



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2018 – 09:36 pm GMT

PDB ID : 5FYI  
Title : Crystal structure of human JMJD2A in complex with pyruvate  
Authors : Nowak, R.; Kopec, J.; Johansson, C.; Szykowska, A.; von Delft, F.; Arrow-smith, C.H.; Bountra, C.; Edwards, A.; Oppermann, U.  
Deposited on : 2016-03-07  
Resolution : 2.10 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk30967  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30967

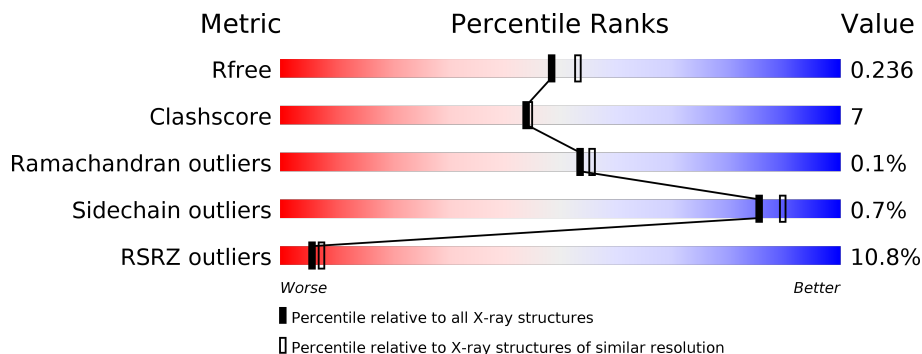
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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))
$R_{free}$	111664	4608 (2.10-2.10)
Clashscore	122126	5109 (2.10-2.10)
Ramachandran outliers	120053	5059 (2.10-2.10)
Sidechain outliers	120020	5060 (2.10-2.10)
RSRZ outliers	108989	4497 (2.10-2.10)

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	381	 9% 79% 13% 8%
1	B	381	 10% 76% 13% 10%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	PYR	A	1356	-	-	-	X
3	PYR	B	1355	-	-	-	X
5	EDO	A	1358	-	-	X	-
5	EDO	B	1357	-	-	X	-
5	EDO	B	1364	-	-	-	X

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 6036 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 LYSINE-SPECIFIC DEMETHYLASE 4A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	351	2844	1844	468	517	15	0	5	0
1	B	344	2756	1791	455	494	16	0	1	0

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-21	MET	-	expression tag	UNP O75164
A	-20	HIS	-	expression tag	UNP O75164
A	-19	HIS	-	expression tag	UNP O75164
A	-18	HIS	-	expression tag	UNP O75164
A	-17	HIS	-	expression tag	UNP O75164
A	-16	HIS	-	expression tag	UNP O75164
A	-15	HIS	-	expression tag	UNP O75164
A	-14	SER	-	expression tag	UNP O75164
A	-13	SER	-	expression tag	UNP O75164
A	-12	GLY	-	expression tag	UNP O75164
A	-11	VAL	-	expression tag	UNP O75164
A	-10	ASP	-	expression tag	UNP O75164
A	-9	LEU	-	expression tag	UNP O75164
A	-8	GLY	-	expression tag	UNP O75164
A	-7	THR	-	expression tag	UNP O75164
A	-6	GLU	-	expression tag	UNP O75164
A	-5	ASN	-	expression tag	UNP O75164
A	-4	LEU	-	expression tag	UNP O75164
A	-3	TYR	-	expression tag	UNP O75164
A	-2	PHE	-	expression tag	UNP O75164
A	-1	GLN	-	expression tag	UNP O75164
A	0	SER	-	expression tag	UNP O75164
B	-21	MET	-	expression tag	UNP O75164
B	-20	HIS	-	expression tag	UNP O75164
B	-19	HIS	-	expression tag	UNP O75164

