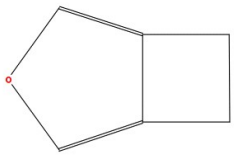
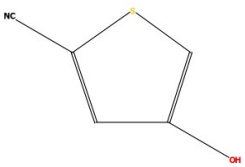
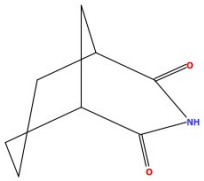
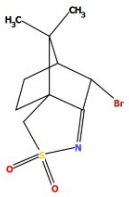
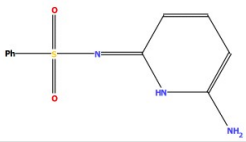
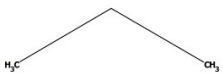
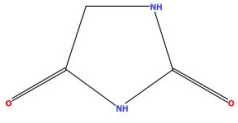
Sources of Symmetry in 'Blind Tests' Crystal Structures

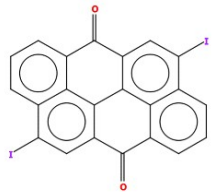
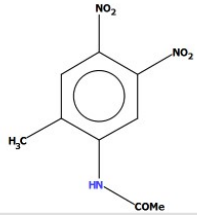
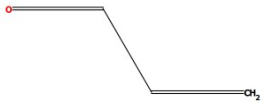
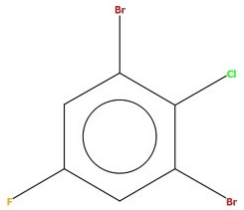
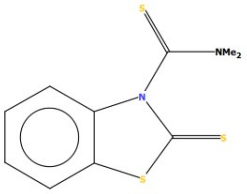
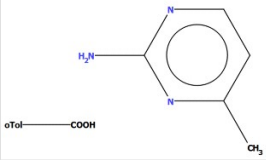
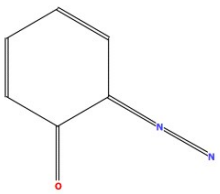
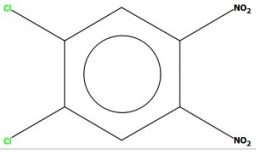
Supporting Information

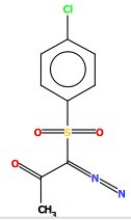
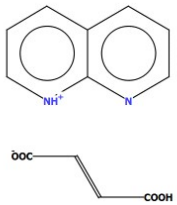
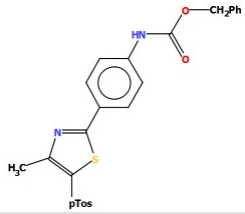
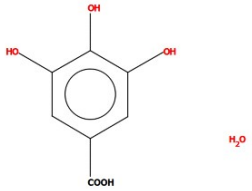
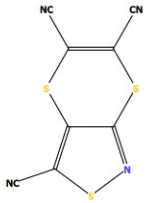
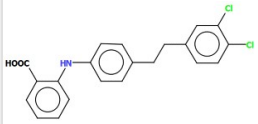
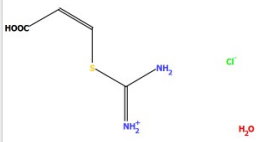
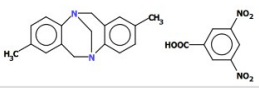
- **Molecular Structures for 31 Compounds in the dataset**
- **Finding and Classifying Aromatic Approaches**
- **Finding and Classifying H-bonds**
- **Complete lists of all 164 aromatic hollows**
- **Complete lists of all 71 H-bonds**
- **Data Sources for Fig. 2**
- **Intermolecular interaction energies**
- **Lattice Energies and rankings**

