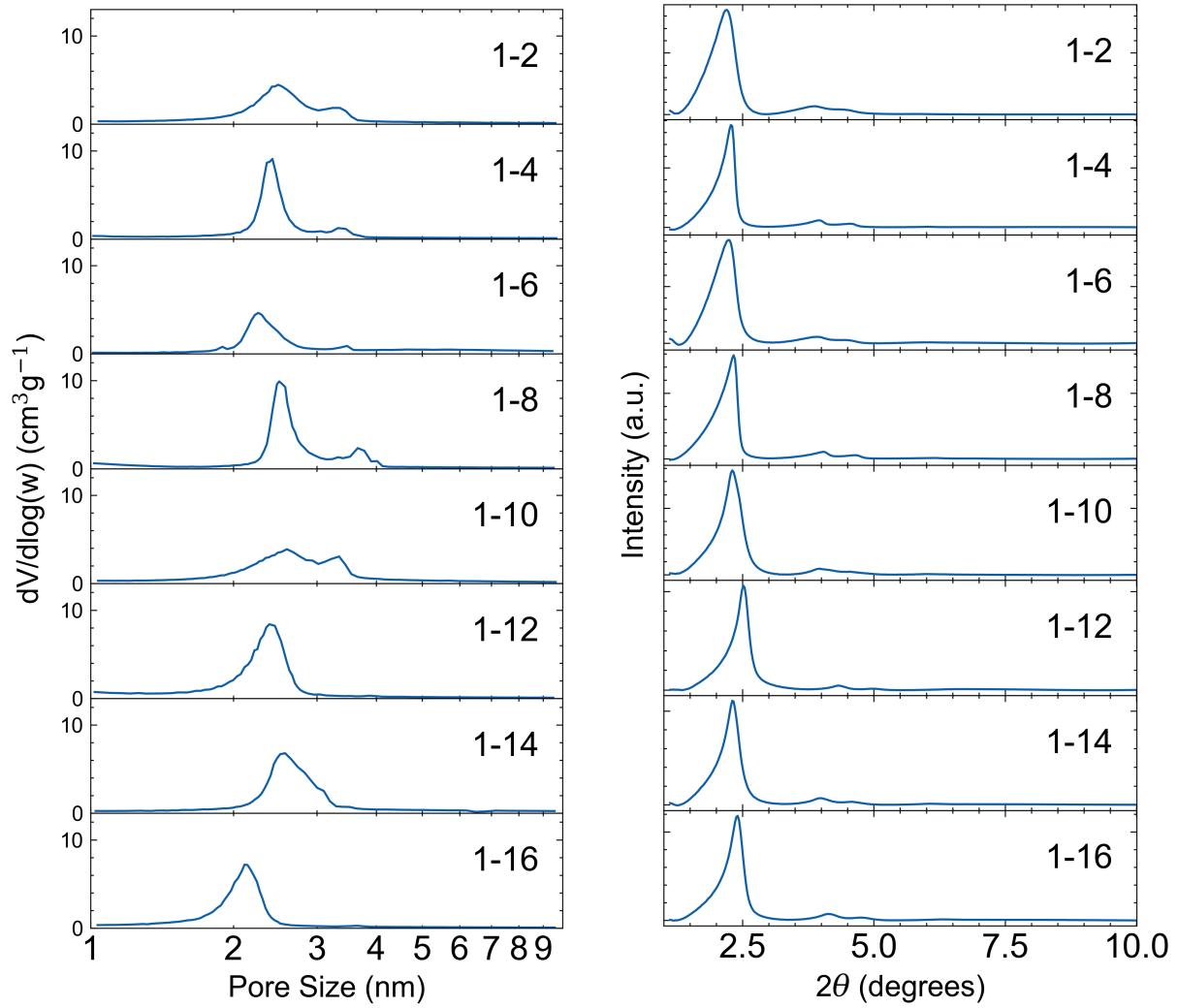


Figure S 5: Pore size distributions (left) and XRD data (right) for two-level four-factor screening samples. For XRD data, baseline intensity was removed manually to allow for easier comparison between peaks.

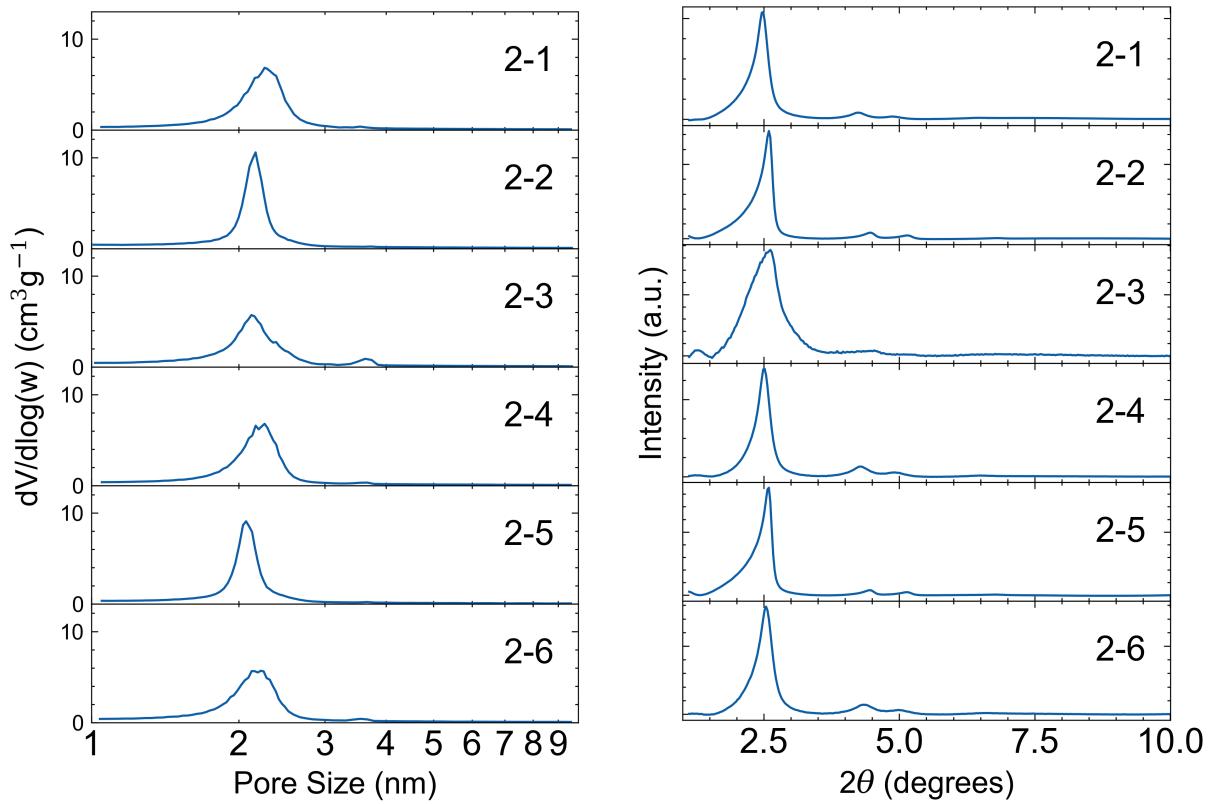


Figure S 6: Pore size distributions (left) and XRD data (right) for component ratio investigation samples. For XRD data, baseline intensity was removed manually to allow for easier comparison between peaks.