*Continued on next page...*

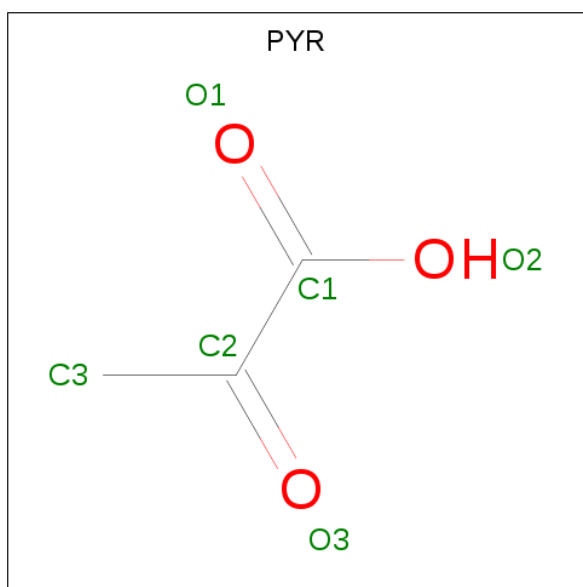
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-18	HIS	-	expression tag	UNP O75164
B	-17	HIS	-	expression tag	UNP O75164
B	-16	HIS	-	expression tag	UNP O75164
B	-15	HIS	-	expression tag	UNP O75164
B	-14	SER	-	expression tag	UNP O75164
B	-13	SER	-	expression tag	UNP O75164
B	-12	GLY	-	expression tag	UNP O75164
B	-11	VAL	-	expression tag	UNP O75164
B	-10	ASP	-	expression tag	UNP O75164
B	-9	LEU	-	expression tag	UNP O75164
B	-8	GLY	-	expression tag	UNP O75164
B	-7	THR	-	expression tag	UNP O75164
B	-6	GLU	-	expression tag	UNP O75164
B	-5	ASN	-	expression tag	UNP O75164
B	-4	LEU	-	expression tag	UNP O75164
B	-3	TYR	-	expression tag	UNP O75164
B	-2	PHE	-	expression tag	UNP O75164
B	-1	GLN	-	expression tag	UNP O75164
B	0	SER	-	expression tag	UNP O75164

- Molecule 2 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

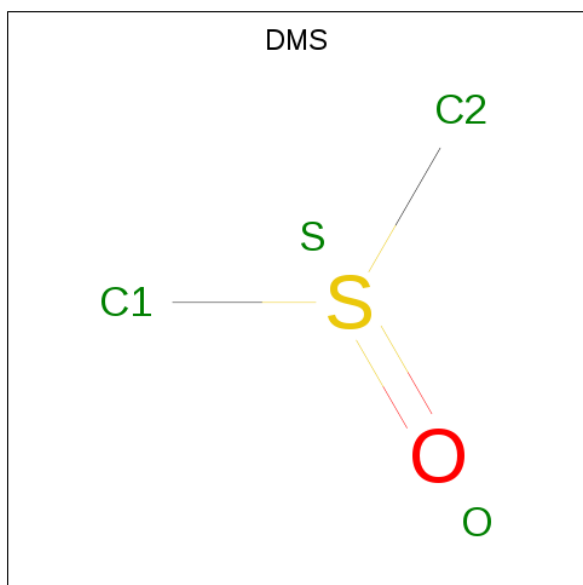
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total Ni 1 1	0	0
2	A	1	Total Ni 1 1	0	0

- Molecule 3 is PYRUVIC ACID (three-letter code: PYR) (formula: C<sub>3</sub>H<sub>4</sub>O<sub>3</sub>).



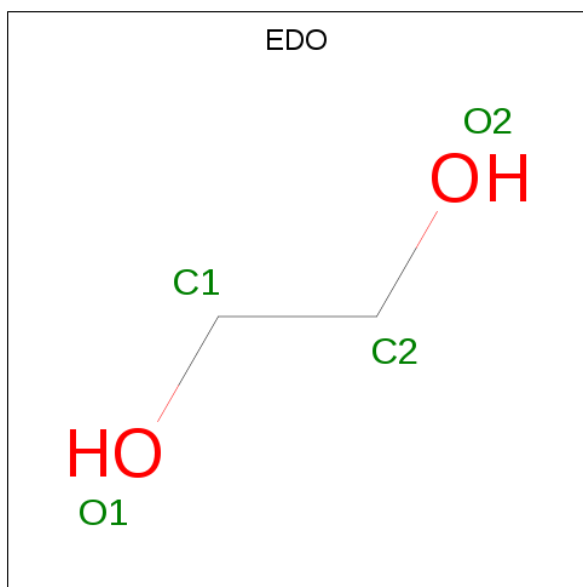
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		
3	A	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula:  $C_2H_6OS$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	O	S	0	0
			4	2	1	1		
4	B	1	Total	C	O	S	0	0
			4	2	1	1		

