



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 30, 2018 – 02:21 PM EDT

PDB ID : 6C8T  
Title : The structure of MppP soaked with the substrate L-Arg  
Authors : Han, L.; Silvaggi, N.R.  
Deposited on : 2018-01-25  
Resolution : 2.20 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 : rb-20031172  
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) : rb-20031172

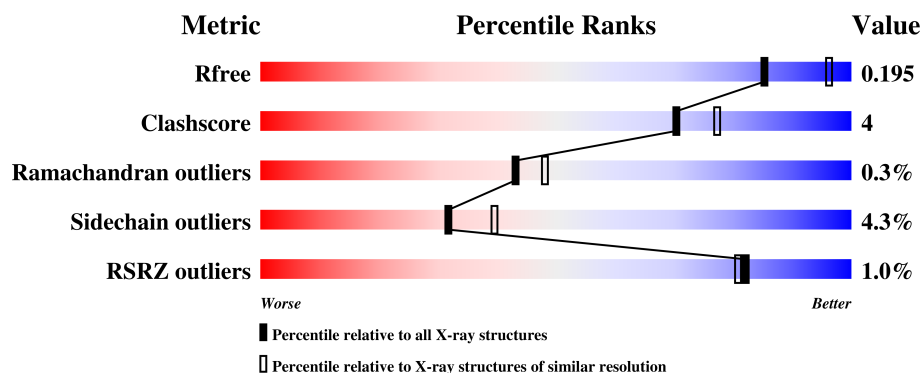
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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	4343 (2.20-2.20)
Clashscore	122126	5027 (2.20-2.20)
Ramachandran outliers	120053	4952 (2.20-2.20)
Sidechain outliers	120020	4953 (2.20-2.20)
RSRZ outliers	108989	4245 (2.20-2.20)

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	376	<div> <div>%</div> <div> <div></div> <div>89%</div> <div>7%</div> <div>..</div> </div> </div>
1	B	376	<div> <div>%</div> <div> <div></div> <div>85%</div> <div>12%</div> <div>..</div> </div> </div>
1	C	376	<div> <div>%</div> <div> <div></div> <div>86%</div> <div>11%</div> <div>..</div> </div> </div>
2	D	376	<div> <div>%</div> <div> <div></div> <div>86%</div> <div>7%</div> <div>•</div> <div>6%</div> </div> </div>

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 23402 atoms, of which 11215 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.

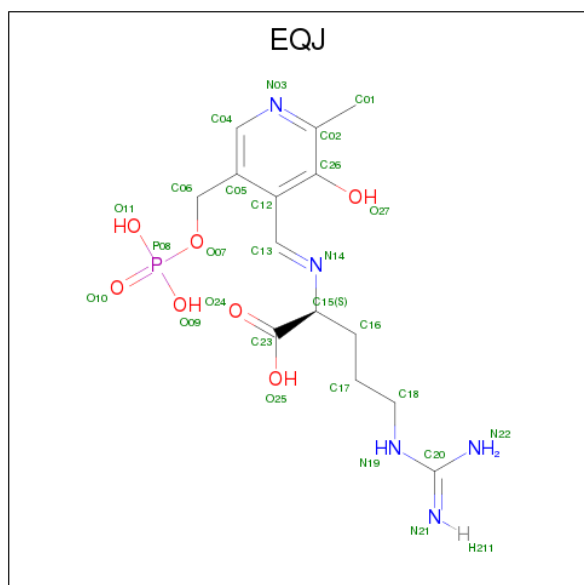
- Molecule 1 is a protein called PLP-Dependent L-Arginine Hydroxylase MppP.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
1	A	368	Total	C	H	N	O	S		6	0	0
			5665	1800	2806	506	546	7				
1	B	368	Total	C	H	N	O	S		6	5	0
			5706	1811	2828	509	551	7				
1	C	368	Total	C	H	N	O	S		6	2	0
			5677	1803	2812	507	548	7				

- Molecule 2 is a protein called PLP-Dependent L-Arginine Hydroxylase MppP.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
2	D	354	Total	C	H	N	O	P	S	0	5	0
			5490	1743	2715	491	533	1	7			