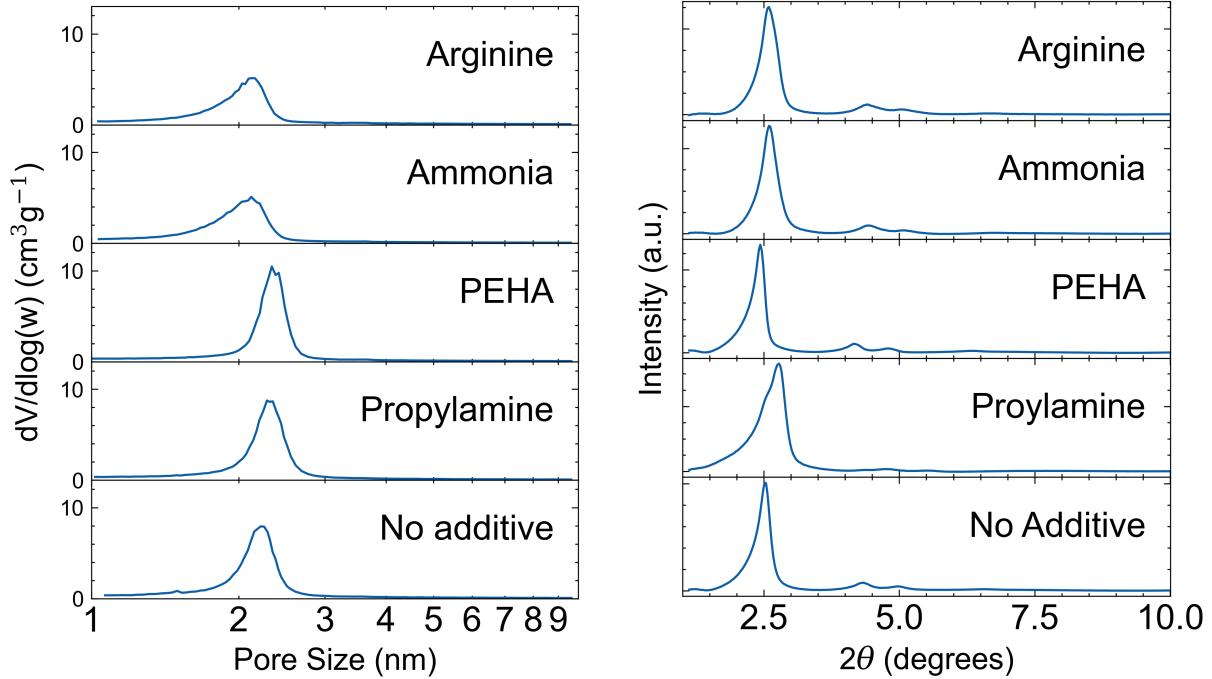


Figure S 7: Pore size distributions (left) and XRD data (right) for investigation with different additives. For XRD data, baseline intensity was removed manually to allow for easier comparison between peaks.

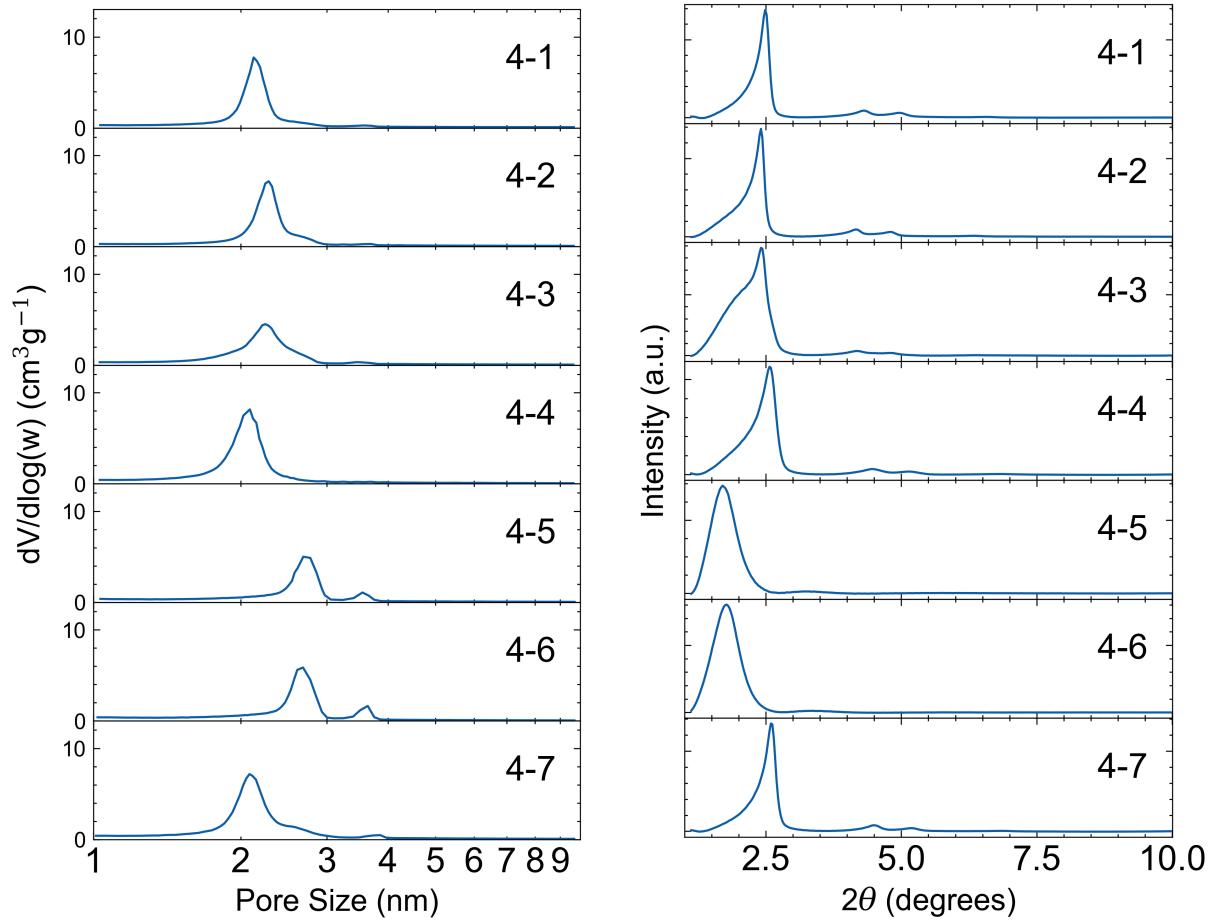


Figure S 8: Pore size distributions (left) and XRD data (right) for samples using PEHA as an additive in synthesis. For XRD data, baseline intensity was removed manually to allow for easier comparison between peaks.

