



wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 09:01 am BST

PDB ID : 6AJL
Title : DOCK7 mutant I1836Y complexed with Cdc42
Authors : Kukimoto-Niino, M.; Shirouzu, M.
Deposited on : 2018-08-28
Resolution : 3.23 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

<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
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

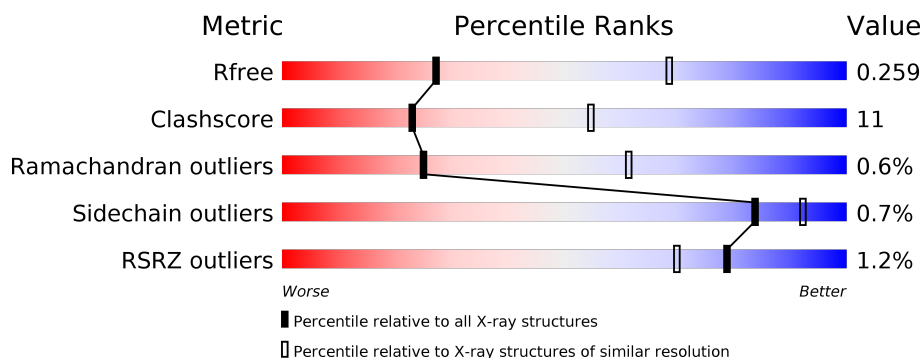
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.23 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}	130704	1619 (3.28-3.20)
Clashscore	141614	1755 (3.28-3.20)
Ramachandran outliers	138981	1728 (3.28-3.20)
Sidechain outliers	138945	1727 (3.28-3.20)
RSRZ outliers	127900	1567 (3.28-3.20)

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 respectively. 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	290	<div> <div></div> <div>68% 27% 5%</div> </div>
1	C	290	<div> <div></div> <div>68% 27% 5%</div> </div>
1	E	290	<div> <div>3%</div> <div>69% 24% 5%</div> </div>
1	G	290	<div> <div></div> <div>68% 26% 5%</div> </div>
2	B	195	<div> <div>2%</div> <div>68% 23% 9%</div> </div>
2	D	195	<div> <div>2%</div> <div>71% 21% 9%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	195	<div><div><div></div><div></div><div></div></div><div><div>2%</div><div>66%</div><div>25%</div><div>9%</div></div></div>
2	H	195	<div><div><div></div><div></div><div></div></div><div><div>3%</div><div>68%</div><div>24%</div><div>9%</div></div></div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 14700 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 Dedicator of cytokinesis protein 7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	275	Total	C	N	O	S	0	0	0
			2287	1463	391	423	10			
1	C	275	Total	C	N	O	S	0	0	0
			2287	1463	391	423	10			
1	E	275	Total	C	N	O	S	0	0	0
			2287	1463	391	423	10			
1	G	275	Total	C	N	O	S	0	0	0
			2287	1463	391	423	10			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1794	GLY	-	expression tag	UNP Q96N67
A	1795	SER	-	expression tag	UNP Q96N67
A	1796	SER	-	expression tag	UNP Q96N67
A	1797	GLY	-	expression tag	UNP Q96N67
A	1798	SER	-	expression tag	UNP Q96N67
A	1799	SER	-	expression tag	UNP Q96N67
A	1800	GLY	-	expression tag	UNP Q96N67
A	1836	TYR	ILE	engineered mutation	UNP Q96N67
C	1794	GLY	-	expression tag	UNP Q96N67
C	1795	SER	-	expression tag	UNP Q96N67
C	1796	SER	-	expression tag	UNP Q96N67
C	1797	GLY	-	expression tag	UNP Q96N67
C	1798	SER	-	expression tag	UNP Q96N67
C	1799	SER	-	expression tag	UNP Q96N67
C	1800	GLY	-	expression tag	UNP Q96N67
C	1836	TYR	ILE	engineered mutation	UNP Q96N67
E	1794	GLY	-	expression tag	UNP Q96N67
E	1795	SER	-	expression tag	UNP Q96N67
E	1796	SER	-	expression tag	UNP Q96N67
E	1797	GLY	-	expression tag	UNP Q96N67
E	1798	SER	-	expression tag	UNP Q96N67

