



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 17, 2020 – 05:45 am BST

PDB ID : 5FOA  
Title : Crystal Structure of Human Complement C3b in complex with DAF (CCP2-4)  
Authors : Forneris, F.; Wu, J.; Xue, X.; Gros, P.  
Deposited on : 2015-11-18  
Resolution : 4.19 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

<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

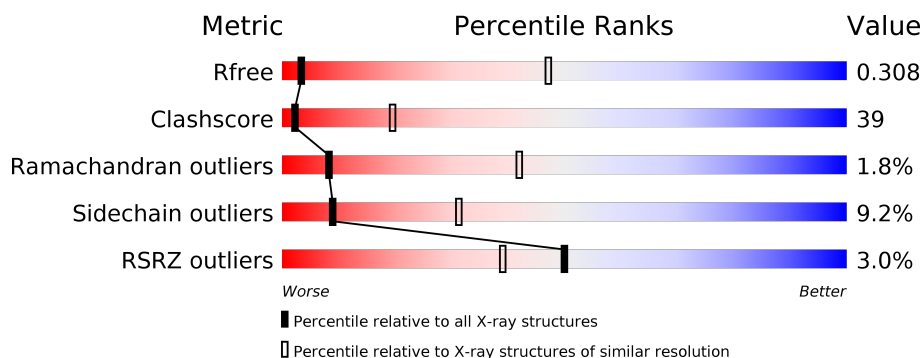
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.19 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	1034 (4.60-3.76)
Clashscore	141614	1030 (4.54-3.80)
Ramachandran outliers	138981	1006 (4.58-3.78)
Sidechain outliers	138945	1037 (4.60-3.76)
RSRZ outliers	127900	1056 (4.66-3.70)

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 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	645	<div> <div>3%</div> <div> <div></div> <div>48%</div> <div>45%</div> <div>6%</div> </div> </div>
1	C	645	<div> <div>2%</div> <div> <div></div> <div>52%</div> <div>41%</div> <div>5%</div> </div> </div>
2	B	915	<div> <div>3%</div> <div> <div></div> <div>43%</div> <div>44%</div> <div>10%</div> </div> </div>
2	D	915	<div> <div>2%</div> <div> <div></div> <div>43%</div> <div>44%</div> <div>10%</div> </div> </div>
3	E	194	<div> <div>6%</div> <div> <div></div> <div>41%</div> <div>43%</div> <div>10%</div> </div> </div>
3	F	194	<div> <div>10%</div> <div> <div></div> <div>41%</div> <div>40%</div> <div>14%</div> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 27275 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 COMPLEMENT C3 BETA CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	642	Total	C	N	O	S	0	0	0
			5002	3185	848	954	15			
1	C	642	Total	C	N	O	S	0	0	0
			5002	3185	848	954	15			

- Molecule 2 is a protein called COMPLEMENT C3B ALPHA CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	902	Total	C	N	O	S	0	0	0
			7177	4550	1205	1384	38			
2	D	902	Total	C	N	O	S	0	0	0
			7164	4543	1204	1380	37			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1013	GLU	GLN	SEE REMARK 999	UNP P01024
D	1013	GLU	GLN	SEE REMARK 999	UNP P01024

- Molecule 3 is a protein called DECAY ACCELERATING FACTOR, CD55.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	189	Total	C	N	O	S	0	0	0
			1465	920	249	283	13			
3	F	189	Total	C	N	O	S	0	0	0
			1465	920	249	283	13			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	95	GLY	-	expression tag	UNP P08174
E	96	SER	-	expression tag	UNP P08174