S.2 Uncertainty Analysis

3

To estimate the uncertainty in results for yield, organic content and material properties of samples in this work, the values obtained for two sets of two samples that were synthesised under the same conditions were compared to give an approximate expected range for results. The first set of samples are 2-1 and 3-1, which were synthesised under the same conditions (see Table 1 of the main paper). The second set of samples (R-1 and R-2) were synthesised under the following conditions: $[Si] = 100$ mM, Si:N = 2, Si:CTAB = 2, pH = 10.8. For this second set of results, only gas adsorption results were available. The estimated uncertainty for each property is given in the final column of Table S3, which is estimated based on the range of results for both sets of samples (where available). Although a more rigorous analysis, based on multiple repeat experiments would have been desirable, we believe the values in Table S3 provide a conservative estimate of the real experimental uncertainty.

Table S 3: Summary of differences in yields, organic contents and material properties obtained for repeat samples synthesised under the same conditions.

Parameter	R-1	R-2	Δ	2-1	3-1	Δ	Estimated Uncertainty (\pm)
Yield (%)	-	-	-	87	80	7	10
Organic Content (%)	-	-	-	56	59	3	5
BET Surface Area ($\text{m}^2 \text{ g}^{-1}$)	1211	1277	66	1159	1149	10	70
Pore Volume ($\text{cm}^3 \text{ g}^{-1}$)	0.958	0.952	0.006	0.843	0.789	0.054	0.06
Pore Diameter (nm)	2.31	2.21	0.1	2.13	2.18	0.05	0.1
Wall Thickness (nm)	-	-	-	1.31	1.21	0.1	0.1
Order Parameter	-	-	-	0.12	0.06	0.06	0.1

S.3 Literature XRD

16

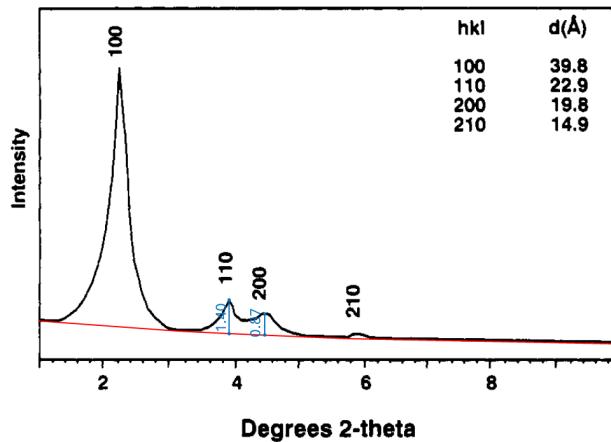


Figure S 9: XRD pattern reported by Beck et al. in the original synthesis of MCM-41, taken from [1]. The red line indicates the baseline reflection from which the relative intensities of peaks (110) and (200) were measured (labelled in blue).

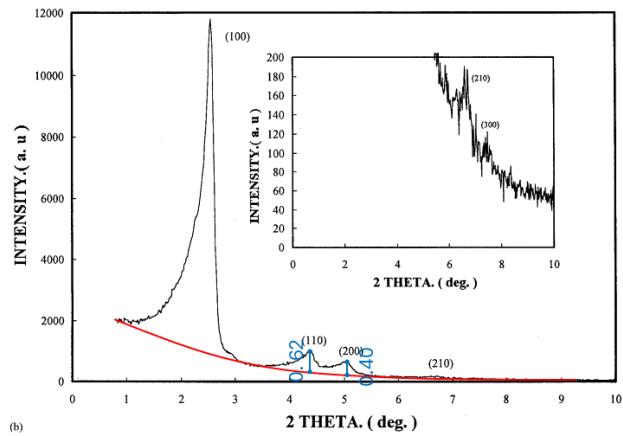


Figure S 10: XRD pattern reported by Cai et al. in their synthesis of MCM-41, taken from [3] for the most well-ordered sample, produced at room temperature. The red line indicates the baseline reflection from which the relative intensities of peaks (110) and (200) were measured (labelled in blue).

S.4 Bio-Inspired Additive Models

17

Arginine has three ionisation points, corresponding to pKa values of 2.17, 9.04 and 12.48 [5]. The most relevant protonation state in this work is shown as IV in Figure S11, which is present in high proportions (>75%) during HLC self-assembly in the early stages of OMS synthesis. This form has an overall negative charge situated on the carboxylate group of the amino acid backbone. The guanidinium group of the arginine sidechain, which is typically positively charged at physiological pH, becomes deprotonated at high pH.

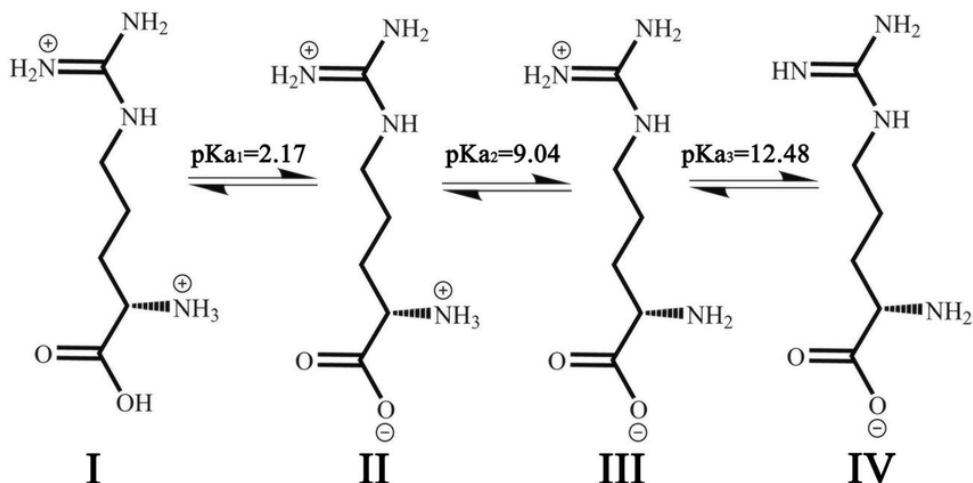


Figure S 11: Diagram of pKa values corresponding to the ionisation of multiple chemical groups in an arginine molecule. Taken from Wang et al. [5].

PEHA's first two ionisation points corresponding to pKa values of 9.7 and 11.0 [6] (shown in Figure S12), refer to protonation of the terminal primary amines. The internal (secondary) amine groups can also protonate, but this happens at pH values that are too low to be relevant in the present context [7]. The most relevant protonation state to this work is when the molecule bears an overall neutral charge with deprotonated amine groups, which occurs at high pH values. The relative composition of different degrees of protonation for both bio-inspired additives between pH 7 and 14 is given in Figure S13.

