



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 28, 2014 – 09:04 PM GMT

PDB ID : 2IRF  
Title : CRYSTAL STRUCTURE OF AN IRF-2/DNA COMPLEX.  
Authors : Fujii, Y.; Shimizu, T.; Kusumoto, M.; Kyogoku, Y.; Taniguchi, T.; Hakoshima, T.  
Deposited on : 1999-05-30  
Resolution : 2.20 Å(reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

---

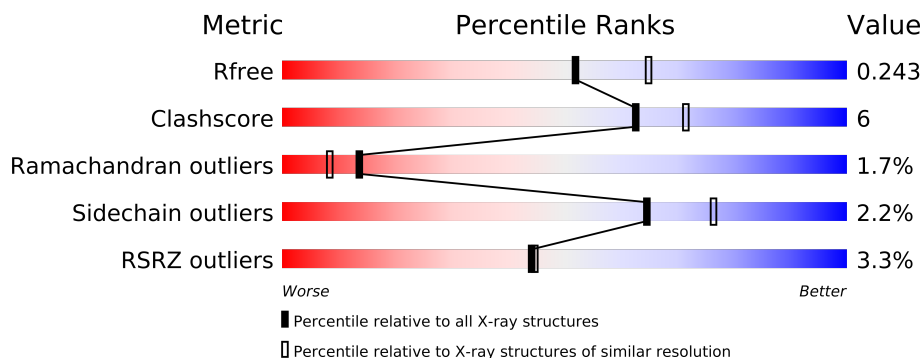
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

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))
$R_{free}$	66092	2938 (2.20-2.20)
Clashscore	79885	3751 (2.20-2.20)
Ramachandran outliers	78287	3681 (2.20-2.20)
Sidechain outliers	78261	3682 (2.20-2.20)
RSRZ outliers	66119	2939 (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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	12	
1	B	12	
1	C	12	
2	D	13	
2	E	13	
2	F	13	
3	G	113	
3	H	113	
3	I	113	
3	J	113	
3	K	113	
3	L	113	

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 7397 atoms, of which 0 are hydrogen and 0 are deuterium.

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 DNA (5'-D(P\*AP\*AP\*GP\*TP\*GP\*AP\*AP\*AP\*GP\*(5IU)P\*GP\*A)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	12	Total	C	I	N	O	P	0	0	0
			254	119	1	54	68	12			
1	B	12	Total	C	I	N	O	P	0	0	0
			254	119	1	54	68	12			
1	C	12	Total	C	I	N	O	P	0	0	0
			254	119	1	54	68	12			

- Molecule 2 is a DNA chain called DNA (5'-D(\*TP\*TP\*CP\*AP\*CP\*TP\*TP\*TP\*CP\*AP\*CP\*(5IU)P\*T)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	D	13	Total	C	I	N	O	P	0	0	0
			255	125	1	36	81	12			
2	E	12	Total	C	I	N	O	P	0	0	0
			238	115	1	34	76	12			
2	F	13	Total	C	I	N	O	P	0	0	0
			255	125	1	36	81	12			

- Molecule 3 is a protein called INTERFERON REGULATORY FACTOR 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	109	Total	C	N	O	S	0	0	0
			911	583	170	152	6			
3	H	109	Total	C	N	O	S	0	0	0
			917	586	173	152	6			
3	I	109	Total	C	N	O	S	0	0	0
			911	583	170	152	6			
3	J	109	Total	C	N	O	S	0	0	0
			917	586	173	152	6			
3	K	109	Total	C	N	O	S	0	0	0
			917	586	173	152	6			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	109	Total	C	N	O	S	0	0	0
			911	583	170	152	6			

- Molecule 4 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	G	1	Total	K	0	0
			1	1		
4	J	1	Total	K	0	0
			1	1		
4	K	1	Total	K	0	0
			1	1		
4	H	1	Total	K	0	0
			1	1		
4	I	1	Total	K	0	0
			1	1		
4	L	1	Total	K	0	0
			1	1		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	21	Total	O	0	0
			21	21		
5	B	20	Total	O	0	0
			20	20		
5	C	15	Total	O	0	0
			15	15		
5	D	26	Total	O	0	0
			26	26		
5	E	24	Total	O	0	0
			24	24		
5	F	29	Total	O	0	0
			29	29		
5	G	63	Total	O	0	0
			63	63		
5	H	47	Total	O	0	0
			47	47		
5	I	47	Total	O	0	0
			47	47		
5	J	14	Total	O	0	0
			14	14		

