



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 24, 2017 – 05:26 PM EDT

PDB ID : 5KYJ  
Title : Brain penetrant liver X receptor (LXR) modulators based on a 2,4,5,6-tetrahydro-*pyrrolo*[3,4-*c*]pyrazole core  
Authors : Chen, G.; McKeever, B.M.  
Deposited on : unknown  
Resolution : 2.80 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

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030345

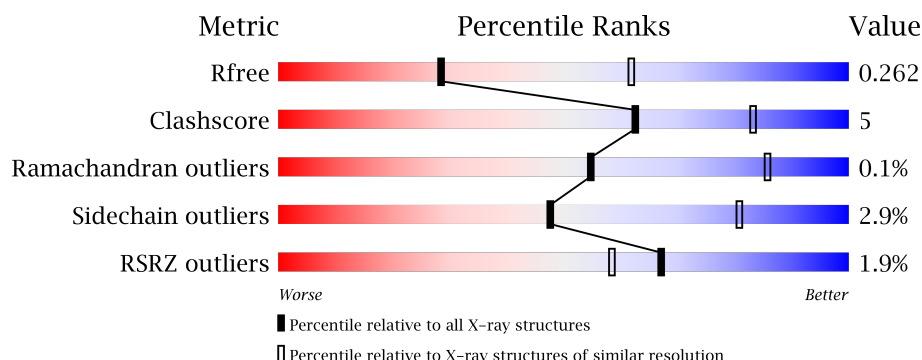
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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))
$R_{free}$	100719	2583 (2.80-2.80)
Clashscore	112137	3033 (2.80-2.80)
Ramachandran outliers	110173	2983 (2.80-2.80)
Sidechain outliers	110143	2985 (2.80-2.80)
RSRZ outliers	101464	2610 (2.80-2.80)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	287	<div> <div>0.1%</div> <div> <div></div> <div>74%</div> <div>13%</div> <div>13%</div> </div> </div>
1	E	287	<div> <div>4%</div> <div> <div></div> <div>72%</div> <div>11%</div> <div>15%</div> </div> </div>
2	B	256	<div> <div>0.1%</div> <div> <div></div> <div>75%</div> <div>8%</div> <div>15%</div> </div> </div>
2	F	256	<div> <div>0.1%</div> <div> <div></div> <div>77%</div> <div>8%</div> <div>15%</div> </div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 7565 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 Oxysterols receptor LXR-beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	251	Total	C	N	O	S	0	0	0
			2028	1297	355	369	7			
1	E	245	Total	C	N	O	S	0	0	0
			1996	1278	351	360	7			

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	194	MET	-	initiating methionine	UNP P55055
A	195	GLY	-	expression tag	UNP P55055
A	196	SER	-	expression tag	UNP P55055
A	197	SER	-	expression tag	UNP P55055
A	198	HIS	-	expression tag	UNP P55055
A	199	HIS	-	expression tag	UNP P55055
A	200	HIS	-	expression tag	UNP P55055
A	201	HIS	-	expression tag	UNP P55055
A	202	HIS	-	expression tag	UNP P55055
A	203	HIS	-	expression tag	UNP P55055
A	204	SER	-	expression tag	UNP P55055
A	205	SER	-	expression tag	UNP P55055
A	206	GLY	-	expression tag	UNP P55055
A	207	LEU	-	expression tag	UNP P55055
A	208	VAL	-	expression tag	UNP P55055
A	209	PRO	-	expression tag	UNP P55055
A	210	ARG	-	expression tag	UNP P55055
A	213	HIS	GLY	engineered mutation	UNP P55055
A	214	MET	GLU	engineered mutation	UNP P55055
A	259	ALA	GLN	engineered mutation	UNP P55055
A	261	GLY	ARG	engineered mutation	UNP P55055
A	262	SER	ASP	engineered mutation	UNP P55055
A	264	SER	ARG	engineered mutation	UNP P55055
A	462	GLY	-	expression tag	UNP P55055
A	463	SER	-	expression tag	UNP P55055

