Full wwPDB X-ray Structure Validation Report (i)

Jun 26, 2024 – 05:35 AM EDT

PDB ID	:	6ZGV
Title	:	Structure of human galactokinase 1 bound with 2-(4-chlorophenyl)-N-(pyrimidin-2-yl)acetamide
Authors	:	Mackinnon, S.R.; Bezerra, G.A.; Zhang, M.; Foster, W.; Krojer, T.; Brandao-Neto, J.; Douangamath, A.; Arrowsmith, C.; Edwards, A.; Bountra, C.; Brennan, P.; Lai, K.; Yue, W.W.
Deposited on	:	2020-06-20
Resolution	:	2.30 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

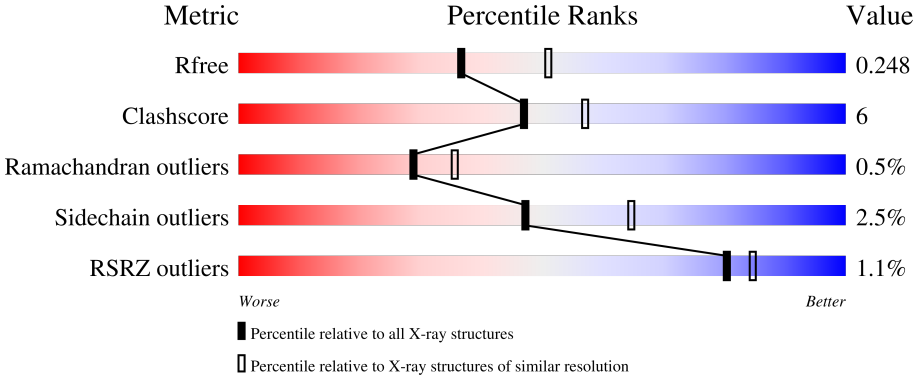
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	399	.% 78% 18% • •
1	B	399	2% 84% 13% • •
1	D	399	2% 79% 13% • 7%
1	E	399	87% 10% • •

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 11450 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Galactokinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	388	2685	1687	474	508	16	0	0	0
1	B	387	2746	1727	484	520	15	0	0	0
1	E	391	2851	1788	507	540	16	0	0	0
1	D	372	2600	1631	462	494	13	0	0	0

There are 40 discrepancies between the modelled and reference sequences:

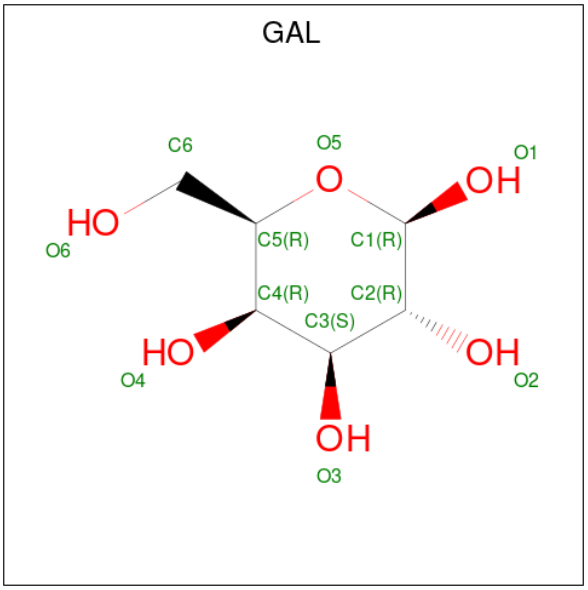
Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	MET	-	initiating methionine	UNP P51570
A	-5	ALA	-	expression tag	UNP P51570
A	-4	HIS	-	expression tag	UNP P51570
A	-3	HIS	-	expression tag	UNP P51570
A	-2	HIS	-	expression tag	UNP P51570
A	-1	HIS	-	expression tag	UNP P51570
A	0	HIS	-	expression tag	UNP P51570
A	1	HIS	-	expression tag	UNP P51570
A	252	ALA	LYS	engineered mutation	UNP P51570
A	253	ALA	GLU	engineered mutation	UNP P51570
B	-6	MET	-	initiating methionine	UNP P51570
B	-5	ALA	-	expression tag	UNP P51570
B	-4	HIS	-	expression tag	UNP P51570
B	-3	HIS	-	expression tag	UNP P51570
B	-2	HIS	-	expression tag	UNP P51570
B	-1	HIS	-	expression tag	UNP P51570
B	0	HIS	-	expression tag	UNP P51570
B	1	HIS	-	expression tag	UNP P51570
B	252	ALA	LYS	engineered mutation	UNP P51570
B	253	ALA	GLU	engineered mutation	UNP P51570
E	-6	MET	-	initiating methionine	UNP P51570

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	-5	ALA	-	expression tag	UNP P51570
E	-4	HIS	-	expression tag	UNP P51570
E	-3	HIS	-	expression tag	UNP P51570
E	-2	HIS	-	expression tag	UNP P51570
E	-1	HIS	-	expression tag	UNP P51570
E	0	HIS	-	expression tag	UNP P51570
E	1	HIS	-	expression tag	UNP P51570
E	252	ALA	LYS	engineered mutation	UNP P51570
E	253	ALA	GLU	engineered mutation	UNP P51570
D	-6	MET	-	initiating methionine	UNP P51570
D	-5	ALA	-	expression tag	UNP P51570
D	-4	HIS	-	expression tag	UNP P51570
D	-3	HIS	-	expression tag	UNP P51570
D	-2	HIS	-	expression tag	UNP P51570
D	-1	HIS	-	expression tag	UNP P51570
D	0	HIS	-	expression tag	UNP P51570
D	1	HIS	-	expression tag	UNP P51570
D	252	ALA	LYS	engineered mutation	UNP P51570
D	253	ALA	GLU	engineered mutation	UNP P51570

- Molecule 2 is beta-D-galactopyranose (three-letter code: GAL) (formula: $C_6H_{12}O_6$).



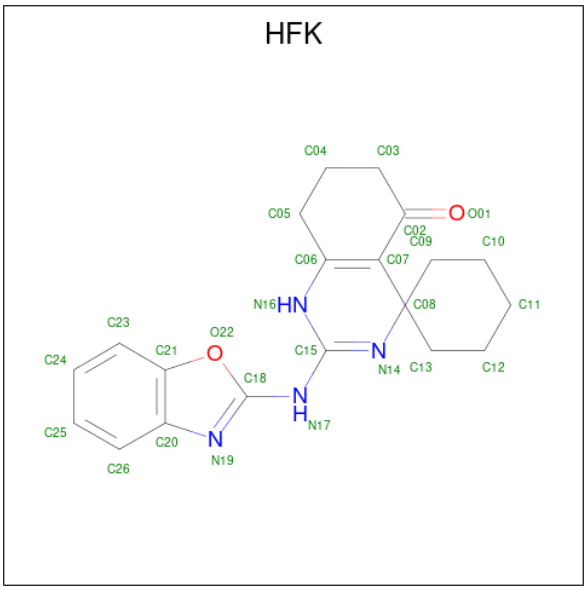
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	12	6	6	0	0
			Total	C	O		
2	B	1	12	6	6	0	0

Continued on next page...

Continued from previous page...

