



# Full wwPDB X-ray Structure Validation Report

Jun 26, 2024 – 06:26 AM EDT

PDB ID : 6ZGX  
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 : 1.86 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](mailto: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  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](#) ①) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtrriage (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

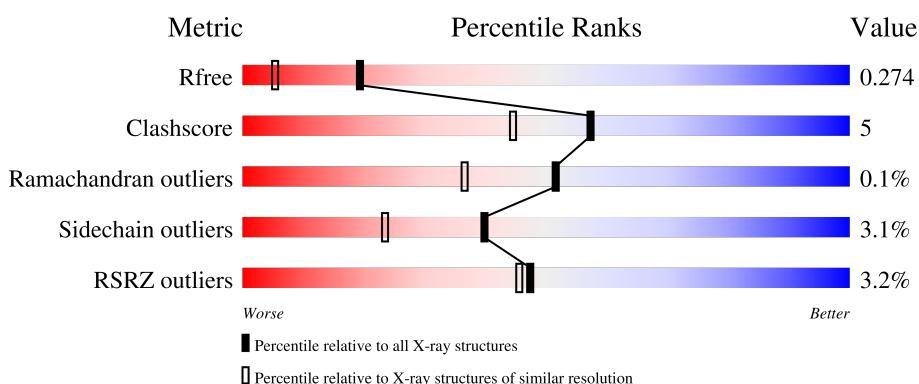
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.86 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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	7% (Poor fit) 85% (0 outliers) 12% (1 outlier) .. (2 outliers)
1	B	399	2% (Poor fit) 90% (0 outliers) 8% (1 outlier) .. (2 outliers)
1	D	399	4% (Poor fit) 79% (0 outliers) 9% (1 outlier) 10% (2 outliers)
1	E	399	91% (0 outliers) 7% (1 outlier) . (2 outliers)

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 11542 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	391	2785	1750	491	528	16	0	0	0
1	B	391	2793	1754	495	529	15	0	0	0
1	D	359	2517	1580	443	481	13	0	0	0
1	E	391	2859	1794	509	540	16	0	0	0

There are 40 discrepancies between the modelled and reference sequences:

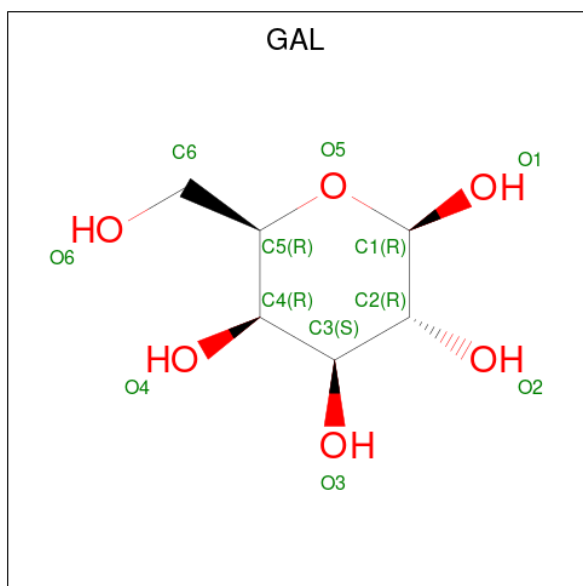
Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	MET	-	initiating methionine	UNP P51570
A	-5	ALA	-	expression tag	UNP P51570
A	-4	HIS	-	expression tag	UNP P51570
A	-3	HIS	-	expression tag	UNP P51570
A	-2	HIS	-	expression tag	UNP P51570
A	-1	HIS	-	expression tag	UNP P51570
A	0	HIS	-	expression tag	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
D	-6	MET	-	initiating methionine	UNP P51570

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
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
E	-6	MET	-	initiating methionine	UNP P51570
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

- Molecule 2 is beta-D-galactopyranose (three-letter code: GAL) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



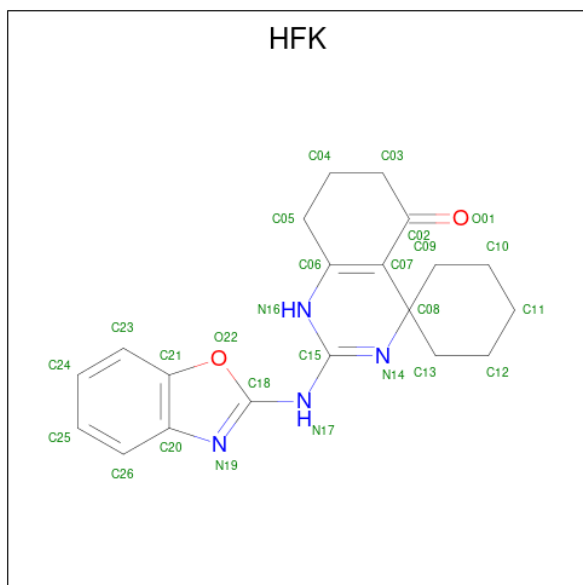
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf		
2	A	1	Total	C	O	0	0
			12	6	6		
2	B	1	Total	C	O	0	0
			12	6	6		

Continued on next page...

Continued from previous page...

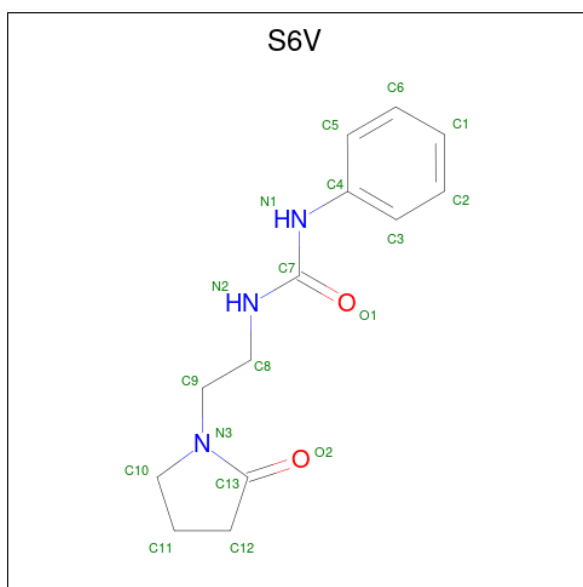
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	E	1	Total	C	O	0	0
			12	6	6		

- 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<sub>20</sub>H<sub>22</sub>N<sub>4</sub>O<sub>2</sub>).