A mapping for arginine residues in proteins was suggested by the authors of the Martini 3 model [8], which is comprised of 3 beads with the amino acid backbone represented by a single bead (see Figure S14, left). While this is practical for models of proteins that contain many connected backbone beads, for forms of arginine that do not have a zwitterionic backbone, (instead having only either a positively charged amine group or a negatively charged carboxylate group) combining the carboxylate group and the primary amine group into a single bead may not be sufficient to capture the full behaviour of the molecule. Therefore, to more accurately capture this behaviour and allow for the representation of different protonation states of arginine (and in particular the high pH form present during OMS self-assembly), an alternative mapping scheme based on 4 beads is proposed, in which the backbone is divided into two beads, one representing the primary amine group and adjoined carbon, whilst the other represents the charged carboxylate group

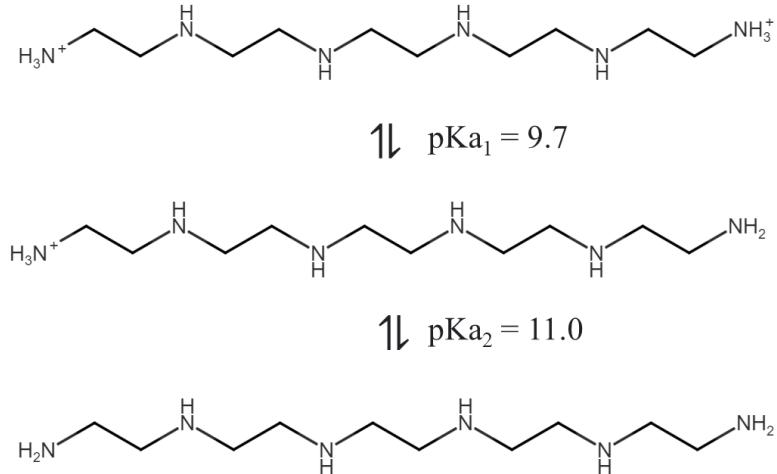


Figure S 12: Diagram of pKa values corresponding to the ionisation of terminal amine groups in PEHA.

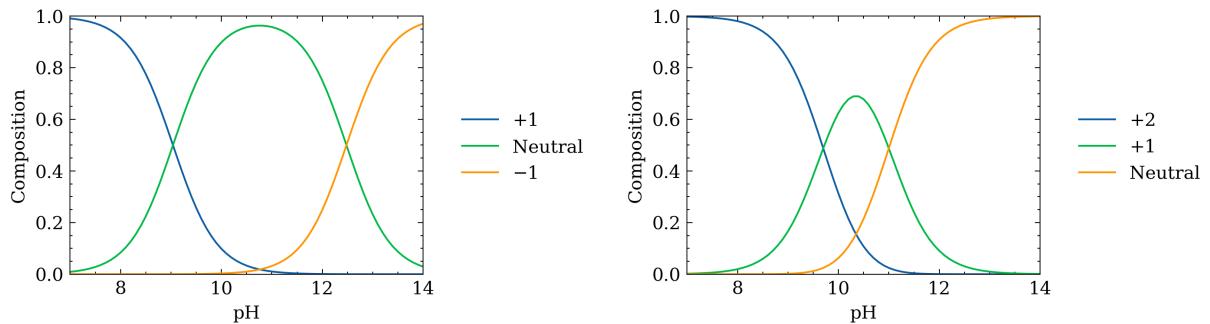


Figure S 13: Distribution of charge states for bio-inspired additives arginine (left) and PEHA (right) between pH 7 and 14. The total molecular charge for each degree of protonation is given in the legend.

(see Figure S14, right). The side chain is mapped with the alkane group represented by one bead with a 3-1 mapping, while the guanidinium group is captured in another bead.

The bead types for the arginine side chain suggested by the Martini 3 authors are the SC3 bead for the alkane chain (ARG_1) and the SQ3p bead for the charged guanidinium group at the end of the side chain (ARG_2). However, to more finely tune the side chain interactions, a recommended approach is to calculate the free energy difference of the side chain beads in different solvents using different candidate bead types, and compare these to experimental values. This procedure was carried out for alternative charged bead types representing the guanidinium group (ARG_2). All small charged bead types were tested using the p label, which represents the hydrogen bonding potential of the group. The two solvents used were water and cyclohexane for which there is experimental data for a neutral analogue of the arginine side chain, *N*-propylguanidine, [9, 10] which gives a free energy of partitioning value of -24.2 kJ mol^{-1} . The free energy of partitioning between water and cyclohexane for the arginine side chain was calculated by thermodynamic integration for each bead. The electrostatic charge of the ARG_2 bead was switched off and only Lennard-

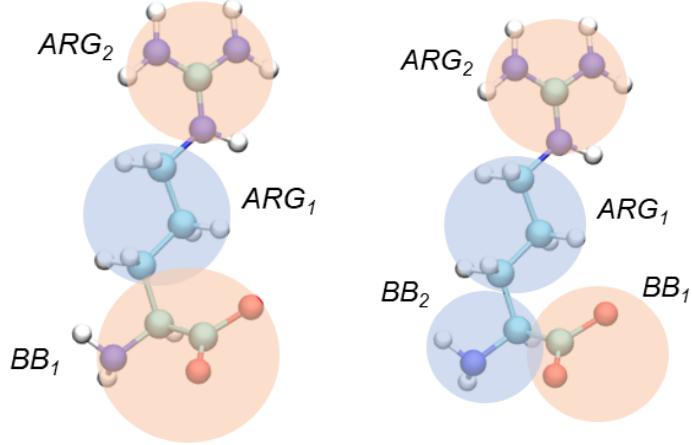


Figure S 14: CG mapping schemes for arginine. The standard Martini 3 mapping scheme is shown on the left while the 4 bead model mapping used in this work is shown on the right.

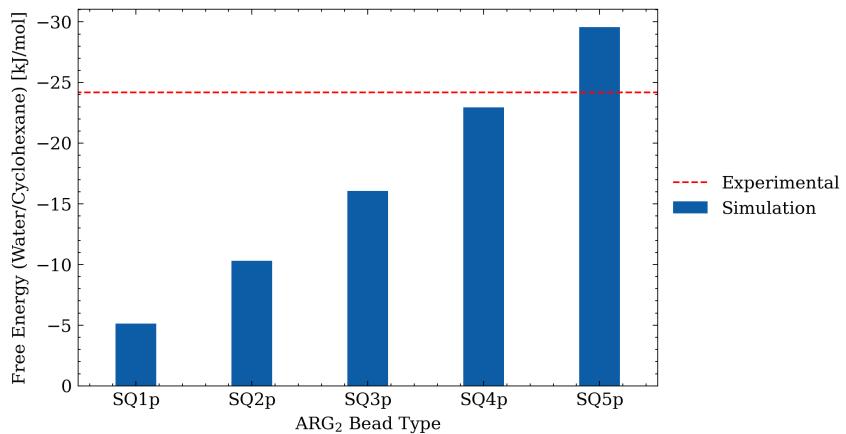


Figure S 15: Partition free energy between water and cyclohexane for different Martini 3 bead types representing the guanidinium group in the arginine side chain analogue.