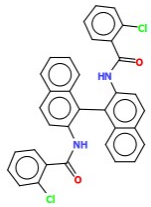
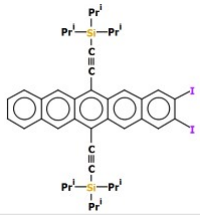
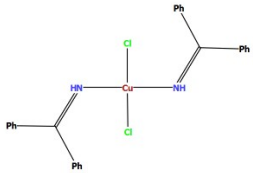
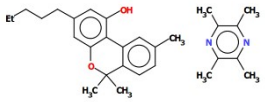
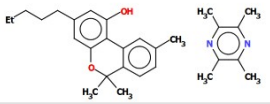
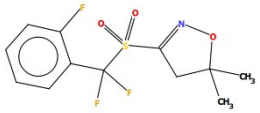
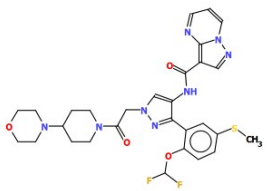
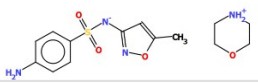
Simon N Black
Department of Chemistry, Durham University, Durham, DH1 3LE, UK
simon.n.black@durham.ac.uk

Molecular Structures for 31 Compounds in the dataset

Test	Molecule		Crystal Structures	REFCODES
1	I		2	XULDUD XULDUD01
1	II		1	GUFJOG
1	III		1	QAMTAZ
2	IV		1	BOQQUT
2	V		1	BOQWIN
2	VI		2	UJIRIO UJIRIO05
1	VII		1	JAYDUI
3	VIII		1	PAHYON01

Test	Molecule		Crystal Structures	REFCODES
3	IX		1	XATMIP
3	X		1	HAMTIZ01
4	XII		1	AXOSOW01
4	XIII		1	SOXLEX01
4	XIV		1	WIDBAO
4	XV		1	WICZUF
5	XVI		1	OBEQUJ
5	XVII		1	OBEQOD

Test	Molecule		Crystal Structures	REFCODES
5	XVIII		1	OBEQET
5	XXI		1	XATJOT
5	XX		1	OBEQIX
5	XIX		4	KONTIQ KONTIQ01 KONTIQ05 KONTIQ06
6	XXII		1	NACJAF
6	XXIII		3	XAFPAY XAFPAY01 XAFPAY03
6	XXIV		1	XAFQON
6	XXV		1	XAFQAZ

Test	Molecule		Crystal Structures	REFCODES
6	XXVI		1	XAFQIH
7	XXVII		1	XIFZOF
7	XXVIII		1	OJIGOG01
7	XXX		1	MIVZEA
7	XXX		1	MIVZIE
7	XXXI		3	ZEHFUR ZEHFUR01 ZEHFUR02
7	XXXII		1	JEKVII
7	XXXIII		2	ZEGWAN ZEGWAN01

Finding and Classifying Aromatic Approaches

The crystal structures were viewed using Mercury, using the 'atom labels' option. In structures containing > 1 aromatic ring, the carbon atom with the lowest number in each ring was used to uniquely identify that ring. The 'short contacts' option was edited to 'Van der Waals radii + 0.2Å' to identify donors for each hollow. As previously, a donor was defined as a heavy (non-carbon) atom within 4 Å of the ring centroid and directly above the ring when viewed exactly perpendicular to the ring plane. Where a hydrogen atom occupied such a position, or a heavy (non-hydrogen) atom was found between 4 and 6 Å above the centroid, the hollow was classified as 'blocked'. Donors were from equivalent rings, different rings on the same molecule, rings on different molecules, conjugated substituents or methyl groups. Approaches were classified as Symthon I, II or III where they met the geometric and symmetry criteria laid out in Black (2024).[4] XAFQAZ contained the only aromatic approach in this dataset between identical aromatic rings that did not meet the criteria for a symthon – the aromatic planes are at an angle of 12.1º to each other and embody glide symmetry, which when repeated = c = 8.236 Å.

Next, asymmetric approaches were identified using the 'colour by symmetry equivalence' option. Thereafter, the symmetry of approaches between identical molecules was identified using the 'colour by symmetry operation' option. The dimensionality of the resulting network was established manually by repeated expansion of the identified close contacts.