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		

*Continued on next page...*

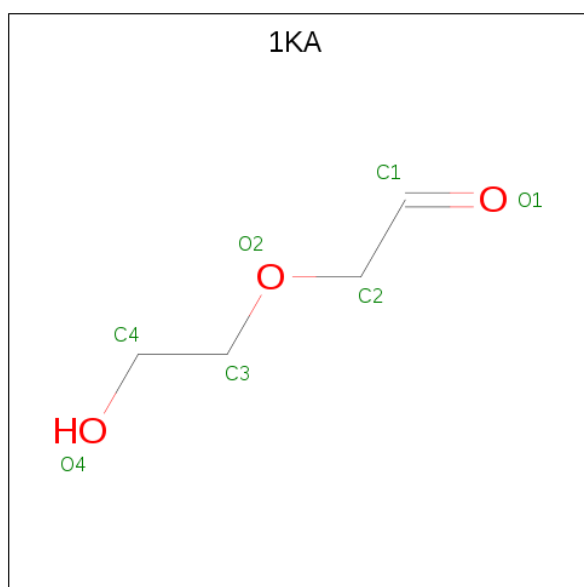
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total C O 4 2 2	0	0
5	B	1	Total C O 4 2 2	0	0
5	B	1	Total C O 4 2 2	0	0
5	B	1	Total C O 4 2 2	0	0
5	B	1	Total C O 4 2 2	0	0
5	B	1	Total C O 4 2 2	0	0
5	B	1	Total C O 4 2 2	0	0

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	B	1	Total Zn 1 1	0	0
6	A	1	Total Zn 1 1	0	0

- Molecule 7 is (2-hydroxyethoxy)acetaldehyde (three-letter code: 1KA) (formula: C<sub>4</sub>H<sub>8</sub>O<sub>3</sub>).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	B	1	Total	C	O	0	0
			7	4	3		

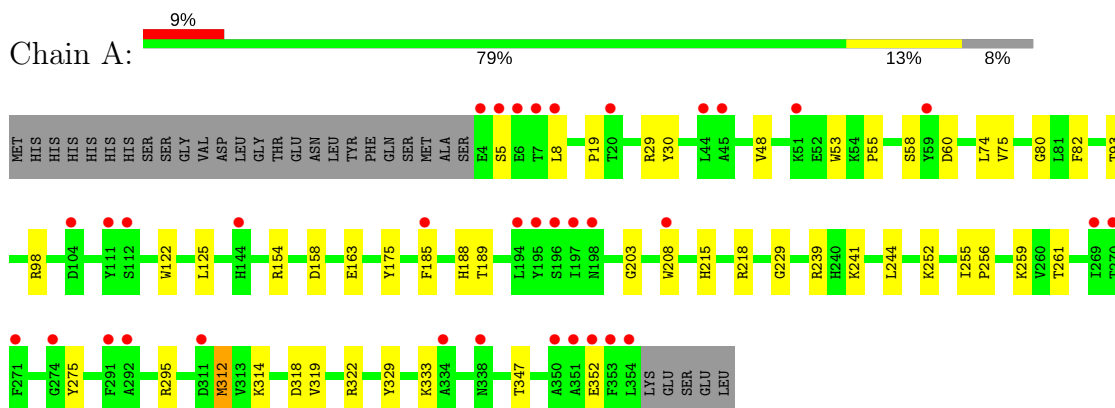
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	159	Total	O	0	0
			159	159		
8	B	166	Total	O	0	0
			166	166		

