



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 14, 2018 – 02:35 am GMT

PDB ID : 1M6X  
Title : Flpe-Holliday Junction Complex  
Authors : Conway, A.B.; Chen, Y.; Rice, P.A.  
Deposited on : 2002-07-17  
Resolution : 2.80 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	:	trunk31020
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)	:	trunk31020

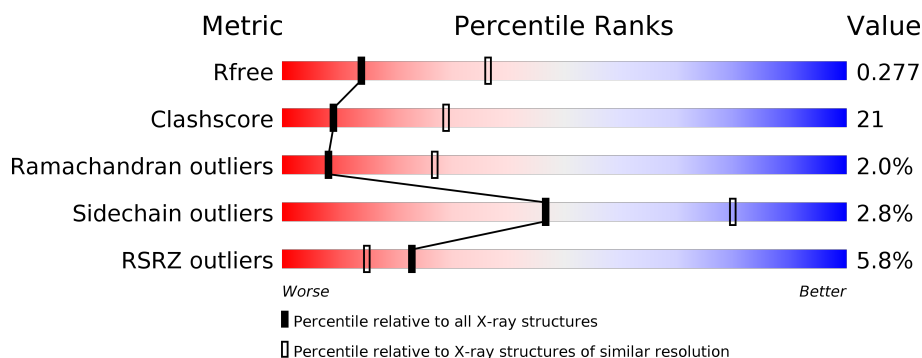
# 1 Overall quality at a glance

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))
$R_{free}$	111664	2792 (2.80-2.80)
Clashscore	122126	3209 (2.80-2.80)
Ramachandran outliers	120053	3158 (2.80-2.80)
Sidechain outliers	120020	3160 (2.80-2.80)
RSRZ outliers	108989	2726 (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	E	13	<div> <div>15%</div> <div>85%</div> </div>
1	F	13	<div> <div>15%</div> <div>77%</div> <div>8%</div> </div>
2	I	20	<div> <div>15%</div> <div>80%</div> <div>5%</div> </div>
2	J	20	<div> <div>30%</div> <div>70%</div> </div>
3	G	33	<div> <div>27%</div> <div>70%</div> <div>.</div> </div>
3	H	33	<div> <div>21%</div> <div>79%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	A	423	<div><div></div><div>14%</div><div>58%</div><div>34%</div><div>6%</div></div>
4	B	423	<div><div></div><div>%</div><div>62%</div><div>30%</div><div>5%</div></div>
5	C	423	<div><div></div><div>8%</div><div>58%</div><div>34%</div><div>5%</div></div>
5	D	423	<div><div></div><div></div><div>62%</div><div>30%</div><div>5%</div></div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 15641 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 DNA chain called Symmetrized FRT site.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	13	Total	C	N	O	P	0	0	0
			259	127	41	79	12			
1	F	13	Total	C	N	O	P	0	0	0
			259	127	41	79	12			

- Molecule 2 is a DNA chain called Symmetrized FRT site.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	I	19	Total	C	N	O	P	0	0	0
			391	189	75	109	18			
2	J	20	Total	C	N	O	P	0	0	0
			395	189	75	112	19			

- Molecule 3 is a DNA chain called Symmetrized FRT site.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	33	Total	C	N	O	P	0	0	0
			656	316	116	192	32			
3	H	33	Total	C	N	O	P	0	0	0
			657	316	116	193	32			

- Molecule 4 is a protein called Flp recombinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	399	Total	C	N	O	S	0	0	0
			3236	2083	545	597	11			
4	B	401	Total	C	N	O	S	0	0	0
			3256	2096	548	601	11			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2	SER	PRO	ENGINEERED	UNP P03870
A	33	SER	LEU	ENGINEERED	UNP P03870
A	108	ASN	TYR	ENGINEERED	UNP P03870
A	294	PRO	SER	ENGINEERED	UNP P03870
B	2	SER	PRO	ENGINEERED	UNP P03870
B	33	SER	LEU	ENGINEERED	UNP P03870
B	108	ASN	TYR	ENGINEERED	UNP P03870
B	294	PRO	SER	ENGINEERED	UNP P03870

