



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 12:47 PM GMT

PDB ID : 3S51
Title : Structure of FANCI
Authors : Pavletich, N.P.
Deposited on : 2011-05-20
Resolution : 3.30 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
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

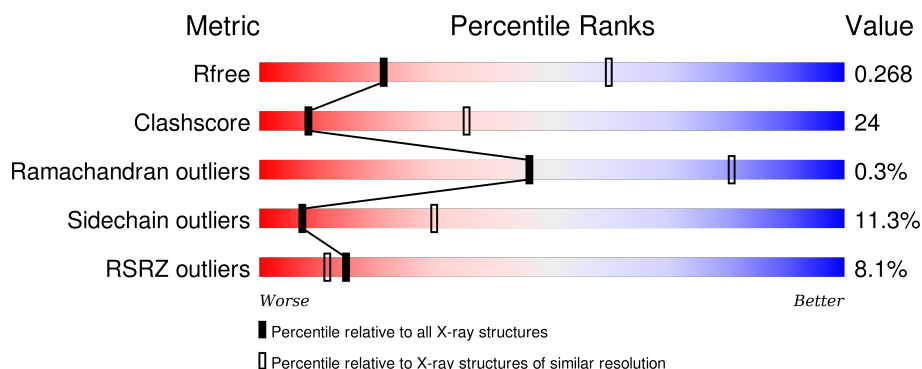
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 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	2060 (3.40-3.20)
Clashscore	102246	1058 (3.38-3.22)
Ramachandran outliers	100387	1038 (3.38-3.22)
Sidechain outliers	100360	1037 (3.38-3.22)
RSRZ outliers	91569	2070 (3.40-3.20)

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 ≥ 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	1308	<div> <div>7%</div> <div>44%</div> <div>37%</div> <div>5%</div> <div>13%</div> </div>
1	B	1308	<div> <div>8%</div> <div>41%</div> <div>36%</div> <div>5%</div> <div>18%</div> </div>
1	C	1308	<div> <div>5%</div> <div>45%</div> <div>37%</div> <div>5%</div> <div>13%</div> </div>
1	D	1308	<div> <div>7%</div> <div>40%</div> <div>34%</div> <div>5%</div> <div>21%</div> </div>

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 34594 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 Fanconi anemia group I protein homolog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1134	Total	C	N	O	S	0	0	0
			8960	5762	1489	1656	53			
1	B	1071	Total	C	N	O	S	0	0	0
			8487	5468	1409	1559	51			
1	C	1134	Total	C	N	O	S	0	0	0
			8960	5762	1489	1656	53			
1	D	1034	Total	C	N	O	S	0	0	0
			8187	5277	1358	1504	48			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1303	HIS	-	EXPRESSION TAG	UNP Q8K368
A	1304	HIS	-	EXPRESSION TAG	UNP Q8K368
A	1305	HIS	-	EXPRESSION TAG	UNP Q8K368
A	1306	HIS	-	EXPRESSION TAG	UNP Q8K368
A	1307	HIS	-	EXPRESSION TAG	UNP Q8K368
A	1308	HIS	-	EXPRESSION TAG	UNP Q8K368
B	1303	HIS	-	EXPRESSION TAG	UNP Q8K368
B	1304	HIS	-	EXPRESSION TAG	UNP Q8K368
B	1305	HIS	-	EXPRESSION TAG	UNP Q8K368
B	1306	HIS	-	EXPRESSION TAG	UNP Q8K368
B	1307	HIS	-	EXPRESSION TAG	UNP Q8K368
B	1308	HIS	-	EXPRESSION TAG	UNP Q8K368
C	1303	HIS	-	EXPRESSION TAG	UNP Q8K368
C	1304	HIS	-	EXPRESSION TAG	UNP Q8K368
C	1305	HIS	-	EXPRESSION TAG	UNP Q8K368
C	1306	HIS	-	EXPRESSION TAG	UNP Q8K368
C	1307	HIS	-	EXPRESSION TAG	UNP Q8K368
C	1308	HIS	-	EXPRESSION TAG	UNP Q8K368
D	1303	HIS	-	EXPRESSION TAG	UNP Q8K368
D	1304	HIS	-	EXPRESSION TAG	UNP Q8K368
D	1305	HIS	-	EXPRESSION TAG	UNP Q8K368

Continued on next page...