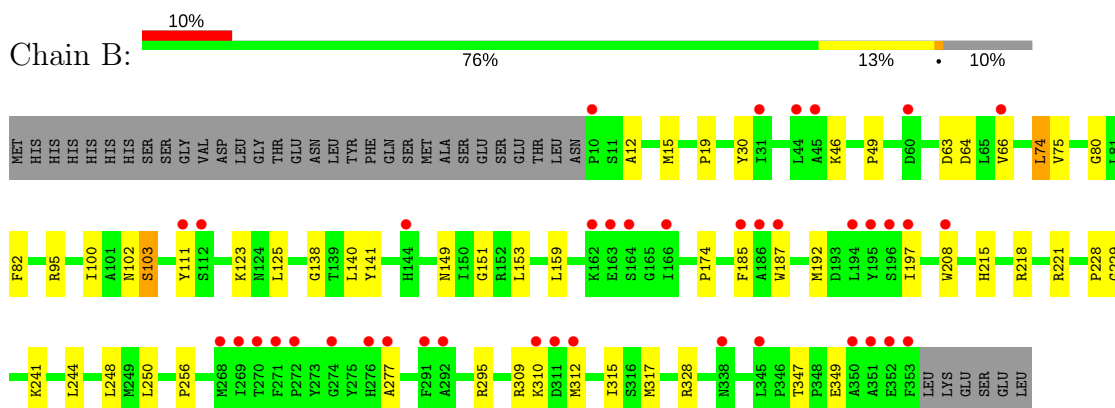
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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: LYSINE-SPECIFIC DEMETHYLASE 4A



- Molecule 1: LYSINE-SPECIFIC DEMETHYLASE 4A



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	100.45Å 148.83Å 57.06Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.77 – 2.10 29.77 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.9 (29.77-2.10) 95.7 (29.77-2.10)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.00 (at 2.10Å)	Xtrriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.197 , 0.236 0.199 , 0.236	Depositor DCC
$R_{free}$ test set	2457 reflections (4.82%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	39.9	Xtrriage
Anisotropy	0.222	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 46.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	6036	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	47.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.73% 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: NI, ZN, 1KA, EDO, DMS, PYR

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	0.40	0/2940	0.54	1/3993 (0.0%)
1	B	0.42	0/2843	0.55	1/3863 (0.0%)
All	All	0.41	0/5783	0.55	2/7856 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	74	LEU	CA-CB-CG	-5.53	102.59	115.30
1	A	74	LEU	CA-CB-CG	-5.37	102.94	115.30

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	2844	0	2702	35	0
1	B	2756	0	2602	41	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	12	0	6	3	0
3	B	12	0	6	4	0
4	A	4	0	6	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	4	0	6	2	0
5	A	36	0	54	10	0
5	B	32	0	48	12	0
6	A	1	0	0	0	0
6	B	1	0	0	0	0
7	B	7	0	8	1	0
8	A	159	0	0	5	0
8	B	166	0	0	6	0
All	All	6036	0	5438	80	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 7.

All (80) 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:B:317[B]:MET:SD	1:B:317[B]:MET:N	2.41	0.92
1:B:277:ALA:HB1	5:B:1357:EDO:H21	1.56	0.88
1:A:93:THR:OG1	8:A:2020:HOH:O	2.01	0.78
1:A:239:ARG:HG2	5:A:1358:EDO:H12	1.71	0.73
1:B:248:LEU:HB2	5:B:1358:EDO:H21	1.71	0.70
1:A:163:GLU:HG3	1:A:319:VAL:HG21	1.75	0.69
1:A:203:GLY:O	8:A:2093:HOH:O	2.12	0.67
1:A:252:LYS:NZ	1:B:15:MET:SD	2.68	0.67
1:A:58:SER:OG	1:A:60:ASP:OD2	2.14	0.66
1:B:111:TYR:OH	8:B:2062:HOH:O	2.12	0.65
1:A:189:THR:H	5:A:1358:EDO:H11	1.62	0.65
1:B:49:PRO:O	8:B:2020:HOH:O	2.16	0.63
1:A:19:PRO:HB3	1:A:30:TYR:CZ	2.35	0.62
1:B:208:TRP:CH2	3:B:1354:PYR:H33	2.35	0.61
3:A:1355:PYR:H32	3:A:1356:PYR:C1	2.31	0.60
1:B:309:ARG:HD2	1:B:312:MET:HE2	1.85	0.57
1:B:151:GLY:HA2	1:B:174:PRO:HG3	1.86	0.57
1:A:208:TRP:CH2	3:A:1355:PYR:H33	2.40	0.56
1:A:241[A]:LYS:NZ	8:A:2087:HOH:O	2.38	0.56
1:B:100:ILE:O	1:B:103:SER:HB3	2.06	0.56
1:B:12:ALA:HB3	7:B:1365:1KA:H7	1.88	0.55
1:A:29:ARG:NE	1:A:352:GLU:OE2	2.39	0.55
1:B:244:LEU:HD22	5:B:1357:EDO:H11	1.88	0.55
1:A:188:HIS:CD2	1:A:241[B]:LYS:HE2	2.43	0.54
1:A:218[A]:ARG:HD2	1:A:255:ILE:HD13	1.90	0.52