*Continued on next page...*

*Continued from previous page...*

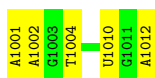
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	K	55	Total 55	O 55	0	0
5	L	36	Total 36	O 36	0	0

### 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 errors 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: DNA (5'-D(P\*AP\*AP\*GP\*TP\*GP\*AP\*AP\*AP\*GP\*(5IU)P\*GP\*A)-3')

Chain A: 



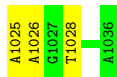
- Molecule 1: DNA (5'-D(P\*AP\*AP\*GP\*TP\*GP\*AP\*AP\*AP\*GP\*(5IU)P\*GP\*A)-3')

Chain B: 



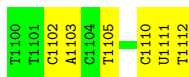
- Molecule 1: DNA (5'-D(P\*AP\*AP\*GP\*TP\*GP\*AP\*AP\*AP\*GP\*(5IU)P\*GP\*A)-3')

Chain C: 



- Molecule 2: DNA (5'-D(\*TP\*TP\*CP\*AP\*CP\*TP\*TP\*TP\*CP\*AP\*CP\*(5IU)P\*T)-3')

Chain D: 



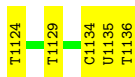
- Molecule 2: DNA (5'-D(\*TP\*TP\*CP\*AP\*CP\*TP\*TP\*TP\*CP\*AP\*CP\*(5IU)P\*T)-3')

Chain E: 



- Molecule 2: DNA (5'-D(\*TP\*TP\*CP\*AP\*CP\*TP\*TP\*TP\*CP\*AP\*CP\*(5IU)P\*T)-3')

Chain F: 



- Molecule 3: INTERFERON REGULATORY FACTOR 2

Chain G:



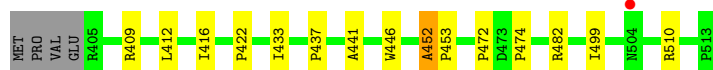
- Molecule 3: INTERFERON REGULATORY FACTOR 2

Chain H:



- Molecule 3: INTERFERON REGULATORY FACTOR 2

Chain I:



- Molecule 3: INTERFERON REGULATORY FACTOR 2

Chain J:



- Molecule 3: INTERFERON REGULATORY FACTOR 2

Chain K:



- Molecule 3: INTERFERON REGULATORY FACTOR 2

Chain L:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	132.30Å 132.30Å 296.50Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	10.00 – 2.20 56.25 – 2.20	Depositor EDS
% Data completeness (in resolution range)	88.4 (10.00-2.20) 90.5 (56.25-2.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.89 (at 2.20Å)	Xtriage
Refinement program	X-PLOR 3.851	Depositor
R, $R_{free}$	0.202 , 0.243 0.206 , 0.243	Depositor DCC
$R_{free}$ test set	3598 reflections (5.35%)	DCC
Wilson B-factor (Å <sup>2</sup> )	22.0	Xtriage
Anisotropy	0.200	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 26.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	4 of 70905 reflections (0.006%)	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	7397	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.98% of the height of the origin peak. No significant pseudotranslation is detected.*

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



## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: K, 5IU

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	0.36	0/264	0.73	0/404
1	B	0.34	0/264	0.74	0/404
1	C	0.38	0/264	0.71	0/404
2	D	0.36	0/259	0.74	0/394
2	E	0.38	0/240	0.74	0/364
2	F	0.34	0/259	0.72	0/394
3	G	0.30	0/940	0.55	0/1268
3	H	0.30	0/946	0.55	0/1275
3	I	0.31	0/940	0.54	0/1268
3	J	0.31	0/946	0.53	0/1275
3	K	0.31	0/946	0.57	0/1275
3	L	0.30	0/940	0.55	0/1268
All	All	0.32	0/7208	0.60	0/9993