Finding and Classifying H-bonds

The default options for hydrogen bonds within Mercury were used. This led to some discrepancies with studies carried out using previous versions of Mercury, which used van der Waals radii from A. Bondi, *J. Phys. Chem*., 1964, 441-451. For example, the radius of a nitrogen atom has changed from 1.55 to 1.66Å. Where the default definition revealed unused H-bond donors, or bifurcated H-bonds, the crystal structures were re-examined using the option 'require H atom to be present'. In this way, all 71 H-bond donors were used, without bifurcation. Each H-bond was identified uniquely by the atom label of the H atom. Functional groups containing both H-bond donors and acceptors were identified by inspection. The symmetry of each H-bond was established by successive use of the 'colour by symmetry equivalence' and 'colour by symmetry operation' options, as for aromatic interactions. The dimensionality of the resulting H-bonded network in each crystal structure was obtained rapidly using the 'expand all' option in 'build networks'. It was noted that this was much easier for H-bonds than for aromatic approaches.

Complete Listing of all 164 Aromatic Hollows

Hollows number	Hollows Total	Molecule	REFCODE	rings	approach	Etter	repeat	axis	classification
1	1	I	XULDUD	C4O1	offset I, C4	R			embody
1	2				Symthon II, C1-C2 edge	C	5.309	a	embody
1	3	I	XULDUD01	C4O1	C4(not)-H5(block)		4.954	a	H block
1	4				C4 (CH2)	C	9.679	c	assist
1	5	II	GUFJOG	C4S1	Symthon III, S1	R	(1+2)		embody
1	6				Symthon III, C3	R	7.715	a	embody
3	9	III	QAMTAZ	C6/C3O1B1 + C6'	C11/C10 + C3-C8; 7.1°	R			assist
2	11			C6/C3O1B1	Symthon III, O2	R	(1+2)	ab	embody
1	12			C6'	Symthon II, C13-C12e	C	7.634	b	embody
1	13	VI	UJRIO	C5N1	Symthon III, C4-C5	R			embody
3	16			C5N1 + C6	blocked				3 blocks
1	17	VI	UJRIO05	C5N1	offset i O1 (s = 3.467)	R			embody
3	20			C5N1 + C6	blocked				3 blocks
12	32	IX	XATMIP	C6 x 6	Symthon I	C	4.202	a	embody
2	34	X	HAMTIZ01	C6	offset trans, O4 + C9	C	4.853	b	embody
2	36	XIII	SOXLEX01	C6	Symthon I, C3 + C6	C	3.894	a	embody
2	38	XIV	WIDBAO	C6/C3N1S1	S1/C2 + C10/C1, skew 20.1°	C	9.335	c	assist
2	40			C6/C3N1S1	blocked				2 blocks
2	42	XV	WICZUF	C6 C4N2	N3/C7+C9/C1, 4.6°	D			asymmetric
1	43			C6	C4-C3e... f	D			asymmetric
1	44			C4N2	H4 block (step, i)				Hblock
1	45	XVI	OBEQUJ	C6	N2, offset i	R			embody
1	46				Symthon II, C6 (+ O1... N1)	C	7.381	b	embody
2	48	XVII	OBEQOD	C6	offset trans block				2 blocks
1	49	XVIII	OBEQET	C6	C7, Symthon III	R			embody
1	50				blocked				block
4	54	XIX	XATJOT	C5N1/C5N1'	Symthon I, N2 +N1 +C4 +C2	C	3.714	b	embody
1	55	XX	OBEQIX	C6	C16-C17e	C	27.632	ac	assist
1	56			C6	Symthon II, C23-C24e	C	6.356	b	embody
2	58			C3N1S1 + C6'	C8 skew 12.8° + H14 block	C	6.356	b	assist + Hblock
1	59			C6"	C18methyl	R			assist
3	62			C6', C6", C3N1S1	blocked				3 blocks