- Molecule 3 is (E)-N 2 -(3-hydroxy-2-methyl-5-[(phosphonooxy)methyl]pyridin-4-yl)methylidene)-L-arginine (three-letter code: EQJ) (formula: C<sub>14</sub>H<sub>22</sub>N<sub>5</sub>O<sub>7</sub>P).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	A	1	Total	C	H	N	O	P	0	0
			45	14	18	5	7	1		
3	B	1	Total	C	H	N	O	P	0	0
			45	14	18	5	7	1		
3	C	1	Total	C	H	N	O	P	0	0
			45	14	18	5	7	1		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	1	Total	Cl	0	0
			1	1		

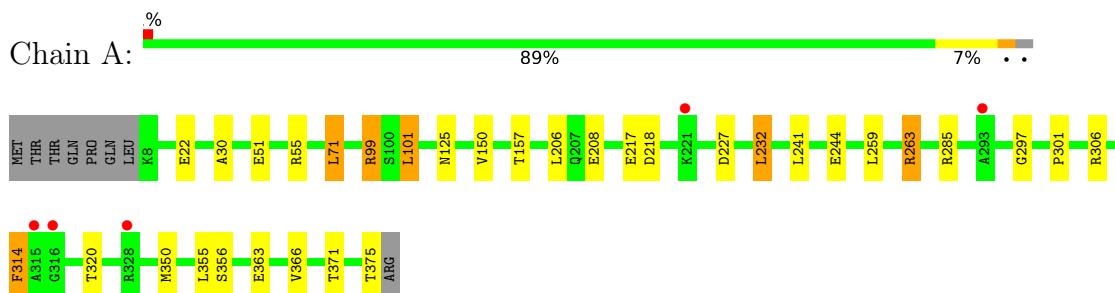
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	197	Total	O	0	0
			197	197		
5	B	187	Total	O	0	0
			187	187		
5	C	169	Total	O	0	0
			169	169		
5	D	175	Total	O	0	0
			175	175		

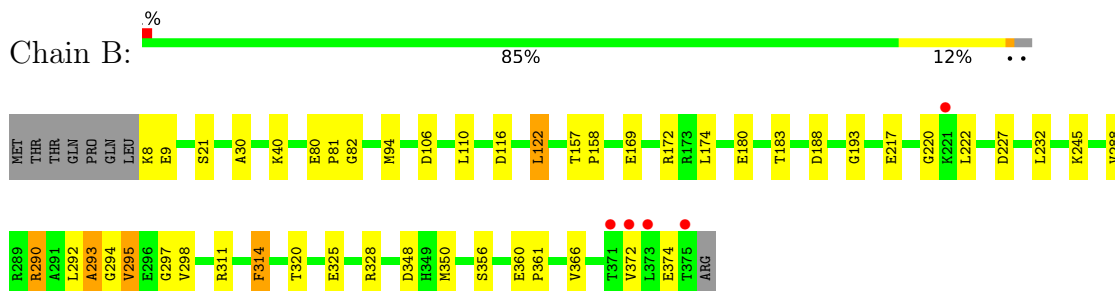
### 3 Residue-property plots [i](#)

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.

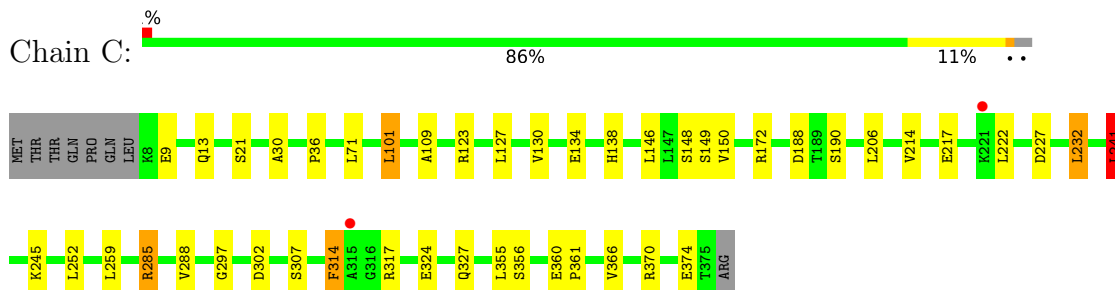
- Molecule 1: PLP-Dependent L-Arginine Hydroxylase MppP