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	464	GLY	-	expression tag	UNP P55055
A	465	SER	-	expression tag	UNP P55055
A	466	GLY	-	expression tag	UNP P55055
A	467	SER	-	expression tag	UNP P55055
A	468	HIS	-	expression tag	UNP P55055
A	469	LYS	-	expression tag	UNP P55055
A	470	ILE	-	expression tag	UNP P55055
A	471	LEU	-	expression tag	UNP P55055
A	472	HIS	-	expression tag	UNP P55055
A	473	ARG	-	expression tag	UNP P55055
A	474	LEU	-	expression tag	UNP P55055
A	475	LEU	-	expression tag	UNP P55055
A	476	GLN	-	expression tag	UNP P55055
A	477	ASP	-	expression tag	UNP P55055
A	478	SER	-	expression tag	UNP P55055
A	479	SER	-	expression tag	UNP P55055
A	480	SER	-	expression tag	UNP P55055
E	194	MET	-	initiating methionine	UNP P55055
E	195	GLY	-	expression tag	UNP P55055
E	196	SER	-	expression tag	UNP P55055
E	197	SER	-	expression tag	UNP P55055
E	198	HIS	-	expression tag	UNP P55055
E	199	HIS	-	expression tag	UNP P55055
E	200	HIS	-	expression tag	UNP P55055
E	201	HIS	-	expression tag	UNP P55055
E	202	HIS	-	expression tag	UNP P55055
E	203	HIS	-	expression tag	UNP P55055
E	204	SER	-	expression tag	UNP P55055
E	205	SER	-	expression tag	UNP P55055
E	206	GLY	-	expression tag	UNP P55055
E	207	LEU	-	expression tag	UNP P55055
E	208	VAL	-	expression tag	UNP P55055
E	209	PRO	-	expression tag	UNP P55055
E	210	ARG	-	expression tag	UNP P55055
E	213	HIS	GLY	engineered mutation	UNP P55055
E	214	MET	GLU	engineered mutation	UNP P55055
E	259	ALA	GLN	engineered mutation	UNP P55055
E	261	GLY	ARG	engineered mutation	UNP P55055
E	262	SER	ASP	engineered mutation	UNP P55055
E	264	SER	ARG	engineered mutation	UNP P55055
E	462	GLY	-	expression tag	UNP P55055
E	463	SER	-	expression tag	UNP P55055

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	464	GLY	-	expression tag	UNP P55055
E	465	SER	-	expression tag	UNP P55055
E	466	GLY	-	expression tag	UNP P55055
E	467	SER	-	expression tag	UNP P55055
E	468	HIS	-	expression tag	UNP P55055
E	469	LYS	-	expression tag	UNP P55055
E	470	ILE	-	expression tag	UNP P55055
E	471	LEU	-	expression tag	UNP P55055
E	472	HIS	-	expression tag	UNP P55055
E	473	ARG	-	expression tag	UNP P55055
E	474	LEU	-	expression tag	UNP P55055
E	475	LEU	-	expression tag	UNP P55055
E	476	GLN	-	expression tag	UNP P55055
E	477	ASP	-	expression tag	UNP P55055
E	478	SER	-	expression tag	UNP P55055
E	479	SER	-	expression tag	UNP P55055
E	480	SER	-	expression tag	UNP P55055

- Molecule 2 is a protein called Retinoic acid receptor RXR-beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	217	Total	C	N	O	S	0	0	0
			1739	1116	304	309	10			
2	F	218	Total	C	N	O	S	0	0	0
			1738	1112	305	311	10			

There are 42 discrepancies between the modelled and reference sequences:

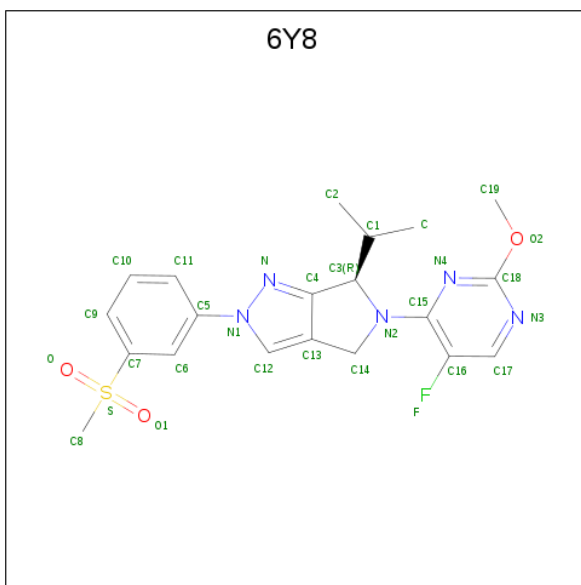
Chain	Residue	Modelled	Actual	Comment	Reference
B	292	HIS	-	expression tag	UNP P28702
B	293	MET	GLY	engineered mutation	UNP P28702
B	529	GLY	-	expression tag	UNP P28702
B	530	SER	-	expression tag	UNP P28702
B	531	GLY	-	expression tag	UNP P28702
B	532	SER	-	expression tag	UNP P28702
B	533	GLY	-	expression tag	UNP P28702
B	534	SER	-	expression tag	UNP P28702
B	535	HIS	-	expression tag	UNP P28702
B	536	LYS	-	expression tag	UNP P28702
B	537	ILE	-	expression tag	UNP P28702
B	538	LEU	-	expression tag	UNP P28702
B	539	HIS	-	expression tag	UNP P28702