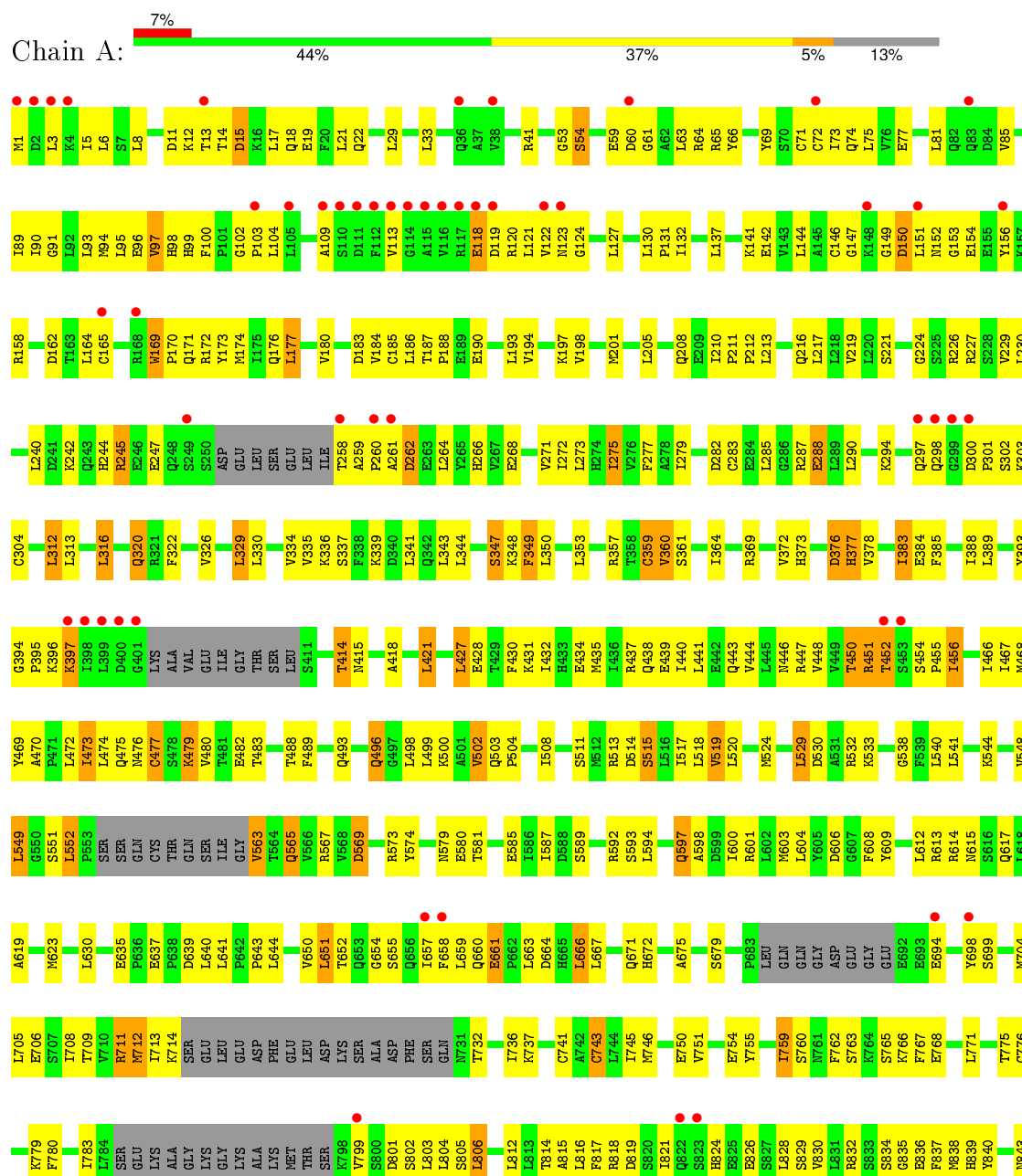
Continued from previous page...