2	64	XXI	KONTIQ	C6	H3block, step, s=3.277 + O1 block	C	4.719	b	Hblock + block
2	66	XXI	KONTIQ01	C6	Symthon I, C6 + C3	C	3.622	b	embody
2	68	XXI	KONTIQ05	C6	Symthon I. C6 + C3	C	3.609	b	embody
2	70	XXI	KONTIQ06	C6	Symthon I, C4-C5 + C2-C3	C	3.641	b	embody
1	71	XXII	NACJAF	C3N1S1	-C3=N1e... C5f	C	14.333	ac	assist
1	72				-C8=N3 block	C	6.696	b	block
2	74	XXIII	XAFPAY	C6 + C6'	C4-C21 + C13-C14; 4.6º	C	16.236	c	assist
2	76			C6" + C6	C8 "T"/C1 + block	R			assist + block
1	77			C6"	C6-C17e/C1f	C	10.53	b	assist
1	78			C6'	C20-C15e/C5f	C	26.258	d	assist
1	79	XXIII	XAFPAY01	C6	C19-C20/C16, Symthon III	R			embody
3	82			C6 + C6' + C6"	C10-C9e + (C13/C16 +Cl2/C8; 7.1°)	C	8.069	ab	assist
1	83			C6'	C4-C3e	C	7.805	b	assist
1	84			C6"	offset I, C7=O2	R			embody
2	86	XXIII	XAFPAY03	C6 +C6'	C14 + C3, 6.5ºskew	C	15.306	ac	assist
2	88			C6" + C6'	C5-C6e + block	R			assist + block
1	89			C6"	C11-C10e	C	10.728	b	assist
1	90			C6	C18 block				block
1	91	XXV	XAFQAZ	C6	Symthon III, C2	R			embody
2	93			C6 C6	C11-C12 + C7-C6, 1.1º	D			asymmetric
3	96			C6' C6	(C16 + C19-C18; 12.1º) + C15 methyl	C	8.126	c	embody
2	98	XXVI	XAFQIH	C6 + C6'	Symthon III C15 + C14-C15e	R			embody + assist
4	102			C6' + C6 +C6"/C6"'	C13T + (C7-C6skew 20.2°) + C13 block x2	R	(1+2)	ab	assist x 2 + 2 blocks
1	103			C6"'	C32 T	C	11.03	b	assist
1	104			C6""	C4-C5e	R	(2+4)	c	assist
1	105			C6"""	Cl1	S			intramolecular
3	108			CC6", C6"", C6"""	blocked				block x 3
5	113	XXVII	XIFZOF	C6 x 5	offset trans: C4,C6,I1, C17,C15	C	9.296	a	embody
3	116			C6 x 3	Symthon III: C20-C21, C18-C19,H8block	R			embody + Hblock
2	118			C6 x 2	block + H52 block				block + Hblock
8	126	XXVIII	OJGOG01	C6 + C6' (+i)	Cl1, H3, H5, H11 blocks				8 blocks
3	129	XXX	MIVZEA	C4N2 + C6 + C6'	(C19-C20 + N1, 5.2°) + C6methyl... C9f	D			asymmetric (3)
2	131			C4N2 + C6	C12 methyl + C7methyl	D	(1+2)	b	asymmetric (2)
1	132			C6'	C17methyl	R			assist
1	133			C6	C50methyl... C40	C	8.763	b	assist
3	136			C6' + C6	blocks, C29 + C30 + ...				3 blocks
1	137	XXX	MIVZIE	C4N2	Symthon III, C24	R			embody
2	139			C4N2 C6	C20 methyl + C26 methyl	D			asymmetric
3	142			C6' + C6	blocked (C26, C29,H22)				2 blocks + 1 Hblock
2	144	XXXI	ZEHFUR	C6	blocked (H11, f... f at 4.03)				2 blocks
1	145	XXXI	ZEHFUR01	C6	C4 methyl	R			assist
1	146				facing void				block/void
1	147	XXXI	ZEHFUR02	C6	Symthon III, C6	R			embody
1	148				O1/C2	C	12.798	c	assist
3	151	XXXII	JEKVII	C4N2/C3N2 + C3N2'	offset inv; C12 + H32block + H11block	R			embody + 2 blocks
2	153			C4N2/C3N2	C27-C26e + block	R	(1+2)		assist + block
1	154			C3N2	blocked H28				Hblock
2	156			C6	blocks, one intra				2 blocks
1	157	XXXIII	ZEGWAN	C6	C3 3.3º (2rot)	R			assist
2	159			C6 + C3N1O1	C10methyl +C1 block (2rot)	R			assist + block
1	160			C3N1O1	block O4				block
2	162	XXXIII	ZEGWAN01	C6 + C3N1O1	C14H2 in C1/C2 pocket	D			asymmetric
1	163			C6	C10 methyl	C	16.844	bc	assist
1	164			C3N1O1	C1-C6e	C	11.471	a	assist

