



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:41 AM GMT

PDB ID : 3FIJ  
Title : Crystal structure of a uncharacterized protein lin1909  
Authors : Sugadev, R.; Burley, S.K.; Swaminathan, S.; New York SGX Research Center  
for Structural Genomics (NYSGXRC)  
Deposited on : 2008-12-11  
Resolution : 2.30 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  
<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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

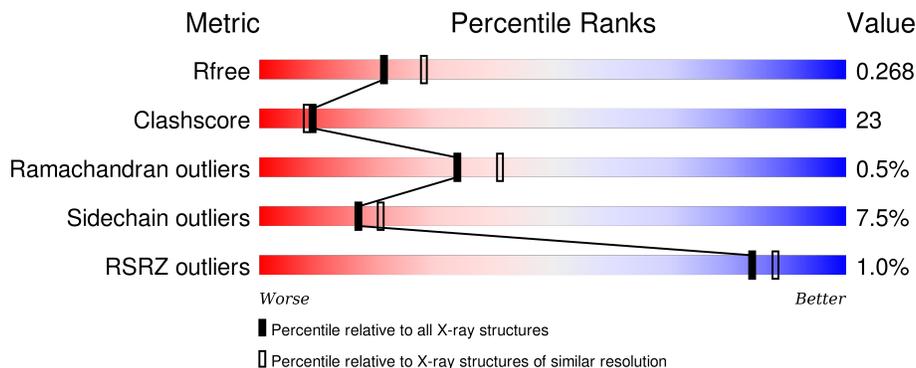
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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.



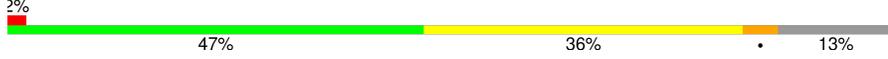
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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	254	 55% 30% 12%
1	B	254	 63% 22% 13%
1	C	254	 53% 30% 13%
1	D	254	 59% 26% 12%
1	E	254	 54% 30% 12%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	254	 <p>52% 31% 12% 2%</p>
1	G	254	 <p>47% 36% 13% 2%</p>
1	H	254	 <p>56% 30% 12%</p>

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 14467 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 Lin1909 protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	224	Total 1739	C 1110	N 292	O 332	S 1	Se 4	0	0	0
1	B	222	Total 1722	C 1100	N 289	O 329	S 1	Se 3	0	0	0
1	C	221	Total 1715	C 1095	N 288	O 328	S 1	Se 3	0	0	0
1	D	223	Total 1731	C 1105	N 291	O 331	S 1	Se 3	0	0	0
1	E	224	Total 1739	C 1110	N 292	O 332	S 1	Se 4	0	0	0
1	F	223	Total 1731	C 1105	N 291	O 331	S 1	Se 3	0	0	0
1	G	222	Total 1723	C 1100	N 290	O 329	S 1	Se 3	0	0	0
1	H	224	Total 1739	C 1110	N 292	O 332	S 1	Se 4	0	0	0