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

- Molecule 4 is 1-[2-(2-oxidanilidenpyrrolidin-1-yl)ethyl]-3-phenyl-urea (three-letter code: S6V) (formula: C<sub>13</sub>H<sub>17</sub>N<sub>3</sub>O<sub>2</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	D	1	18	13	3	2	0	0

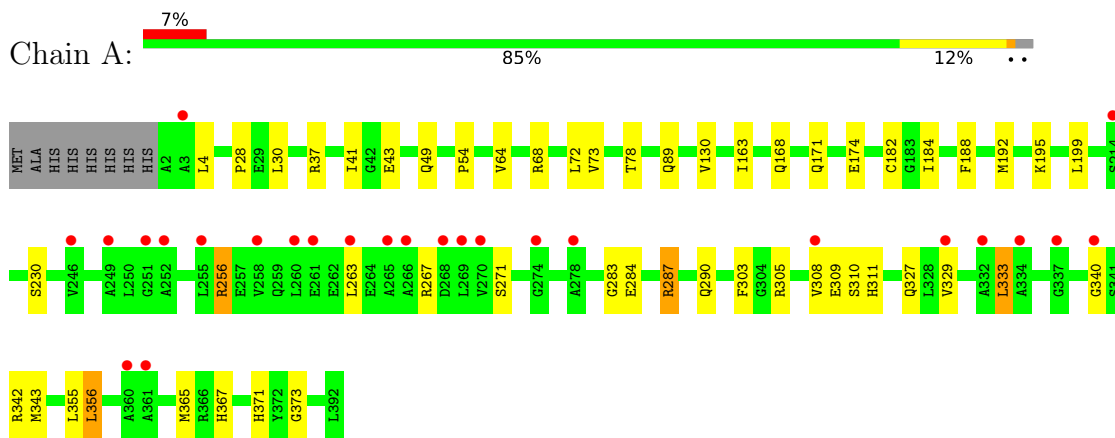
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	107	Total	O	0	0
			107	107		
5	B	102	Total	O	0	0
			102	102		
5	D	87	Total	O	0	0
			87	87		
5	E	160	Total	O	0	0
			160	160		

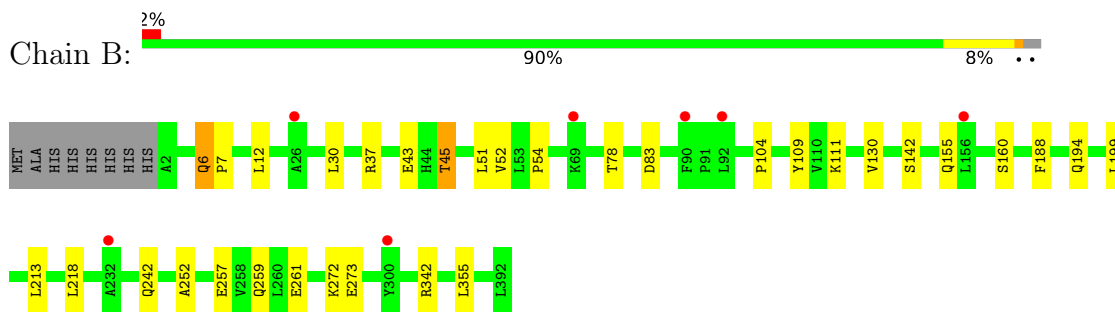
### 3 Residue-property plots [i](#)

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.

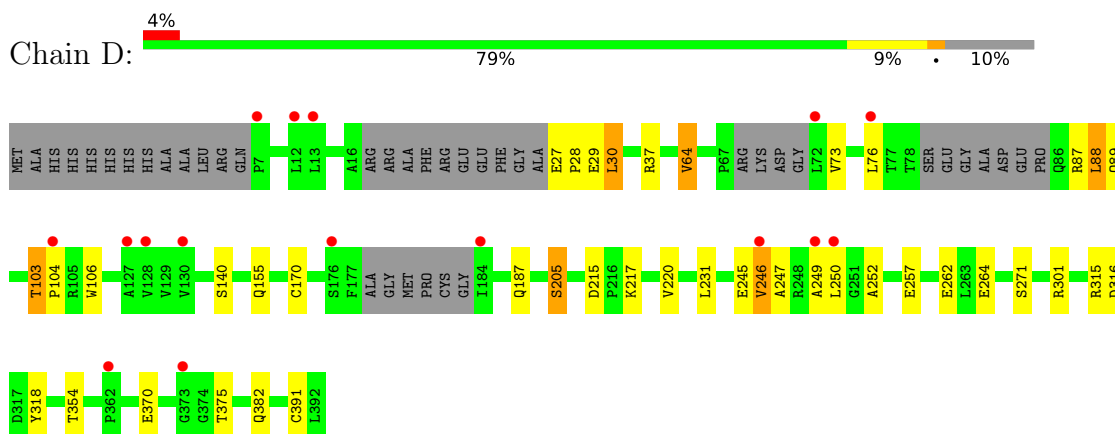
- Molecule 1: Galactokinase



- Molecule 1: Galactokinase


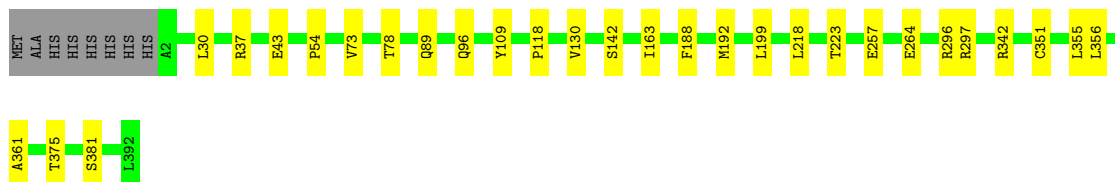


- Molecule 1: Galactokinase





## ● Molecule 1: Galactokinase

Chain E:  91% 7%

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.77Å 114.45Å 120.92Å 90.00° 100.55° 90.00°	Depositor
Resolution (Å)	82.59 – 1.86 82.45 – 1.86	Depositor EDS
% Data completeness (in resolution range)	99.4 (82.59-1.86) 99.4 (82.45-1.86)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.62 (at 1.86Å)	Xtrriage
Refinement program	REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.249 , 0.274 0.249 , 0.274	Depositor DCC
$R_{free}$ test set	8248 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	30.2	Xtrriage
Anisotropy	0.590	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 43.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	11542	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.77% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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, S6V, GAL

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	0.37	0/2839	0.74	1/3875 (0.0%)
1	B	0.36	0/2847	0.68	0/3884
1	D	0.42	0/2556	0.77	1/3486 (0.0%)
1	E	0.38	0/2913	0.70	0/3965
All	All	0.38	0/11155	0.72	2/15210 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	D	0	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	246	VAL	CB-CA-C	5.40	121.66	111.40
1	A	287	ARG	CG-CD-NE	-5.27	100.74	111.80

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	205	SER	Mainchain

## 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	2785	0	2662	42	0
1	B	2793	0	2671	24	0
1	D	2517	0	2389	27	0
1	E	2859	0	2801	18	0
2	A	12	0	12	0	0
2	B	12	0	12	0	0
2	E	12	0	12	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	18	0	0	1	0
5	A	107	0	0	2	0
5	B	102	0	0	1	0
5	D	87	0	0	3	0
5	E	160	0	0	4	0
All	All	11542	0	10559	107	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 5.