Complete Listing of all 71 H-bonds

Test		Molecule	Form	REFCODE	H-bonds	type	Etter	repeat	axis	notes
1	1	II		GUFJOG	O1-H2... N1	assist	C	8.332	b	
2	2	IV	form I	BOQQUT	**N1-H1... O1**	**embody**	**C**	**7.705**	**a**	**cis amide chain**
2	3	VI	form I	UJRIO	N2-H5... O1	assist	C	8.958	b	
2	4				N2-H6... N3	assist	C	8.958	b	combined = chain
2	5				N1-H1... O2	assist	C	8.958	b	combined = chain
2	6	VI	form III	UJRIO05	N2-H11... O1	assist	C	9.138	b	
2	7				N2-H10... N3	assist	C	9.138	b	combined = chain
2	8				N1-H9... O2	assist	C	9.138	b	combined = chain
3	9	VIII		PAHYON01	**N1-H4... O1**	**embody**	**R**	**10.28**	**d**	**cis-amide ring**
3	10	VIII		PAHYON01	**N2-H3... O2**	**embody**	**R**	**combined**		**cis-amide ring**
3	11	X		HAMTIZ01	**N1-H1... O1**	**embody**	**C**	**4.583**	**b**	**trans amide chain**
4	12	XV	co-crystal	WICZUF	**N2-H2... N1**	**embody**	**R**			**pyrimidine ring**
4	13				N2-H1... O2	asymmetric	D			pyrimidine... acid pairing
4	14				O1-H15... N3	asymmetric	D			combined
5	15	XIX	salt	XATJOT	N1-H1... O1	asymmetric	D			charge assisted
	16				O4-H8... O3	assist	C	12.654	c	
5	17	XX		OBEQIX	N2-H5... O4	assist	R			
	18	XXI	hydrate I	KONTIQ	**O5-H6... O6**	**embody**	**R**			**carboxylic acid ring**
	19	XXI			O2-H3... O6	assist	C	7.473	d	
	20				O1-H1... O3	asymmetric	D			O1 = water DDAA
	21				O1-H2... O4	asymmetric	D			"
	22				O3-H4... O1	asymmetric	D			"
	23	XXI			O4-H5... O1	asymmetric	D			"
5	24	XXI	hydrate II	KONTIQ01	O5-H5... O2	assist	R			
	25	XXI			-O4... O5	assist	R			ring: 2-fold rotation
	26				O1-H1... O6	asymmetric	D			O6 = water DDAA
	27				O3-H3... O6	asymmetric	D			"
	28				O6-H7... O4	asymmetric	D			"
	29	XXI			O6-H8... O3	asymmetric	D			"
5	30	XXI	hydrate IV	KONTIQ05	O2-H4... O3	assist	C	3.609	b	chain with Symthon I
	31	XXI			O3-H5... O5	assist	C	3.609	b	chain with Symthon I
	32				O6-H7... O2	asymmetric	D			O6 = water DDAA
	33				O6-H8... O1	asymmetric	D			"
	34				O1-O3... O6	asymmetric	D			"