Jones contributions to the free energy were considered, as only these interactions are affected by the choice of bead type and the available experimental data for the side chain is for an uncharged residue. The results are presented in Figure S15. These results show that the SQ4p bead provides the closest match to experimental data with a calculated value of $-22.95 \text{ kJ mol}^{-1}$, suggesting that this bead may be more appropriate to represent the guanidinium group in arginine. Therefore, this bead type was used for our CG model. 59
60
61
62
63
64

For the backbone beads (BB_1 and BB_2), the bead types suggested by the Martini 3 authors for a carboxylate group and primary amine group are Q5n and N6d, respectively. Therefore, the SQ5n bead type is used to represent the carboxylate group (a small bead size is used due to the 3-1 heavy atom to bead mapping) and the TN6d bead is used for the primary amine group (a tiny bead size is used due to the 2-1 heavy atom to bead mapping). 65
66
67
68
69
70

Pentaethylenehexamine is made up of 16 heavy atoms. Each end of the molecule terminates in a primary amine group, which is connected by repeating secondary amine 71
72

groups interspersed with two carbons. To maintain the symmetry of the molecule, each of these repeating units is represented by a single small Martini bead, while the end groups are represented by a tiny Martini bead representing the primary amine and nearest bonded carbon atom (as well as bonded hydrogen). The mapping scheme is presented in Figure S16. 73
74
75
76
77

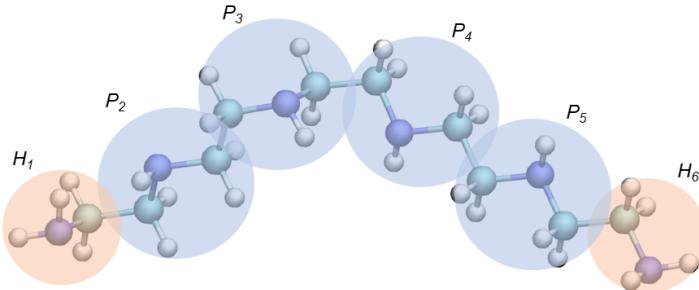


Figure S 16: Martini 3 mapping scheme for the PEHA molecule. Terminal primary amine groups with adjacent carbon are given the label H_1 and H_6 while secondary amine groups have labels P_{2-5} . 78
79
80
81
82
83
84
85
86
87
88

The bead type assignment for the chosen mapping is straightforward, since both primary and secondary amines have bead types recommended by the authors of the Martini 3 force-field [8]. The assignment is therefore only dependent on the charge state of the bead. For neutral primary amines and adjacent carbon (H_1 and H_6), with a 2-1 mapping, the TN6d bead type is recommended with the label representing the hydrogen bond donor characteristic of the amine group. When the terminal primary amine group is positively charged, the TQ5p bead is used instead. For the secondary amine groups and adjoining carbons (P_{2-5}), the SN4 bead type is recommended for neutral amine groups. While not explored in this work, the SQ2p bead could be used to represent positively charged secondary amine groups in PEHA, which occur at low pH [7]. The bead types assigned for both bio-inspired additives are summarised in Table S4. 89
90
91
92
93
94
95

Bonded parameters for the bio-inspired additives are determined from atomistic reference simulations as described in Section 2.4. Bonded parameters for bonds and angles are provided in Tables S5 and S6, respectively. For PEHA, calculated bond lengths and force constants for $H-P$ and $P-P$ bonds, and $H-P-P$ and $P-P-P$ angles, are averaged to a single value. Bonded parameters were generated using a single PEHA molecule with an overall neutral charge and the parameters are also used for CG models of charged PEHA molecules. 96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
5510
5511
5512
5513
5514
5515
5516
5517
5518
5519
5520
5521
5522
5523
5524
5525
5526
5527
5528
5529
5530
5531
5532
5533
5534
5535
5536
5537
5538
5539
5540
5541
5542
5543
5544
5545
5546
5547
5548
5549
55410
55411
55412
55413
55414
55415
55416
55417
55418
55419
55420
55421
55422
55423
55424
55425
55426
55427
55428
55429
55430
55431
55432
55433
55434
55435
55436
55437
55438
55439
55440
55441
55442
55443
55444
55445
55446
55447
55448
55449
55450
55451
55452
55453
55454
55455
55456
55457
55458
55459
55460
55461
55462
55463
55464
55465
55466
55467
55468
55469
55470
55471
55472
55473
55474
55475
55476
55477
55478
55479
55480
55481
55482
55483
55484
55485
55486
55487
55488
55489
55490
55491
55492
55493
55494
55495
55496
55497
55498
55499
554100
554101
554102
554103
554104
554105
554106
554107
554108
554109
554110
554111
554112
554113
554114
554115
554116
554117
554118
554119
554120
554121
554122
554123
554124
554125
554126
554127
554128
554129
554130
554131
554132
554133
554134
554135
554136
554137
554138
554139
554140
554141
554142
554143
554144
554145
554146
554147
554148
554149
554150
554151
554152
554153
554154
554155
554156
554157
554158
554159
554160
554161
554162
554163
554164
554165
554166
554167
554168
554169
554170
554171
554172
554173
554174
554175
554176
554177
554178
554179
554180
554181
554182
554183
554184
554185
554186
554187
554188
554189
554190
554191
554192
554193
554194
554195
554196
554197
554198
554199
554200
554201
554202
554203
554204
554205
554206
554207
554208
554209
554210
554211
554212
554213
554214
554215
554216
554217
554218
554219
554220
554221
554222
554223
554224
554225
554226
554227
554228
554229
554230
554231
554232
554233
554234
554235
554236
554237
554238
554239
554240
554241
554242
554243
554244
554245
554246
554247
554248
554249
554250
554251
554252
554253
554254
554255
554256
554257
554258
554259
554260
554261
554262
554263
554264
554265
554266
554267
554268
554269
554270
554271
554272
554273
554274
554275
554276
554277
554278
554279
554280
554281
554282
554283
554284
554285
554286
554287
554288
554289
554290
554291
554292
554293
554294
554295
554296
554297
554298
554299
554300
554301
554302
554303
554304
554305
554306
554307
554308
554309
554310
554311
554312
554313
554314
554315
554316
554317
554318
554319
554320
554321
554322
554323
554324
554325
554326
554327
554328
554329
554330
554331
554332
554333
554334
554335
554336
554337
554338
554339
554340
554341
554342
554343
554344
554345
554346
554347
554348
554349
554350
554351
554352
554353
554354
554355
554356
554357
554358
554359
554360
554361
554362
554363
554364
554365
554366
554367
554368
554369
554370
554371
554372
554373
554374
554375
554376
554377
554378
554379
554380
554381
554382
554383
554384
554385
554386
554387
554388
554389
554390
554391
554392
554393
554394
554395
554396
554397
554398
554399
554400
554401
554402
554403
554404
554405
554406
554407
554408
554409
554410
554411
554412
554413
554414
554415
554416
554417
554418
554419
554420
554421
554422
554423
554424
554425
554426
554427
554428
554429
554430
554431
554432
554433
554434
554435
554436
554437
554438
554439
554440
554441
554442
554443
554444
554445
554446
554447
554448
554449
554450
554451
554452
554453
554454
554455
554456
554457
554458
554459
554460
554461
554462
554463
554464
554465
554466
554467
554468
554469
554470
554471
554472
554473
554474
554475
554476
554477
554478
554479
554480
554481
554482
554483
554484
554485
554486
554487
554488
554489
554490
554491
554492
554493
554494
554495
554496
554497
554498
554499
554500
554501
554502
554503
554504
554505
554506
554507
554508
554509
554510
554511
554512
554513
554514
554515
554516
554517
554518
554519
554520
554521
554522
554523
554524
554525
554526
554527
554528
554529
554530
554531
554532
554533
554534
554535
554536
554537
554538
554539
554540
554541
554542
554543
554544
554545
554546
554547
554548
554549
554550
554551
554552
554553
554554
554555
554556
554557
554558
554559
554560
554561
554562
554563
554564
554565
554566
554567
554568
554569
554570
554571
554572
554573
554574
554575
554576
554577
554578
554579
554580
554581
554582
554583
554584
554585
554586
554587
554588
554589
554590
554591
554592
554593
554594
554595
554596
554597
554598
554599
554600
554601
554602
554603
554604
554605
554606
554607
554608
554609
554610
554611
554612
554613
554614
554615
554616
554617
554618
554619
554620
554621
554622
554623
554624
554625
554626
554627
554628
554629
554630
554631
554632
554633
554634
554635
554636
554637
554638
554639
554640
554641
554642
554643
554644
554645
554646
554647
554648
554649
554650
554651
554652
554653
554654
554655
554656
554657
554658
554659
554660
554661
554662
554663
554664
554665
554666
554667
554668
554669
554670
554671
554672
554673
554674
554675
554676
554677
554678
554679
554680
554681
554682
554683
554684
554685
554686
554687
554688
554689
554690
554691
554692
554693
554694
554695
554696
554697
554698
554699
554700
554701
554702
554703
554704
554705
554706
554707
554708
554709
554710
554711
554712
554713
554714
554715
554716
554717
554718
554719
554720
554721
554722
554723
554724
554725
554726
554727
554728
554729
554730
554731
554732
554733
554734
554735
554736
554737
554738
554739
554740
554741
554742
554743
554744
554745
554746
554747
554748
554749
554750
554751
554752
554753
554754
554755
554756
554757
554758
554759
554760
554761
554762
554763
554764
554765
554766
554767
554768
554769
554770
554771
554772
554773
554774
554775
554776
554777
554778
554779
554780
554781
554782
554783
554784
554785
554786
554787
554788
554789
554790
554791
554792
554793
554794
554795
554796
554797
554798
554799
554800
554801
554802
554803
554804
554805
554806
554807
554808
554809
554810
554811
554812
554813
554814
554815
554816
554817
554818
554819
554820
554821
554822
554823
554824
554825
554826
554827
554828
554829
554830
554831
554832
554833
554834
554835
554836
554837
554838
554839
554840
554841
554842
554843
554844
554845
554846
554847
554848
554849
554850
554851
554852
554853
554854
554855
554856
554857
554858
554859
554860
554861
554862
554863
554864
554865
554866
554867
554868
554869
554870
554871
554872
554873
554874
554875
554876
554877
554878
554879
554880
554881
554882
554883
554884
554885
554886
554887
554888
554889
554890
554891
554892
554893
554894
554895
554896
554897
554898
554899
554900
554901
554902
554903
554904
554905
554906
554907
554908
554909
554910
554911
554912
554913
554914
554915
554916
554917
554918
554919
554920
554921
554922
554923
554924
554925
554926
554927
554928
554929
554930
554931
554932
554933
554934
554935
554936
554937
554938
554939
554940
554941
554942
554943
554944
554945
554946
554947
554948
554949
554950
554951
554952
554953
554954
554955
554956
554957
554958
554959
554960
554961
554962
554963
554964
554965
554966
554967
554968
554969
554970
554971
554972
554973
554974
554975
554976
554977
554978
554979
554980
554981
554982
554983
554984
554985
554986
554987
554988
554989
554990
554991
55499