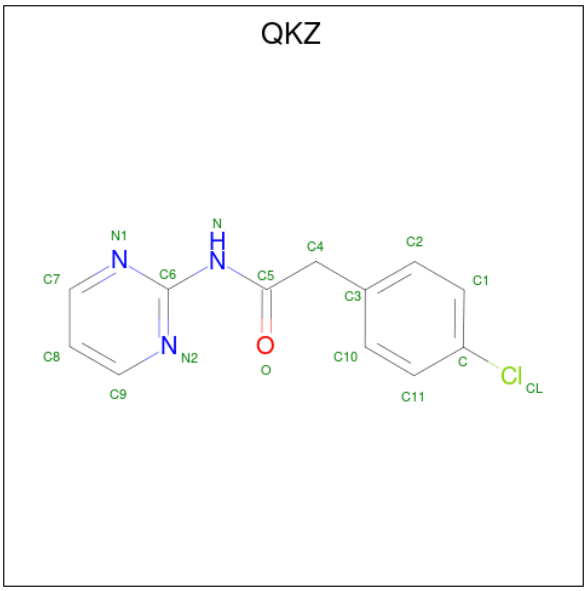
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	E	1	12	6	6	0	0

- Molecule 3 is 2-(1,3-benzoxazol-2-ylamino)spiro[1,6,7,8-tetrahydroquinazoline-4,1'-cyclohexane]-5-one (three-letter code: HFK) (formula: $C_{20}H_{22}N_4O_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	26	20	4	2	0	0
			Total	C	N	O		
3	B	1	26	20	4	2	0	0
			Total	C	N	O		
3	E	1	26	20	4	2	0	0

- Molecule 4 is 2-(4-chlorophenyl)- {N}-pyrimidin-2-yl-ethanamide (three-letter code: QKZ) (formula: $C_{12}H_{10}ClN_3O$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	Cl	N	O		
4	D	1	17	12	1	3	1	0	0

- Molecule 5 is water.

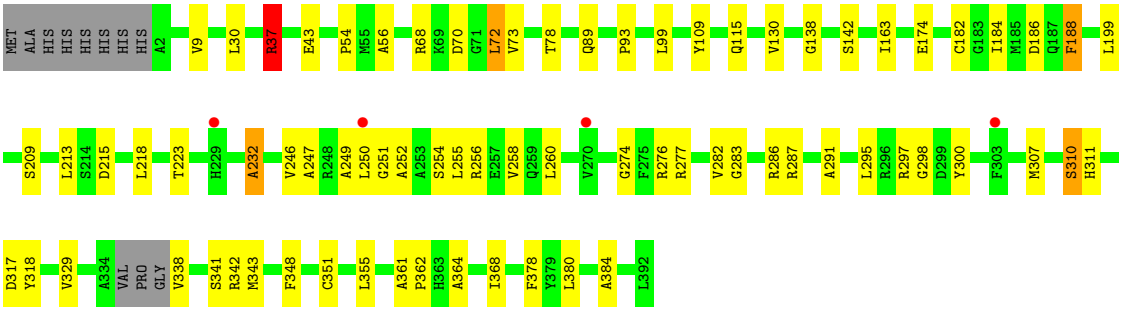
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	90	90	90	0	0
			Total	O		
5	B	82	82	82	0	0
			Total	O		
5	E	157	157	157	0	0
			Total	O		
5	D	108	108	108	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

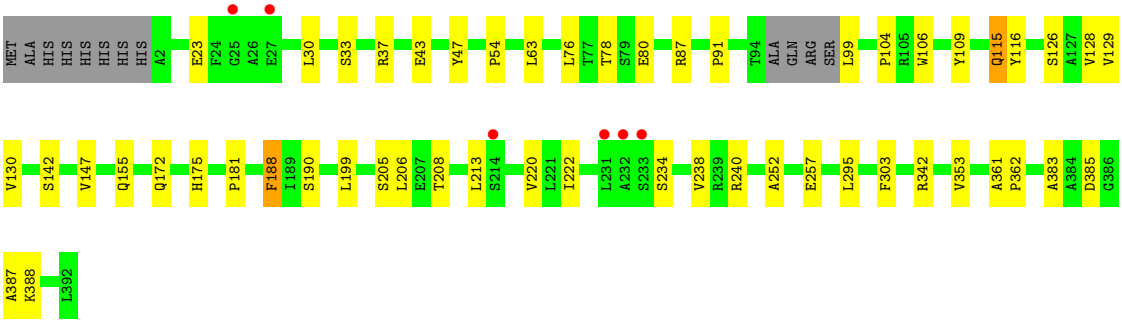
- Molecule 1: Galactokinase

Chain A:



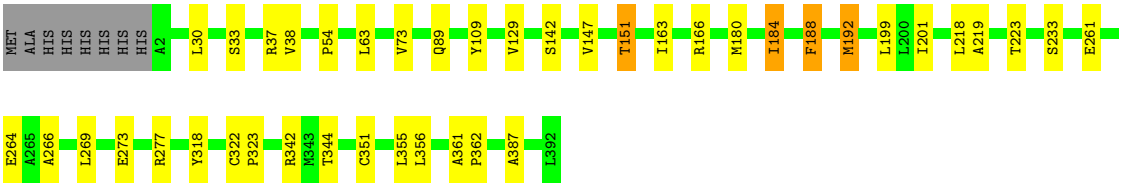
- Molecule 1: Galactokinase

Chain B:



- Molecule 1: Galactokinase

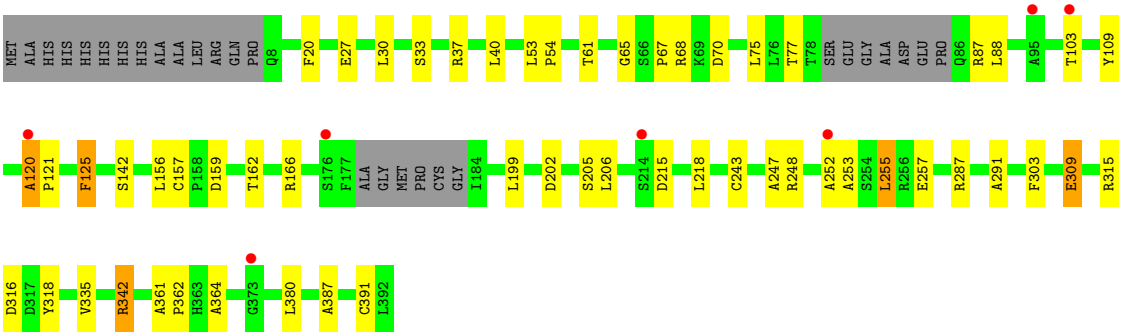
Chain E:



- Molecule 1: Galactokinase

Chain D:

2%
79%
13%
7%


MET ALA HIS HIS HIS HIS HIS ALA ALA LEU ARG GLN PRO Q8 F20 E27 L30 S33 R37 L40 L53 P54 T61 G65 S66 P67 R68 K69 D70 L75 L76 T77 T78 SER GLU GLY ALA ASP GLU PRO Q86 R87 L88 A95 T103 Y109
A120 P121 F125 S142 L156 C157 P158 D159 T162 R166 S176 F177 ALA GLY MET PRO CYS GLY I184 L199 D202 S205 L206 S214 D215 L218 C243 A247 R248 A252 A253 S254 L255 R256 E257 R287 A291 F303 E309 R315
D316 D317 Y318 V335 R342 A361 P362 H363 A364 G373 L380 A387 C391 L392