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	540	ARG	-	expression tag	UNP P28702
B	541	LEU	-	expression tag	UNP P28702
B	542	LEU	-	expression tag	UNP P28702
B	543	GLN	-	expression tag	UNP P28702
B	544	ASP	-	expression tag	UNP P28702
B	545	SER	-	expression tag	UNP P28702
B	546	SER	-	expression tag	UNP P28702
B	547	SER	-	expression tag	UNP P28702
F	292	HIS	-	expression tag	UNP P28702
F	293	MET	GLY	engineered mutation	UNP P28702
F	529	GLY	-	expression tag	UNP P28702
F	530	SER	-	expression tag	UNP P28702
F	531	GLY	-	expression tag	UNP P28702
F	532	SER	-	expression tag	UNP P28702
F	533	GLY	-	expression tag	UNP P28702
F	534	SER	-	expression tag	UNP P28702
F	535	HIS	-	expression tag	UNP P28702
F	536	LYS	-	expression tag	UNP P28702
F	537	ILE	-	expression tag	UNP P28702
F	538	LEU	-	expression tag	UNP P28702
F	539	HIS	-	expression tag	UNP P28702
F	540	ARG	-	expression tag	UNP P28702
F	541	LEU	-	expression tag	UNP P28702
F	542	LEU	-	expression tag	UNP P28702
F	543	GLN	-	expression tag	UNP P28702
F	544	ASP	-	expression tag	UNP P28702
F	545	SER	-	expression tag	UNP P28702
F	546	SER	-	expression tag	UNP P28702
F	547	SER	-	expression tag	UNP P28702

- Molecule 3 is (6 {R})-5-(5-fluoranyl-2-methoxy-pyrimidin-4-yl)-2-(3-methylsulfonylphenyl)-6-propan-2-yl-4,6-dihydropyrrolo[3,4-c]pyrazole (three-letter code: 6Y8) (formula: C<sub>20</sub>H<sub>22</sub>FN<sub>5</sub>O<sub>3</sub>S).



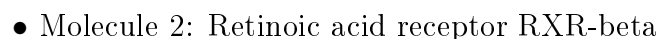
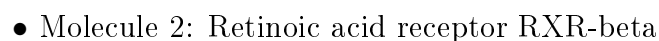
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	A	1	Total	C	F	N	O	S	0	0
			30	20	1	5	3	1		
3	E	1	Total	C	F	N	O	S	0	0
			30	20	1	5	3	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	3	Total	O	0	0
			3	3		
4	E	1	Total	O	0	0
			1	1		



- Molecule 1: Oxysterols receptor LXR-beta





Q477	
I490	
I512	
I513	
O514	
ASP	
THR	
PRO	
ILE	
D519	
T520	
F521	
L522	
M523	
A528	
GLY	
SER	
GLY	
SER	
H533	
L538	
H539	
O543	
ASP	
SER	
SER	
SER	

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	69.92Å 100.93Å 145.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.81 – 2.80 43.81 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.8 (43.81-2.80) 99.9 (43.81-2.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.68 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.8.0071	Depositor
R, $R_{free}$	0.199 , 0.253 0.204 , 0.262	Depositor DCC
$R_{free}$ test set	1290 reflections (5.21%)	DCC
Wilson B-factor (Å <sup>2</sup> )	56.2	Xtriage
Anisotropy	0.022	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 37.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	7565	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.78% 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 ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: 6Y8

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	0.59	0/2067	0.68	0/2796
1	E	0.59	0/2034	0.69	1/2748 (0.0%)
2	B	0.54	0/1769	0.75	2/2383 (0.1%)
2	F	0.55	0/1767	0.74	1/2379 (0.0%)
All	All	0.57	0/7637	0.71	4/10306 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	477	GLN	CB-CA-C	-6.08	98.23	110.40
2	F	477	GLN	CB-CA-C	-6.06	98.28	110.40
2	B	464	ARG	CG-CD-NE	5.95	124.29	111.80
1	E	391	ARG	NE-CZ-NH1	5.03	122.81	120.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts ⓘ