Table S 4: Summary of bead type assignments for PEHA and arginine.

Molecule	Bead	Charge	Bead Type
Arginine	BB_1	-1	SQ5n
	BB_1	0	SP2
	BB_2	0	TN6d
	BB_2	+1	TQ4p
	ARG_1	0	SC3
	ARG_2	+1	SQ4p
	ARG_2	0	SQ4p
PEHA	H	0	TN6d
	H	+1	TQ5p
	P	0	SN4
	P	+1	SQ2p

Table S 5: Bonded parameters for bio-inspired additive Martini 3 models. Bead names refer to labels in Figures S16 and S14. b_{ij} is the bond length and k_{ij} is the bond force constant.

Molecule	Bond	b_{ij} (nm)	k_{ij} (kJ mol $^{-1}$)
Arginine	BB_1-BB_2	0.253	37000
	BB_2-ARG_1	0.363	4500
	ARG_1-ARG_2	0.387	9600
PEHA	$H-P$	0.308	6900
	$P-P$	0.368	9400

Table S 6: Angle type parameters for bio-inspired additive Martini 3 models. Bead names refer to labels in Figures S16 and S14. θ_{ijk} is the angle between beads and k_{ijk} is the angle force constant.

Molecule	Beads	θ_{ijk} (degrees)	Calculated k_{ijk} (kJ mol $^{-1}$)
Arginine	$BB_1-BB_2-ARG_1$	84	60
	$BB_2-ARG_1-ARG_2$	141	44
PEHA	$H-P-P$	144	238
	$P-P-P$	146	384

S.5 Ionisation Model for Silicates

96

In order to account for the deprotonation of hydroxyl groups of silica oligomers at varying pH, a series of pKa values for each subsequent deprotonation needs to be determined. Each pKa value corresponds to the deprotonation of an exposed hydroxyl group bonded to a different silicon atom. The number of pKa values required is therefore equal to the number of silicon atoms in the oligomer (or in the case of the coarse-grained model, the number of silica beads). Note that this ignores the possibility of multiple hydroxyl groups bonded to a single silicon atoms becoming deprotonated which can occur in small oligomers. However, this only occurs at extremely high pH values [11], and therefore for simplicity this is not accounted for in this model.

Only very limited experimental data is available for the pKa values of silicate species. The most comprehensive review on the topic gives the required pKa values only for monomers and dimers [11]. However, the pKa value of a silica surface (that is, the surface of a significantly condensed silica network) is also known [12]. Therefore, a relationship can be devised between the number of silicate units in an oligomer and the pKa value of the first deprotonation. This function takes the form:

$$pKa = \frac{a}{b \times N_{Si} + 1} + 1.7i + 6.8 \quad (1)$$

where N_{Si} is the number of silicon units in the oligomer, i corresponds to the degree of deprotonation (i.e. $i = 1$ gives the pKa of the first deprotonation, $i = 2$ gives the pKa of the second deprotonation and so on) and a and b are parameters that are fitted to the known values for the first deprotonation of silica monomers and dimers. This function tends towards a minimum value of 6.8, the pKa of a silica surface, as the number of silicon units (N_{Si}) increases. The prefactor before the term i is the difference in pKa between the first and second deprotonation in dimers, and it is assumed that this difference is maintained for subsequent deprotonations in higher oligomers. The pKa values determined by this method for silica oligomers containing up to 8 silicon atoms (i.e. octamers) are given in Table S7 and Figure S17.