	35	XXI			O4-H6... O6	asymmetric	D			"
5	36	XXI	hydrate V	KONTIQ06	**O4-H4... O5**	**embody**	R			**carboxylic acid ring**
	37	XXI			O1-H1... O5	assist	C	7.607	a	
	38	XXI			O2-H2... O3	assist	C	3.641	b	
	39	XXI			**O6-H8... O6**	**embody**	C	**3.641**	**b**	**water chain**
	40				O3-H3... O6	asymmetric	D			O6 water DDAA
	41	XXI			O6-H7... O2	asymmetric	D			
6	42	XXII	form A	XAFPAY	**O2-H17... O1**	**embody**	R			**carboxylic acid ring**
6	43				N1-H16... O1	intramolecular	I			
6	44	XXII	form B	XAFPAY01	**O1-H17... O2**	**embody**	R			**carboxylic acid ring**
6	45				N1-H16... O2	intramolecular	I			
6	46	XXII	form D	XAFPAY03	**O2-H17... O1**	**embody**	R			**carboxylic acid ring**
	47				N1-H16... O2	intramolecular	I			
6	48	XXIV	hydrated salt	XAFQON	N1-H1... O1	assist	R			
	49				N1-H2... Cl1	asymmetric	D			chloride ion 5A
	50				N2-H4... Cl1	asymmetric	D			2+3 = 3.991 = a
	51				N2-H5... Cl1	asymmetric	D			3+4 = g x2 = 10.101 = c
	52				O2-H6... O3	asymmetric	D			O3 water DDA
	53				O3-H8... Cl1	asymmetric	D			
	54				O3-H9... Cl1	asymmetric	D			
6	55	XXV	cocrystal	XAFQAZ	O1-H3... N3	asymmetric	D			cocrystal pairing
6	56	XXVI		XAFQIH	N1-H1... Cl1	intramolecular	I			
	57				N2-H2... O1	assist	R			
7	58	XXVIII	molecule on i	OJIGOG01	**N1-H1... Cl1**	**embody**	R	**5.261**	**a**	**combined**
	59				**N1-H1A-Cl1A**	**"**	R	**"**	**"**	**"**
7	60	XXX	1:1 cocrystal	MIVZIE	O2-H1... N2	asymmetric	D			1:1 complex
7	61	XXX	2:1 cocrystal	MIVZEA	O2-H25... N1	asymmetric	D			2:1 complex
7	62				O4-H54... N2	asymmetric	D			"
7	63	XXXII	Form A	JEKVII	N3-H32... N4	intramolecular	I			
7	64	XXXIII	salt Form I	ZEGWAN	N4-H20... O1	asymmetric	D			
7	65				N4-H19... N1	asymmetric	D			
7	66				N2-H9... N3	assist	R			
7	67				N2-H10... O1	assist	R			2-fold rotation
7	68		salt Form II	ZEGWAN01	N4-H20... N3	asymmetric	D			
7	69				N4-H19... N1	asymmetric	D			1 +2 = g x 2 = 11.741 = a
7	70				N2-H3... O3	assist	C	11.378	c	
7	71				N2-H4... O2	assist	C	16.844	d (=cb)	