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	2028	0	2062	24	0
1	E	1996	0	2030	29	0
2	B	1739	0	1791	16	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	F	1738	0	1783	10	0
3	A	30	0	0	0	0
3	E	30	0	0	0	0
4	A	3	0	0	0	0
4	E	1	0	0	0	0
All	All	7565	0	7666	76	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 (76) 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:A:301:ILE:CG2	1:A:305:LYS:HE3	1.77	1.14
1:A:301:ILE:HG22	1:A:305:LYS:HE3	1.38	0.99
2:B:540:ARG:HH11	2:B:540:ARG:HG3	1.34	0.92
1:E:455:GLU:OE2	1:E:468:HIS:HA	1.81	0.80
1:E:455:GLU:OE2	1:E:468:HIS:C	2.20	0.79
1:E:455:GLU:OE2	1:E:468:HIS:CA	2.34	0.76
2:B:542:LEU:O	2:B:543:GLN:HG3	1.87	0.74
2:F:519:ASP:O	2:F:523:MET:HB2	1.88	0.73
2:B:542:LEU:O	2:B:543:GLN:CG	2.38	0.72
1:E:270:HIS:HE2	1:E:335:TYR:HH	1.37	0.72
1:E:423:MET:HG3	2:F:468:TYR:OH	1.91	0.70
1:E:455:GLU:OE2	1:E:469:LYS:N	2.24	0.70
1:A:301:ILE:O	1:A:305:LYS:HG3	1.92	0.70
2:B:540:ARG:HH11	2:B:540:ARG:CG	2.06	0.69
1:A:301:ILE:HG23	1:A:305:LYS:HE3	1.72	0.67
2:F:389:ILE:O	2:F:429:ARG:NH1	2.26	0.67
1:A:270:HIS:HE2	1:A:335:TYR:HH	1.43	0.66
2:B:389:ILE:O	2:B:429:ARG:NH1	2.28	0.66
1:A:301:ILE:HG22	1:A:305:LYS:CE	2.22	0.65
1:A:423:MET:HG3	2:B:468:TYR:OH	1.98	0.64
1:A:339:ASP:OD1	1:A:342:ARG:NH2	2.33	0.61
2:B:369:VAL:HG13	2:B:538:LEU:HD23	1.81	0.61
2:F:369:VAL:HG13	2:F:538:LEU:HD23	1.82	0.60
1:E:356:PHE:CE1	1:E:360:MET:HE1	2.36	0.60
2:F:512:LEU:C	2:F:513:ILE:HG13	2.23	0.59
1:E:294:GLN:O	1:E:391:ARG:NH2	2.39	0.56
1:A:455:GLU:OE1	1:A:470:ILE:HG13	2.05	0.56
2:B:444:ILE:O	2:B:464:ARG:NH1	2.39	0.56