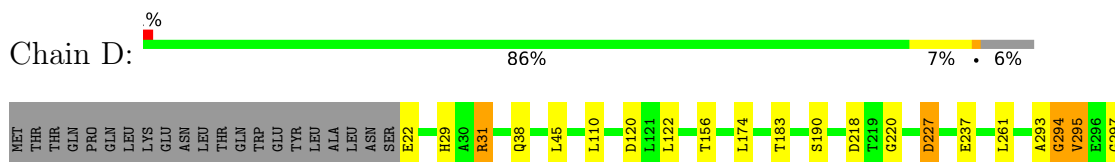
- Molecule 1: PLP-Dependent L-Arginine Hydroxylase MppP

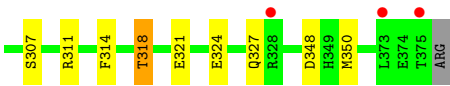


- Molecule 1: PLP-Dependent L-Arginine Hydroxylase MppP



- Molecule 2: PLP-Dependent L-Arginine Hydroxylase MppP





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	85.79Å 108.42Å 196.22Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.05 – 2.20 49.05 – 2.20	Depositor EDS
% Data completeness (in resolution range)	91.6 (49.05-2.20) 91.6 (49.05-2.20)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.78 (at 2.20Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.146 , 0.195 0.146 , 0.195	Depositor DCC
$R_{free}$ test set	1999 reflections (2.33%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.2	Xtriage
Anisotropy	0.355	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.43 , 57.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	23402	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.19% 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: EQJ, LLP, CL

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	0.48	0/2920	0.71	8/3976 (0.2%)
1	B	0.47	0/2957	0.65	1/4026 (0.0%)
1	C	0.51	1/2932 (0.0%)	0.69	4/3992 (0.1%)
2	D	0.49	0/2826	0.65	0/3847
All	All	0.49	1/11635 (0.0%)	0.68	13/15841 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	134	GLU	CB-CG	-5.29	1.42	1.52

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	263	ARG	NE-CZ-NH2	-8.96	115.82	120.30
1	A	263	ARG	NE-CZ-NH1	8.70	124.65	120.30
1	C	285	ARG	NE-CZ-NH2	-7.95	116.32	120.30
1	A	55	ARG	NE-CZ-NH2	-6.96	116.82	120.30
1	A	285	ARG	NE-CZ-NH2	-6.73	116.94	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	2859	2806	2806	13	0
1	B	2878	2828	2812	34	0
1	C	2865	2812	2806	25	0
2	D	2775	2715	2702	16	0
3	A	27	18	0	0	0
3	B	27	18	0	1	0
3	C	27	18	0	1	0
4	D	1	0	0	0	0
5	A	197	0	0	2	0
5	B	187	0	0	12	1
5	C	169	0	0	10	1
5	D	175	0	0	6	2
All	All	12187	11215	11126	87	2

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 4.

The worst 5 of 87 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:C:370:ARG:NH1	1:C:374:GLU:OE2	1.99	0.95
2:D:220:GLY:O	5:D:501:HOH:O	1.90	0.88
1:C:138:HIS:NE2	5:C:503:HOH:O	2.06	0.88
1:B:180:GLU:OE1	5:B:501:HOH:O	1.92	0.86
1:B:220:GLY:O	5:B:502:HOH:O	1.94	0.84

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:B:681:HOH:O	5:D:664:HOH:O[1_455]	1.98	0.22
5:C:506:HOH:O	5:D:625:HOH:O[4_497]	2.15	0.05

## 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	366/376 (97%)	360 (98%)	6 (2%)	0	100	100
1	B	371/376 (99%)	359 (97%)	10 (3%)	2 (0%)	31	33
1	C	368/376 (98%)	360 (98%)	8 (2%)	0	100	100
2	D	356/376 (95%)	347 (98%)	6 (2%)	3 (1%)	21	20
All	All	1461/1504 (97%)	1426 (98%)	30 (2%)	5 (0%)	43	48

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	293	ALA
1	B	295	VAL
2	D	293	ALA
2	D	294	GLY
2	D	295	VAL