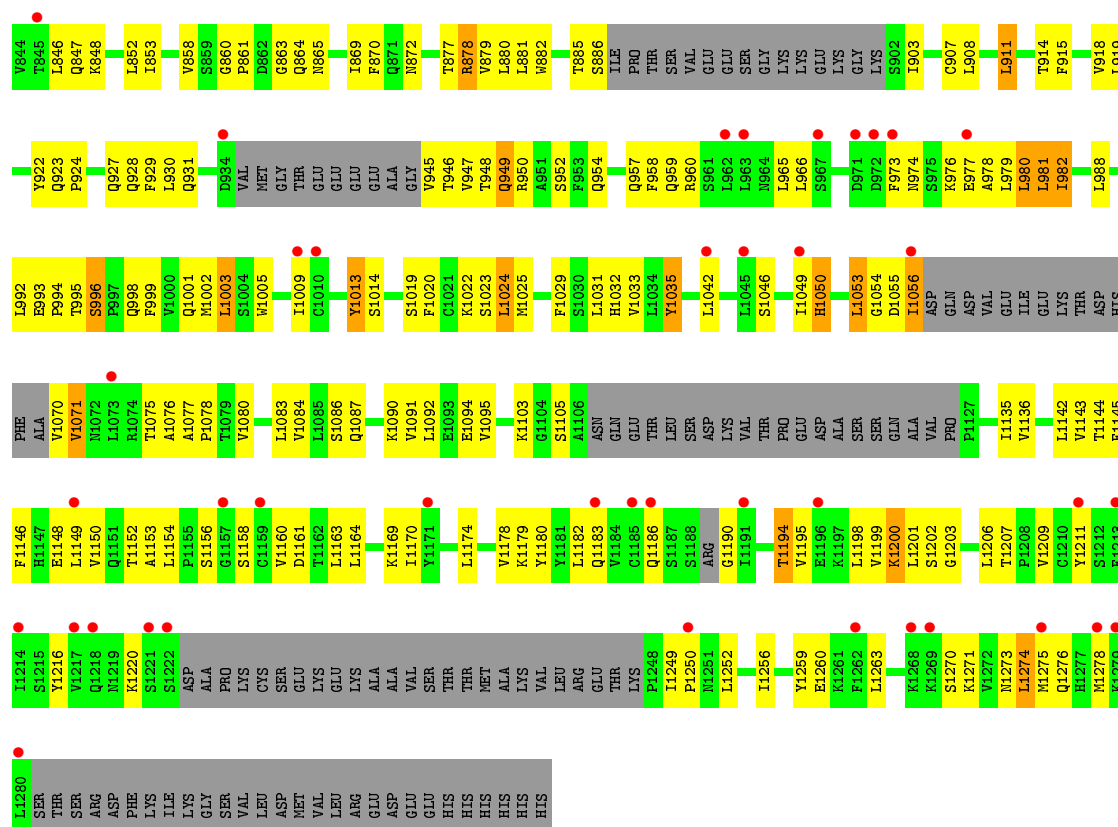
Chain	Residue	Modelled	Actual	Comment	Reference
D	1306	HIS	-	EXPRESSION TAG	UNP Q8K368
D	1307	HIS	-	EXPRESSION TAG	UNP Q8K368
D	1308	HIS	-	EXPRESSION TAG	UNP Q8K368