All (107) 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:315:ARG:HD2	1:D:316:ASP:OD1	1.67	0.95
1:B:45:THR:HG21	1:B:51:LEU:O	1.67	0.94
1:D:87:ARG:O	1:D:88:LEU:HD13	1.69	0.93
1:E:163:ILE:HG21	1:E:192:MET:SD	2.20	0.81
1:D:301:ARG:CB	5:D:576:HOH:O	2.28	0.81
1:B:12:LEU:HD22	1:B:130:VAL:HG23	1.66	0.78
1:B:45:THR:CG2	1:B:51:LEU:O	2.34	0.76
1:D:103:THR:CB	1:D:104:PRO:CD	2.64	0.76
1:E:257:GLU:HG3	5:E:561:HOH:O	1.84	0.76
1:A:356:LEU:O	1:A:356:LEU:HD13	1.87	0.75
1:A:4:LEU:HB2	1:A:365:MET:HE1	1.68	0.74
1:A:311:HIS:HD2	1:A:343:MET:H	1.33	0.74

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:287:ARG:NH2	1:A:310:SER:HA	2.04	0.72
1:B:12:LEU:HD22	1:B:130:VAL:CG2	2.19	0.72
1:A:263:LEU:CB	5:A:596:HOH:O	2.36	0.72
1:D:27:GLU:N	1:D:28:PRO:HD2	2.06	0.70
1:A:163:ILE:HD11	1:B:194:GLN:CD	2.12	0.70
1:D:103:THR:CB	1:D:104:PRO:HD3	2.21	0.70
1:A:4:LEU:HB2	1:A:365:MET:CE	2.21	0.70
1:A:284:GLU:HA	1:A:287:ARG:HD3	1.78	0.66
1:A:327:GLN:NE2	1:A:373:GLY:H	1.95	0.64
1:A:163:ILE:HG21	1:A:192:MET:SD	2.40	0.62
1:A:290:GLN:NE2	5:A:501:HOH:O	2.35	0.60
1:D:205:SER:O	1:D:205:SER:OG	2.05	0.60
1:D:30:LEU:HD21	1:D:155:GLN:HG3	1.85	0.59
1:A:184:ILE:HD12	1:A:184:ILE:C	2.24	0.58
1:B:43:GLU:HB2	1:B:342:ARG:NH2	2.19	0.58
1:B:30:LEU:HD21	1:B:155:GLN:HB3	1.86	0.57
1:D:27:GLU:N	1:D:28:PRO:CD	2.67	0.57
1:A:43:GLU:HB2	1:A:342:ARG:NH2	2.18	0.57
1:E:96:GLN:N	1:E:96:GLN:OE1	2.36	0.57
1:D:262:GLU:OE1	5:D:502:HOH:O	2.18	0.55
1:D:315:ARG:CD	1:D:316:ASP:OD1	2.49	0.55
1:B:218:LEU:HD11	1:B:355:LEU:HG	1.89	0.55
1:D:30:LEU:HD21	1:D:155:GLN:CB	2.37	0.55
1:D:382:GLN:OE1	5:D:501:HOH:O	2.17	0.54
1:A:174:GLU:HG2	1:A:182:CYS:SG	2.48	0.54
1:A:72:LEU:C	1:A:72:LEU:HD12	2.27	0.53
1:A:305:ARG:O	1:A:309:GLU:HB2	2.06	0.53
1:A:49:GLN:O	1:A:256:ARG:HG3	2.09	0.53
1:A:163:ILE:HD11	1:B:194:GLN:OE1	2.08	0.53
1:B:45:THR:HG22	1:B:52:VAL:HG12	1.90	0.52
1:A:356:LEU:HD13	1:A:356:LEU:C	2.29	0.52
1:A:308:VAL:HG21	1:A:333:LEU:HD21	1.91	0.52
1:E:43:GLU:HB2	1:E:342:ARG:NH2	2.24	0.52
1:B:6:GLN:N	1:B:7:PRO:HD3	2.24	0.52
1:B:83:ASP:HB3	1:B:104:PRO:HB2	1.92	0.52
1:A:78:THR:OG1	1:A:130:VAL:HG12	2.10	0.52
1:A:28:PRO:CB	1:A:64:VAL:HG13	2.40	0.51
1:D:391:CYS:HB2	1:E:30:LEU:CD2	2.41	0.51
1:A:43:GLU:HB2	1:A:342:ARG:HH22	1.77	0.49
1:D:28:PRO:CB	1:D:64:VAL:HG13	2.41	0.49
1:E:296:ARG:NH1	5:E:504:HOH:O	2.44	0.49