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:5:SER:HA	1:A:8:LEU:HD12	1.92	0.52
1:A:322:ARG:HG3	1:A:329:TYR:CE1	2.45	0.51
1:A:19:PRO:HD2	1:A:48:VAL:O	2.11	0.50
1:A:239:ARG:HG2	5:A:1358:EDO:C1	2.41	0.50
1:A:229:GLY:H	4:A:1357:DMS:H21	1.78	0.49
1:B:19:PRO:HB3	1:B:30:TYR:CZ	2.48	0.49
1:B:46:LYS:NZ	5:B:1363:EDO:H22	2.26	0.49
1:B:218:ARG:HG2	1:B:221:ARG:NH2	2.27	0.49
1:B:328:ARG:NE	8:B:2155:HOH:O	2.45	0.48
1:A:80:GLY:HA2	5:A:1359:EDO:H11	1.94	0.48
1:B:141:TYR:CE2	1:B:149:ASN:HA	2.48	0.48
1:A:189:THR:OG1	5:A:1358:EDO:H11	2.14	0.48
1:B:295:ARG:HB2	1:B:347:THR:HA	1.95	0.48
1:B:228:PRO:HD2	4:B:1356:DMS:H21	1.95	0.47
3:B:1354:PYR:H32	3:B:1355:PYR:C2	2.45	0.47
1:B:63:ASP:HB3	1:B:95:ARG:HB2	1.96	0.47
1:B:185:PHE:HB2	5:B:1364:EDO:H11	1.96	0.47
1:A:185:PHE:HB2	5:A:1365:EDO:H22	1.96	0.46
3:B:1354:PYR:H32	3:B:1355:PYR:C1	2.45	0.46
1:B:80:GLY:HA2	5:B:1358:EDO:H11	1.97	0.46
1:A:75:VAL:HG11	1:A:125:LEU:HG	1.98	0.46
1:B:64:ASP:OD1	1:B:95:ARG:NH2	2.43	0.45
1:B:215:HIS:CD2	1:B:256:PRO:HG2	2.51	0.45
1:A:295:ARG:HB2	1:A:347:THR:HA	1.98	0.45
3:B:1355:PYR:H32	5:B:1364:EDO:H22	1.98	0.45
1:B:102:ASN:HD22	5:B:1361:EDO:H21	1.81	0.45
1:B:349:GLU:OE2	8:B:2165:HOH:O	2.20	0.45
1:A:259:LYS:HB2	8:A:2095:HOH:O	2.16	0.44
1:B:66:VAL:O	8:B:2027:HOH:O	2.20	0.44
1:B:187:TRP:CZ2	1:B:250:LEU:HD11	2.53	0.44
1:A:98:ARG:NH1	8:A:2043:HOH:O	2.26	0.43
1:A:154:ARG:HA	1:A:158:ASP:OD2	2.18	0.43
1:A:53:TRP:CZ3	1:A:55:PRO:HD3	2.54	0.43
1:B:123:LYS:O	8:B:2069:HOH:O	2.21	0.43
1:B:82:PHE:HB2	1:B:244:LEU:HB2	2.01	0.43
3:A:1356:PYR:H32	5:A:1365:EDO:H12	2.01	0.43
1:A:318[B]:ASP:OD2	1:A:333:LYS:NZ	2.41	0.43
1:B:244:LEU:HD22	5:B:1357:EDO:C1	2.47	0.42
1:A:312:MET:CE	1:A:314:LYS:HB2	2.49	0.42
1:B:241:LYS:HD3	5:B:1364:EDO:H12	2.01	0.42
1:A:229:GLY:H	4:A:1357:DMS:C2	2.32	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:82:PHE:HB2	1:A:244:LEU:HB2	2.01	0.42
1:B:192:MET:HG3	1:B:315:ILE:HB	2.02	0.42
1:A:261:THR:HG21	5:A:1362:EDO:H12	2.02	0.42
1:A:215:HIS:CD2	1:A:256:PRO:HG2	2.55	0.41
1:B:153:LEU:HD11	1:B:197:ILE:HG21	2.02	0.41
1:B:159:LEU:HA	1:B:159:LEU:HD23	1.82	0.41
1:A:275:TYR:CZ	5:A:1358:EDO:H21	2.56	0.41
1:A:122:TRP:HB3	5:A:1359:EDO:H22	2.01	0.41
1:B:229:GLY:H	4:B:1356:DMS:C2	2.32	0.41
1:B:75:VAL:HG11	1:B:125:LEU:HG	2.02	0.41
1:B:140:LEU:HD23	1:B:140:LEU:HA	1.84	0.40
1:B:244:LEU:HD22	5:B:1357:EDO:H22	2.02	0.40
1:B:46:LYS:HZ3	5:B:1363:EDO:H22	1.86	0.40
1:B:138:GLY:O	1:B:174:PRO:HB2	2.20	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	354/381 (93%)	348 (98%)	6 (2%)	0	100	100
1	B	343/381 (90%)	337 (98%)	5 (2%)	1 (0%)	43	43
All	All	697/762 (92%)	685 (98%)	11 (2%)	1 (0%)	53	55