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	-	expression tag	UNP Q92AL3
A	2	SER	-	expression tag	UNP Q92AL3
A	3	LEU	-	expression tag	UNP Q92AL3
A	247	GLU	-	expression tag	UNP Q92AL3
A	248	GLY	-	expression tag	UNP Q92AL3
A	249	HIS	-	expression tag	UNP Q92AL3
A	250	HIS	-	expression tag	UNP Q92AL3
A	251	HIS	-	expression tag	UNP Q92AL3
A	252	HIS	-	expression tag	UNP Q92AL3
A	253	HIS	-	expression tag	UNP Q92AL3
A	254	HIS	-	expression tag	UNP Q92AL3
B	1	MSE	-	expression tag	UNP Q92AL3
B	2	SER	-	expression tag	UNP Q92AL3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	3	LEU	-	expression tag	UNP Q92AL3
B	247	GLU	-	expression tag	UNP Q92AL3
B	248	GLY	-	expression tag	UNP Q92AL3
B	249	HIS	-	expression tag	UNP Q92AL3
B	250	HIS	-	expression tag	UNP Q92AL3
B	251	HIS	-	expression tag	UNP Q92AL3
B	252	HIS	-	expression tag	UNP Q92AL3
B	253	HIS	-	expression tag	UNP Q92AL3
B	254	HIS	-	expression tag	UNP Q92AL3
C	1	MSE	-	expression tag	UNP Q92AL3
C	2	SER	-	expression tag	UNP Q92AL3
C	3	LEU	-	expression tag	UNP Q92AL3
C	247	GLU	-	expression tag	UNP Q92AL3
C	248	GLY	-	expression tag	UNP Q92AL3
C	249	HIS	-	expression tag	UNP Q92AL3
C	250	HIS	-	expression tag	UNP Q92AL3
C	251	HIS	-	expression tag	UNP Q92AL3
C	252	HIS	-	expression tag	UNP Q92AL3
C	253	HIS	-	expression tag	UNP Q92AL3
C	254	HIS	-	expression tag	UNP Q92AL3
D	1	MSE	-	expression tag	UNP Q92AL3
D	2	SER	-	expression tag	UNP Q92AL3
D	3	LEU	-	expression tag	UNP Q92AL3
D	247	GLU	-	expression tag	UNP Q92AL3
D	248	GLY	-	expression tag	UNP Q92AL3
D	249	HIS	-	expression tag	UNP Q92AL3
D	250	HIS	-	expression tag	UNP Q92AL3
D	251	HIS	-	expression tag	UNP Q92AL3
D	252	HIS	-	expression tag	UNP Q92AL3
D	253	HIS	-	expression tag	UNP Q92AL3
D	254	HIS	-	expression tag	UNP Q92AL3
E	1	MSE	-	expression tag	UNP Q92AL3
E	2	SER	-	expression tag	UNP Q92AL3
E	3	LEU	-	expression tag	UNP Q92AL3
E	247	GLU	-	expression tag	UNP Q92AL3
E	248	GLY	-	expression tag	UNP Q92AL3
E	249	HIS	-	expression tag	UNP Q92AL3
E	250	HIS	-	expression tag	UNP Q92AL3
E	251	HIS	-	expression tag	UNP Q92AL3
E	252	HIS	-	expression tag	UNP Q92AL3
E	253	HIS	-	expression tag	UNP Q92AL3
E	254	HIS	-	expression tag	UNP Q92AL3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	1	MSE	-	expression tag	UNP Q92AL3
F	2	SER	-	expression tag	UNP Q92AL3
F	3	LEU	-	expression tag	UNP Q92AL3
F	247	GLU	-	expression tag	UNP Q92AL3
F	248	GLY	-	expression tag	UNP Q92AL3
F	249	HIS	-	expression tag	UNP Q92AL3
F	250	HIS	-	expression tag	UNP Q92AL3
F	251	HIS	-	expression tag	UNP Q92AL3
F	252	HIS	-	expression tag	UNP Q92AL3
F	253	HIS	-	expression tag	UNP Q92AL3
F	254	HIS	-	expression tag	UNP Q92AL3
G	1	MSE	-	expression tag	UNP Q92AL3
G	2	SER	-	expression tag	UNP Q92AL3
G	3	LEU	-	expression tag	UNP Q92AL3
G	247	GLU	-	expression tag	UNP Q92AL3
G	248	GLY	-	expression tag	UNP Q92AL3
G	249	HIS	-	expression tag	UNP Q92AL3
G	250	HIS	-	expression tag	UNP Q92AL3
G	251	HIS	-	expression tag	UNP Q92AL3
G	252	HIS	-	expression tag	UNP Q92AL3
G	253	HIS	-	expression tag	UNP Q92AL3
G	254	HIS	-	expression tag	UNP Q92AL3
H	1	MSE	-	expression tag	UNP Q92AL3
H	2	SER	-	expression tag	UNP Q92AL3
H	3	LEU	-	expression tag	UNP Q92AL3
H	247	GLU	-	expression tag	UNP Q92AL3
H	248	GLY	-	expression tag	UNP Q92AL3
H	249	HIS	-	expression tag	UNP Q92AL3
H	250	HIS	-	expression tag	UNP Q92AL3
H	251	HIS	-	expression tag	UNP Q92AL3
H	252	HIS	-	expression tag	UNP Q92AL3
H	253	HIS	-	expression tag	UNP Q92AL3
H	254	HIS	-	expression tag	UNP Q92AL3

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	G	1	Total Mn 1 1	0	0
2	D	1	Total Mn 1 1	0	0
2	E	1	Total Mn 1 1	0	0