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:311:HIS:CD2	1:A:343:MET:H	2.21	0.49
1:A:283:GLY:O	1:A:287:ARG:CD	2.60	0.49
1:A:68:ARG:CZ	1:A:72:LEU:HD11	2.42	0.48
1:A:287:ARG:NH2	1:A:309:GLU:O	2.47	0.48
1:D:245:GLU:C	1:D:247:ALA:H	2.17	0.48
1:A:303:PHE:CD1	1:A:355:LEU:HD22	2.48	0.48
1:E:218:LEU:HD11	1:E:355:LEU:HG	1.94	0.48
1:B:78:THR:OG1	1:B:130:VAL:HG12	2.14	0.48
1:B:43:GLU:HB2	1:B:342:ARG:HH22	1.79	0.48
1:A:283:GLY:O	1:A:287:ARG:HD3	2.14	0.47
1:A:287:ARG:HH21	1:A:309:GLU:C	2.17	0.47
1:E:78:THR:OG1	1:E:130:VAL:HG12	2.14	0.47
1:D:30:LEU:HD21	1:D:155:GLN:CG	2.45	0.47
1:A:327:GLN:HE22	1:A:373:GLY:H	1.63	0.46
1:B:111:LYS:HE2	5:B:534:HOH:O	2.13	0.46
1:B:259:GLN:OE1	1:B:261:GLU:HB2	2.16	0.46
1:B:272:LYS:NZ	1:E:257:GLU:O	2.47	0.46
1:E:297:ARG:HD3	5:E:569:HOH:O	2.16	0.46
1:A:163:ILE:HG21	1:A:192:MET:CG	2.47	0.45
1:B:6:GLN:NE2	1:B:6:GLN:H	2.15	0.45
1:D:249:ALA:O	1:D:250:LEU:CB	2.64	0.45
1:A:311:HIS:CD2	1:A:329:VAL:HG21	2.52	0.45
1:A:356:LEU:O	1:A:356:LEU:CD1	2.62	0.45
1:E:356:LEU:HD11	1:E:361:ALA:HA	1.99	0.44
1:D:170:CYS:HB2	1:D:187:GLN:HG2	1.98	0.44
1:A:195:LYS:HE3	1:B:160:SER:O	2.18	0.43
1:A:73:VAL:O	1:A:89:GLN:HA	2.19	0.43
1:E:54:PRO:HD2	1:E:199:LEU:O	2.18	0.43
1:A:28:PRO:HG3	1:A:64:VAL:HG13	2.01	0.43
1:D:215:ASP:OD1	1:D:217:LYS:HG2	2.18	0.43
1:B:242:GLN:NE2	1:B:273:GLU:OE1	2.52	0.43
1:D:88:LEU:N	1:D:88:LEU:HD22	2.34	0.43
1:A:54:PRO:HD2	1:A:199:LEU:O	2.19	0.43
1:A:367:HIS:O	1:A:371:HIS:HD2	2.01	0.43
1:D:87:ARG:C	1:D:88:LEU:HD22	2.39	0.43
1:E:109:TYR:CE1	1:E:142:SER:HB2	2.53	0.42
1:E:73:VAL:O	1:E:89:GLN:HA	2.19	0.42
1:B:218:LEU:HD11	1:B:355:LEU:CG	2.49	0.42
1:E:43:GLU:HB2	1:E:342:ARG:HH22	1.84	0.42
4:D:401:S6V:O1	4:D:401:S6V:C5	2.67	0.42
1:B:109:TYR:CE1	1:B:142:SER:HB2	2.55	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:104:PRO:O	1:D:106:TRP:N	2.49	0.42
1:D:73:VAL:O	1:D:89:GLN:HA	2.20	0.41
1:B:252:ALA:HB1	1:B:257:GLU:HB2	2.01	0.41
1:B:54:PRO:HD2	1:B:199:LEU:O	2.21	0.41
1:A:308:VAL:HG23	1:A:340:GLY:HA2	2.03	0.41
1:E:118:PRO:HG3	5:E:576:HOH:O	2.20	0.41
1:D:220:VAL:HA	1:D:354:THR:O	2.20	0.41
1:E:223:THR:O	1:E:351:CYS:HA	2.21	0.40
1:A:4:LEU:HB2	1:A:365:MET:HE3	2.01	0.40
1:D:76:LEU:HD23	1:D:76:LEU:C	2.42	0.40
1:D:252:ALA:HB1	1:D:257:GLU:HB2	2.03	0.40
1:A:171:GLN:NE2	1:A:184:ILE:HG22	2.37	0.40
1:E:218:LEU:HD11	1:E:355:LEU:CG	2.52	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	389/399 (98%)	384 (99%)	5 (1%)	0	100	100
1	B	389/399 (98%)	382 (98%)	7 (2%)	0	100	100
1	D	349/399 (88%)	337 (97%)	10 (3%)	2 (1%)	25	12
1	E	389/399 (98%)	384 (99%)	5 (1%)	0	100	100
All	All	1516/1596 (95%)	1487 (98%)	27 (2%)	2 (0%)	51	36

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	103	THR
1	D	246	VAL

### 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	270/315 (86%)	259 (96%)	11 (4%)	30	13
1	B	270/315 (86%)	265 (98%)	5 (2%)	57	43
1	D	240/315 (76%)	228 (95%)	12 (5%)	24	9
1	E	288/315 (91%)	283 (98%)	5 (2%)	60	47
All	All	1068/1260 (85%)	1035 (97%)	33 (3%)	40	23

All (33) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	30	LEU
1	A	37	ARG
1	A	41	ILE
1	A	168	GLN
1	A	188	PHE
1	A	230	SER
1	A	256	ARG
1	A	267	ARG
1	A	271	SER
1	A	333	LEU
1	A	356	LEU
1	B	6	GLN
1	B	37	ARG
1	B	45	THR
1	B	188	PHE
1	B	213	LEU
1	D	29	GLU
1	D	30	LEU
1	D	37	ARG
1	D	64	VAL
1	D	88	LEU
1	D	140	SER
1	D	231	LEU
1	D	264	GLU

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	D	271	SER
1	D	318	TYR
1	D	370	GLU
1	D	375	THR
1	E	37	ARG
1	E	188	PHE
1	E	264	GLU
1	E	375	THR
1	E	381	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	6	GLN
1	A	168	GLN
1	A	171	GLN
1	A	226	ASN
1	A	290	GLN
1	A	311	HIS
1	A	327	GLN
1	A	371	HIS
1	B	6	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	2.56	7 (26%)	24,44,44	1.62	7 (29%)
3	HFK	A	402	-	26,30,30	2.81	7 (26%)	24,44,44	1.61	5 (20%)
2	GAL	B	401	-	12,12,12	0.66	0	17,17,17	1.33	2 (11%)
2	GAL	E	401	-	12,12,12	0.93	0	17,17,17	3.15	5 (29%)
2	GAL	A	401	-	12,12,12	0.70	0	17,17,17	2.61	4 (23%)
3	HFK	E	402	-	26,30,30	2.52	6 (23%)	24,44,44	1.66	6 (25%)
4	S6V	D	401	-	19,19,19	1.66	4 (21%)	24,24,24	1.18	3 (12%)

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
3	HFK	A	402	-	-	2/2/43/43	0/5/5/5
2	GAL	B	401	-	-	1/2/22/22	0/1/1/1
2	GAL	E	401	-	-	1/2/22/22	0/1/1/1
2	GAL	A	401	-	-	2/2/22/22	0/1/1/1
3	HFK	E	402	-	-	2/2/43/43	0/5/5/5
4	S6V	D	401	-	-	1/10/20/20	0/2/2/2