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	310	LYS

### 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	294/335 (88%)	292 (99%)	2 (1%)	85	89
1	B	279/335 (83%)	277 (99%)	2 (1%)	85	89
All	All	573/670 (86%)	569 (99%)	4 (1%)	85	89

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	175	TYR
1	A	312	MET
1	B	74	LEU
1	B	103	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	102	ASN

### 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

Of 28 ligands modelled in this entry, 4 are monoatomic - leaving 24 for Mogul analysis.

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	PYR	A	1355	2	2,5,5	0.23	0	2,6,6	0.74	0
3	PYR	A	1356	-	2,5,5	0.16	0	2,6,6	0.74	0
4	DMS	A	1357	-	3,3,3	0.58	0	3,3,3	0.61	0
5	EDO	A	1358	-	3,3,3	0.56	0	2,2,2	0.34	0
5	EDO	A	1359	-	3,3,3	0.44	0	2,2,2	0.54	0
5	EDO	A	1360	-	3,3,3	0.48	0	2,2,2	0.44	0
5	EDO	A	1361	-	3,3,3	0.44	0	2,2,2	0.43	0
5	EDO	A	1362	-	3,3,3	0.50	0	2,2,2	0.34	0
5	EDO	A	1363	-	3,3,3	0.54	0	2,2,2	0.19	0
5	EDO	A	1364	-	3,3,3	0.53	0	2,2,2	0.28	0
5	EDO	A	1365	-	3,3,3	0.56	0	2,2,2	0.10	0
5	EDO	A	3001	-	3,3,3	0.50	0	2,2,2	0.29	0
3	PYR	B	1354	2	2,5,5	0.27	0	2,6,6	0.36	0
3	PYR	B	1355	-	2,5,5	0.03	0	2,6,6	0.27	0
4	DMS	B	1356	-	3,3,3	0.58	0	3,3,3	0.71	0
5	EDO	B	1357	-	3,3,3	0.44	0	2,2,2	0.31	0
5	EDO	B	1358	-	3,3,3	0.45	0	2,2,2	0.38	0
5	EDO	B	1359	-	3,3,3	0.56	0	2,2,2	0.34	0
5	EDO	B	1360	-	3,3,3	0.70	0	2,2,2	0.23	0
5	EDO	B	1361	-	3,3,3	0.54	0	2,2,2	0.22	0
5	EDO	B	1362	-	3,3,3	0.68	0	2,2,2	0.16	0
5	EDO	B	1363	-	3,3,3	0.44	0	2,2,2	0.58	0
5	EDO	B	1364	-	3,3,3	0.65	0	2,2,2	0.24	0
7	1KA	B	1365	-	6,6,6	1.98	2 (33%)	5,5,5	2.65	1 (20%)