- Molecule 5 is a protein called Flp recombinase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
5	C	400	Total	C	N	O	P	S	0	0	0
			3266	2102	552	600	1	11			
5	D	400	Total	C	N	O	P	S	0	0	0
			3266	2102	552	600	1	11			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	2	SER	PRO	ENGINEERED	UNP P03870
C	33	SER	LEU	ENGINEERED	UNP P03870
C	108	ASN	TYR	ENGINEERED	UNP P03870
C	294	PRO	SER	ENGINEERED	UNP P03870
C	343	PTR	TYR	MODIFIED RESIDUE	UNP P03870
D	2	SER	PRO	ENGINEERED	UNP P03870
D	33	SER	LEU	ENGINEERED	UNP P03870
D	108	ASN	TYR	ENGINEERED	UNP P03870
D	294	PRO	SER	ENGINEERED	UNP P03870
D	343	PTR	TYR	MODIFIED RESIDUE	UNP P03870

### 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.

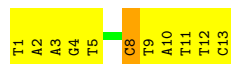
- Molecule 1: Symmetrized FRT site

Chain E: 



- Molecule 1: Symmetrized FRT site

Chain F: 

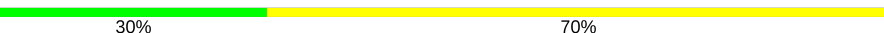


- Molecule 2: Symmetrized FRT site

Chain I: 



- Molecule 2: Symmetrized FRT site

Chain J: 



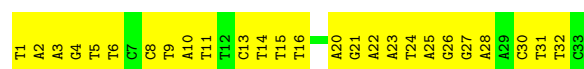
- Molecule 3: Symmetrized FRT site

Chain G: 