All (24) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	402	HFK	C06-C07	10.61	1.50	1.37
3	B	402	HFK	C06-C07	9.51	1.49	1.37
3	E	402	HFK	C06-C07	8.87	1.48	1.37
3	A	402	HFK	C15-N14	6.08	1.46	1.30
3	B	402	HFK	C15-N14	5.73	1.45	1.30
3	E	402	HFK	C15-N14	5.59	1.44	1.30
4	D	401	S6V	C13-N3	3.49	1.43	1.34

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	402	HFK	C15-N17	3.48	1.42	1.36
4	D	401	S6V	C5-C4	3.40	1.45	1.39
3	E	402	HFK	C15-N17	3.23	1.42	1.36
3	B	402	HFK	C15-N17	3.11	1.42	1.36
4	D	401	S6V	C8-N2	2.98	1.53	1.46
3	A	402	HFK	C05-C06	2.97	1.54	1.49
3	B	402	HFK	C05-C06	2.88	1.54	1.49
3	E	402	HFK	C05-C06	2.84	1.54	1.49
3	A	402	HFK	C08-N14	2.70	1.49	1.46
3	B	402	HFK	C03-C02	2.59	1.54	1.50
3	E	402	HFK	C13-C08	2.19	1.57	1.53
3	E	402	HFK	C15-N16	2.14	1.40	1.36
3	A	402	HFK	C18-N17	2.13	1.40	1.36
3	B	402	HFK	C15-N16	2.11	1.40	1.36
3	A	402	HFK	C15-N16	2.09	1.40	1.36
4	D	401	S6V	C6-C5	2.01	1.43	1.38
3	B	402	HFK	C18-N17	2.01	1.40	1.36

All (32) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	401	GAL	C1-O5-C5	-9.35	96.03	113.66
2	A	401	GAL	C1-O5-C5	-6.96	100.53	113.66
2	E	401	GAL	C1-C2-C3	-6.09	97.68	110.31
2	A	401	GAL	C1-C2-C3	-5.61	98.68	110.31
2	E	401	GAL	O1-C1-C2	4.98	123.06	109.03
2	A	401	GAL	O1-C1-C2	4.26	121.03	109.03
3	E	402	HFK	N17-C15-N14	4.13	124.67	117.81
3	A	402	HFK	N17-C15-N14	3.36	123.39	117.81
3	B	402	HFK	C02-C07-C06	-3.32	116.01	119.17
3	B	402	HFK	N17-C15-N14	3.32	123.32	117.81
3	B	402	HFK	C09-C10-C11	3.10	116.17	111.37
3	A	402	HFK	C18-N17-C15	-3.06	119.70	126.49
3	A	402	HFK	C13-C12-C11	2.94	115.93	111.37
2	A	401	GAL	C3-C4-C5	2.93	115.47	110.24
3	E	402	HFK	C02-C07-C06	-2.85	116.45	119.17
3	B	402	HFK	C18-N17-C15	-2.75	120.39	126.49
2	B	401	GAL	C1-C2-C3	-2.74	104.64	110.31
2	E	401	GAL	C6-C5-C4	2.69	119.31	113.00
4	D	401	S6V	C8-N2-C7	-2.66	116.42	121.75
3	B	402	HFK	O01-C02-C07	-2.57	118.00	121.54
3	E	402	HFK	C04-C03-C02	2.50	118.02	113.58

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	401	S6V	C4-N1-C7	-2.40	121.70	126.61
2	B	401	GAL	O5-C1-C2	-2.35	106.09	110.28
3	B	402	HFK	C08-C07-C06	-2.27	119.19	121.75
3	E	402	HFK	O01-C02-C07	-2.24	118.45	121.54
3	A	402	HFK	C02-C07-C06	-2.23	117.05	119.17
3	B	402	HFK	C05-C06-N16	2.14	118.63	115.34
2	E	401	GAL	C3-C4-C5	2.10	113.99	110.24
3	E	402	HFK	C03-C02-C07	2.09	121.26	117.36
3	E	402	HFK	C18-N17-C15	-2.06	121.92	126.49
4	D	401	S6V	C3-C4-C5	2.04	121.82	119.03
3	A	402	HFK	O01-C02-C07	-2.00	118.78	121.54

There are no chirality outliers.

All (11) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	401	GAL	O5-C5-C6-O6
2	A	401	GAL	C4-C5-C6-O6
2	E	401	GAL	O5-C5-C6-O6
2	B	401	GAL	O5-C5-C6-O6
4	D	401	S6V	C8-C9-N3-C10
3	A	402	HFK	N14-C15-N17-C18
3	A	402	HFK	N16-C15-N17-C18
3	B	402	HFK	N14-C15-N17-C18
3	B	402	HFK	N16-C15-N17-C18
3	E	402	HFK	N16-C15-N17-C18
3	E	402	HFK	N14-C15-N17-C18