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	1799	SER	-	expression tag	UNP Q96N67
E	1800	GLY	-	expression tag	UNP Q96N67
E	1836	TYR	ILE	engineered mutation	UNP Q96N67
G	1794	GLY	-	expression tag	UNP Q96N67
G	1795	SER	-	expression tag	UNP Q96N67
G	1796	SER	-	expression tag	UNP Q96N67
G	1797	GLY	-	expression tag	UNP Q96N67
G	1798	SER	-	expression tag	UNP Q96N67
G	1799	SER	-	expression tag	UNP Q96N67
G	1800	GLY	-	expression tag	UNP Q96N67
G	1836	TYR	ILE	engineered mutation	UNP Q96N67

- Molecule 2 is a protein called Cell division control protein 42 homolog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	178	Total	C	N	O	S	0	0	0
			1388	893	221	267	7			
2	D	178	Total	C	N	O	S	0	0	0
			1388	893	221	267	7			
2	F	178	Total	C	N	O	S	0	0	0
			1388	893	221	267	7			
2	H	178	Total	C	N	O	S	0	0	0
			1388	893	221	267	7			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-6	GLY	-	expression tag	UNP P60953
B	-5	SER	-	expression tag	UNP P60953
B	-4	SER	-	expression tag	UNP P60953
B	-3	GLY	-	expression tag	UNP P60953
B	-2	SER	-	expression tag	UNP P60953
B	-1	SER	-	expression tag	UNP P60953
B	0	GLY	-	expression tag	UNP P60953
B	188	SER	CYS	engineered mutation	UNP P60953
D	-6	GLY	-	expression tag	UNP P60953
D	-5	SER	-	expression tag	UNP P60953
D	-4	SER	-	expression tag	UNP P60953
D	-3	GLY	-	expression tag	UNP P60953
D	-2	SER	-	expression tag	UNP P60953
D	-1	SER	-	expression tag	UNP P60953
D	0	GLY	-	expression tag	UNP P60953

Continued on next page...

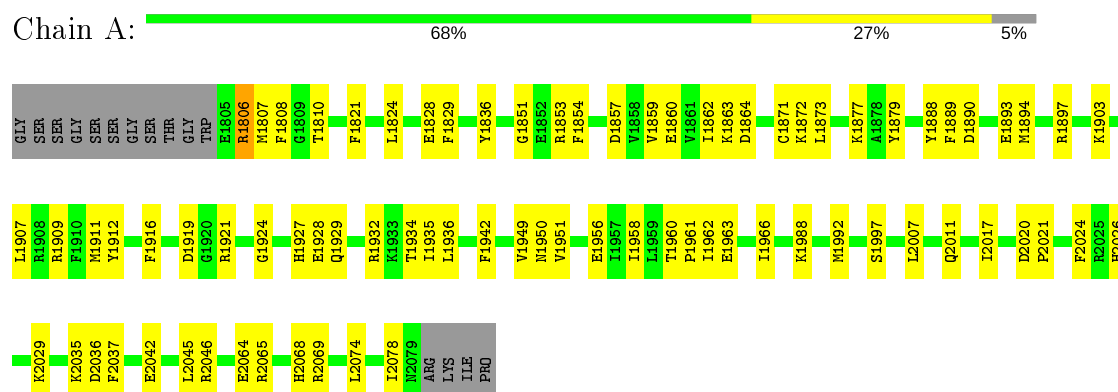
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	188	SER	CYS	engineered mutation	UNP P60953
F	-6	GLY	-	expression tag	UNP P60953
F	-5	SER	-	expression tag	UNP P60953
F	-4	SER	-	expression tag	UNP P60953
F	-3	GLY	-	expression tag	UNP P60953
F	-2	SER	-	expression tag	UNP P60953
F	-1	SER	-	expression tag	UNP P60953
F	0	GLY	-	expression tag	UNP P60953
F	188	SER	CYS	engineered mutation	UNP P60953
H	-6	GLY	-	expression tag	UNP P60953
H	-5	SER	-	expression tag	UNP P60953
H	-4	SER	-	expression tag	UNP P60953
H	-3	GLY	-	expression tag	UNP P60953
H	-2	SER	-	expression tag	UNP P60953
H	-1	SER	-	expression tag	UNP P60953
H	0	GLY	-	expression tag	UNP P60953
H	188	SER	CYS	engineered mutation	UNP P60953