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:407:THR:HG21	1:E:418:PHE:HB2	1.88	0.54
2:B:447:PHE:O	2:B:464:ARG:NH1	2.42	0.53
1:E:456:ILE:HG23	1:E:457:TRP:CD1	2.44	0.52
1:A:456:ILE:HG23	1:A:457:TRP:CD1	2.46	0.51
1:E:356:PHE:CZ	1:E:360:MET:CE	2.94	0.51
1:E:356:PHE:CZ	1:E:360:MET:HE1	2.46	0.51
1:E:403:LEU:O	1:E:407:THR:HB	2.12	0.50
2:B:384:PHE:C	2:B:384:PHE:CD1	2.85	0.49
1:E:423:MET:HE3	2:F:490:LEU:HD22	1.94	0.49
1:E:455:GLU:CG	1:E:468:HIS:HA	2.44	0.48
1:E:262:SER:HA	1:E:265:GLN:HE21	1.78	0.48
1:A:283:VAL:HG22	1:A:304:LEU:HD13	1.97	0.47
1:A:299:ASP:O	1:A:303:LEU:HG	2.16	0.46
1:E:319:ARG:HD3	1:E:328:THR:O	2.15	0.46
2:F:335:PRO:HG3	2:F:407:SER:OG	2.15	0.45
1:A:319:ARG:HD3	1:A:328:THR:O	2.15	0.45
1:E:455:GLU:CD	1:E:468:HIS:HA	2.36	0.45
1:E:305:LYS:HD2	1:E:468:HIS:CE1	2.51	0.45
2:F:301:VAL:HG13	2:F:470:SER:HB3	1.98	0.45
2:B:540:ARG:NH1	2:B:540:ARG:CG	2.73	0.45
2:B:519:ASP:O	2:B:523:MET:HB2	2.17	0.44
1:E:363:LEU:HB3	1:E:365:LEU:HD13	2.00	0.44
1:E:366:ASP:OD1	1:E:366:ASP:C	2.56	0.44
1:E:356:PHE:CE1	1:E:360:MET:CE	3.01	0.43
2:B:361:SER:HA	2:B:368:GLN:NE2	2.33	0.43
1:E:397:GLN:N	1:E:398:PRO:HD2	2.33	0.43
1:A:397:GLN:N	1:A:398:PRO:HD2	2.34	0.43
1:A:363:LEU:HB3	1:A:365:LEU:HD13	2.01	0.43
1:A:272:THR:CG2	1:A:450:PRO:HG3	2.49	0.43
1:A:221:THR:HG22	1:A:223:ALA:N	2.34	0.42
1:A:245:ASP:O	1:A:248:LYS:HB2	2.19	0.42
1:A:355:GLU:HG3	1:A:358:ARG:NH2	2.35	0.42
1:A:366:ASP:C	1:A:366:ASP:OD1	2.57	0.42
2:F:512:LEU:O	2:F:513:ILE:HG13	2.19	0.42
1:A:380:SER:O	1:A:383:ARG:HG2	2.19	0.42
2:B:542:LEU:O	2:B:543:GLN:HG2	2.20	0.41
1:E:221:THR:HG22	1:E:223:ALA:N	2.36	0.41
1:E:245:ASP:O	1:E:248:LYS:HB2	2.20	0.41
1:E:272:THR:CG2	1:E:450:PRO:HG3	2.50	0.41
1:A:349:PHE:CZ	1:A:353:ILE:HD11	2.56	0.41
2:B:351:VAL:CG1	2:B:542:LEU:HD23	2.51	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:270:HIS:NE2	1:E:335:TYR:OH	2.31	0.41
2:B:413:VAL:O	2:B:413:VAL:CG1	2.69	0.41
1:E:337:LYS:O	1:E:340:PHE:HB2	2.21	0.40
1:A:470:ILE:O	1:A:473:ARG:N	2.54	0.40
1:A:221:THR:HG22	1:A:224:GLN:H	1.86	0.40
1:E:380:SER:O	1:E:383:ARG:HG2	2.21	0.40
2:F:413:VAL:CG1	2:F:413:VAL:O	2.70	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	247/287 (86%)	239 (97%)	8 (3%)	0	100	100
1	E	239/287 (83%)	232 (97%)	7 (3%)	0	100	100
2	B	209/256 (82%)	201 (96%)	8 (4%)	0	100	100
2	F	210/256 (82%)	204 (97%)	5 (2%)	1 (0%)	32	67
All	All	905/1086 (83%)	876 (97%)	28 (3%)	1 (0%)	55	86

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	513	ILE

### 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	222/250 (89%)	216 (97%)	6 (3%)	50	83
1	E	219/250 (88%)	214 (98%)	5 (2%)	56	86
2	B	191/217 (88%)	184 (96%)	7 (4%)	39	73
2	F	190/217 (88%)	184 (97%)	6 (3%)	44	78
All	All	822/934 (88%)	798 (97%)	24 (3%)	48	81

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	245	ASP
1	A	271	PHE
1	A	272	THR
1	A	338	ASP
1	A	397	GLN
1	A	433	SER
2	B	384	PHE
2	B	434	ASP
2	B	458	SER
2	B	521	PHE
2	B	535	HIS
2	B	540	ARG
2	B	541	LEU
1	E	245	ASP
1	E	271	PHE
1	E	272	THR
1	E	338	ASP
1	E	407	THR
2	F	361	SER
2	F	384	PHE
2	F	434	ASP
2	F	458	SER
2	F	521	PHE
2	F	539	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	239	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	385	ASN
1	A	396	GLN
1	E	237	GLN
1	E	239	ASN
1	E	265	GLN
1	E	288	GLN
1	E	396	GLN
1	E	415	GLN
2	F	359	HIS

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

2 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	6Y8	A	501	-	28,33,33	3.22	4 (14%)	32,50,50	3.21	14 (43%)
3	6Y8	E	501	-	28,33,33	3.32	5 (17%)	32,50,50	3.21	14 (43%)