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	73.56Å 114.46Å 121.01Å 90.00° 100.39° 90.00°	Depositor
Resolution (Å)	72.46 – 2.30 72.35 – 2.14	Depositor EDS
% Data completeness (in resolution range)	99.1 (72.46-2.30) 99.1 (72.35-2.14)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.90 (at 2.14Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.207 , 0.246 0.211 , 0.248	Depositor DCC
R_{free} test set	5314 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å^2)	37.7	Xtriage
Anisotropy	0.576	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 36.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	11450	wwPDB-VP
Average B, all atoms (Å^2)	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.78% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: HFK, GAL, QKZ

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\|>5$	RMSZ	$\#\|Z\|>5$
1	A	0.87	5/2733 (0.2%)	0.98	6/3736 (0.2%)
1	B	0.77	0/2799	0.92	0/3821
1	D	0.79	2/2644 (0.1%)	0.94	1/3607 (0.0%)
1	E	0.79	0/2904	0.92	1/3953 (0.0%)
All	All	0.81	7/11080 (0.1%)	0.94	8/15117 (0.1%)

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	282	VAL	C-N	7.36	1.46	1.33
1	D	125	PHE	C-O	5.93	1.34	1.23
1	A	286	ARG	C-O	5.73	1.34	1.23
1	A	298	GLY	C-O	5.66	1.32	1.23
1	A	282	VAL	C-O	5.40	1.33	1.23
1	A	310	SER	C-O	5.30	1.33	1.23
1	D	309	GLU	CD-OE1	5.06	1.31	1.25

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	37	ARG	NE-CZ-NH2	-8.10	116.25	120.30
1	A	37	ARG	CG-CD-NE	-6.51	98.12	111.80
1	A	37	ARG	NE-CZ-NH1	6.28	123.44	120.30
1	E	342	ARG	NE-CZ-NH2	-5.84	117.38	120.30
1	A	256	ARG	NE-CZ-NH2	-5.72	117.44	120.30
1	A	256	ARG	NE-CZ-NH1	5.26	122.93	120.30
1	A	297	ARG	NE-CZ-NH1	5.05	122.83	120.30
1	D	342	ARG	NE-CZ-NH1	5.01	122.81	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2685	0	2511	42	0
1	B	2746	0	2601	30	0
1	D	2600	0	2433	38	0
1	E	2851	0	2787	24	0
2	A	12	0	11	0	0
2	B	12	0	12	0	0
2	E	12	0	11	0	0
3	A	26	0	0	0	0
3	B	26	0	0	0	0
3	E	26	0	0	0	0
4	D	17	0	0	0	0
5	A	90	0	0	1	0
5	B	82	0	0	1	0
5	D	108	0	0	0	0
5	E	157	0	0	1	0
All	All	11450	0	10366	133	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 6.

All (133) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:87:ARG:O	1:D:88:LEU:HD23	1.65	0.96
1:A:274:GLY:O	1:A:276:ARG:O	1.97	0.83
1:D:248:ARG:HA	1:D:252:ALA:HB2	1.61	0.82
1:D:159:ASP:O	1:D:159:ASP:OD1	1.99	0.81
1:B:23:GLU:OE1	1:B:87:ARG:NH2	2.16	0.79
1:E:147:VAL:O	1:E:151:THR:HG23	1.86	0.75
1:A:311:HIS:HD2	1:A:343:MET:H	1.34	0.73
1:D:120:ALA:HB3	1:D:121:PRO:HD3	1.73	0.70
1:B:188:PHE:CE1	1:B:208:THR:HG21	2.26	0.70