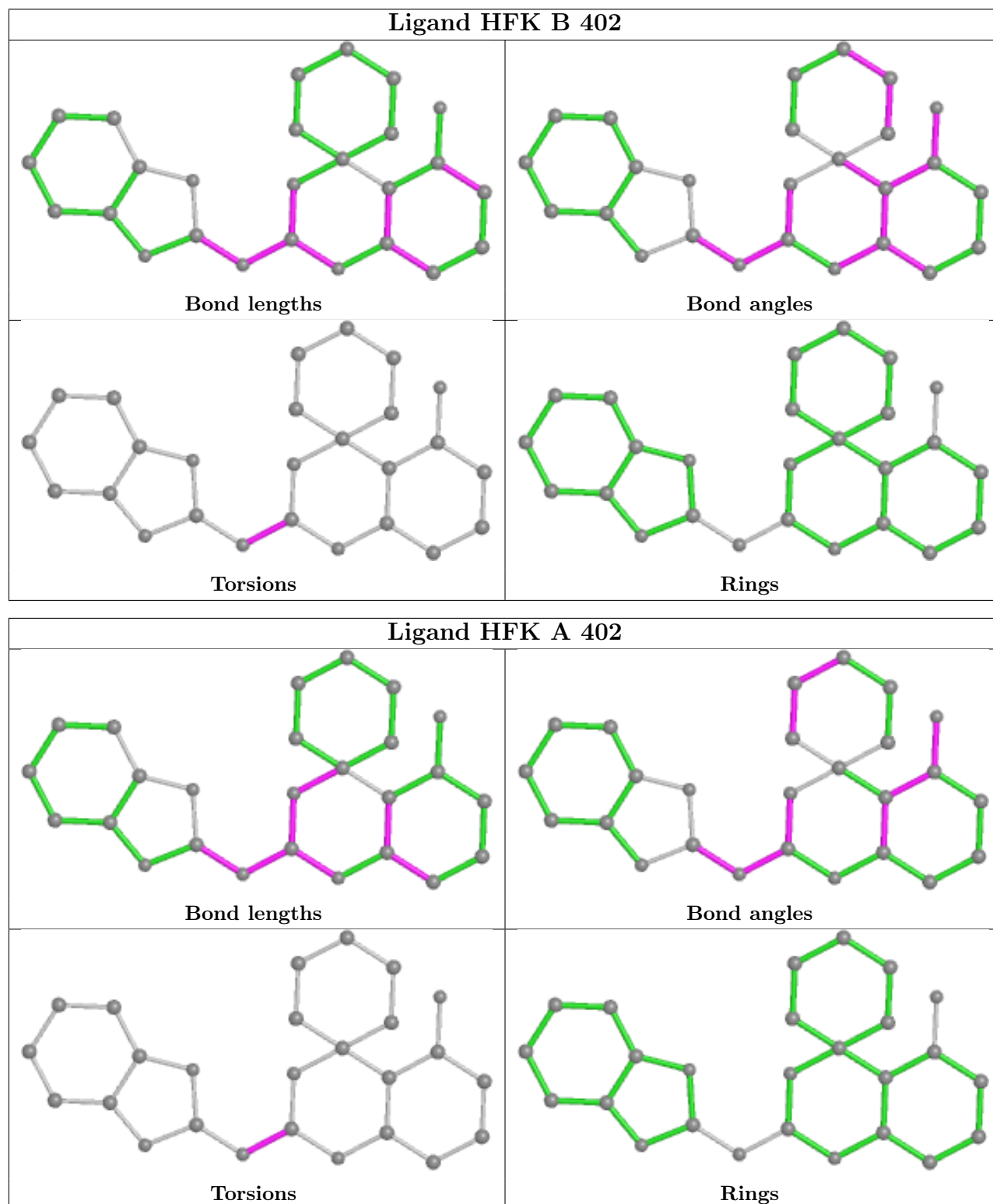
There are no ring outliers.

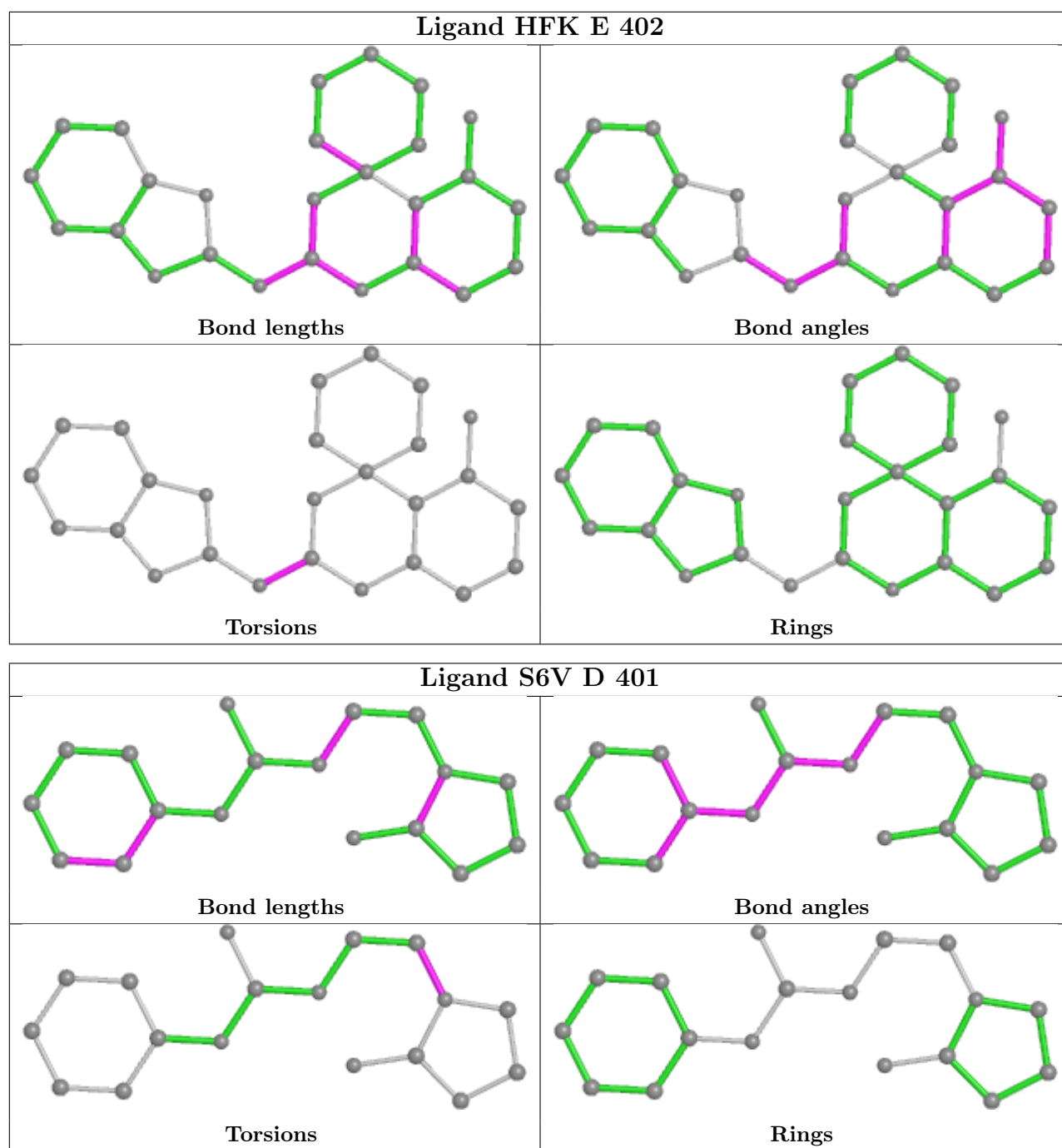
1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	D	401	S6V	1	0

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 than 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.