*Continued on next page...*

*Continued from previous page...*

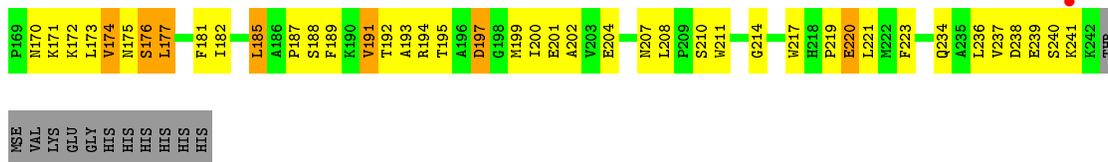
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	H	1	Total 1	Mn 1	0	0
2	B	1	Total 1	Mn 1	0	0
2	C	1	Total 1	Mn 1	0	0
2	A	1	Total 1	Mn 1	0	0
2	F	1	Total 1	Mn 1	0	0

- Molecule 3 is water.

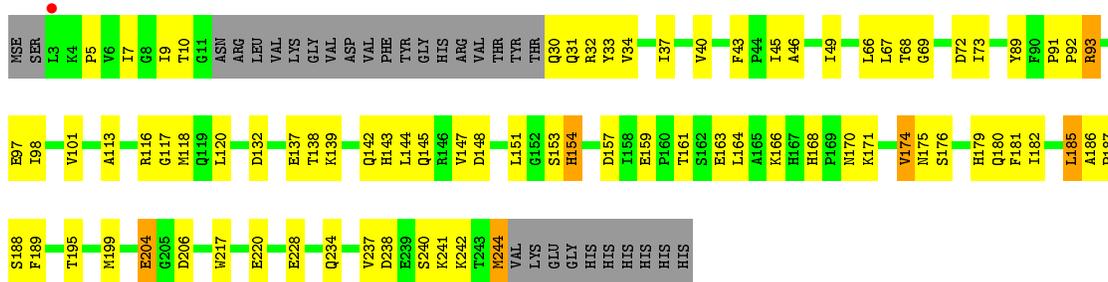
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	100	Total 100	O 100	0	0
3	B	85	Total 85	O 85	0	0
3	C	85	Total 85	O 85	0	0
3	D	92	Total 92	O 92	0	0
3	E	72	Total 72	O 72	0	0
3	F	73	Total 73	O 73	0	0
3	G	56	Total 56	O 56	0	0
3	H	57	Total 57	O 57	0	0







• Molecule 1: Lin1909 protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	91.06Å 131.21Å 93.29Å 90.00° 108.79° 90.00°	Depositor
Resolution (Å)	43.10 – 2.30 43.10 – 2.30	Depositor EDS
% Data completeness (in resolution range)	95.4 (43.10-2.30) 95.2 (43.10-2.30)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.85 (at 2.29Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.217 , 0.269 0.217 , 0.268	Depositor DCC
$R_{free}$ test set	3520 reflections (4.01%)	DCC
Wilson B-factor (Å <sup>2</sup> )	22.0	Xtriage
Anisotropy	0.228	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 43.0	EDS
Estimated twinning fraction	0.024 for l,-k,h	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 91934 reflections	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	14467	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.93% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: MN

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	0.38	0/1772	0.68	0/2402
1	B	0.36	0/1756	0.65	0/2383
1	C	0.38	0/1749	0.66	0/2374
1	D	0.36	0/1765	0.65	0/2395
1	E	0.37	0/1772	0.67	0/2402
1	F	0.35	0/1765	0.63	0/2395
1	G	0.39	0/1757	0.66	2/2384 (0.1%)
1	H	0.34	0/1772	0.63	0/2402
All	All	0.37	0/14108	0.65	2/19137 (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	G	177	LEU	N-CA-CB	5.82	122.05	110.40
1	G	176	SER	CB-CA-C	-5.29	100.04	110.10

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	1739	0	1739	89	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1722	0	1723	60	0
1	C	1715	0	1710	92	0
1	D	1731	0	1730	66	0
1	E	1739	0	1739	75	0
1	F	1731	0	1730	104	0
1	G	1723	0	1721	102	0
1	H	1739	0	1739	83	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0
3	A	100	0	0	10	0
3	B	85	0	0	5	0
3	C	85	0	0	9	0
3	D	92	0	0	2	0
3	E	72	0	0	3	0
3	F	73	0	0	7	0
3	G	56	0	0	4	0
3	H	57	0	0	2	0
All	All	14467	0	13831	638	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 23.