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
2	D	0	1
2	E	0	1
2	F	0	1
All	All	0	3

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	1105	DT	Sidechain
2	E	1117	DT	Sidechain
2	F	1129	DT	Sidechain

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	254	0	132	9	0
1	B	254	0	132	10	0
1	C	254	0	132	5	0
2	D	255	0	149	4	0
2	E	238	0	136	3	0
2	F	255	0	149	3	0
3	G	911	0	906	6	0
3	H	917	0	917	13	0
3	I	911	0	906	8	0
3	J	917	0	917	13	0
3	K	917	0	917	9	0
3	L	911	0	906	10	0
4	G	1	0	0	0	0
4	H	1	0	0	0	0
4	I	1	0	0	0	0
4	J	1	0	0	0	0
4	K	1	0	0	0	0
4	L	1	0	0	0	0
5	A	21	0	0	0	0
5	B	20	0	0	0	0
5	C	15	0	0	0	0
5	D	26	0	0	0	0
5	E	24	0	0	0	0
5	F	29	0	0	1	0
5	G	63	0	0	1	0
5	H	47	0	0	2	0
5	I	47	0	0	1	0
5	J	14	0	0	0	0
5	K	55	0	0	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	L	36	0	0	0	0
All	All	7397	0	6299	75	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 6.

All (75) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:L:2100:LYS:HA	3:L:2105:ALA:HA	1.42	1.00
2:F:1124:DT:H4'	3:J:701:LYS:O	1.81	0.81
1:A:1001:DA:H2''	1:A:1002:DA:C8	2.24	0.71
1:B:1013:DA:H2''	1:B:1014:DA:C8	2.27	0.69
3:G:39:MET:HE2	3:G:104:ASN:H	1.57	0.68
3:K:888:LEU:HD23	3:K:891:ILE:HD12	1.77	0.67
3:L:2088:LEU:HD23	3:L:2091:ILE:HD12	1.77	0.66
1:A:1012:DA:O3'	1:B:1013:DA:H5'	1.96	0.65
2:D:1103:DA:H5''	3:L:2007:ARG:CZ	2.27	0.65
3:H:238:TRP:CD2	3:H:307:ARG:HD3	2.32	0.63
3:J:698:SER:HB3	3:J:706:PHE:O	2.00	0.61
1:B:1024:DA:O3'	1:C:1025:DA:P	2.60	0.59
1:C:1025:DA:H2''	1:C:1026:DA:C8	2.39	0.58
1:B:1022:5IU:H2'	3:K:882:ARG:CZ	2.36	0.55
1:C:1026:DA:C8	3:L:2075:LYS:HE2	2.40	0.55
3:J:699:ILE:O	3:J:705:ALA:HA	2.07	0.55
3:J:647:ASP:HB3	3:J:650:LYS:HB3	1.91	0.53
3:G:52:ALA:N	3:G:53:PRO:HD2	2.24	0.53
2:F:1134:DC:H2'	2:F:1135:5IU:I5	2.80	0.52
1:B:1016:DT:H2''	1:B:1017:DG:C8	2.44	0.52
1:A:1010:5IU:H2'	3:I:482:ARG:CZ	2.39	0.52
3:L:2052:ALA:N	3:L:2053:PRO:HD2	2.25	0.52
1:C:1028:DT:H2'	3:L:2082:ARG:CZ	2.40	0.51
1:B:1016:DT:H2'	3:J:682:ARG:CZ	2.41	0.51
3:J:652:ALA:N	3:J:653:PRO:CD	2.74	0.51
2:E:1124:DT:H2'	5:I:3159:HOH:O	2.11	0.50
3:K:852:ALA:N	3:K:853:PRO:HD2	2.26	0.50
1:A:1010:5IU:H2'	3:I:482:ARG:NH1	2.27	0.50
3:L:2052:ALA:N	3:L:2053:PRO:CD	2.75	0.50
3:H:300:LYS:HB3	5:H:3223:HOH:O	2.12	0.49
1:B:1016:DT:H2'	3:J:682:ARG:NH1	2.28	0.49
1:C:1028:DT:H2'	3:L:2082:ARG:NH1	2.29	0.48
3:G:52:ALA:N	3:G:53:PRO:CD	2.76	0.48
3:K:852:ALA:N	3:K:853:PRO:CD	2.76	0.48