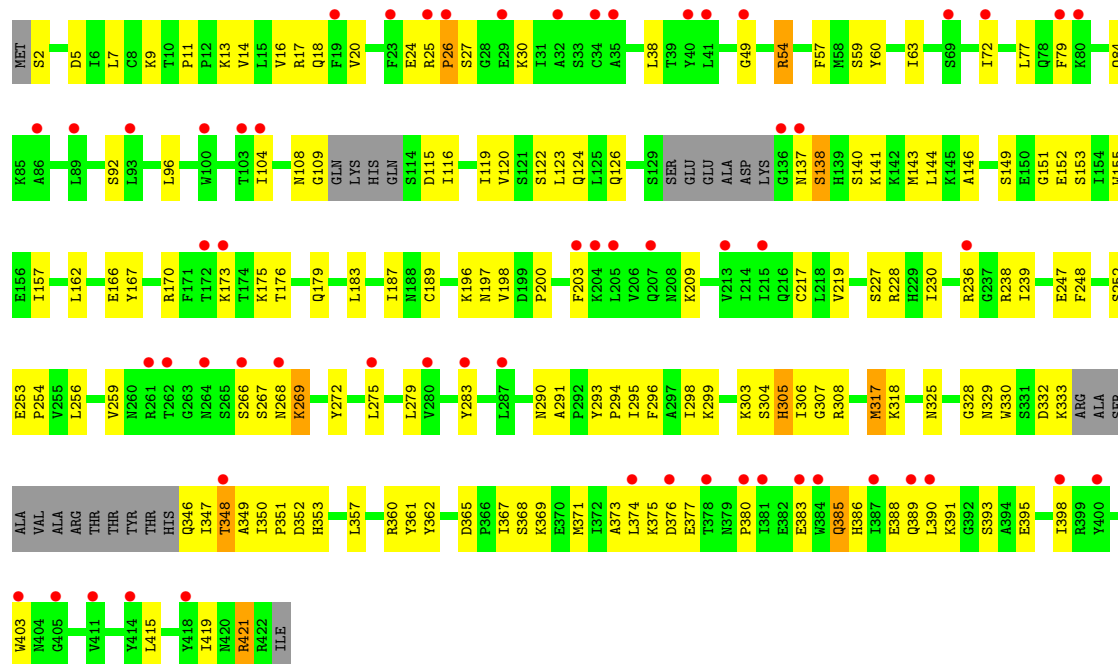


- Molecule 3: Symmetrized FRT site

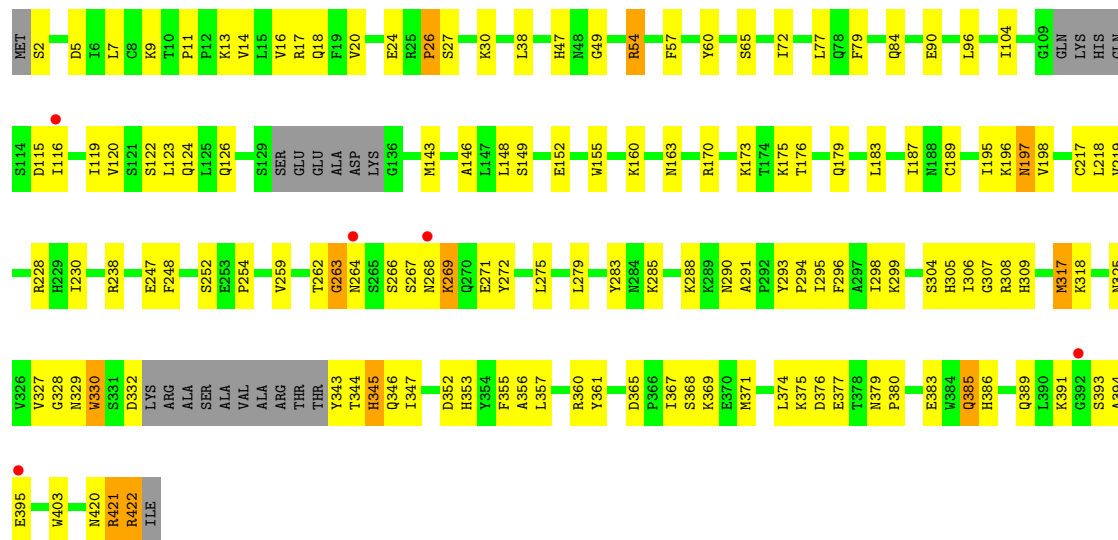
Chain H: 