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:315:ARG:NH1	1:D:316:ASP:OD1	2.23	0.70
1:D:248:ARG:CA	1:D:252:ALA:HB2	2.21	0.69
1:D:253:ALA:HB3	1:D:257:GLU:HG3	1.75	0.67
1:A:72:LEU:C	1:A:72:LEU:HD12	2.16	0.66
1:D:202:ASP:O	1:D:205:SER:O	2.15	0.65
1:E:361:ALA:HB3	1:E:362:PRO:HD3	1.80	0.64
1:A:213:LEU:CD1	1:A:295:LEU:HD11	2.28	0.63
1:A:54:PRO:HD2	1:A:199:LEU:O	1.99	0.62
1:B:99:LEU:CB	1:B:115:GLN:HE21	2.12	0.62
1:A:223:THR:O	1:A:351:CYS:HA	2.00	0.61
1:E:147:VAL:O	1:E:151:THR:CG2	2.47	0.61
1:A:78:THR:OG1	1:A:130:VAL:HG12	2.01	0.60
1:D:248:ARG:N	1:D:252:ALA:HB2	2.16	0.60
1:B:43:GLU:HB2	1:B:342:ARG:NH2	2.17	0.59
1:D:159:ASP:OD1	1:D:159:ASP:C	2.40	0.59
1:D:162:THR:O	1:D:166:ARG:HG3	2.03	0.58
1:E:163:ILE:HG21	1:E:192:MET:SD	2.42	0.58
1:A:247:ALA:HB1	1:A:252:ALA:O	2.03	0.58
1:A:218:LEU:HD11	1:A:355:LEU:HD11	1.85	0.58
1:D:247:ALA:HB1	1:D:252:ALA:HA	1.86	0.57
1:B:252:ALA:HB1	1:B:257:GLU:HB2	1.87	0.57
1:D:54:PRO:HD2	1:D:199:LEU:O	2.05	0.57
1:D:252:ALA:O	1:D:253:ALA:C	2.39	0.56
1:D:361:ALA:HB3	1:D:362:PRO:HD3	1.88	0.56
1:E:151:THR:HG21	1:E:387:ALA:HB1	1.87	0.55
1:E:273:GLU:O	1:E:277:ARG:HG2	2.07	0.55
1:E:218:LEU:HD11	1:E:355:LEU:HG	1.89	0.55
1:A:249:ALA:C	1:A:251:GLY:H	2.11	0.54
1:B:23:GLU:CD	1:B:87:ARG:NH2	2.61	0.54
1:B:30:LEU:HD21	1:B:155:GLN:HB3	1.90	0.53
1:E:261:GLU:OE1	5:E:501:HOH:O	2.19	0.52
1:B:99:LEU:CB	1:B:115:GLN:NE2	2.73	0.52
1:B:385:ASP:OD2	1:B:388:LYS:NZ	2.37	0.52
1:A:68:ARG:CZ	1:A:72:LEU:HD11	2.39	0.52
1:A:364:ALA:O	1:A:368:ILE:HG12	2.10	0.52
1:A:277:ARG:HA	1:A:318:TYR:HD1	1.76	0.51
1:D:67:PRO:HB3	1:D:156:LEU:HD22	1.92	0.51
1:A:368:ILE:HB	1:A:378:PHE:HZ	1.75	0.51
1:A:260:LEU:HD12	1:A:260:LEU:O	2.11	0.50
1:A:37:ARG:NH2	1:A:138:GLY:O	2.45	0.50
1:B:147:VAL:HG21	1:B:190:SER:HB3	1.94	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:361:ALA:HB3	1:D:362:PRO:CD	2.41	0.50
1:B:80:GLU:CB	5:B:549:HOH:O	2.59	0.50
1:D:215:ASP:CG	1:D:218:LEU:HD13	2.32	0.49
1:E:54:PRO:HD2	1:E:199:LEU:O	2.12	0.49
1:D:68:ARG:NH1	1:D:70:ASP:OD2	2.42	0.49
1:E:30:LEU:CD2	1:D:391:CYS:HB2	2.43	0.49
1:E:109:TYR:CE1	1:E:142:SER:HB2	2.47	0.48
1:B:33:SER:O	1:B:387:ALA:HA	2.13	0.48
1:A:277:ARG:HA	1:A:318:TYR:CD1	2.49	0.48
1:A:250:LEU:HD12	1:A:250:LEU:N	2.29	0.47
1:A:218:LEU:HD22	1:A:300:TYR:CE1	2.49	0.47
1:B:222:ILE:HD12	1:B:222:ILE:N	2.30	0.47
1:E:266:ALA:HB1	1:E:269:LEU:HD12	1.97	0.47
1:A:184:ILE:O	1:A:188:PHE:HB2	2.14	0.47
1:A:307:MET:O	1:A:310:SER:HB3	2.15	0.47
1:A:311:HIS:CD2	1:A:343:MET:H	2.23	0.47
1:A:109:TYR:CE1	1:A:142:SER:HB2	2.50	0.47
1:E:322:CYS:HB2	1:E:323:PRO:HD2	1.97	0.47
1:D:335:VAL:HG21	1:D:364:ALA:HA	1.97	0.47
1:B:109:TYR:CE1	1:B:142:SER:HB2	2.50	0.46
1:D:87:ARG:O	1:D:88:LEU:CD2	2.52	0.46
1:D:315:ARG:NH1	1:D:316:ASP:CG	2.69	0.46
1:E:356:LEU:HD23	1:E:356:LEU:N	2.31	0.45
1:E:38:VAL:HG23	1:E:344:THR:HG21	1.97	0.45
1:A:56:ALA:HB3	1:A:384:ALA:O	2.16	0.45
1:B:234:SER:O	1:B:238:VAL:HG23	2.17	0.44
1:A:213:LEU:HD13	1:A:295:LEU:HD21	1.98	0.44
1:D:243:CYS:HB3	1:D:255:LEU:HD22	1.99	0.44
1:A:361:ALA:HB1	1:A:380:LEU:HD11	1.98	0.44
1:B:175:HIS:CD2	1:B:181:PRO:HA	2.52	0.44
1:A:277:ARG:N	1:A:317:ASP:OD1	2.50	0.44
1:B:43:GLU:HB2	1:B:342:ARG:HH22	1.81	0.44
1:E:63:LEU:CD1	1:E:129:VAL:HG22	2.48	0.44
1:B:213:LEU:HD12	1:B:383:ALA:HB2	1.99	0.43
1:D:20:PHE:CE2	1:D:65:GLY:HA2	2.53	0.43
1:A:37:ARG:HD3	1:A:186:ASP:OD1	2.18	0.43
1:A:70:ASP:OD1	1:A:72:LEU:HG	2.18	0.43
1:B:47:TYR:CE1	1:B:240:ARG:HG2	2.52	0.43
1:D:121:PRO:HD2	1:D:157:CYS:SG	2.57	0.43
1:B:54:PRO:HD2	1:B:199:LEU:O	2.18	0.43
1:B:205:SER:O	1:B:206:LEU:HB2	2.18	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:247:ALA:C	1:D:252:ALA:HB2	2.39	0.43
1:B:361:ALA:N	1:B:362:PRO:CD	2.82	0.43
1:A:93:PRO:HB3	1:A:115:GLN:HE22	1.84	0.43
1:B:78:THR:OG1	1:B:130:VAL:HG12	2.18	0.43
1:D:68:ARG:HG2	1:D:125:PHE:N	2.34	0.43
1:D:315:ARG:NH1	1:D:316:ASP:OD2	2.52	0.43
1:D:33:SER:HA	1:D:61:THR:O	2.18	0.43
1:A:255:LEU:O	1:A:258:VAL:N	2.47	0.42
1:D:291:ALA:HB1	1:D:303:PHE:CE1	2.54	0.42
1:E:184:ILE:O	1:E:188:PHE:HB2	2.19	0.42
1:E:223:THR:O	1:E:351:CYS:HA	2.19	0.42
1:A:218:LEU:HD22	1:A:300:TYR:CZ	2.53	0.42
1:B:295:LEU:HD13	1:B:303:PHE:CD2	2.55	0.42
1:D:53:LEU:HD12	1:D:53:LEU:HA	1.88	0.42
1:A:43:GLU:HB2	1:A:342:ARG:NH2	2.34	0.42
1:E:163:ILE:HD13	1:E:166:ARG:HH11	1.83	0.42
1:A:246:VAL:O	1:A:250:LEU:HD13	2.20	0.42
1:B:63:LEU:CD1	1:B:129:VAL:HG22	2.50	0.42
1:D:109:TYR:CE1	1:D:142:SER:HB2	2.55	0.42
1:D:40:LEU:O	1:D:342:ARG:HD3	2.20	0.42
1:A:99:LEU:H	1:A:115:GLN:NE2	2.18	0.41
1:A:163:ILE:HG13	5:A:551:HOH:O	2.19	0.41
1:B:76:LEU:O	1:B:128:VAL:HA	2.20	0.41
1:E:199:LEU:HD23	1:E:201:ILE:HD11	2.01	0.41
1:D:202:ASP:O	1:D:206:LEU:HA	2.20	0.41
1:A:232:ALA:HA	1:A:348:PHE:CD2	2.54	0.41
1:A:174:GLU:HG2	1:A:182:CYS:SG	2.61	0.41
1:A:361:ALA:N	1:A:362:PRO:CD	2.83	0.41
1:B:116:TYR:CE2	1:B:172:GLN:HG2	2.56	0.41
1:D:380:LEU:HD23	1:D:380:LEU:HA	1.88	0.41
1:B:63:LEU:HD11	1:B:129:VAL:HG22	2.02	0.41
1:E:73:VAL:O	1:E:89:GLN:HA	2.20	0.41
1:D:287:ARG:HD3	1:D:309:GLU:HB3	2.03	0.41
1:D:33:SER:O	1:D:387:ALA:HA	2.21	0.41
1:A:287:ARG:O	1:A:291:ALA:HB2	2.21	0.41
1:B:104:PRO:HB2	1:B:106:TRP:CD1	2.55	0.41
1:E:33:SER:O	1:E:387:ALA:HA	2.21	0.40
1:A:43:GLU:HB2	1:A:342:ARG:HH22	1.87	0.40
1:A:73:VAL:O	1:A:89:GLN:HA	2.20	0.40
1:B:220:VAL:CG1	1:B:353:VAL:HG23	2.52	0.40
1:E:180:MET:HE1	1:E:233:SER:HA	2.04	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:219:ALA:O	1:E:356:LEU:HD23	2.22	0.40