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:J:612:LEU:O	3:J:616:ILE:HG13	2.14	0.47
1:A:1012:DA:O3'	1:B:1013:DA:P	2.73	0.46
3:L:2012:LEU:O	3:L:2016:ILE:HG13	2.15	0.46
3:I:452:ALA:N	3:I:453:PRO:CD	2.78	0.46
3:J:628:ASN:OD1	3:J:631:LYS:HB2	2.16	0.46
2:F:1136:DT:H2'	5:G:3077:HOH:O	2.16	0.46
1:B:1022:5IU:H2'	3:K:882:ARG:NH1	2.30	0.46
3:I:433:ILE:HG12	3:I:510:ARG:HG3	1.99	0.45
2:D:1112:DT:O3'	2:E:1113:DT:H5'	2.16	0.45
1:A:1004:DT:H2'	3:H:282:ARG:CZ	2.47	0.45
3:J:622:PRO:O	3:J:637:PRO:HG2	2.17	0.45
3:I:422:PRO:O	3:I:437:PRO:HG2	2.17	0.45
2:D:1110:DC:H2'	2:D:1111:5IU:I5	2.87	0.44
5:F:3147:HOH:O	3:H:287:SER:HB3	2.18	0.44
3:I:412:LEU:O	3:I:416:ILE:HG13	2.19	0.43
3:H:252:ALA:N	3:H:253:PRO:CD	2.82	0.43
3:J:652:ALA:N	3:J:653:PRO:HD2	2.34	0.43
3:H:307:ARG:HD2	5:H:3223:HOH:O	2.19	0.42
3:J:691:ILE:HG12	3:J:711:MET:HG2	2.00	0.42
3:K:832:LYS:HB3	3:K:911:MET:HB2	2.01	0.42
1:A:1001:DA:H2''	1:A:1002:DA:H8	1.80	0.42
1:A:1002:DA:H2'	3:H:275:LYS:HG3	2.01	0.42
3:I:441:ALA:HA	3:I:446:TRP:CG	2.54	0.42
3:K:833:ILE:HG12	3:K:910:ARG:HG3	2.02	0.42
2:D:1102:DC:H2''	2:D:1103:DA:C8	2.54	0.42
3:J:700:LYS:HA	3:J:705:ALA:HB1	2.02	0.42
3:H:266:GLN:HB2	3:H:269:ILE:HB	2.01	0.41
3:H:243:ARG:HH22	3:H:301:LYS:NZ	2.18	0.41
3:L:2104:ASN:H	3:L:2104:ASN:ND2	2.18	0.41
3:K:891:ILE:HG12	3:K:911:MET:HG2	2.02	0.41
3:K:815:GLN:HG2	5:K:3056:HOH:O	2.21	0.41
1:A:1002:DA:C8	3:H:275:LYS:HE3	2.56	0.41
3:H:284:ALA:O	3:H:288:LEU:HG	2.20	0.41
3:G:88:LEU:HD12	3:G:91:ILE:HD12	2.03	0.40
1:B:1022:5IU:H2''	1:B:1023:DG:C8	2.56	0.40
3:H:291:ILE:HG12	3:H:311:MET:HG2	2.03	0.40
2:E:1122:DC:H2'	2:E:1123:5IU:I5	2.91	0.40
3:G:21:ILE:HG21	3:G:54:LEU:HB2	2.03	0.40
3:H:212:LEU:O	3:H:216:ILE:HG13	2.21	0.40
3:I:472:PRO:C	3:I:474:PRO:HD3	2.42	0.40
3:G:100:LYS:HA	3:G:105:ALA:HB1	2.04	0.40