### 5.3.2 Protein sidechains ⓘ

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	307/315 (98%)	293 (95%)	14 (5%)	29	37
1	B	312/315 (99%)	301 (96%)	11 (4%)	39	49
1	C	309/315 (98%)	294 (95%)	15 (5%)	27	33
2	D	298/314 (95%)	286 (96%)	12 (4%)	34	43
All	All	1226/1259 (97%)	1174 (96%)	52 (4%)	32	41

5 of 52 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	314	PHE
1	C	150	VAL
2	D	307	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	348	ASP
1	C	71	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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
2	LLP	D	221	2	24,24,25	1.44	3 (12%)	28,32,34	1.70	7 (25%)

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
2	LLP	D	221	2	-	0/15/17/19	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	221	LLP	C6-C5	2.00	1.41	1.37
2	D	221	LLP	O-C	3.80	1.35	1.19
2	D	221	LLP	C3-C2	4.79	1.44	1.40

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	221	LLP	OP2-P-OP4	-4.43	94.95	106.73
2	D	221	LLP	C2'-C2-C3	-2.60	117.87	120.96
2	D	221	LLP	O-C-CA	-2.18	118.98	124.96
2	D	221	LLP	CB-CA-C	-2.04	107.96	111.85
2	D	221	LLP	OP2-P-OP1	2.72	121.23	110.60

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 1 is monoatomic - leaving 3 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	EQJ	A	401	-	24,27,27	2.16	5 (20%)	30,37,37	1.81	5 (16%)
3	EQJ	B	401	-	24,27,27	2.29	6 (25%)	30,37,37	1.25	3 (10%)
3	EQJ	C	401	-	24,27,27	2.40	7 (29%)	30,37,37	1.72	6 (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	EQJ	A	401	-	-	1/18/22/22	0/1/1/1
3	EQJ	B	401	-	-	0/18/22/22	0/1/1/1
3	EQJ	C	401	-	-	0/18/22/22	0/1/1/1

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	401	EQJ	C26-C02	-3.09	1.38	1.40
3	B	401	EQJ	C26-C02	-3.03	1.38	1.40
3	B	401	EQJ	C06-C05	2.33	1.57	1.50
3	C	401	EQJ	C06-C05	2.49	1.57	1.50
3	B	401	EQJ	C20-N21	2.60	1.42	1.32

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	401	EQJ	C05-C04-N03	-2.45	119.68	123.83
3	C	401	EQJ	C05-C04-N03	-2.38	119.80	123.83
3	A	401	EQJ	C05-C04-N03	-2.37	119.82	123.83
3	B	401	EQJ	C01-C02-C26	-2.25	118.28	120.96
3	C	401	EQJ	C04-N03-C02	2.06	123.16	119.19

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	401	EQJ	C16-C15-N14-C13

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	401	EQJ	1	0
3	C	401	EQJ	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

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, 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	368/376 (97%)	-0.48	5 (1%) 75 73	14, 28, 63, 110	1 (0%)
1	B	368/376 (97%)	-0.56	5 (1%) 75 73	14, 27, 68, 120	1 (0%)
1	C	368/376 (97%)	-0.55	2 (0%) 90 90	14, 28, 63, 98	1 (0%)
2	D	353/376 (93%)	-0.54	3 (0%) 86 85	14, 27, 71, 120	0
All	All	1457/1504 (96%)	-0.53	15 (1%) 82 81	14, 27, 68, 120	3 (0%)

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	375	THR	5.7
1	A	315	ALA	4.8
1	C	221	LYS	4.2
2	D	375	THR	3.4
2	D	373	LEU	3.1

### 6.2 Non-standard residues in protein, DNA, RNA chains [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
2	LLP	D	221	24/25	0.99	0.10	12,24,32,36	0

### 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(Å <sup>2</sup> )	Q<0.9
4	CL	D	401	1/1	0.97	0.15	35,35,35,35	0
3	EQJ	B	401	27/27	0.97	0.13	14,29,43,50	0
3	EQJ	A	401	27/27	0.98	0.15	17,30,46,54	0
3	EQJ	C	401	27/27	0.98	0.14	17,29,49,59	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.