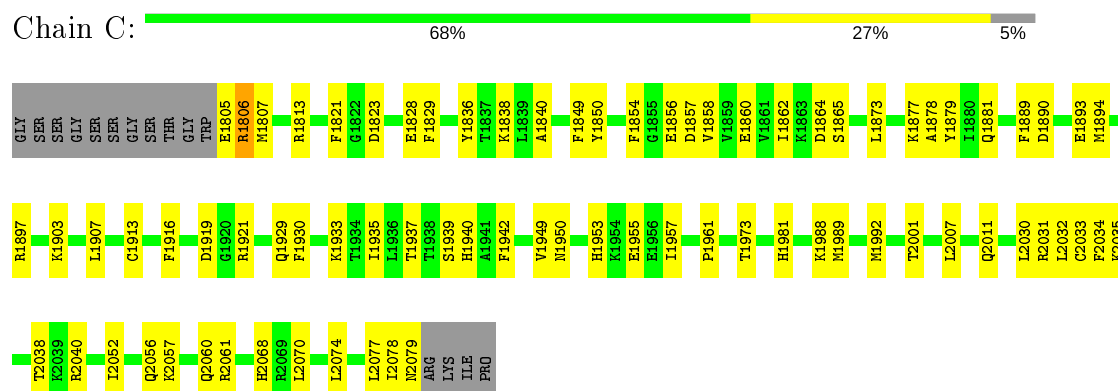
3 Residue-property plots [i](#)

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.

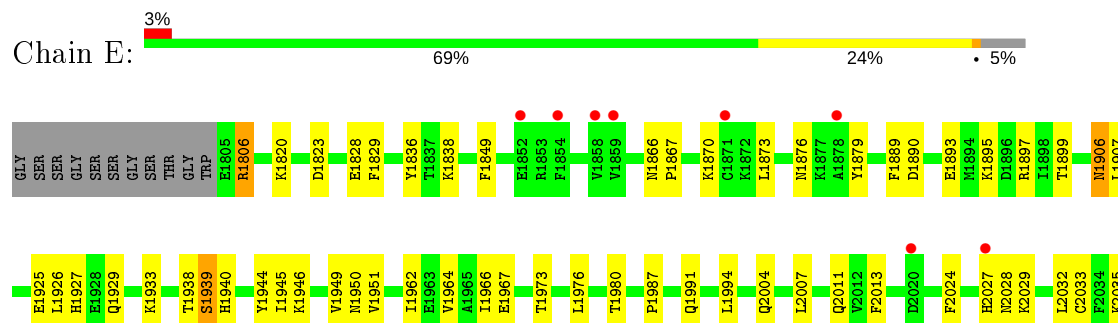
- Molecule 1: Dedicator of cytokinesis protein 7

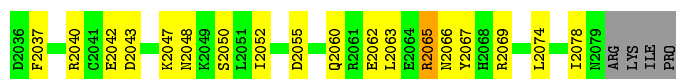


- Molecule 1: Dedicator of cytokinesis protein 7



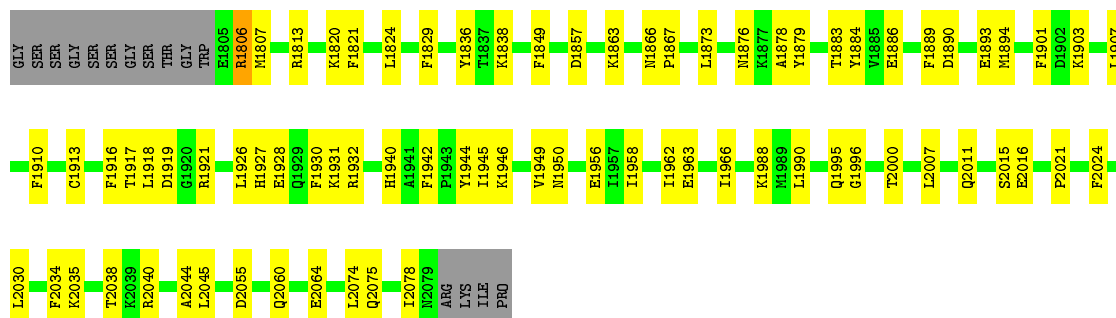
- Molecule 1: Dedicator of cytokinesis protein 7