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	
3	G	107/113 (95%)	102 (95%)	4 (4%)	1 (1%)	25	21
3	H	107/113 (95%)	104 (97%)	1 (1%)	2 (2%)	12	7
3	I	107/113 (95%)	103 (96%)	2 (2%)	2 (2%)	12	7
3	J	107/113 (95%)	99 (92%)	4 (4%)	4 (4%)	5	2
3	K	107/113 (95%)	102 (95%)	4 (4%)	1 (1%)	25	21
3	L	107/113 (95%)	99 (92%)	7 (6%)	1 (1%)	25	21
All	All	642/678 (95%)	609 (95%)	22 (3%)	11 (2%)	14	8

All (11) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	J	703	ASN
3	J	696	ASP
3	H	252	ALA
3	I	452	ALA
3	J	652	ALA
3	H	299	ILE
3	G	52	ALA
3	J	699	ILE
3	K	852	ALA
3	L	2052	ALA
3	I	499	ILE

### 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	
3	G	96/101 (95%)	93 (97%)	3 (3%)	52	63
3	H	97/101 (96%)	95 (98%)	2 (2%)	66	78
3	I	96/101 (95%)	95 (99%)	1 (1%)	85	93
3	J	97/101 (96%)	95 (98%)	2 (2%)	66	78
3	K	97/101 (96%)	94 (97%)	3 (3%)	52	63
3	L	96/101 (95%)	94 (98%)	2 (2%)	66	78
All	All	579/606 (96%)	566 (98%)	13 (2%)	64	76

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	G	7	ARG
3	G	9	ARG
3	G	12	LEU
3	H	212	LEU
3	H	247	ASP
3	I	409	ARG
3	J	607	ARG
3	J	612	LEU
3	K	812	LEU
3	K	888	LEU
3	K	894	VAL
3	L	2012	LEU
3	L	2104	ASN

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

Mol	Chain	Res	Type
3	G	15	GLN
3	G	57	ASN
3	H	215	GLN
3	H	257	ASN
3	I	444	HIS
3	J	665	HIS
3	K	815	GLN
3	K	857	ASN
3	L	2015	GLN
3	L	2057	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

6 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$
1	5IU	A	1010	1,2	19,21,22	0.82	0	22,30,33	1.66	2 (9%)
1	5IU	B	1022	1,2	19,21,22	0.88	1 (5%)	22,30,33	1.68	3 (13%)
1	5IU	C	1034	1,2	19,21,22	0.95	1 (5%)	22,30,33	1.55	3 (13%)
2	5IU	D	1111	1,2	19,21,22	0.96	1 (5%)	22,30,33	1.67	3 (13%)
2	5IU	E	1123	1,2	19,21,22	0.84	1 (5%)	22,30,33	1.70	3 (13%)
2	5IU	F	1135	1,2	19,21,22	0.87	0	22,30,33	1.51	2 (9%)

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
1	5IU	A	1010	1,2	-	0/5/21/22	0/2/2/2
1	5IU	B	1022	1,2	-	0/5/21/22	0/2/2/2
1	5IU	C	1034	1,2	-	0/5/21/22	0/2/2/2
2	5IU	D	1111	1,2	-	0/5/21/22	0/2/2/2
2	5IU	E	1123	1,2	-	0/5/21/22	0/2/2/2
2	5IU	F	1135	1,2	-	0/5/21/22	0/2/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	1123	5IU	C6-N1	2.06	1.39	1.34
2	D	1111	5IU	C6-N1	2.05	1.39	1.34

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	1022	5IU	P-OP1	2.05	1.49	1.46
1	C	1034	5IU	C4-C5	2.00	1.45	1.40

