



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 1, 2017 – 01:31 AM EDT

PDB ID : 1LJ7
Title : Crystal structure of calcium-depleted human C-reactive protein from perfectly twinned data
Authors : Ramadan, M.A.; Shrive, A.K.; Holden, D.; Myles, D.A.; Volanakis, J.E.; DeLucas, L.J.; Greenhough, T.J.
Deposited on : unknown
Resolution : 3.15 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
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20030345
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	rb-20030345

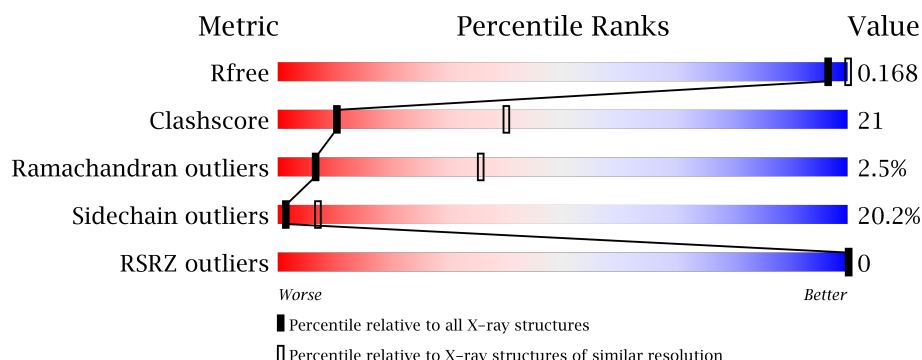
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.15 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}	100719	1259 (3.20-3.12)
Clashscore	112137	1397 (3.20-3.12)
Ramachandran outliers	110173	1368 (3.20-3.12)
Sidechain outliers	110143	1367 (3.20-3.12)
RSRZ outliers	101464	1264 (3.20-3.12)

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	206	
1	B	206	
1	C	206	
1	D	206	
1	E	206	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	206	 49% 40% 11%
1	G	206	 50% 41% 9%
1	H	206	 51% 37% 9% •
1	I	206	 50% 40% 7% •
1	J	206	 49% 43% 8%

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 15986 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.

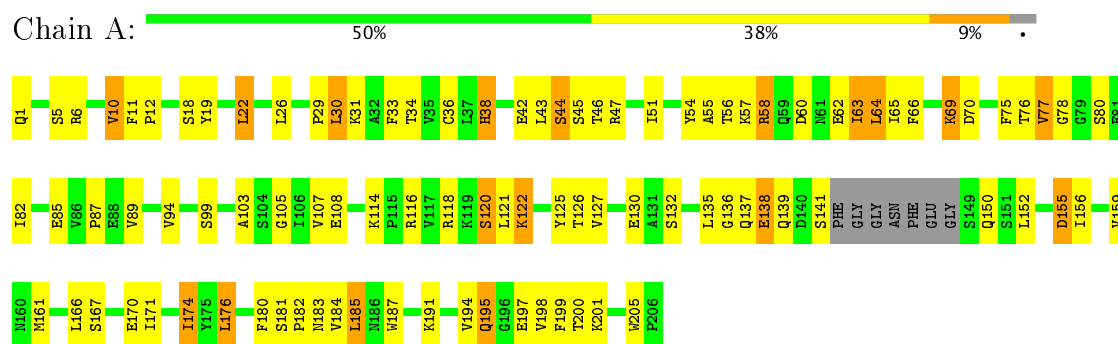
- Molecule 1 is a protein called C-reactive protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	199	Total	C	N	O	S	0	0	0
			1580	1025	253	298	4			
1	B	206	Total	C	N	O	S	0	0	0
			1631	1058	261	308	4			
1	C	199	Total	C	N	O	S	0	0	0
			1579	1025	253	297	4			
1	D	197	Total	C	N	O	S	0	0	0
			1566	1018	251	293	4			
1	E	197	Total	C	N	O	S	0	0	0
			1566	1018	251	293	4			
1	F	206	Total	C	N	O	S	0	0	0
			1631	1058	261	308	4			
1	G	206	Total	C	N	O	S	0	0	0
			1631	1058	261	308	4			
1	H	200	Total	C	N	O	S	0	0	0
			1591	1034	254	299	4			
1	I	199	Total	C	N	O	S	0	0	0
			1580	1025	253	298	4			
1	J	206	Total	C	N	O	S	0	0	0
			1631	1058	261	308	4			