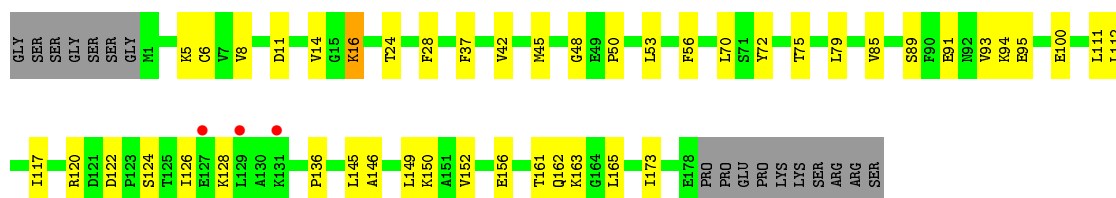
- Molecule 1: Dedicator of cytokinesis protein 7

Chain G: 68% 26% 5%



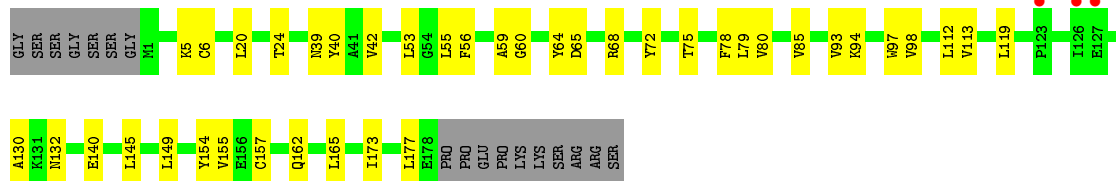
- Molecule 2: Cell division control protein 42 homolog

Chain B: 68% 23% 9%



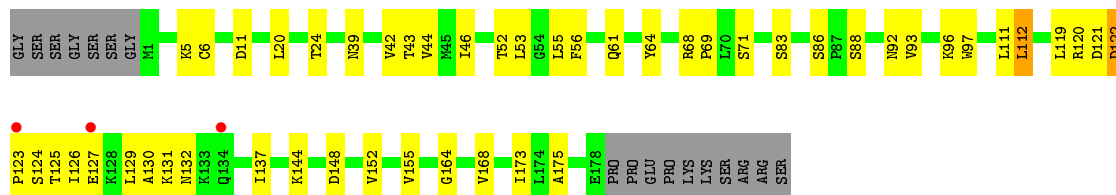
- Molecule 2: Cell division control protein 42 homolog

Chain D: 71% 21% 9%

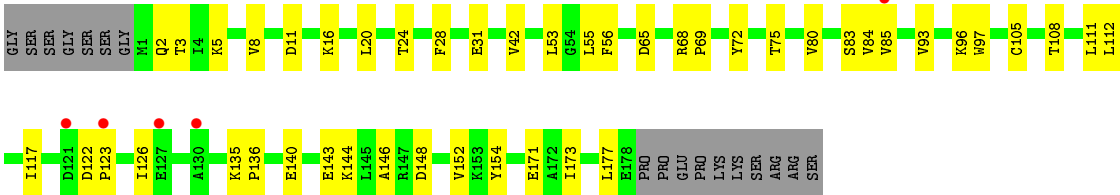


- Molecule 2: Cell division control protein 42 homolog

Chain F: 66% 25% 9%



- Molecule 2: Cell division control protein 42 homolog



4 Data and refinement statistics

Property	Value	Source
Space group	P 2 ₁ 2 ₁ 2	Depositor
Cell constants a, b, c, α , β , γ	194.55Å 105.29Å 117.89Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.64 – 3.23 48.64 – 3.23	Depositor EDS
% Data completeness (in resolution range)	99.4 (48.64-3.23) 99.5 (48.64-3.23)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.54 (at 3.25Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, R_{free}	0.184 , 0.259 0.184 , 0.259	Depositor DCC
R_{free} test set	1996 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å ²)	79.8	Xtriage
Anisotropy	0.431	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 58.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	14700	wwPDB-VP
Average B, all atoms (Å ²)	77.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.60% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	0.55	0/2342	0.69	0/3157
1	C	0.55	0/2342	0.74	0/3157
1	E	0.53	0/2342	0.69	0/3157
1	G	0.54	0/2342	0.70	0/3157
2	B	0.47	0/1418	0.68	1/1930 (0.1%)
2	D	0.50	0/1418	0.68	0/1930
2	F	0.49	0/1418	0.66	1/1930 (0.1%)
2	H	0.49	0/1418	0.68	0/1930
All	All	0.52	0/15040	0.69	2/20348 (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	A	0	1
1	C	0	1
1	E	0	1
1	G	0	1
All	All	0	4