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	1123	5IU	C6-N1-C2	-5.55	120.83	122.41
2	D	1111	5IU	C6-N1-C2	-5.11	120.95	122.41
1	A	1010	5IU	C6-N1-C2	-5.02	120.98	122.41
1	A	1010	5IU	C5-C6-N1	4.90	123.33	120.11
1	B	1022	5IU	C5-C6-N1	4.77	123.24	120.11
1	B	1022	5IU	C6-N1-C2	-4.65	121.09	122.41
1	C	1034	5IU	C5-C6-N1	4.53	123.08	120.11
2	E	1123	5IU	C5-C6-N1	4.35	122.97	120.11
2	F	1135	5IU	C6-N1-C2	-4.32	121.18	122.41
2	D	1111	5IU	C5-C6-N1	4.32	122.95	120.11
2	F	1135	5IU	C5-C6-N1	3.85	122.64	120.11
1	C	1034	5IU	C6-N1-C2	-3.64	121.37	122.41
1	C	1034	5IU	C4-C5-I5	3.13	122.29	121.31
1	B	1022	5IU	C4-C5-I5	3.03	122.26	121.31
2	E	1123	5IU	C2'-C1'-N1	2.20	119.81	114.08
2	D	1111	5IU	C2'-C1'-N1	2.07	119.47	114.08

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 6 ligands modelled in this entry, 6 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.



There are no ring outliers.

## 5.7 Other polymers ⓘ

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 ⓘ

### 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	12/12 (100%)	-0.68	0 100 100	9, 12, 25, 26	0
1	B	12/12 (100%)	-0.68	0 100 100	11, 15, 24, 29	0
1	C	12/12 (100%)	-0.69	0 100 100	11, 14, 24, 28	0
2	D	13/13 (100%)	-0.58	0 100 100	12, 15, 21, 46	0
2	E	12/13 (92%)	-0.73	0 100 100	12, 15, 22, 23	0
2	F	13/13 (100%)	-0.69	0 100 100	10, 12, 23, 45	0
3	G	109/113 (96%)	-0.45	0 100 100	9, 20, 44, 52	0
3	H	109/113 (96%)	-0.07	7 (6%) 19 18	9, 23, 75, 90	0
3	I	109/113 (96%)	-0.43	1 (0%) 81 82	11, 25, 44, 58	0
3	J	109/113 (96%)	0.15	7 (6%) 19 18	13, 35, 70, 80	0
3	K	109/113 (96%)	-0.45	1 (0%) 81 82	11, 24, 48, 63	0
3	L	109/113 (96%)	0.14	8 (7%) 15 14	11, 25, 80, 94	0
All	All	728/753 (96%)	-0.23	24 (3%) 44 45	9, 24, 55, 94	0

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	L	2098	SER	11.8
3	L	2101	LYS	10.4
3	L	2102	GLY	8.8
3	J	698	SER	8.6
3	J	704	ASN	8.0
3	L	2099	ILE	7.6
3	L	2103	ASN	6.6
3	H	302	GLY	6.4
3	H	303	ASN	6.4
3	H	304	ASN	6.1
3	L	2097	ARG	5.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
3	L	2104	ASN	5.7
3	J	699	ILE	5.7
3	J	697	ARG	5.7
3	J	703	ASN	5.5
3	H	298	SER	5.2
3	L	2100	LYS	4.7
3	H	299	ILE	4.1
3	J	700	LYS	4.0
3	H	296	ASP	3.8
3	I	504	ASN	3.7
3	J	696	ASP	3.7
3	H	297	ARG	3.6
3	K	904	ASN	2.0

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

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
1	5IU	A	1010	20/21	0.11	2.01	8,14,20,28	0
2	5IU	F	1135	20/21	0.10	1.81	5,11,18,20	0
2	5IU	D	1111	20/21	0.10	1.03	10,17,21,22	0
2	5IU	E	1123	20/21	0.10	0.79	8,14,21,24	0
1	5IU	C	1034	20/21	0.10	0.20	4,12,21,27	0
1	5IU	B	1022	20/21	0.09	-0.28	6,13,18,27	0

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	K	J	4005	1/1	0.10	-0.50	35,35,35,35	0
4	K	K	4004	1/1	0.09	-0.50	27,27,27,27	0
4	K	I	4003	1/1	0.08	-0.57	28,28,28,28	0
4	K	L	4002	1/1	0.08	-0.85	28,28,28,28	0
4	K	G	4001	1/1	0.07	-1.09	25,25,25,25	0
4	K	H	4006	1/1	0.08	-1.96	31,31,31,31	0

## 6.5 Other polymers

There are no such residues in this entry.