This model can be applied to approximate the proportion of silicate species of each degree of deprotonation for a given system pH. This is shown in Figure 18 for silica octamers between pH 7 and 15.

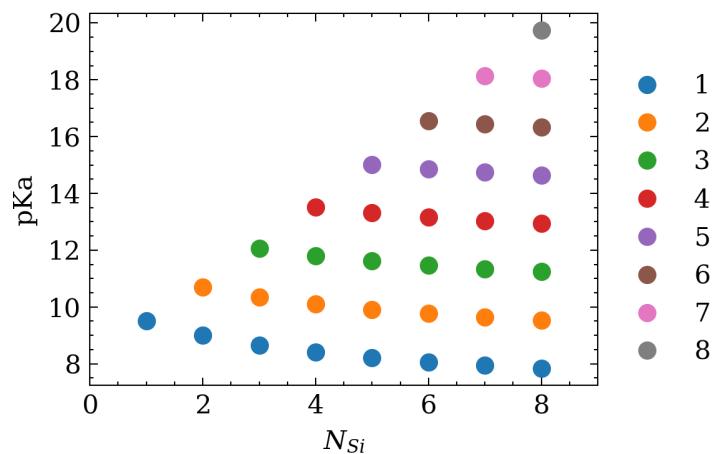


Figure S 17: Estimated pKa values for silica oligomers with up to 8 silicon units. The legend labels refer to the value of i , which is described by Eq.1.

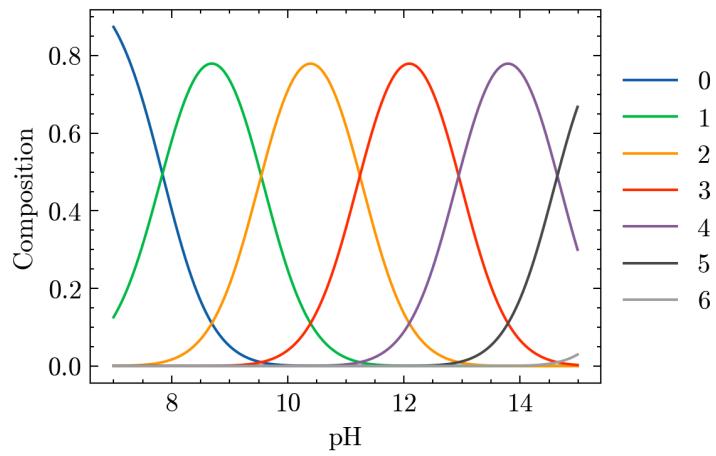


Figure S 18: Charge composition for silica octamers between pH 7 and 15. The legend labels refer to the total (negative) molecular charge.

Table S 7: Table of pKa values for silica oligomers with up to 8 silicon units. N_{Si} gives the number of silicon units while i corresponds to the degree of deprotonation. pKa values are calculated from Eq. 1.

		i							
		1	2	3	4	5	6	7	8
N_{Si}	1	9.50	-	-	-	-	-	-	-
	2	9.00	10.70	-	-	-	-	-	-
	3	8.68	10.38	12.08	-	-	-	-	-
	4	8.47	10.17	11.87	13.57	-	-	-	-
	5	8.30	10.00	11.70	13.40	15.10	-	-	-
	6	8.16	9.86	11.56	13.26	14.96	16.66	-	-
	7	8.05	9.75	11.45	13.15	14.85	16.55	18.25	-
	8	7.95	9.65	11.35	13.05	14.75	16.45	18.15	19.85

S.6 Simulation Details

125

Table S 8: Details of coarse-grained simulations carried out to investigate the effect of Si:CTAB ratio and pH on self-assembly. N is the number of molecules of each species (denoted by the subscript) present in the simulation. Note that each water bead represents 4 water molecules. L is the box length in the x, y and z direction. t is the total simulation time.

Si:CTAB	pH	Box Type	N_{CTAB}	N_{SN1}	N_{SI1}	N_{SN8c}	$N_{SI8c(1-)}$	$N_{SI8c(2-)}$	$N_{SI8c(3-)}$	$N_{SI8c(4-)}$	N_{Br}	N_{TMA}	N_{Water}	L_x (nm)	L_y (nm)	L_z (nm)	t (μs)
0.5	13	Slab	1000	0	20	0	0	0	30	30	770	0	26000	22.4	22.4	7.5	3.0
1	13		1000	0	40	0	0	0	60	60	540	0	52000	31.6	31.6	6.9	3.0
2	13		1000	0	80	0	0	0	120	120	80	0	26000	22.4	22.4	7.5	3.0
4	13		1000	0	160	0	0	0	240	240	0	840	52000	31.6	31.6	7.1	3.0
8	13		1000	0	320	0	0	0	480	480	0	2680	104000	44.7	44.7	6.8	3.0
16	13		1000	0	640	0	0	0	960	960	0	6360	208000	63.3	63.3	6.7	3.0
2	10		1000	61	19	0	62	178	0	0	563	0	26000	22.4	22.4	6.8	3.0
4	10		1000	122	38	0	125	355	0	0	127	0	52000	31.6	31.6	7.1	3.0
8	10		1000	243	77	0	250	710	0	0	0	747	104000	44.7	44.7	6.8	3.0
16	10		1000	486	154	0	499	1421	0	0	0	2495	208000	63.3	63.3	6.7	3.0
2	7		1000	80	0	209	31	0	0	0	969	0	26000	22.4	22.4	6.8	3.0
4	7		1000	160	0	418	62	0	0	0	938	0	52000	31.6	31.6	7.1	3.0
8	7		1000	320	0	835	125	0	0	0	875	0	104000	44.7	44.7	6.8	3.0
16	7		1000	640	0	1670	250	0	0	0	750	0	208000	63.3	63.3	6.6	3.0

20

Table S 9: Details of coarse-grained simulations carried out with bio-inspired additives. N is the number of molecules of each species (denoted by the subscript) present in the simulation. Note that each water bead represents 4 water molecules. L is the box length in the x , y and z direction. t is the total simulation time.

Simulation	Box Type	N_{CTAB}	N_{SI2}	N_{Arg}	N_{PEHA}	N_{Water}	L_x (nm)	L_y (nm)	L_z (nm)	t (μs)
SI2	Elongated	1000	500	0	0	10000	8.2	8.2	27.3	6.0
Arg + SI2	Elongated	1000	500	250	0	10000	8.3	8.3	27.5	6.0
PEHA + SI2	Elongated	1000	500	0	250	13000	8.8	8.8	29.3	6.0

Table S 10: Details of atomistic simulations carried out with bio-inspired additives to determine coarse-grained model parameters. N is the number of molecules of each species (denoted by the subscript) present in the simulation. L is the box length in the x , y and z direction. t is the total simulation time.