Data Sources for Fig. 2

Supramolecular Synthons: data from Reference 1: J. D. Dunitz & A. Gavezotti, *Cryst. Growth Des.* 2012, **12**, 5873−5877.

H-bond pairs: Table 2/synthons 1,2,5,7,9-13

Single H-bonds: Table 2/synthons 3,6,8,14-16

C-H…O/Cl pairs: Table 3/synthons 19-21 & Table 4/synthon 24

Aromatic Symthons: data from Figure 4 in Reference 4: S. N. Black and R. J. Davey, *CrystEngComm*, 2024, **26**, 4498-4508.

Table: Intermolecular interaction energies for close approaches shown in Figures 3-13.

Fig.	REFCODE	Symmetry	Method	E_ele	E_pol	E_dis	E_rep	E_tot	Description
				kJ/mol	kJ/mol	kJ/mol	kJ/mol		
3	XATMIP	translation	H	-14.1	-1.2	-130.9	68.9	-77.3	Symthon I
4	XIFZOF	translation	H	-6.7	-4.1	-111.1	53.5	-66.3	Symthon I offset
4	XIFZOF	inversion	H	-12.1	-4.8	-109.2	0	-113.8	Symthon III
5,10	HAMTIZ01	translation	B	-33.4	-13.8	-45.6	63.5	-46	H-bond chain + Symthon I offset
6	XULDUD	glide	B	-4.4	-0.7	-19.1	14.9	-12.5	Symthon II
6	GUFJOG	inversion	B	-14.7	-0.8	-30.7	22.5	-29	Symthon III
6	GUFJOG	inversion	B	-14.1	-0.8	-22	21	-21.7	Symthon III
7	QAMTAZ	inversion	B	-3.3	-1.7	-66.8	40.7	-37.8	Symthon III offset
8	XAFPAY01	translation	B	-2.9	-0.5	-22.2	9.7	-16.8	asymmetric edge... face assist
9	XAFQAZ	asymmetric	B	-7.9	-3	-40.7	21.8	-32.5	face... face
10	KONTIQ06	screw	B	-17.5	-5.1	-3.6	23.6	-10.8	H-bond chain
10	BOQQUT	glide	B	-31.5	-7.7	-14	31.5	-31.7	H-bond glide
10	KONTIQ	inversion	B	-120.6	-26.9	-12.2	137.1	-73.3	H-bond pair
11	PAHYON01	inversion	B	-70.8	-14.2	-10.7	54.1	-61.2	H-bond pair
11	PAHYON01	inversion	B	-72.6	-17.9	-14.1	83.7	-50.6	H-bond pair
11	OJIGOG01	translation	H	-93.6	-46.7	-109	87.2	-153.3	H-bond pair
12	XAFQON	inversion	B	-119.8	-53	-18.5	69.7	-138.9	2 x H-bond assist
12	GUFJOG	translation	B	-49.8	-10.6	-5.8	50.6	-34.3	H-bond chain
13	WICZUF	asymmetric	B	-119.3	-27.6	-19.4	148.9	-71.4	H-bond pair
13	WICZUF	inversion	B	-71.4	-16.2	-15.5	78.3	-52.6	H-bond pair

Calculations were performed using Crystal Explorer 21, [6,7] downloaded on 23/7/25[16], with either the B3LYP/6-31G(d,p) force field (method B) or, where this method failed, HF/3-21G (method H). Light grey shading indicates a large contribution from $E_{\_dis}$. Dark grey shading indicates indicates a large contribution from $E_{\_ele}$.

Table: Lattice Energies and rankings

	N	symmetry	E_tot	E/molecule	Elatt	% lattice	description
			kJ/mol	kJ	kJ/mol		
SOXLEX01 first	2	translation	-26	-26	-76.15	34	Symthon I
XATMIP first	2	translation	-77.3	-77.3	-177.6	44	Symthon I
HAMTIZ01 first	2	translation	-46	-46	-160.8	29	H-bond + symthon I offset
XULDUD first	2	glide	-12.5	-12.5	-76.3	16	Symthon II
XULDUD second	1	inversion	-21.6	-10.8		14	Symthon III offset
GUFJOG first	2	translation	-34.3	-34.3	-88.7	39	H-bond assist
GUFJOG second	1	inversion	-29	-14.5	-88.7	16	Symthon III
GUFJOG third	1	inversion	-21.7	-10.85	-88.7	12	Symthon III
QAMTAZ first	1	inversion	-37.8	-18.9	-115.3	16	assisted inversion
QAMTAZ second	2	translation	-18.9	-18.9	-115.3	16	side... side
QAMTAZ third	1	inversion	-30.5	-15.25	-115.3	13	Symthon III
QAMTAZ fourth	1	inversion	-25.9	-12.95	-115.3	11	step
QAMTAZ fifth	2	screw	-12.8	-12.8	-115.3	11	Symthon II
UJIRIO first	2	screw	-64.1	-64.1	-257.5	25	2 x H-bond assists
UJIRIO second	2	translation	-63.6	-63.6	-257.5	25	H-bond assist ++
UJIRIO third	1	inversion	-82.1	-41.05	-257.5	16	Symthon III
UJIRIO05 first	2	translation	-66.6	-66.6	-253.65	26	H-bond assist ++
UJIRIO05 second	2	glide	-64.4	-64.4	-253.65	25	2 x H-bond assists
UJIRIO05 third	1	inversion	-67.3	-33.65	-253.65	13	Symthon III offset

Prior to ranking, the interaction energies were expressed per mole of molecules. Lattice Energies were calculated following the procedure outlined in the Crystal Explorer on-line manual.[16]