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	384/399 (96%)	354 (92%)	26 (7%)	4 (1%)	15	17
1	B	383/399 (96%)	368 (96%)	13 (3%)	2 (0%)	29	35
1	D	366/399 (92%)	339 (93%)	25 (7%)	2 (0%)	29	35
1	E	389/399 (98%)	381 (98%)	8 (2%)	0	100	100
All	All	1522/1596 (95%)	1442 (95%)	72 (5%)	8 (0%)	29	35

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	103	THR
1	A	254	SER
1	A	283	GLY
1	B	115	GLN
1	A	232	ALA
1	B	91	PRO
1	D	120	ALA
1	A	215	ASP

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	247/315 (78%)	238 (96%)	9 (4%)	35	49
1	B	262/315 (83%)	259 (99%)	3 (1%)	73	86
1	D	241/315 (76%)	234 (97%)	7 (3%)	42	58
1	E	286/315 (91%)	279 (98%)	7 (2%)	49	66
All	All	1036/1260 (82%)	1010 (98%)	26 (2%)	47	65

All (26) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	9	VAL
1	A	30	LEU
1	A	37	ARG
1	A	72	LEU
1	A	188	PHE
1	A	209	SER
1	A	329	VAL
1	A	338	VAL
1	A	341	SER
1	B	37	ARG
1	B	126	SER
1	B	188	PHE
1	E	37	ARG
1	E	151	THR
1	E	184	ILE
1	E	188	PHE
1	E	192	MET
1	E	264	GLU
1	E	318	TYR
1	D	27	GLU
1	D	30	LEU
1	D	37	ARG
1	D	75	LEU
1	D	77	THR
1	D	255	LEU
1	D	318	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	115	GLN
1	A	311	HIS
1	B	86	GLN
1	B	115	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

7 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	HFK	B	402	-	26,30,30	2.86	6 (23%)	24,44,44	1.25	3 (12%)
4	QKZ	D	401	-	18,18,18	1.46	1 (5%)	23,23,23	1.09	1 (4%)
2	GAL	A	401	-	12,12,12	0.61	0	17,17,17	2.20	5 (29%)
3	HFK	E	402	-	26,30,30	2.39	7 (26%)	24,44,44	1.74	6 (25%)
3	HFK	A	402	-	26,30,30	2.78	6 (23%)	24,44,44	1.64	7 (29%)
2	GAL	E	401	-	12,12,12	1.07	0	17,17,17	2.81	7 (41%)
2	GAL	B	401	-	12,12,12	1.08	1 (8%)	17,17,17	1.72	5 (29%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral

centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	HFK	B	402	-	-	2/2/43/43	0/5/5/5
4	QKZ	D	401	-	-	0/8/8/8	0/2/2/2
2	GAL	A	401	-	-	2/2/22/22	0/1/1/1
3	HFK	E	402	-	-	2/2/43/43	0/5/5/5
3	HFK	A	402	-	-	2/2/43/43	0/5/5/5
2	GAL	E	401	-	-	2/2/22/22	0/1/1/1
2	GAL	B	401	-	-	1/2/22/22	0/1/1/1

All (21) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	402	HFK	C06-C07	10.22	1.50	1.37
3	A	402	HFK	C06-C07	10.18	1.50	1.37
3	E	402	HFK	C06-C07	6.58	1.45	1.37
3	A	402	HFK	C15-N14	5.76	1.45	1.30
3	B	402	HFK	C15-N14	5.51	1.44	1.30
3	E	402	HFK	C15-N14	5.39	1.44	1.30
3	B	402	HFK	C15-N17	4.67	1.45	1.36
4	D	401	QKZ	C6-N2	4.39	1.40	1.34
3	A	402	HFK	C15-N17	4.34	1.44	1.36
3	E	402	HFK	C15-N16	4.04	1.43	1.36
3	E	402	HFK	C15-N17	3.80	1.43	1.36
3	B	402	HFK	C15-N16	3.70	1.42	1.36
3	B	402	HFK	C05-C06	3.70	1.55	1.49
3	A	402	HFK	C18-N17	3.65	1.43	1.36
3	B	402	HFK	C18-N17	3.29	1.42	1.36
3	E	402	HFK	C05-C06	3.12	1.55	1.49
3	E	402	HFK	C08-N14	2.48	1.49	1.46
3	E	402	HFK	C18-N17	2.39	1.41	1.36
3	A	402	HFK	C05-C06	2.38	1.53	1.49
3	A	402	HFK	C15-N16	2.25	1.40	1.36
2	B	401	GAL	C4-C3	2.15	1.57	1.52

All (34) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	E	401	GAL	C1-O5-C5	-7.36	99.78	113.66
2	A	401	GAL	C1-O5-C5	-5.89	102.55	113.66

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	E	401	GAL	C1-C2-C3	-5.67	98.54	110.31
2	A	401	GAL	C1-C2-C3	-4.51	100.96	110.31
3	E	402	HFK	C09-C10-C11	3.87	117.35	111.37
2	E	401	GAL	O1-C1-C2	3.84	119.84	109.03
3	E	402	HFK	N17-C15-N14	3.79	124.10	117.81
2	B	401	GAL	C1-C2-C3	-3.34	103.39	110.31
2	E	401	GAL	C3-C4-C5	3.28	116.09	110.24
3	E	402	HFK	C05-C06-N16	3.21	120.28	115.34
3	A	402	HFK	C02-C07-C06	-3.20	116.13	119.17
3	A	402	HFK	N17-C15-N14	2.97	122.74	117.81
3	A	402	HFK	C04-C03-C02	-2.93	108.37	113.58
3	E	402	HFK	C13-C12-C11	2.92	115.89	111.37
2	A	401	GAL	O1-C1-C2	2.76	116.80	109.03
2	B	401	GAL	O5-C5-C6	2.74	113.25	106.44
2	B	401	GAL	C1-O5-C5	-2.63	108.70	113.66
3	A	402	HFK	C13-C12-C11	2.59	115.38	111.37
2	B	401	GAL	O2-C2-C1	2.57	115.12	109.16
3	A	402	HFK	C18-N17-C15	-2.54	120.84	126.49
3	B	402	HFK	C08-C07-C06	-2.48	118.95	121.75
3	A	402	HFK	C09-C10-C11	2.43	115.12	111.37
3	E	402	HFK	C15-N16-C06	2.38	123.52	120.67
3	E	402	HFK	C18-N17-C15	-2.35	121.27	126.49
2	A	401	GAL	O5-C1-C2	-2.34	106.11	110.28
2	E	401	GAL	O6-C6-C5	-2.34	103.28	111.29
4	D	401	QKZ	C2-C1-C	-2.29	116.82	119.24
2	E	401	GAL	O1-C1-O5	-2.17	103.87	110.38
3	A	402	HFK	C08-C07-C06	-2.12	119.35	121.75
2	A	401	GAL	O5-C5-C6	2.11	111.68	106.44
3	B	402	HFK	C02-C07-C06	-2.08	117.19	119.17
2	B	401	GAL	O6-C6-C5	-2.07	104.19	111.29
3	B	402	HFK	C05-C06-N16	2.06	118.52	115.34
2	E	401	GAL	O3-C3-C2	2.01	114.99	110.35

There are no chirality outliers.