Simulation	Box Type	N_{Arg}	N_{PEHA}	N_{Water}	L_x (nm)	L_y (nm)	L_z (nm)	t (ns)
Arginine	Elongated	1	0	874	3.0	3.0	3.0	50
PEHA	Elongated	0	1	2146	4.0	4.0	4.0	50

S.7 Additional Experimental Results

126

The organics content of the precipitate is strongly dependent on the Si:CTAB ratio, as 127 shown in Figure S19. Synthesis carried out at a low Si:CTAB ratio, i.e. with a high 128 concentration of surfactant species compared with the silica concentration, produced a 129 precipitate with a high content of organic species, suggesting that additional surfactant 130 molecules are bound to condensed silica when compared with syntheses where less sur- 131 factant is present. Sample 1-3 appears to be an outlier with a high organic content despite 132 the high Si:CTAB ratio; however this can be explained by the extremely low quantity of 133 material that was obtained under these synthesis conditions, which resulted in no measur- 134 able quantity of silica being present after calcination. Therefore, this sample is not shown 135 in the plot of Figure S19. 136

A significantly higher organics content present in samples at a low Si:CTAB ratio 137 (prior to calcination) indicates a greater surfactant uptake in the precipitated solid phase 138 when the concentration of surfactant is increased. However, since this is not accompanied 139 by an increase in yield of silica, this is more likely to be the result of surfactant rich 140 assemblies that are bound to silica particles but do not possess sufficient silica to result 141 in a porous solid being obtained after calcination. This is generally unfavourable, as the 142 quantity of surfactant needed for these syntheses is much higher, without resulting in an 143 increase in the quantity of mesoporous silica obtained. This is particularly pertinent when 144 using calcination to remove the surfactant template from silica, as this bound surfactant 145 is destroyed. Thus, a high ratio of Si:CTAB (and therefore a low concentration of CTAB) 146 is desirable, particularly in terms of process scalability and economics. 147

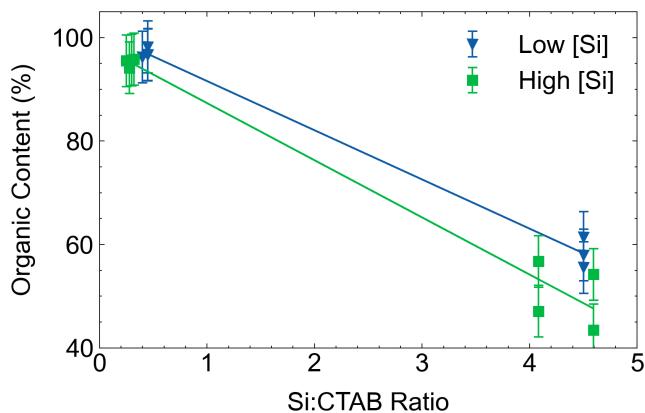


Figure S 19: Dependence of organic content on Si:CTAB ratio for all samples produced in the two-level four-factor factorial design. The line is a guide to the eye.

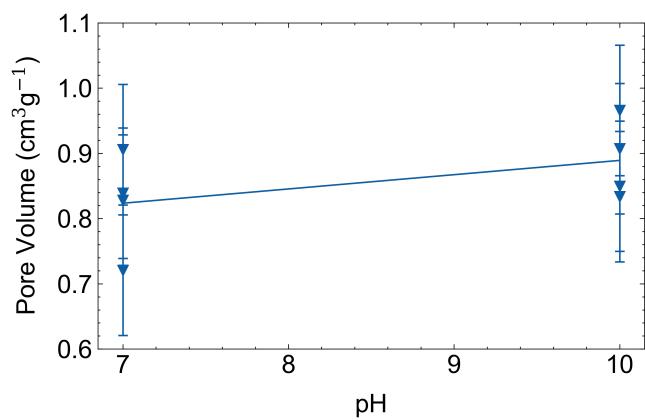


Figure S 20: Dependence of pore volume on pH. The line is a guide to the eye.

References

148

[1] J. S. Beck et al., ‘A new family of mesoporous molecular sieves prepared with liquid crystal templates’, *Journal of the American Chemical Society* **1992**, *114*, 10834–10843. 149
150
151

[2] C. T. Kresge, M. E. Leonowicz, W. J. Roth, J. C. Vartuli, J. S. Beck, ‘Ordered mesoporous molecular sieves synthesized by a liquid—crystal template mechanism’, *Nature* **1992**, *359*, 710–712. 152
153
154

[3] Q. Cai, W.-Y. Lin, F.-S. Xiao, W.-Q. Pang, X.-H. Chen, B.-S. Zou, ‘The preparation of highly ordered MCM—41 with extremely low surfactant concentration’, *Microporous and Mesoporous Materials* **1999**, *32*, 1–15. 155
156
157

[4] E. V. Vyshegorodtseva, Y. Larichev, G. V. Mamontov, ‘The influence of CTAB/Si ratio on the textural properties of MCM—41 prepared from sodium silicate’, *Journal of Sol-Gel Science and Technology* **2019**, *92*, 496–505. 158
159
160

[5] Y. Wang, L. Jiang, Q. Shen, J. Shen, Y. Han, H. Zhang, ‘Investigation on the self—assembled behaviors of C 18 unsaturated fatty acids in arginine aqueous solution’, *RSC Adv.* **2017**, *7*, 41561–41572. 161
162
163

[6] E. J. Shepherd, J. A. Kitchener, ‘474. The ionization of ethyleneimine and poly-ethyleneimine’, *Journal of the Chemical Society (Resumed)* **1956**, 2448–2452. 164
165

[7] J. R. H. Manning, T. W. S. Yip, A. Centi, M. Jorge, S. V. Patwardhan, ‘An Eco—Friendly, Tunable and Scalable Method for Producing Porous Functional Nano-materials Designed Using Molecular Interactions’, *ChemSusChem* **2017**, *10*, 1683–1691. 166
167
168
169

[8] P. C. T. Souza et al., ‘Martini 3: a general purpose force field for coarse—grained molecular dynamics’, *Nature Methods* **2021**, *18*, 382–388. 170
171

[9] R. Wolfenden, L. Andersson, P. M. Cullis, C. C. B. Southgate, ‘Affinities of amino acid side chains for solvent water’, *Biochemistry* **1981**, *20*, 849–855. 172
173

[10] A. Radzicka, R. Wolfenden, ‘Comparing the polarities of the amino acids: side—chain distribution coefficients between the vapor phase, cyclohexane, 1—octanol, and neutral aqueous solution’, *Biochemistry* **1988**, *27*, 1664–1670. 174
175
176

[11] J. Šefčík, A. V. McCormick, ‘Thermochemistry of aqueous silicate solution precursors to ceramics’, *AIChE Journal* **1997**, *43*, 2773–2784. 177
178

[12] S. Sjöberg, ‘Silica in aqueous environments’, *Journal of Non-Crystalline Solids* **1996**, *196*, 51–57. 179
180