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	112	LEU	CA-CB-CG	-5.88	101.79	115.30
2	B	16	LYS	CD-CE-NZ	5.26	123.80	111.70

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1806	ARG	Peptide
1	C	1806	ARG	Peptide
1	E	1806	ARG	Peptide
1	G	1806	ARG	Peptide

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	A	2287	0	2255	63	0
1	C	2287	0	2255	51	0
1	E	2287	0	2255	49	0
1	G	2287	0	2255	62	0
2	B	1388	0	1405	31	0
2	D	1388	0	1405	30	0
2	F	1388	0	1405	31	0
2	H	1388	0	1405	29	0
All	All	14700	0	14640	333	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 11.

The worst 5 of 333 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:2057:LYS:HE3	1:C:2061:ARG:HH22	1.32	0.94
1:G:2007:LEU:HD22	1:G:2011:GLN:HE21	1.43	0.82
1:A:1806:ARG:HH11	1:A:1836:TYR:HA	1.46	0.81
2:D:6:CYS:SG	2:D:79:LEU:HD22	2.22	0.80
1:E:1890:ASP:HB2	1:E:1893:GLU:H	1.44	0.79

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	273/290 (94%)	247 (90%)	24 (9%)	2 (1%)	22	58
1	C	273/290 (94%)	255 (93%)	16 (6%)	2 (1%)	22	58
1	E	273/290 (94%)	253 (93%)	19 (7%)	1 (0%)	34	68
1	G	273/290 (94%)	253 (93%)	20 (7%)	0	100	100
2	B	176/195 (90%)	158 (90%)	16 (9%)	2 (1%)	14	48
2	D	176/195 (90%)	166 (94%)	10 (6%)	0	100	100
2	F	176/195 (90%)	159 (90%)	16 (9%)	1 (1%)	25	61
2	H	176/195 (90%)	156 (89%)	18 (10%)	2 (1%)	14	48
All	All	1796/1940 (93%)	1647 (92%)	139 (8%)	10 (1%)	25	61

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	1981	HIS
1	E	1926	LEU
2	F	39	ASN
2	B	162	GLN
2	B	50	PRO

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	250/261 (96%)	249 (100%)	1 (0%)	91	95

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	250/261 (96%)	247 (99%)	3 (1%)	71	86
1	E	250/261 (96%)	245 (98%)	5 (2%)	55	78
1	G	250/261 (96%)	249 (100%)	1 (0%)	91	95
2	B	158/172 (92%)	158 (100%)	0	100	100
2	D	158/172 (92%)	158 (100%)	0	100	100
2	F	158/172 (92%)	157 (99%)	1 (1%)	86	93
2	H	158/172 (92%)	157 (99%)	1 (1%)	86	93
All	All	1632/1732 (94%)	1620 (99%)	12 (1%)	84	92

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	1906	ASN
1	E	1939	SER
2	F	122	ASP
1	E	1823	ASP
1	E	2065	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
2	F	92	ASN
1	G	2011	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 ⓘ

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	275/290 (94%)	-0.20	0 100 100	41, 71, 119, 175	0
1	C	275/290 (94%)	-0.18	0 100 100	40, 64, 107, 133	0
1	E	275/290 (94%)	0.06	8 (2%) 51 40	46, 81, 127, 161	0
1	G	275/290 (94%)	-0.09	0 100 100	42, 71, 118, 161	0
2	B	178/195 (91%)	0.10	3 (1%) 70 60	47, 70, 128, 179	0
2	D	178/195 (91%)	0.11	3 (1%) 70 60	43, 67, 132, 175	0
2	F	178/195 (91%)	-0.07	3 (1%) 70 60	57, 78, 126, 166	0
2	H	178/195 (91%)	0.12	5 (2%) 53 41	54, 80, 135, 167	0
All	All	1812/1940 (93%)	-0.04	22 (1%) 79 70	40, 73, 123, 179	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	127	GLU	4.4
2	B	127	GLU	4.4
2	D	123	PRO	4.1
2	H	123	PRO	4.0
2	F	127	GLU	3.5

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.