3 Residue-property plots

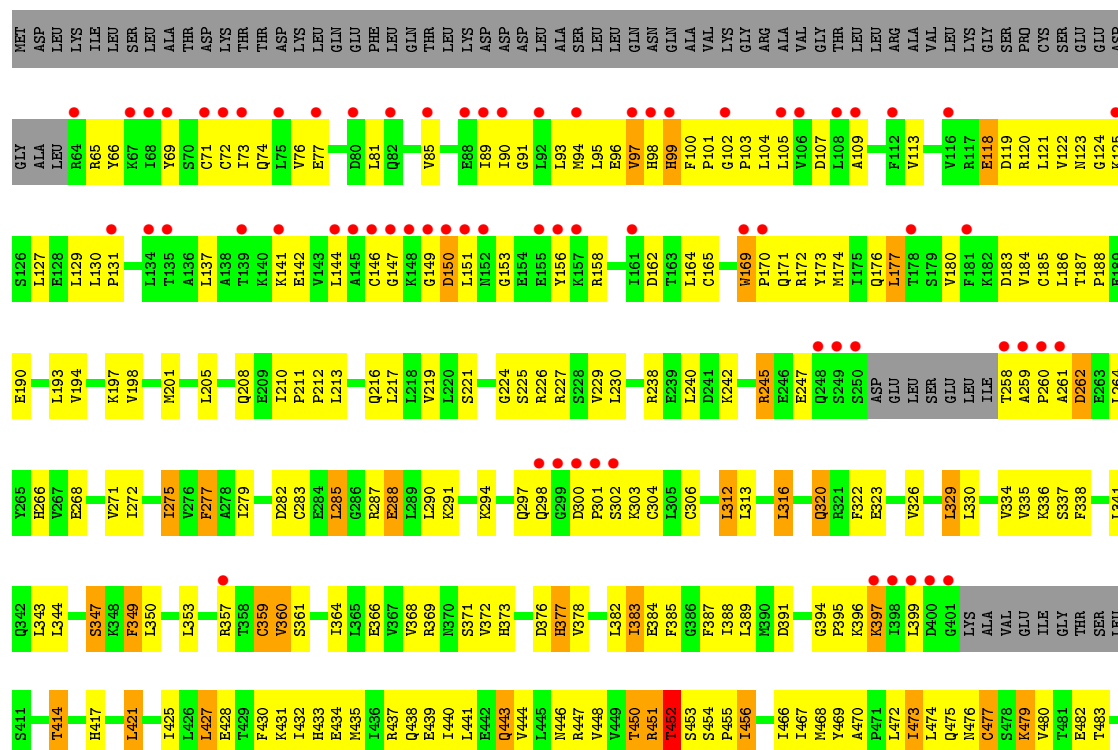
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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: Fanconi anemia group I protein homolog





• Molecule 1: Fanconi anemia group I protein homolog









THR	THR	L1164	V1091
MET	GLY	K1169	L1092
ALA	LYS	T1170	E1093
VAL	VAL	Y1171	E1094
LEU	LEU		V1095
ARG	ARG	L1174	K1103
GLU	GLU	T1175	G1104
THR	THR		S1105
LYS	LYS	V1178	A1106
P1248	GLY	K1179	ASN
I1249	LYS	Y1180	GLN
P1250	LYS	Y1181	GLU
H1251	LYS	L1182	THR
L1252	LYS	Q1183	LEU
	LYS	V1184	SER
I1256	LYS	C1185	ASP
	LYS	Q1186	LYS
Y1259	LYS	Q1187	VAL
E1260	LYS	S1187	THR
	LYS	S1188	PRO
L1263	LYS	ARG	GLU
I1264	LYS	G1190	ASP
	LYS	I1191	ALA
S1270	LYS	T1194	SER
K1271	LYS	V1195	SER
V1272	LYS		GLN
M1273	LYS	L1198	ALA
L1274	LYS	V1199	VAL
M1275	LYS	K1200	PRO
Q1276	LYS	L1201	P1127
H1277	LYS	S1202	T1128
M1278	LYS	G1203	L1129
K1279	LYS	L1206	L1130
L1280	LYS	T1207	I1131
SER	LYS		I1135
THR	LYS	Y1211	
SER	LYS	S1212	L1142
ARG	LYS	F1213	V1143
ASP	LYS		T1144
PHE	LYS	Y1216	F1145
LYS	LYS	V1217	F1146
ILE	LYS		H1147
LYS	LYS	S1222	E1148
GLY	LYS	ASP	L1149
SER	LYS	ALA	V1150
VAL	LYS	PRO	Q1151
LEU	LYS	LYS	T1152
ASP	LYS	CYS	A1153
MET	LYS	SER	L1154
VAL	LYS	GLU	P1155
LEU	LYS	LYS	S1156
ARG	LYS	GLU	G1157
GLU	LYS	GLU	S1158
ASP	LYS	ALA	C1159
GLU	LYS	ALA	V1160
GLU	LYS	VAL	D1161
HIS	LYS	SER	T1162
HIS	LYS	THR	L1163
HIS	LYS		