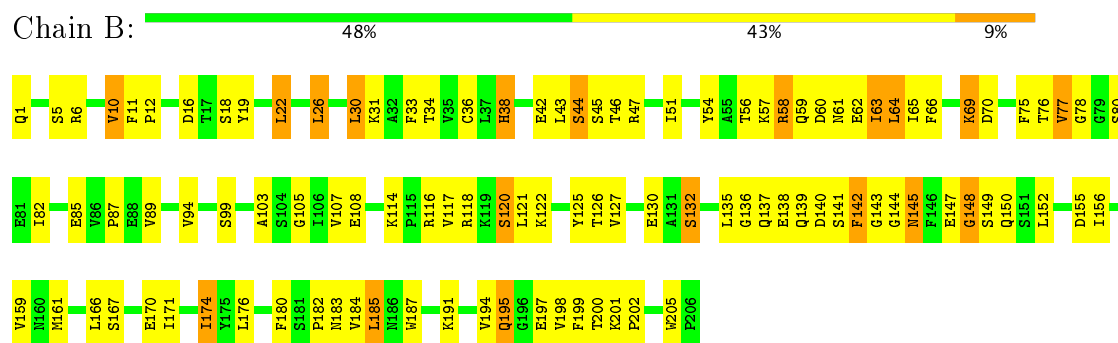
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: C-reactive protein