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	6Y8	A	501	-	-	0/19/32/32	0/4/4/4
3	6Y8	E	501	-	-	0/19/32/32	0/4/4/4

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	501	6Y8	C7-S	-15.51	1.59	1.77
3	E	501	6Y8	C7-S	-15.29	1.59	1.77
3	E	501	6Y8	C5-N1	-5.63	1.32	1.44
3	A	501	6Y8	C5-N1	-4.12	1.35	1.44
3	E	501	6Y8	C8-S	-2.74	1.63	1.75
3	A	501	6Y8	C8-S	-2.10	1.66	1.75
3	E	501	6Y8	C18-N3	2.52	1.35	1.33
3	E	501	6Y8	C15-C16	2.91	1.47	1.42
3	A	501	6Y8	C15-C16	3.98	1.48	1.42

All (28) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	501	6Y8	N3-C18-N4	-9.43	121.61	128.40
3	A	501	6Y8	N3-C18-N4	-7.96	122.68	128.40
3	E	501	6Y8	C11-C5-C6	-6.47	115.14	121.55
3	A	501	6Y8	C11-C5-C6	-6.29	115.32	121.55
3	E	501	6Y8	C13-C14-N2	-4.34	100.29	102.26
3	A	501	6Y8	C13-C14-N2	-4.33	100.30	102.26
3	E	501	6Y8	O1-S-O	-3.60	111.22	117.84
3	E	501	6Y8	C13-C12-N1	-3.48	103.47	107.58
3	E	501	6Y8	C4-C3-C1	-3.48	103.37	113.01
3	A	501	6Y8	C13-C12-N1	-3.08	103.94	107.58
3	A	501	6Y8	O-S-C7	-2.65	106.15	108.30
3	A	501	6Y8	C9-C7-S	-2.22	117.44	119.60
3	A	501	6Y8	C-C1-C3	2.23	113.74	110.71
3	E	501	6Y8	F-C16-C15	2.48	121.55	119.73
3	E	501	6Y8	C6-C5-N1	2.54	122.22	119.13
3	E	501	6Y8	C10-C11-C5	2.54	121.89	118.66
3	E	501	6Y8	O1-S-C8	2.72	112.53	108.49
3	A	501	6Y8	C5-C6-C7	3.11	121.82	119.24
3	E	501	6Y8	C19-O2-C18	3.66	123.84	117.56
3	A	501	6Y8	C6-C7-S	3.73	122.37	119.09

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	501	6Y8	C8-S-C7	3.89	109.21	104.58
3	A	501	6Y8	C8-S-C7	4.00	109.34	104.58
3	A	501	6Y8	C10-C11-C5	4.05	123.81	118.66
3	A	501	6Y8	C6-C5-N1	4.97	125.19	119.13
3	E	501	6Y8	C5-C6-C7	5.57	123.87	119.24
3	E	501	6Y8	C17-N3-C18	5.87	121.22	114.92
3	A	501	6Y8	C17-N3-C18	6.50	121.90	114.92
3	A	501	6Y8	F-C16-C15	7.09	124.93	119.73

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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	251/287 (87%)	-0.23	2 (0%) 86 81	29, 51, 92, 112	0
1	E	245/287 (85%)	-0.09	11 (4%) 34 24	24, 50, 96, 133	0
2	B	217/256 (84%)	-0.20	2 (0%) 84 79	33, 55, 85, 114	0
2	F	218/256 (85%)	-0.18	3 (1%) 75 69	32, 54, 89, 120	0
All	All	931/1086 (85%)	-0.18	18 (1%) 67 58	24, 53, 90, 133	0

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	218	VAL	4.3
1	E	242	SER	3.9
1	E	254	LEU	3.7
1	A	219	GLN	3.4
1	E	468	HIS	3.4
1	E	243	PHE	3.3
2	B	518	ILE	3.0
1	E	244	SER	2.7
1	E	246	GLN	2.7
1	E	265	GLN	2.6
2	B	401	LEU	2.5
2	F	297	GLU	2.4
1	E	251	PRO	2.4
1	E	262	SER	2.4
1	E	247	PRO	2.3
1	E	241	ARG	2.2
2	F	344	ASP	2.1
2	F	402	HIS	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 carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	6Y8	A	501	30/30	0.98	0.17	0.20	39,43,49,51	0
3	6Y8	E	501	30/30	0.97	0.17	-0.27	34,44,55,60	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.