## 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

### 6.1 Protein, DNA and RNA chains

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	391/399 (97%)	0.37	26 (6%) 18 17	21, 41, 78, 101	0
1	B	391/399 (97%)	0.03	7 (1%) 68 68	21, 39, 70, 100	0
1	D	359/399 (89%)	0.33	16 (4%) 33 32	19, 42, 77, 113	0
1	E	391/399 (97%)	-0.22	0 100 100	17, 31, 52, 69	0
All	All	1532/1596 (95%)	0.12	49 (3%) 47 45	17, 38, 73, 113	0

All (49) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	250	LEU	7.0
1	A	263	LEU	6.0
1	A	249	ALA	5.5
1	A	337	GLY	4.3
1	D	373	GLY	4.0
1	A	268	ASP	3.7
1	B	232	ALA	3.7
1	D	128	VAL	3.7
1	A	266	ALA	3.6
1	A	258	VAL	3.5
1	A	260	LEU	3.3
1	A	308	VAL	3.3
1	A	278	ALA	3.3
1	D	249	ALA	3.2
1	A	274	GLY	3.2
1	A	334	ALA	3.2
1	D	76	LEU	3.1
1	A	246	VAL	3.1
1	A	270	VAL	3.0
1	A	251	GLY	2.7
1	A	265	ALA	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	26	ALA	2.6
1	D	130	VAL	2.6
1	D	104	PRO	2.5
1	D	13	LEU	2.5
1	D	7	PRO	2.5
1	B	92	LEU	2.4
1	A	214	SER	2.4
1	A	361	ALA	2.4
1	D	12	LEU	2.3
1	B	90	PHE	2.3
1	A	269	LEU	2.3
1	A	3	ALA	2.3
1	A	340	GLY	2.3
1	D	246	VAL	2.3
1	D	176	SER	2.3
1	D	127	ALA	2.2
1	A	329	VAL	2.2
1	B	69	LYS	2.2
1	B	156	LEU	2.1
1	A	255	LEU	2.1
1	D	72	LEU	2.1
1	A	252	ALA	2.1
1	D	184	ILE	2.1
1	A	332	ALA	2.1
1	A	360	ALA	2.0
1	B	300	TYR	2.0
1	D	362	PRO	2.0
1	A	261	GLU	2.0

## 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](#)

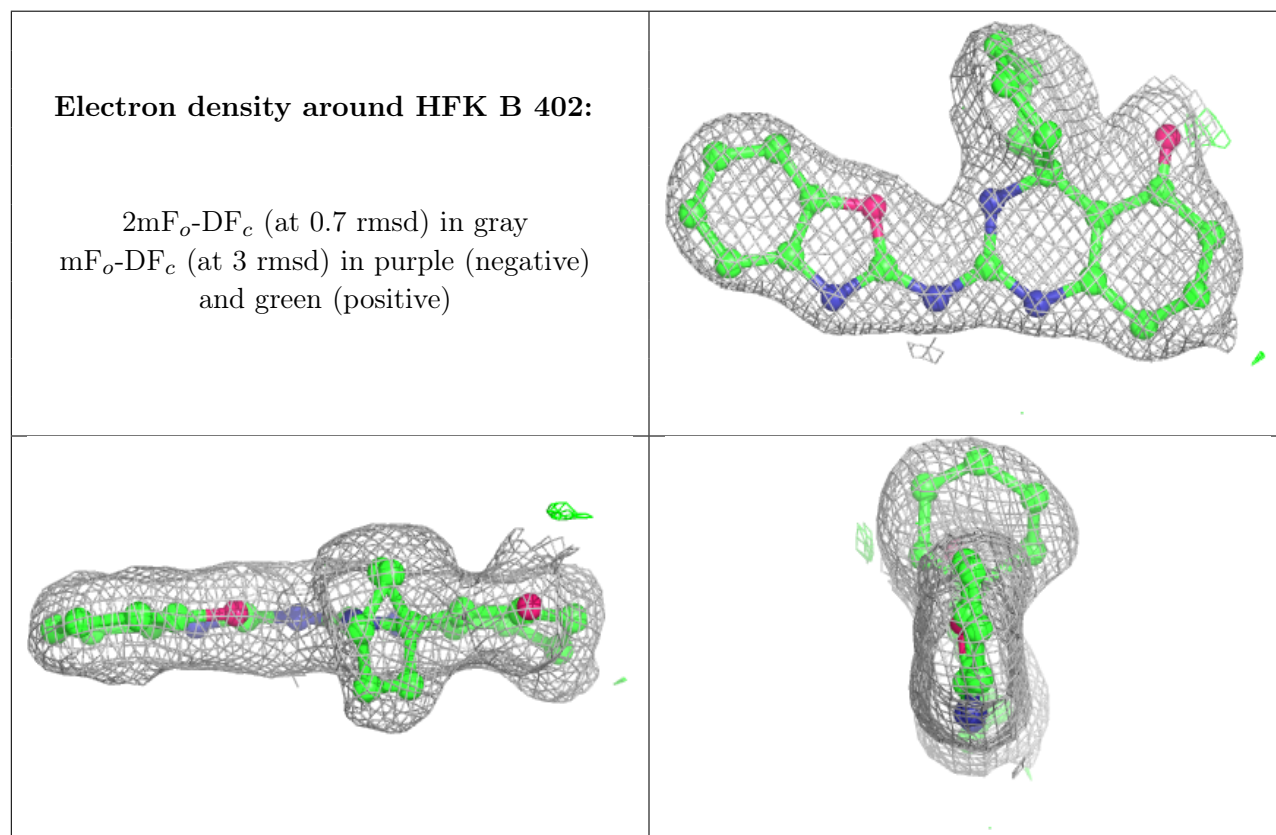
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum,