• Molecule 4: Flp recombinase

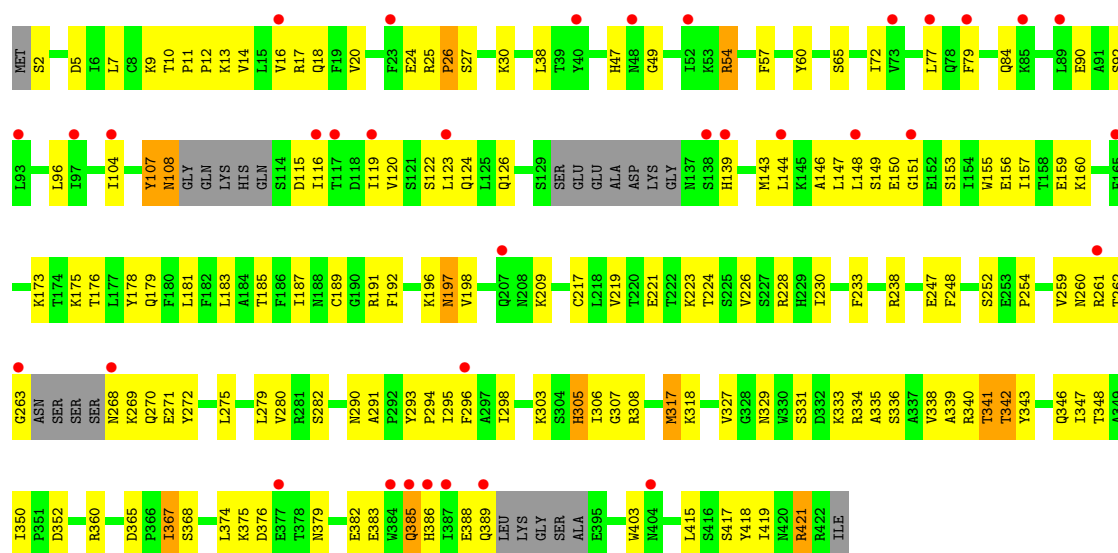


• Molecule 4: Flp recombinase



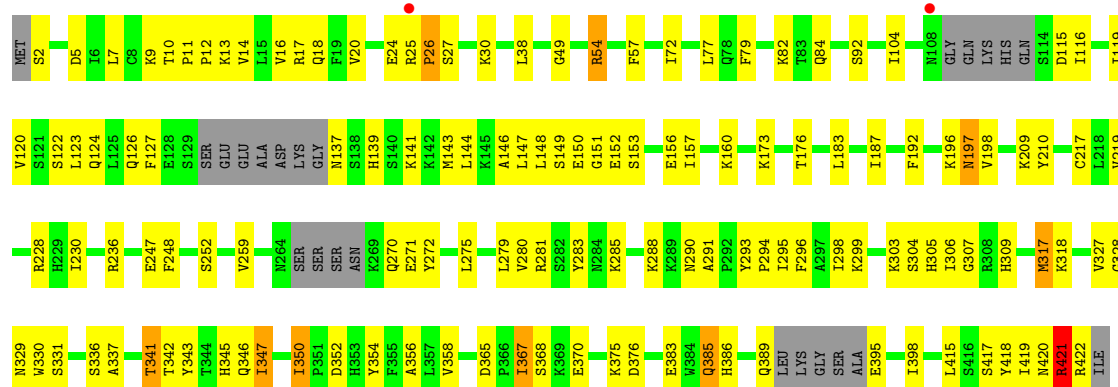
• Molecule 5: Flp recombinase





### • Molecule 5: Flp recombinase

Chain D: 62% 30% 5%





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	79.78Å 116.51Å 142.43Å 90.00° 97.26° 90.00°	Depositor
Resolution (Å)	37.47 – 2.80 89.89 – 2.78	Depositor EDS
% Data completeness (in resolution range)	75.5 (37.47-2.80) 74.2 (89.89-2.78)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.09 (at 2.77Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.239 , 0.281 0.235 , 0.277	Depositor DCC
$R_{free}$ test set	4855 reflections (9.31%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	44.7	Xtriage
Anisotropy	0.467	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 53.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	15641	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.39% 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: PTR

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	E	0.51	0/288	0.90	0/442
1	F	0.70	0/288	0.92	0/442
2	I	0.45	0/440	0.94	1/678 (0.1%)
2	J	0.58	0/444	0.95	0/685
3	G	0.56	0/735	0.86	0/1134
3	H	0.55	0/736	0.89	0/1135
4	A	0.36	0/3306	0.58	0/4463
4	B	0.44	0/3328	0.64	1/4495 (0.0%)
5	C	0.37	0/3320	0.60	0/4482
5	D	0.44	0/3320	0.64	1/4482 (0.0%)
All	All	0.43	0/16205	0.69	3/22438 (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	F	0	1
3	G	0	1
All	All	0	2

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	345	HIS	N-CA-C	-5.81	95.31	111.00
5	D	350	ILE	N-CA-C	-5.68	95.67	111.00
2	I	16	DT	C1'-O4'-C4'	-5.03	105.07	110.10

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	F	8	DC	Sidechain
3	G	23	DA	Sidechain

## 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	E	259	0	150	22	0
1	F	259	0	150	15	0
2	I	391	0	217	29	0
2	J	395	0	216	26	0
3	G	656	0	362	34	0
3	H	657	0	365	51	0
4	A	3236	0	3271	156	0
4	B	3256	0	3281	131	0
5	C	3266	0	3296	138	0
5	D	3266	0	3296	127	0
All	All	15641	0	14604	648	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 21.