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	PYR	A	1355	2	-	0/0/4/4	0/0/0/0
3	PYR	A	1356	-	-	0/0/4/4	0/0/0/0
4	DMS	A	1357	-	-	0/0/0/0	0/0/0/0
5	EDO	A	1358	-	-	0/1/1/1	0/0/0/0
5	EDO	A	1359	-	-	0/1/1/1	0/0/0/0
5	EDO	A	1360	-	-	0/1/1/1	0/0/0/0
5	EDO	A	1361	-	-	0/1/1/1	0/0/0/0
5	EDO	A	1362	-	-	0/1/1/1	0/0/0/0
5	EDO	A	1363	-	-	0/1/1/1	0/0/0/0
5	EDO	A	1364	-	-	0/1/1/1	0/0/0/0
5	EDO	A	1365	-	-	0/1/1/1	0/0/0/0
5	EDO	A	3001	-	-	0/1/1/1	0/0/0/0
3	PYR	B	1354	2	-	0/0/4/4	0/0/0/0
3	PYR	B	1355	-	-	0/0/4/4	0/0/0/0
4	DMS	B	1356	-	-	0/0/0/0	0/0/0/0
5	EDO	B	1357	-	-	0/1/1/1	0/0/0/0
5	EDO	B	1358	-	-	0/1/1/1	0/0/0/0
5	EDO	B	1359	-	-	0/1/1/1	0/0/0/0
5	EDO	B	1360	-	-	0/1/1/1	0/0/0/0
5	EDO	B	1361	-	-	0/1/1/1	0/0/0/0
5	EDO	B	1362	-	-	0/1/1/1	0/0/0/0
5	EDO	B	1363	-	-	0/1/1/1	0/0/0/0
5	EDO	B	1364	-	-	0/1/1/1	0/0/0/0
7	1KA	B	1365	-	-	0/3/4/4	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	B	1365	1KA	C3-C4	2.12	1.60	1.49
7	B	1365	1KA	O1-C1	4.25	1.45	1.19

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	B	1365	1KA	O1-C1-C2	-5.83	107.11	126.38

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

16 monomers are involved in 32 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1355	PYR	2	0
3	A	1356	PYR	2	0
4	A	1357	DMS	2	0
5	A	1358	EDO	5	0
5	A	1359	EDO	2	0
5	A	1362	EDO	1	0
5	A	1365	EDO	2	0
3	B	1354	PYR	3	0
3	B	1355	PYR	3	0
4	B	1356	DMS	2	0
5	B	1357	EDO	4	0
5	B	1358	EDO	2	0
5	B	1361	EDO	1	0
5	B	1363	EDO	2	0
5	B	1364	EDO	3	0
7	B	1365	1KA	1	0

## 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	351/381 (92%)	0.54	35 (9%) <b>7</b> <b>9</b>	27, 45, 71, 90	0
1	B	344/381 (90%)	0.61	40 (11%) <b>4</b> <b>6</b>	28, 45, 67, 87	0
All	All	695/762 (91%)	0.57	75 (10%) <b>6</b> <b>7</b>	27, 45, 70, 90	0

All (75) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	353	PHE	8.6
1	A	8	LEU	5.5
1	A	7	THR	5.5
1	A	354	LEU	5.4
1	B	311	ASP	5.3
1	B	166	ILE	5.0
1	B	351	ALA	4.8
1	A	311	ASP	4.4
1	A	269	ILE	4.3
1	B	162	LYS	4.0
1	A	5	SER	4.0
1	A	6	GLU	4.0
1	A	351	ALA	3.9
1	A	270	THR	3.9
1	B	269	ILE	3.9
1	B	112	SER	3.8
1	A	291	PHE	3.7
1	A	197	ILE	3.6
1	B	195	TYR	3.6
1	B	45	ALA	3.6
1	A	4	GLU	3.5
1	B	163	GLU	3.5
1	B	270	THR	3.4
1	A	195	TYR	3.4