4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	114.70 Å 136.50 Å 149.70 Å 115.90° 106.00° 95.00°	Depositor
Resolution (Å)	39.82 – 3.30 39.82 – 3.28	Depositor EDS
% Data completeness (in resolution range)	82.8 (39.82-3.30) 73.3 (39.82-3.28)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.52 (at 3.25 Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.1_743)	Depositor
R, R_{free}	0.259 , 0.278 0.244 , 0.268	Depositor DCC
R_{free} test set	1931 reflections (2.00%)	DCC
Wilson B-factor (Å ²)	93.7	Xtriage
Anisotropy	0.486	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 119.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 110174 reflections	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	34594	wwPDB-VP
Average B, all atoms (Å ²)	192.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.11% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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

5.1 Standard geometry

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	0.26	2/9099 (0.0%)	0.46	0/12286
1	B	0.29	4/8624 (0.0%)	0.46	0/11646
1	C	0.26	2/9099 (0.0%)	0.46	0/12286
1	D	0.26	2/8319 (0.0%)	0.46	0/11234
All	All	0.27	10/35141 (0.0%)	0.46	0/47452

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	A	0	1
1	B	0	1
1	C	0	1
1	D	0	1
All	All	0	4

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	959	GLN	CD-NE2	-9.98	1.07	1.32
1	B	959	GLN	CD-OE1	-7.80	1.06	1.24
1	C	320	GLN	CD-NE2	-6.76	1.16	1.32
1	D	320	GLN	CD-NE2	-6.59	1.16	1.32
1	A	320	GLN	CD-NE2	-6.15	1.17	1.32

There are no bond angle outliers.

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1053	LEU	Peptide
1	B	1053	LEU	Peptide
1	C	1053	LEU	Peptide
1	D	1053	LEU	Peptide

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	8960	0	9275	454	0
1	B	8487	0	8776	449	0
1	C	8960	0	9275	435	0
1	D	8187	0	8476	411	0
All	All	34594	0	35802	1705	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 24.

The worst 5 of 1705 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:799:VAL:O	1:C:847:GLN:NE2	1.87	1.07
1:A:799:VAL:O	1:A:847:GLN:NE2	1.88	1.07
1:B:799:VAL:O	1:B:847:GLN:NE2	1.88	1.06
1:D:799:VAL:O	1:D:847:GLN:NE2	1.88	1.05
1:A:489:PHE:HB3	1:C:450:THR:HG21	1.35	1.04

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	1108/1308 (85%)	1023 (92%)	82 (7%)	3 (0%)	46	81
1	B	1045/1308 (80%)	969 (93%)	72 (7%)	4 (0%)	39	76
1	C	1108/1308 (85%)	1022 (92%)	83 (8%)	3 (0%)	46	81
1	D	1008/1308 (77%)	933 (93%)	71 (7%)	4 (0%)	39	76
All	All	4269/5232 (82%)	3947 (92%)	308 (7%)	14 (0%)	46	81

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	122	VAL
1	A	150	ASP
1	B	122	VAL
1	B	150	ASP
1	C	122	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	1032/1188 (87%)	917 (89%)	115 (11%)	8	31
1	B	979/1188 (82%)	867 (89%)	112 (11%)	7	29
1	C	1032/1188 (87%)	917 (89%)	115 (11%)	8	31
1	D	945/1188 (80%)	838 (89%)	107 (11%)	7	30
All	All	3988/4752 (84%)	3539 (89%)	449 (11%)	7	30

5 of 449 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	878	ARG
1	C	288	GLU
1	D	765	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	981	LEU
1	C	19	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	433	HIS
1	B	672	HIS
1	D	433	HIS
1	B	377	HIS
1	D	377	HIS

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 [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 [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, 95th 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 ²)	Q<0.9
1	A	1134/1308 (86%)	0.26	92 (8%) 15 11	85, 192, 306, 410	0
1	B	1071/1308 (81%)	0.28	109 (10%) 9 7	87, 188, 306, 415	0
1	C	1134/1308 (86%)	0.26	67 (5%) 26 20	81, 191, 306, 410	0
1	D	1034/1308 (79%)	0.29	87 (8%) 14 11	84, 185, 302, 416	0
All	All	4373/5232 (83%)	0.27	355 (8%) 15 11	81, 189, 304, 416	0

The worst 5 of 355 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	109	ALA	12.4
1	D	135	THR	12.0
1	D	250	SER	9.6
1	D	299	GLY	8.5
1	D	101	PRO	8.0

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.