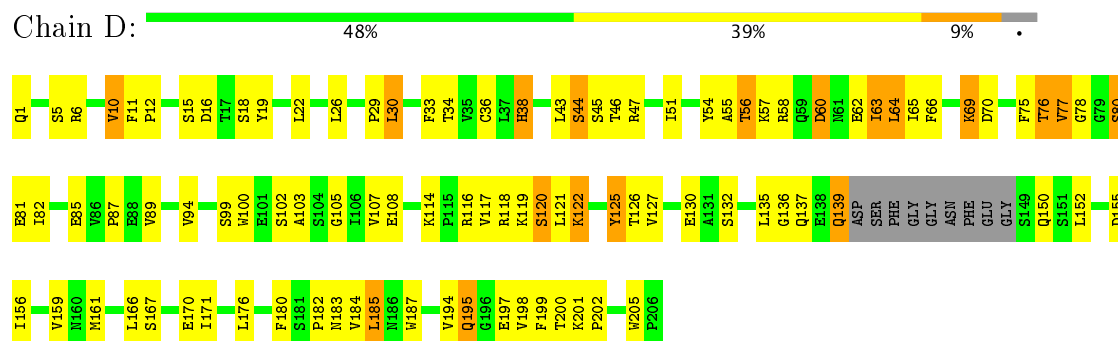
• Molecule 1: C-reactive protein



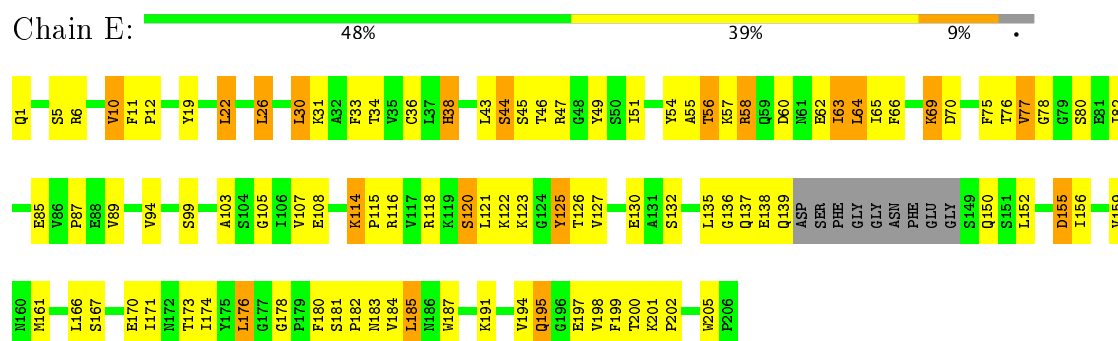
• Molecule 1: C-reactive protein



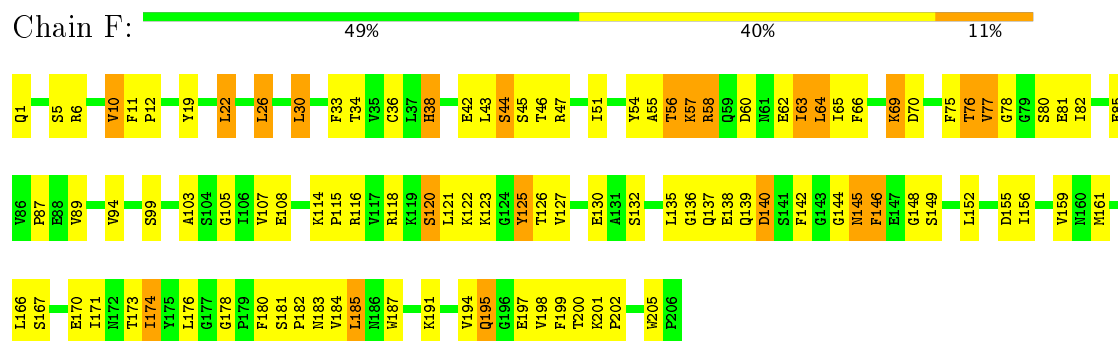
• Molecule 1: C-reactive protein



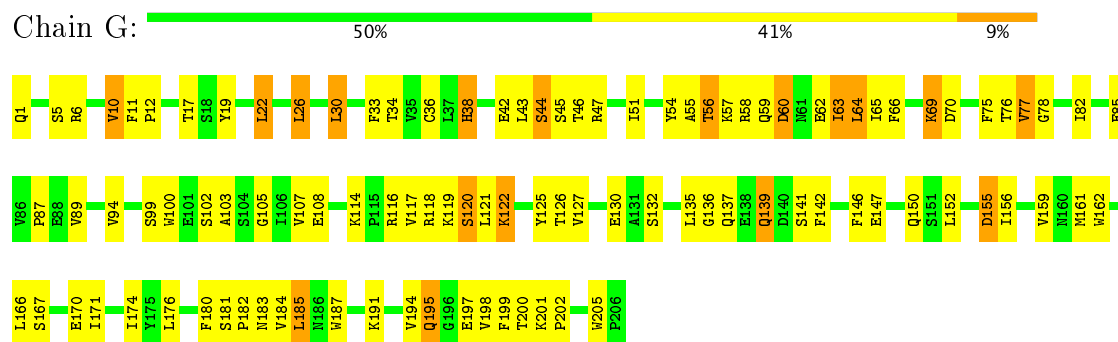
- Molecule 1: C-reactive protein



- Molecule 1: C-reactive protein

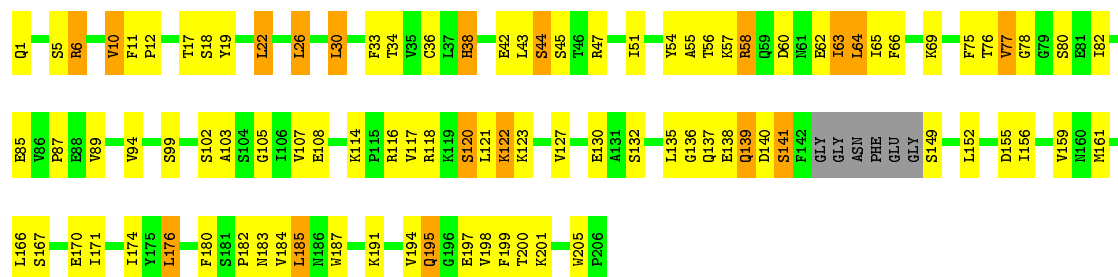


- Molecule 1: C-reactive protein



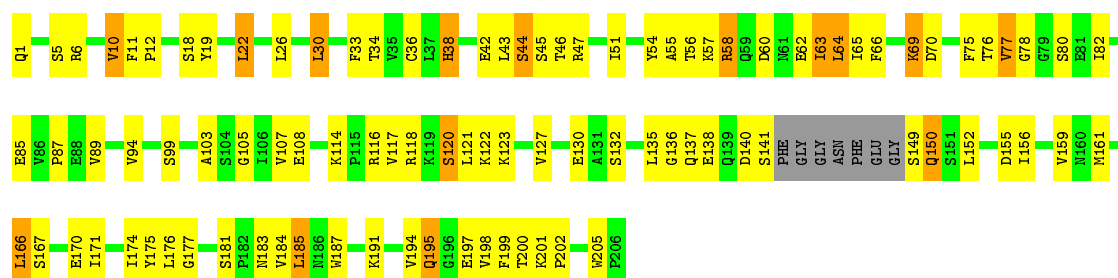
- Molecule 1: C-reactive protein

Chain H:  51% 37% 9% .



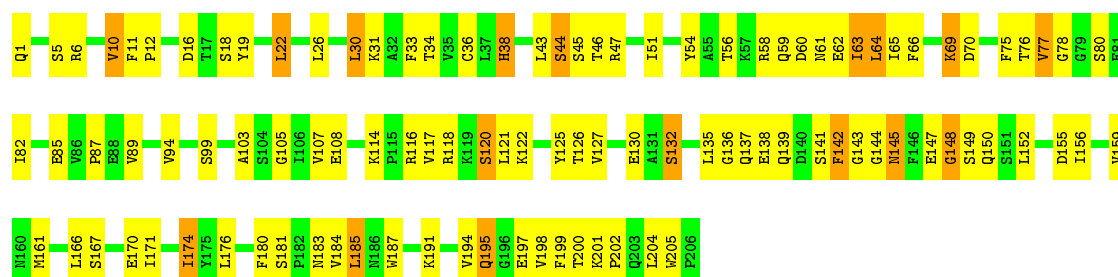
• Molecule 1: C-reactive protein

Chain I:  50% 40% 7% .



• Molecule 1: C-reactive protein

Chain J:  49% 43% 8% .



4 Data and refinement statistics

Property	Value	Source
Space group	P 43	Depositor
Cell constants a, b, c, α , β , γ	102.31Å 102.31Å 309.23Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.15 19.97 – 3.15	Depositor EDS
% Data completeness (in resolution range)	81.2 (20.00-3.15) 92.1 (19.97-3.15)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.50 (at 3.15Å)	Xtriage
Refinement program	X-PLOR 3.851	Depositor
R, R_{free}	0.186 , 0.199 0.158 , 0.168	Depositor DCC
R_{free} test set	2533 reflections (5.32%)	DCC
Wilson B-factor (Å ²)	70.0	Xtriage
Anisotropy	0.376	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.24 , -9.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.40$, $\langle L^2 \rangle = 0.23$	Xtriage
Estimated twinning fraction	0.499 for h,-k,-l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	15986	wwPDB-VP
Average B, all atoms (Å ²)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.66% 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	0.39	0/1623	0.57	0/2206
1	B	0.37	0/1677	0.58	0/2279
1	C	0.38	0/1622	0.56	0/2204
1	D	0.44	0/1609	0.58	0/2187
1	E	0.42	0/1609	0.58	0/2187
1	F	0.38	0/1677	0.57	0/2279
1	G	0.41	0/1677	0.57	0/2279
1	H	0.42	0/1635	0.58	0/2222
1	I	0.37	0/1623	0.56	0/2206
1	J	0.37	0/1677	0.57	0/2279
All	All	0.40	0/16429	0.57	0/22328

There are no bond length outliers.

There are no bond angle outliers.