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	A	196	SER	3.4
1	A	51	LYS	3.3
1	A	144	HIS	3.2
1	B	350	ALA	3.2
1	B	44	LEU	3.1
1	B	353	PHE	3.1
1	B	194	LEU	3.1
1	B	310	LYS	3.1
1	B	271	PHE	3.0
1	B	345	LEU	3.0
1	B	352	GLU	3.0
1	A	111	TYR	2.9
1	B	31	ILE	2.9
1	B	268	MET	2.9
1	B	196	SER	2.9
1	B	312	MET	2.8
1	B	208	TRP	2.8
1	B	197	ILE	2.8
1	A	194	LEU	2.7
1	A	271	PHE	2.7
1	A	352	GLU	2.6
1	B	60	ASP	2.6
1	B	291	PHE	2.6
1	A	112	SER	2.5
1	A	350	ALA	2.5
1	B	164	SER	2.5
1	B	338	ASN	2.5
1	B	10	PRO	2.4
1	A	104	ASP	2.4
1	A	185	PHE	2.4
1	A	44	LEU	2.4
1	B	111	TYR	2.4
1	B	144	HIS	2.4
1	B	277	ALA	2.3
1	A	208	TRP	2.3
1	B	185	PHE	2.2
1	B	187	TRP	2.2
1	A	198	ASN	2.2
1	A	20	THR	2.2
1	A	338	ASN	2.1
1	A	292	ALA	2.1
1	B	276	HIS	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	45	ALA	2.1
1	B	292	ALA	2.1
1	A	274	GLY	2.1
1	B	274	GLY	2.1
1	A	334	ALA	2.0
1	B	66	VAL	2.0
1	B	186	ALA	2.0
1	B	272	PRO	2.0
1	A	59	TYR	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. 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( $\text{\AA}^2$ )	Q<0.9
5	EDO	B	1364	4/4	0.38	0.66	54,56,59,66	0
5	EDO	A	3001	4/4	0.64	0.29	68,71,73,80	0
3	PYR	A	1356	6/6	0.67	0.64	45,46,49,51	6
5	EDO	B	1360	4/4	0.71	0.35	42,50,51,57	0
5	EDO	A	1365	4/4	0.72	0.37	56,57,59,62	0
5	EDO	B	1361	4/4	0.75	0.19	58,66,71,72	0
3	PYR	B	1355	6/6	0.76	0.64	43,46,49,50	6
7	1KA	B	1365	7/7	0.78	0.21	57,64,66,66	0
5	EDO	A	1362	4/4	0.81	0.43	53,64,66,66	0
5	EDO	B	1362	4/4	0.85	0.28	41,44,53,62	0
5	EDO	A	1363	4/4	0.85	0.46	53,56,60,63	0
5	EDO	A	1360	4/4	0.85	0.44	50,52,62,67	0
5	EDO	A	1364	4/4	0.90	0.20	55,61,62,65	0
5	EDO	A	1361	4/4	0.90	0.36	39,39,47,60	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	PYR	A	1355	6/6	0.90	0.26	33,39,43,44	6
5	EDO	A	1358	4/4	0.91	0.21	38,41,43,46	0
5	EDO	B	1358	4/4	0.92	0.24	46,50,53,67	0
5	EDO	B	1359	4/4	0.92	0.22	37,43,45,54	0
5	EDO	B	1363	4/4	0.92	0.28	49,54,59,59	0
5	EDO	B	1357	4/4	0.94	0.20	35,35,37,47	0
5	EDO	A	1359	4/4	0.95	0.25	39,42,42,45	0
4	DMS	B	1356	4/4	0.96	0.13	51,54,58,58	0
3	PYR	B	1354	6/6	0.96	0.19	29,37,41,43	0
4	DMS	A	1357	4/4	0.96	0.14	42,52,54,56	0
6	ZN	B	502	1/1	0.99	0.06	40,40,40,40	0
2	NI	B	501	1/1	0.99	0.12	42,42,42,42	0
6	ZN	A	3000	1/1	0.99	0.06	34,34,34,34	0
2	NI	A	501	1/1	1.00	0.10	42,42,42,42	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.