The worst 5 of 648 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:E:4:DG:H2''	1:E:5:DT:H5''	1.26	1.09
3:H:8:DC:H2'	3:H:9:DT:H71	1.33	1.05
2:J:25:DA:H2''	2:J:26:DG:H5'	1.38	1.05
2:I:25:DA:H2''	2:I:26:DG:H5'	1.38	1.04
1:E:8:DC:OP2	4:A:2:SER:HB2	1.65	0.96

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	
4	A	391/423 (92%)	330 (84%)	52 (13%)	9 (2%)	7	23
4	B	393/423 (93%)	339 (86%)	47 (12%)	7 (2%)	9	30
5	C	389/423 (92%)	342 (88%)	38 (10%)	9 (2%)	7	23
5	D	389/423 (92%)	340 (87%)	43 (11%)	6 (2%)	11	36
All	All	1562/1692 (92%)	1351 (86%)	180 (12%)	31 (2%)	8	27

5 of 31 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	152	GLU
4	A	269	LYS
4	A	375	LYS
4	B	269	LYS
4	B	375	LYS

### 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	
4	A	363/383 (95%)	355 (98%)	8 (2%)	55	85
4	B	365/383 (95%)	353 (97%)	12 (3%)	41	75
5	C	364/382 (95%)	353 (97%)	11 (3%)	44	78
5	D	364/382 (95%)	354 (97%)	10 (3%)	48	81
All	All	1456/1530 (95%)	1415 (97%)	41 (3%)	47	80

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	B	421	ARG
5	C	108	ASN
5	D	352	ASP
4	B	422	ARG
5	C	54	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
4	B	197	ASN
5	C	61	ASN
5	D	197	ASN
4	B	264	ASN
4	B	268	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues 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	$\# Z  > 2$
5	PTR	C	343	1,5	12,15,17	1.47	1 (8%)	15,19,24	1.11	0
5	PTR	D	343	1,5	12,15,17	1.75	5 (41%)	15,19,24	1.05	1 (6%)

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
5	PTR	C	343	1,5	-	0/6/10/13	0/1/1/1
5	PTR	D	343	1,5	-	0/6/10/13	0/1/1/1

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	D	343	PTR	CE1-CZ	2.00	1.42	1.38
5	D	343	PTR	CE1-CD1	2.14	1.42	1.38
5	D	343	PTR	CA-C	2.21	1.53	1.50
5	D	343	PTR	CE2-CD2	2.22	1.42	1.38
5	C	343	PTR	OH-CZ	2.52	1.44	1.40

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	D	343	PTR	CD2-CE2-CZ	-2.23	116.98	119.73

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	343	PTR	4	0
5	D	343	PTR	2	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 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(Å²)	Q<0.9	
1	E	13/13 (100%)	-0.14	0	100	100	47, 82, 100, 101	0
1	F	13/13 (100%)	-0.49	0	100	100	22, 25, 40, 47	0
2	I	19/20 (95%)	-0.42	0	100	100	43, 65, 88, 90	0
2	J	20/20 (100%)	-0.67	0	100	100	14, 36, 63, 77	0
3	G	33/33 (100%)	-0.53	0	100	100	14, 61, 84, 88	0
3	H	33/33 (100%)	-0.38	0	100	100	20, 52, 114, 117	0
4	A	399/423 (94%)	0.86	59 (14%)	2	1	46, 96, 121, 121	0
4	B	401/423 (94%)	-0.10	5 (1%)	79	72	10, 37, 90, 108	0
5	C	399/423 (94%)	0.42	35 (8%)	10	5	39, 74, 115, 121	0
5	D	399/423 (94%)	-0.19	2 (0%)	90	88	12, 38, 70, 98	0
All	All	1729/1824 (94%)	0.20	101 (5%)	23	15	10, 59, 115, 121	0

The worst 5 of 101 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	B	392	GLY	6.9
5	C	268	ASN	5.1
4	A	387	ILE	5.1
4	A	137	ASN	4.7
4	A	266	SER	4.5

### 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( $\text{\AA}^2$ )	Q<0.9
5	PTR	D	343	15/17	0.94	0.13	52,57,62,62	0
5	PTR	C	343	15/17	0.96	0.13	31,33,37,40	0

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.