All (11) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	E	401	GAL	O5-C5-C6-O6
2	A	401	GAL	O5-C5-C6-O6
2	A	401	GAL	C4-C5-C6-O6
2	B	401	GAL	O5-C5-C6-O6
2	E	401	GAL	C4-C5-C6-O6

Continued on next page...

Continued from previous page...

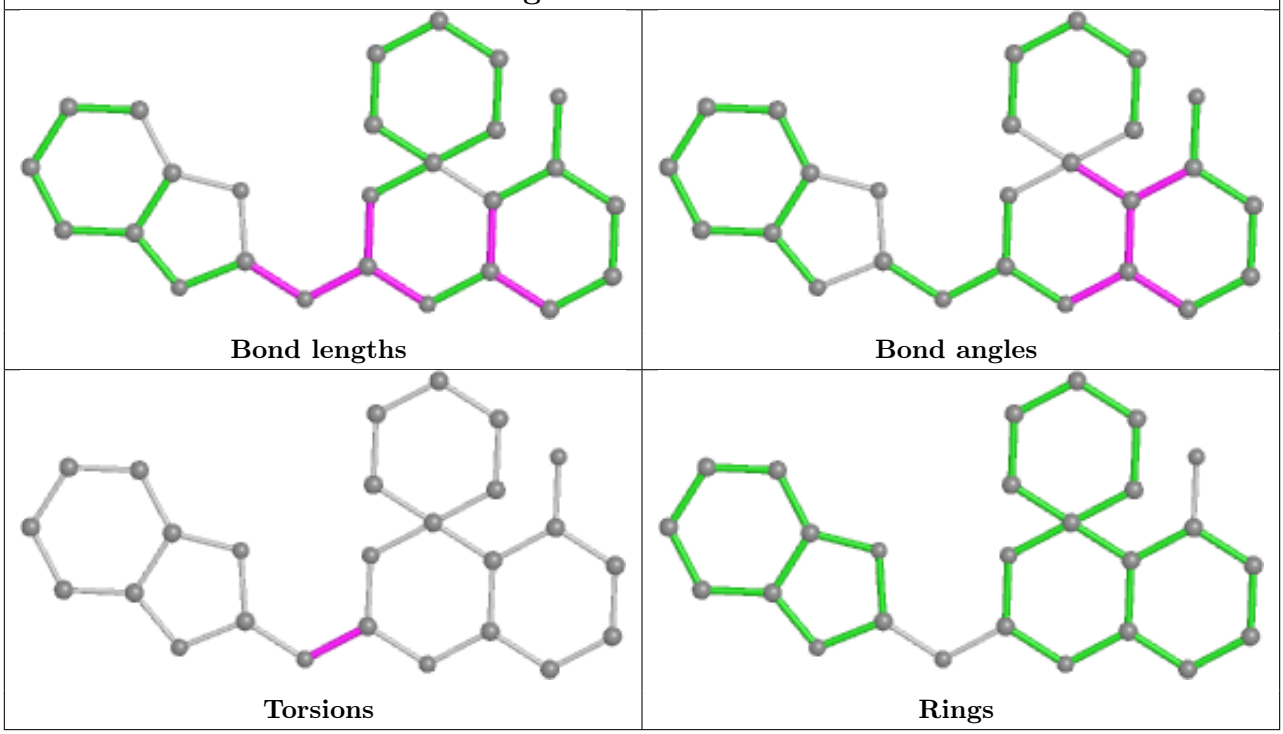
Mol	Chain	Res	Type	Atoms
3	A	402	HFK	N14-C15-N17-C18
3	A	402	HFK	N16-C15-N17-C18
3	B	402	HFK	N16-C15-N17-C18
3	E	402	HFK	N16-C15-N17-C18
3	B	402	HFK	N14-C15-N17-C18
3	E	402	HFK	N14-C15-N17-C18

There are no ring outliers.

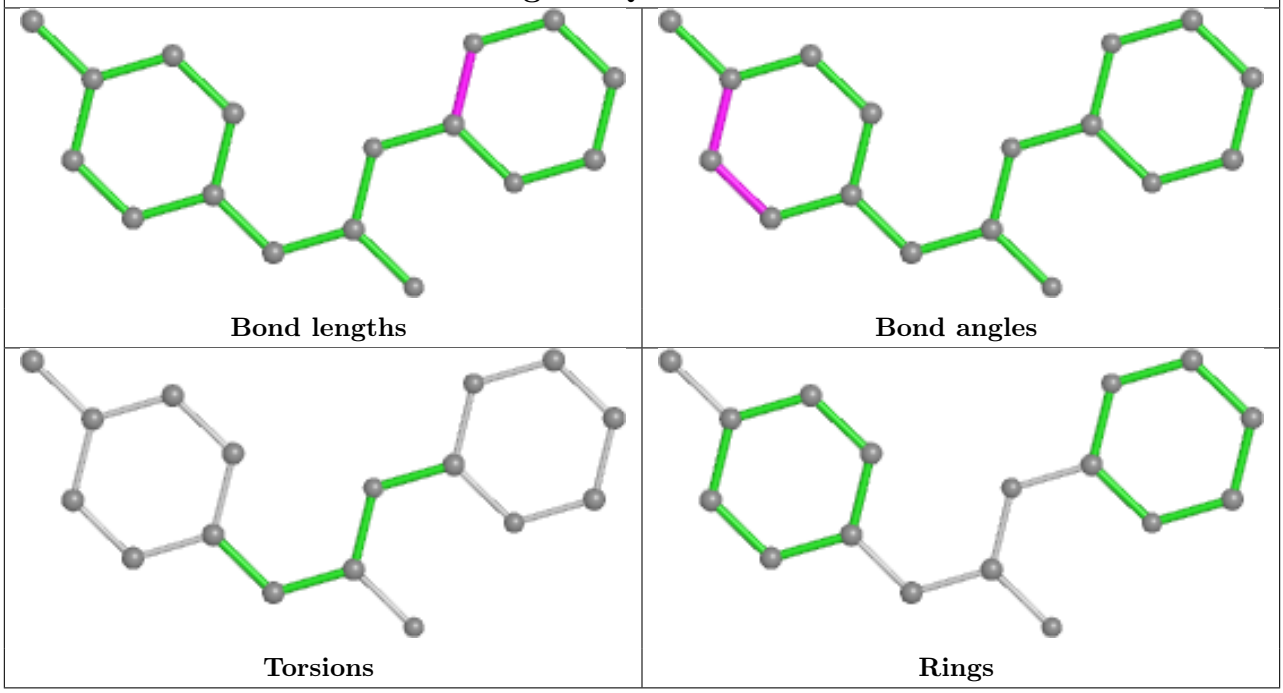
No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