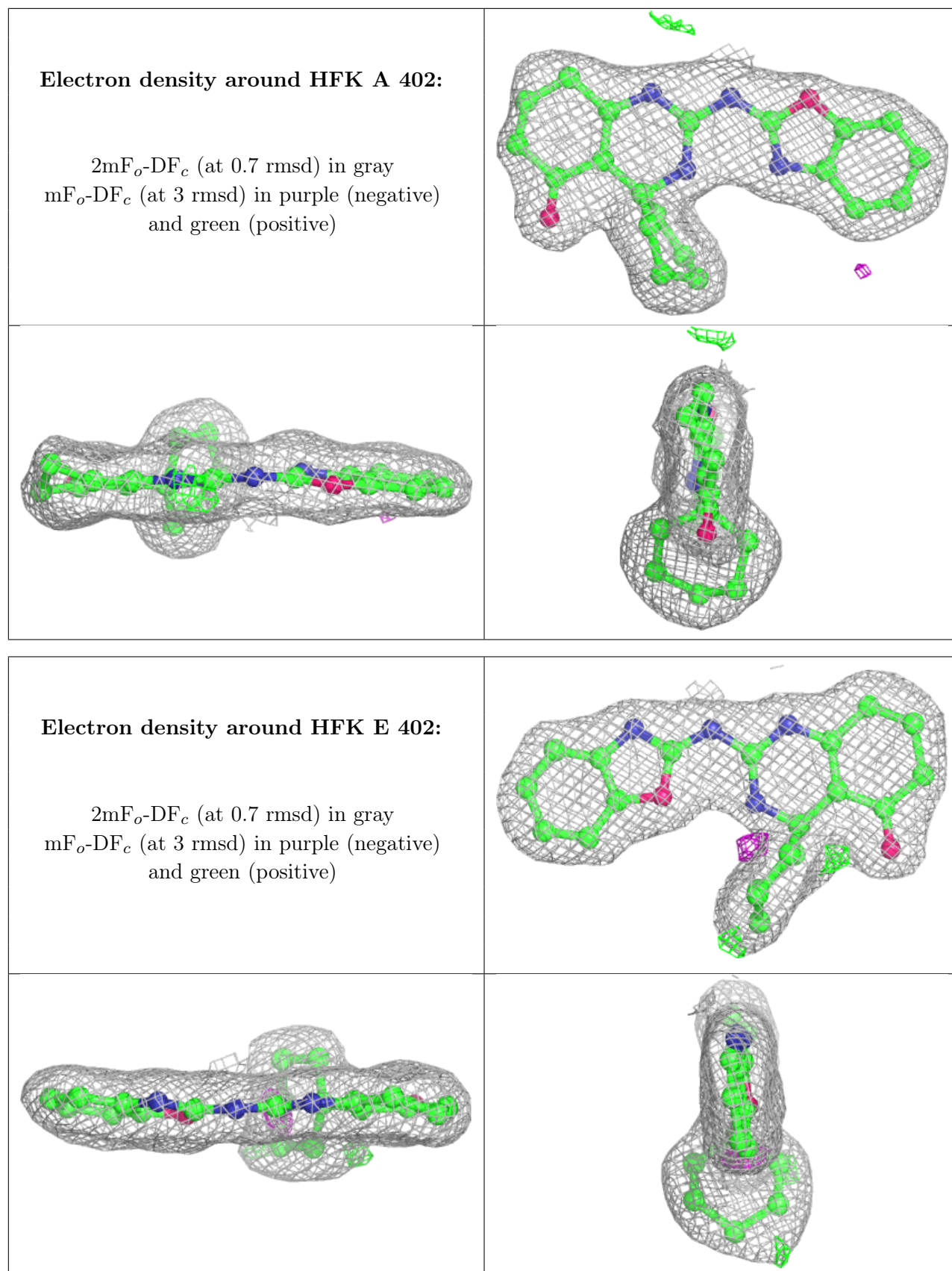


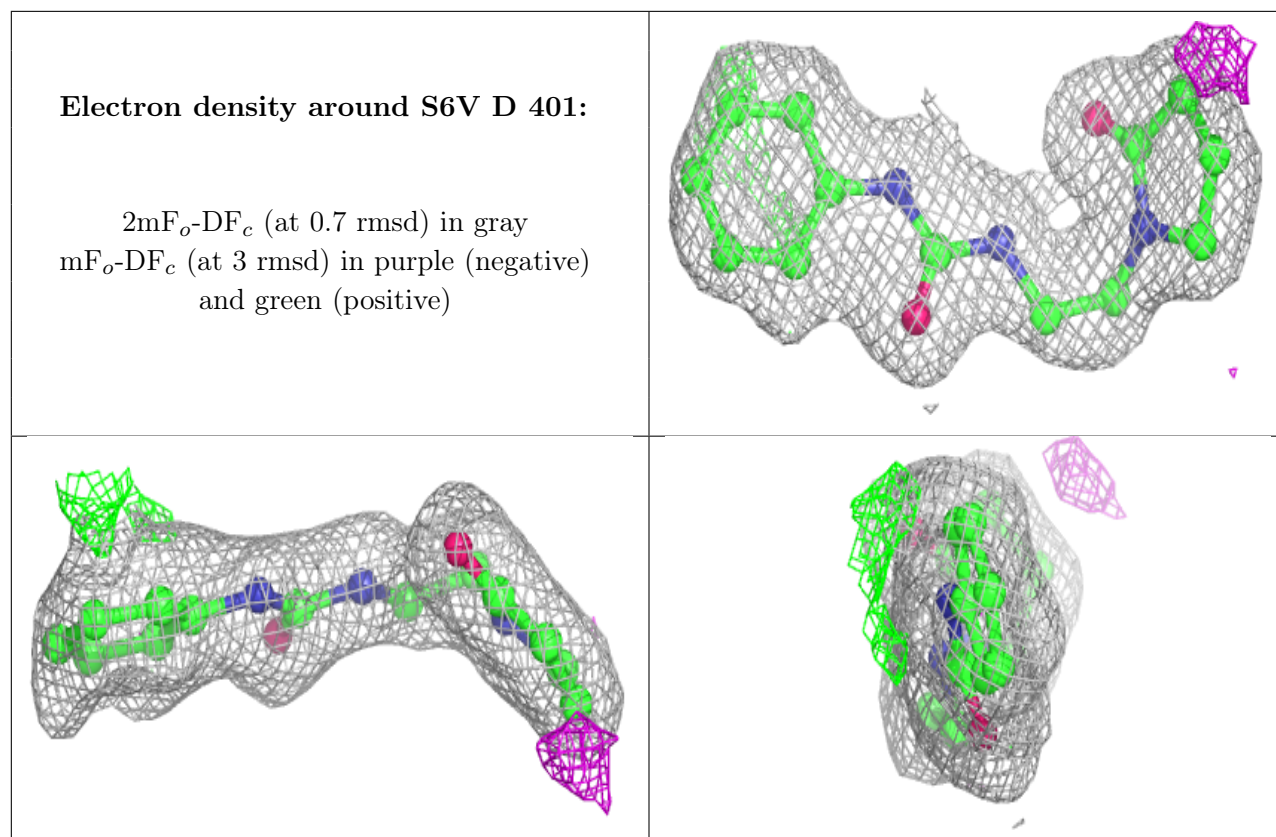
median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	HFK	B	402	26/26	0.86	0.13	40,48,52,56	0
3	HFK	A	402	26/26	0.92	0.11	25,39,46,52	0
2	GAL	A	401	12/12	0.92	0.09	25,30,35,40	0
2	GAL	B	401	12/12	0.93	0.11	27,34,37,53	0
3	HFK	E	402	26/26	0.93	0.10	22,31,36,36	0
4	S6V	D	401	18/18	0.94	0.10	20,27,43,46	0
2	GAL	E	401	12/12	0.96	0.10	18,21,23,35	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.







## 6.5 Other polymers [i](#)

There are no such residues in this entry.