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	1580	0	1553	58	0
1	B	1631	0	1593	73	1
1	C	1579	0	1553	74	0
1	D	1566	0	1544	68	0
1	E	1566	0	1544	72	0
1	F	1631	0	1593	70	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	G	1631	0	1593	69	1
1	H	1591	0	1562	61	0
1	I	1580	0	1553	60	0
1	J	1631	0	1593	70	0
All	All	15986	0	15681	659	1

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 659 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:58:ARG:HB3	1:C:58:ARG:NH1	1.70	1.06
1:F:140:ASP:HB2	1:F:145:ASN:O	1.58	1.01
1:D:55:ALA:HB3	1:D:139:GLN:NE2	1.87	0.89
1:C:58:ARG:HH11	1:C:58:ARG:HB3	1.34	0.88
1:G:55:ALA:HB3	1:G:139:GLN:NE2	1.88	0.86

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:140:ASP:OD2	1:G:58:ARG:NH2[1_655]	2.16	0.04

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	195/206 (95%)	174 (89%)	16 (8%)	5 (3%)	6	34
1	B	204/206 (99%)	180 (88%)	17 (8%)	7 (3%)	4	26

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	195/206 (95%)	176 (90%)	16 (8%)	3 (2%)	12	49
1	D	193/206 (94%)	172 (89%)	16 (8%)	5 (3%)	6	34
1	E	193/206 (94%)	173 (90%)	16 (8%)	4 (2%)	8	40
1	F	204/206 (99%)	179 (88%)	20 (10%)	5 (2%)	6	35
1	G	204/206 (99%)	179 (88%)	19 (9%)	6 (3%)	5	31
1	H	196/206 (95%)	178 (91%)	14 (7%)	4 (2%)	9	42
1	I	195/206 (95%)	175 (90%)	16 (8%)	4 (2%)	8	40
1	J	204/206 (99%)	179 (88%)	18 (9%)	7 (3%)	4	26
All	All	1983/2060 (96%)	1765 (89%)	168 (8%)	50 (2%)	6	35

5 of 50 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	44	SER
1	B	44	SER
1	B	145	ASN
1	C	44	SER
1	D	44	SER

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	176/180 (98%)	139 (79%)	37 (21%)	1	6
1	B	180/180 (100%)	145 (81%)	35 (19%)	1	8
1	C	175/180 (97%)	139 (79%)	36 (21%)	1	6
1	D	174/180 (97%)	142 (82%)	32 (18%)	2	10
1	E	174/180 (97%)	139 (80%)	35 (20%)	1	7
1	F	180/180 (100%)	142 (79%)	38 (21%)	1	6
1	G	180/180 (100%)	145 (81%)	35 (19%)	1	8
1	H	177/180 (98%)	140 (79%)	37 (21%)	1	6

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	I	176/180 (98%)	137 (78%)	39 (22%)	1	5
1	J	180/180 (100%)	146 (81%)	34 (19%)	2	9
All	All	1772/1800 (98%)	1414 (80%)	358 (20%)	1	7

5 of 358 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	120	SER
1	F	139	GLN
1	J	31	LYS
1	E	166	LEU
1	F	38	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 48 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	139	GLN
1	F	139	GLN
1	J	137	GLN
1	E	150	GLN
1	F	61	ASN

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	199/206 (96%)	-0.75	0 100 100	4, 20, 42, 61	0
1	B	206/206 (100%)	-0.69	0 100 100	3, 19, 38, 62	0
1	C	199/206 (96%)	-0.71	0 100 100	6, 19, 42, 56	0
1	D	197/206 (95%)	-0.72	0 100 100	4, 19, 41, 54	0
1	E	197/206 (95%)	-0.78	0 100 100	5, 20, 42, 51	0
1	F	206/206 (100%)	-0.71	0 100 100	5, 19, 42, 53	0
1	G	206/206 (100%)	-0.65	0 100 100	2, 18, 46, 58	0
1	H	200/206 (97%)	-0.70	0 100 100	5, 19, 42, 72	0
1	I	199/206 (96%)	-0.78	0 100 100	6, 20, 42, 60	0
1	J	206/206 (100%)	-0.71	0 100 100	5, 20, 40, 59	0
All	All	2015/2060 (97%)	-0.72	0 100 100	2, 19, 42, 72	0

There are no RSRZ outliers to report.

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.