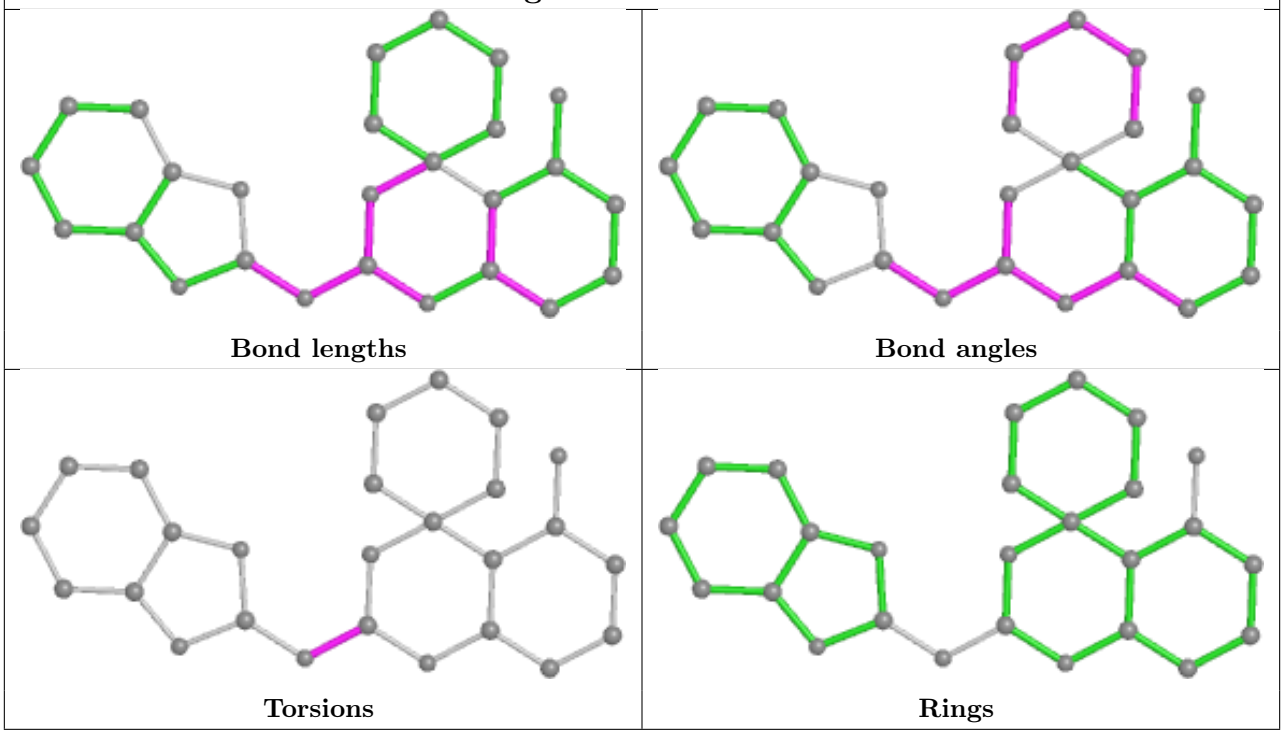
Ligand HFK B 402



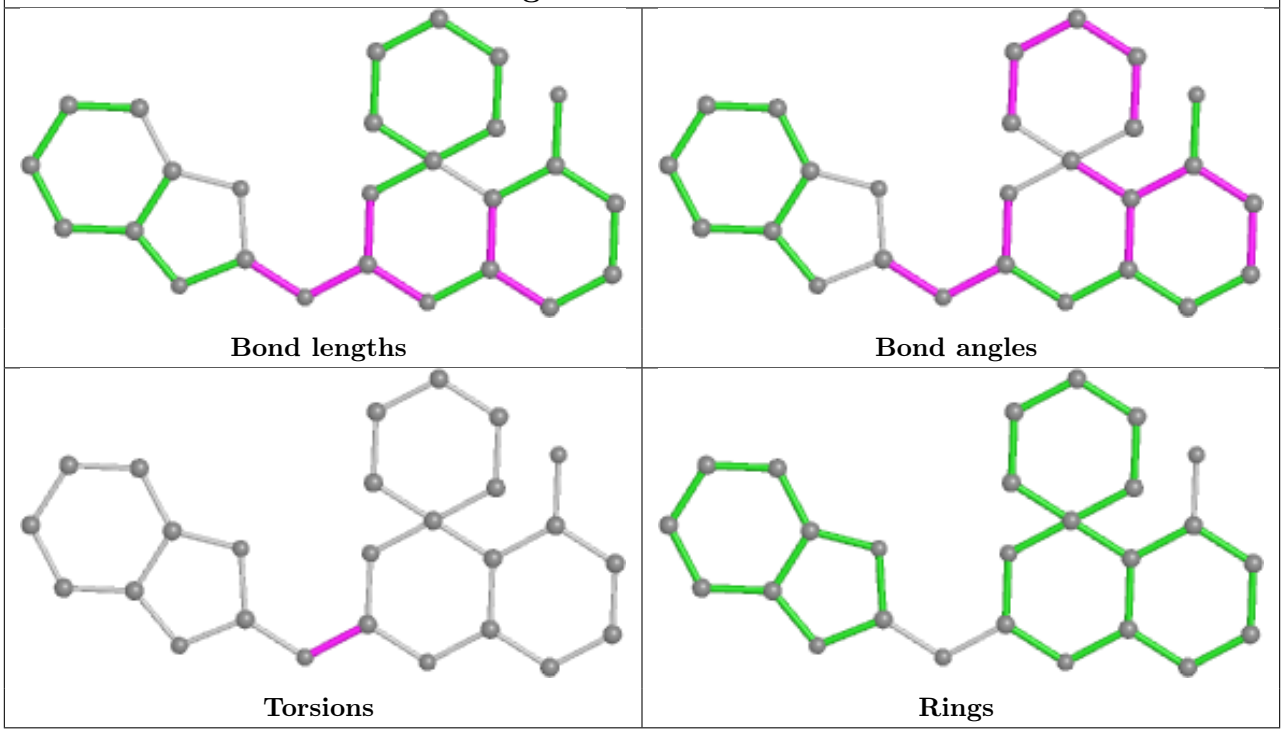
Ligand QKZ D 401



Ligand HFK E 402



Ligand HFK A 402



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	388/399 (97%)	-0.03	4 (1%) 82 86	26, 52, 101, 138	0
1	B	387/399 (96%)	-0.07	6 (1%) 72 77	28, 51, 89, 106	0
1	D	372/399 (93%)	0.00	7 (1%) 66 73	25, 53, 101, 138	0
1	E	391/399 (97%)	-0.25	0 100 100	24, 39, 63, 76	0
All	All	1538/1596 (96%)	-0.09	17 (1%) 80 85	24, 48, 95, 138	0

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	120	ALA	4.2
1	D	252	ALA	4.2
1	B	232	ALA	4.1
1	B	214	SER	3.2
1	A	303	PHE	3.1
1	D	95	ALA	2.9
1	D	214	SER	2.9
1	B	233	SER	2.8
1	A	270	VAL	2.7
1	D	373	GLY	2.7
1	A	229	HIS	2.6
1	D	103	THR	2.4
1	D	176	SER	2.3
1	B	231	LEU	2.3
1	B	25	GLY	2.3
1	A	250	LEU	2.2
1	B	27	GLU	2.2