The worst 5 of 638 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:128:THR:HG22	1:C:183:LYS:HB3	1.28	1.15
1:A:241:LYS:HA	1:A:244:MSE:HE3	1.11	1.05
1:G:182:ILE:HD12	1:G:200:ILE:HD11	1.41	1.00
1:A:241:LYS:HA	1:A:244:MSE:CE	1.92	1.00
1:G:176:SER:O	1:G:177:LEU:HG	1.62	0.99

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	220/254 (87%)	208 (94%)	9 (4%)	3 (1%)	14	13
1	B	218/254 (86%)	211 (97%)	6 (3%)	1 (0%)	34	41
1	C	217/254 (85%)	208 (96%)	9 (4%)	0	100	100
1	D	219/254 (86%)	207 (94%)	11 (5%)	1 (0%)	34	41
1	E	220/254 (87%)	210 (96%)	8 (4%)	2 (1%)	21	24
1	F	219/254 (86%)	208 (95%)	10 (5%)	1 (0%)	34	41
1	G	218/254 (86%)	206 (94%)	12 (6%)	0	100	100
1	H	220/254 (87%)	210 (96%)	10 (4%)	0	100	100
All	All	1751/2032 (86%)	1668 (95%)	75 (4%)	8 (0%)	34	41

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	188	SER
1	E	149	GLU
1	A	149	GLU
1	E	196	ALA
1	F	242	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	190/212 (90%)	177 (93%)	13 (7%)	20	25

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	188/212 (89%)	172 (92%)	16 (8%)	13	16
1	C	187/212 (88%)	172 (92%)	15 (8%)	15	18
1	D	189/212 (89%)	178 (94%)	11 (6%)	25	33
1	E	190/212 (90%)	175 (92%)	15 (8%)	15	19
1	F	189/212 (89%)	174 (92%)	15 (8%)	15	19
1	G	188/212 (89%)	171 (91%)	17 (9%)	12	14
1	H	190/212 (90%)	178 (94%)	12 (6%)	22	29
All	All	1511/1696 (89%)	1397 (92%)	114 (8%)	17	21

5 of 114 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	155	THR
1	E	159	GLU
1	H	138	THR
1	D	192	THR
1	E	93	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 62 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	71	GLN
1	E	145	GLN
1	H	119	GLN
1	E	119	GLN
1	E	234	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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](#)

Of 8 ligands modelled in this entry, 8 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.

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 [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	220/254 (86%)	-0.12	1 (0%) 91 94	11, 21, 36, 53	0
1	B	219/254 (86%)	-0.09	1 (0%) 91 94	11, 20, 36, 51	0
1	C	218/254 (85%)	0.05	3 (1%) 78 83	15, 22, 36, 43	0
1	D	220/254 (86%)	-0.13	3 (1%) 78 83	11, 20, 36, 57	0
1	E	220/254 (86%)	-0.03	3 (1%) 78 83	15, 24, 38, 59	0
1	F	220/254 (86%)	0.13	2 (0%) 85 89	15, 26, 41, 56	0
1	G	219/254 (86%)	0.16	4 (1%) 71 78	14, 31, 44, 60	0
1	H	220/254 (86%)	0.10	1 (0%) 91 94	12, 30, 43, 60	0
All	All	1756/2032 (86%)	0.01	18 (1%) 84 88	11, 24, 41, 60	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	3	LEU	4.7
1	D	243	THR	4.1
1	F	3	LEU	3.3
1	H	3	LEU	2.9
1	G	40	VAL	2.7

### 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(Å <sup>2</sup> )	Q<0.9
2	MN	D	300	1/1	1.00	0.12	-0.02	15,15,15,15	0
2	MN	F	300	1/1	1.00	0.12	-0.07	19,19,19,19	0
2	MN	C	300	1/1	1.00	0.11	-0.45	16,16,16,16	0
2	MN	A	300	1/1	1.00	0.10	-1.04	15,15,15,15	0
2	MN	G	300	1/1	1.00	0.10	-1.08	18,18,18,18	0
2	MN	B	300	1/1	1.00	0.09	-1.24	16,16,16,16	0
2	MN	E	300	1/1	0.99	0.07	-2.82	17,17,17,17	0
2	MN	H	300	1/1	0.99	0.07	-5.78	17,17,17,17	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.