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

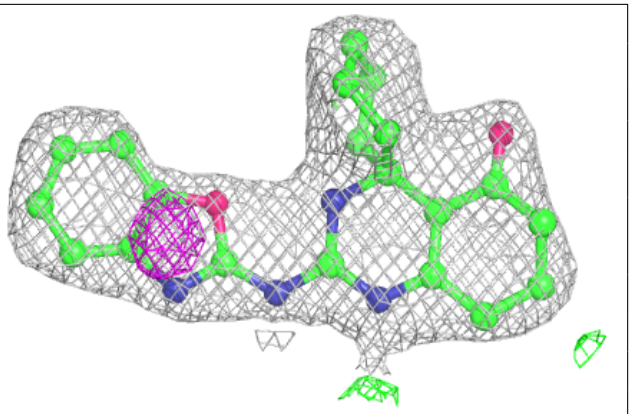
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

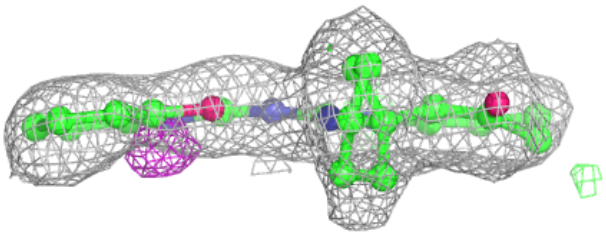
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	HFK	B	402	26/26	0.78	0.18	59,64,71,72	0
2	GAL	B	401	12/12	0.80	0.20	37,46,49,54	0
3	HFK	A	402	26/26	0.90	0.16	46,57,63,65	0
4	QKZ	D	401	17/17	0.92	0.14	42,48,69,70	0
3	HFK	E	402	26/26	0.93	0.12	36,40,46,48	0
2	GAL	A	401	12/12	0.96	0.12	37,42,46,46	0
2	GAL	E	401	12/12	0.96	0.12	23,27,32,39	0

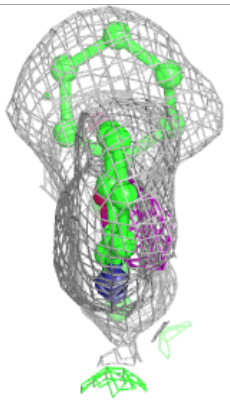
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around HFK B 402:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

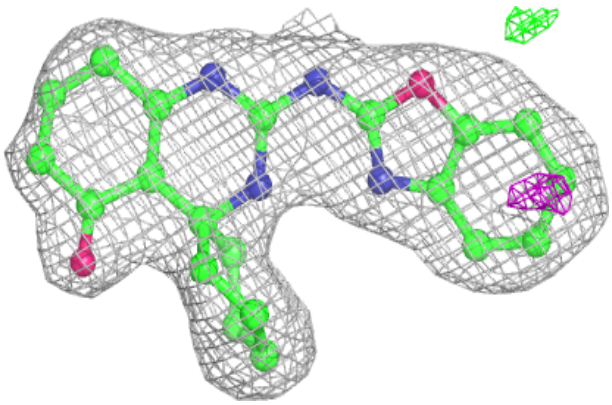


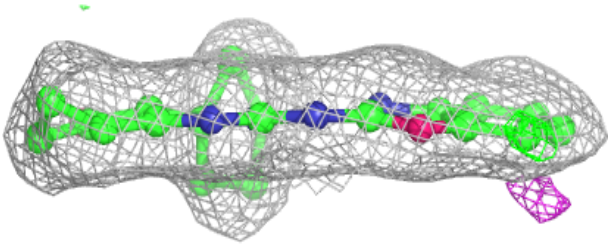


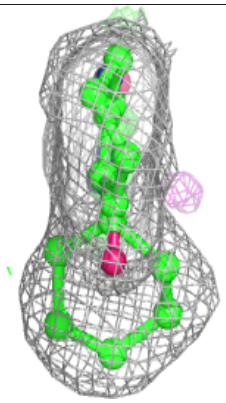


Electron density around HFK A 402:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

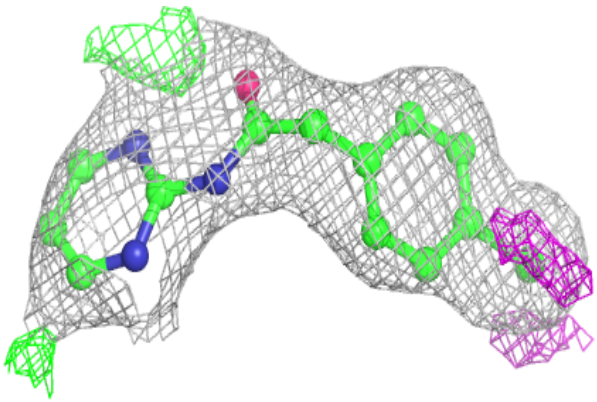


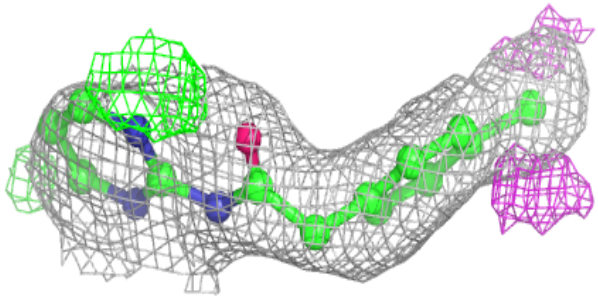


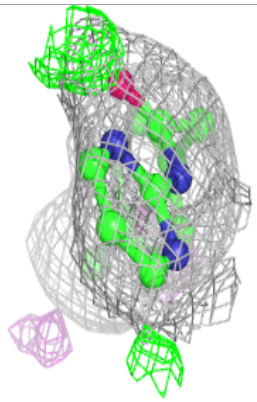


Electron density around QKZ D 401:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

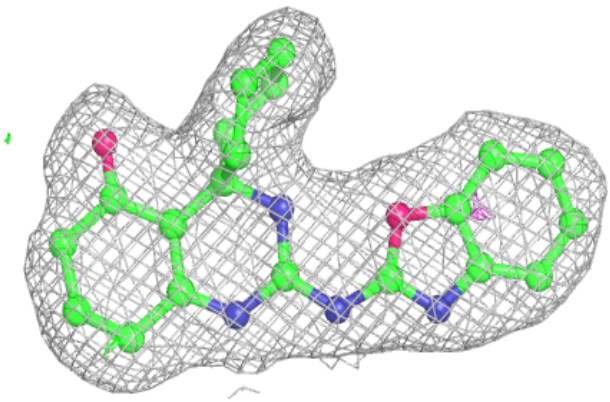


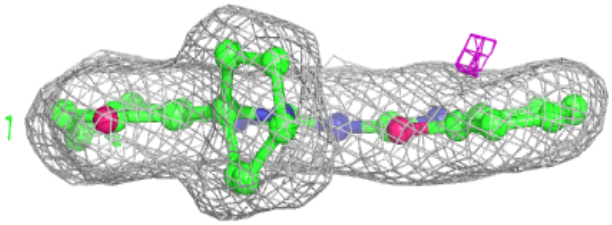


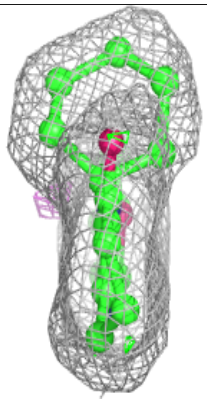


Electron density around HFK E 402:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)







6.5 Other polymers (i)

There are no such residues in this entry.