*Continued on next page...*

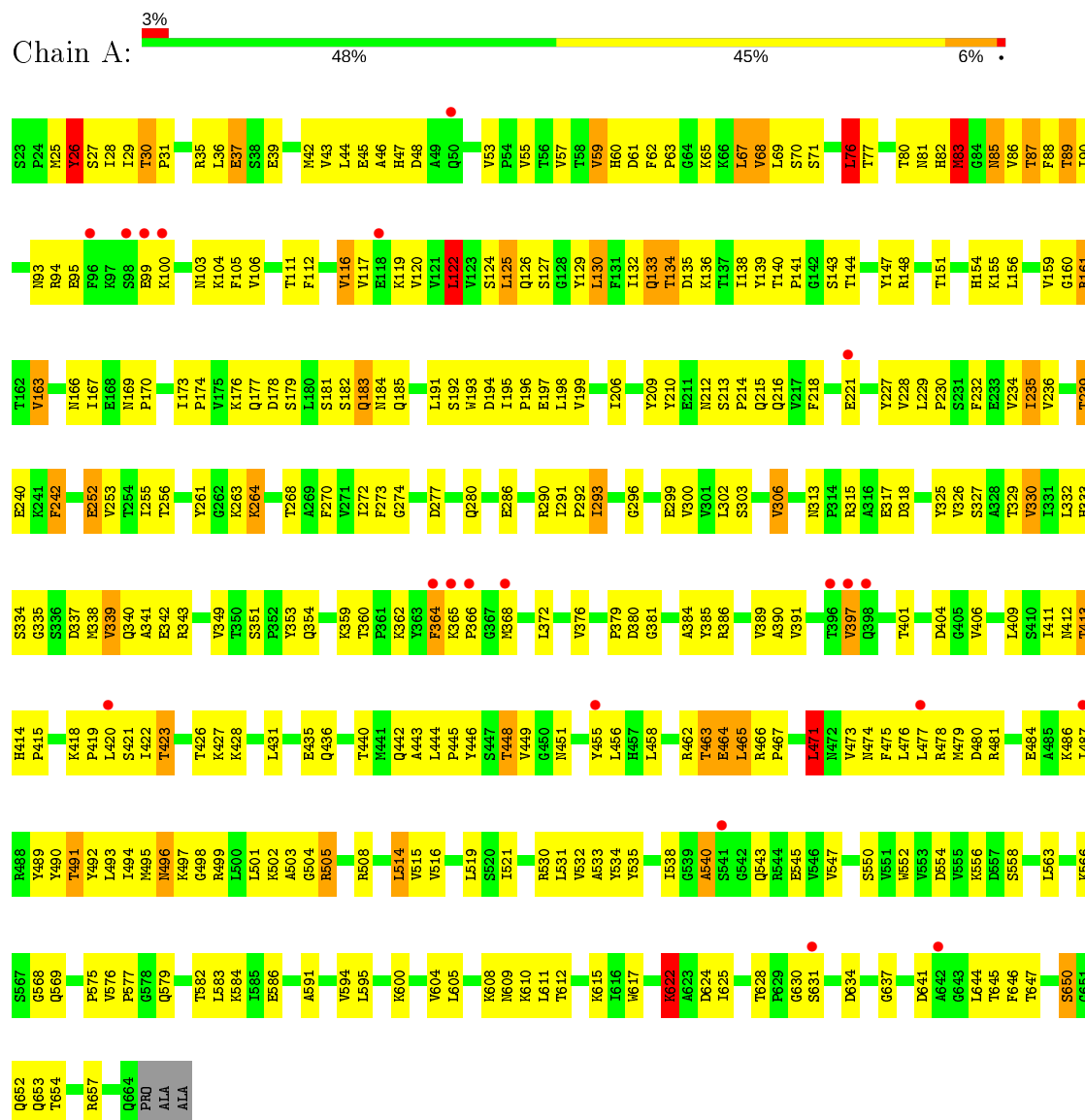
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	286	ALA	-	expression tag	UNP P08174
E	287	ALA	-	expression tag	UNP P08174
E	288	ALA	-	expression tag	UNP P08174
F	95	GLY	-	expression tag	UNP P08174
F	96	SER	-	expression tag	UNP P08174
F	286	ALA	-	expression tag	UNP P08174
F	287	ALA	-	expression tag	UNP P08174
F	288	ALA	-	expression tag	UNP P08174

### 3 Residue-property plots

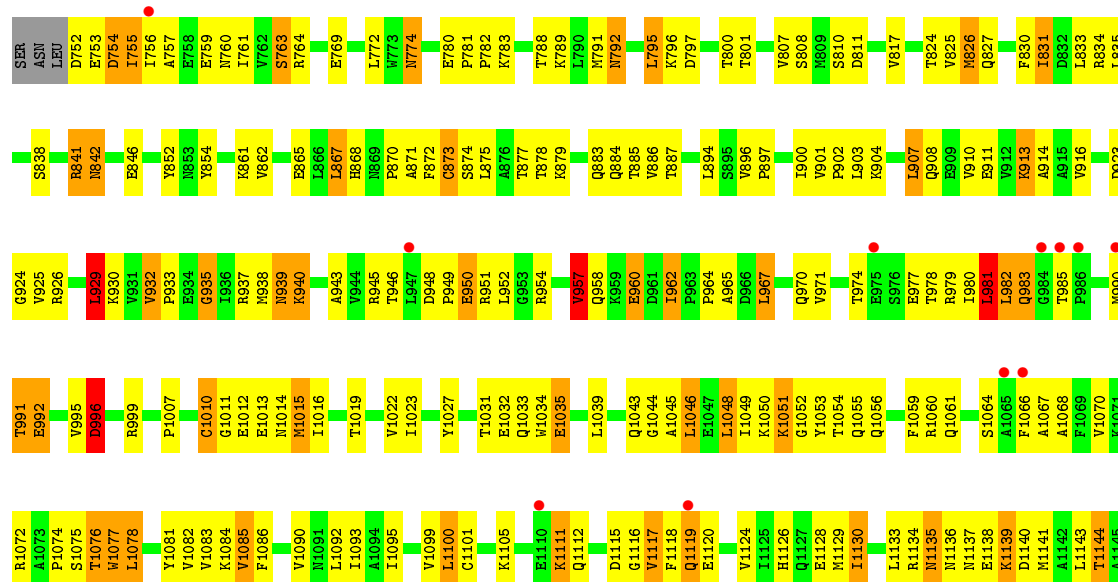
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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.

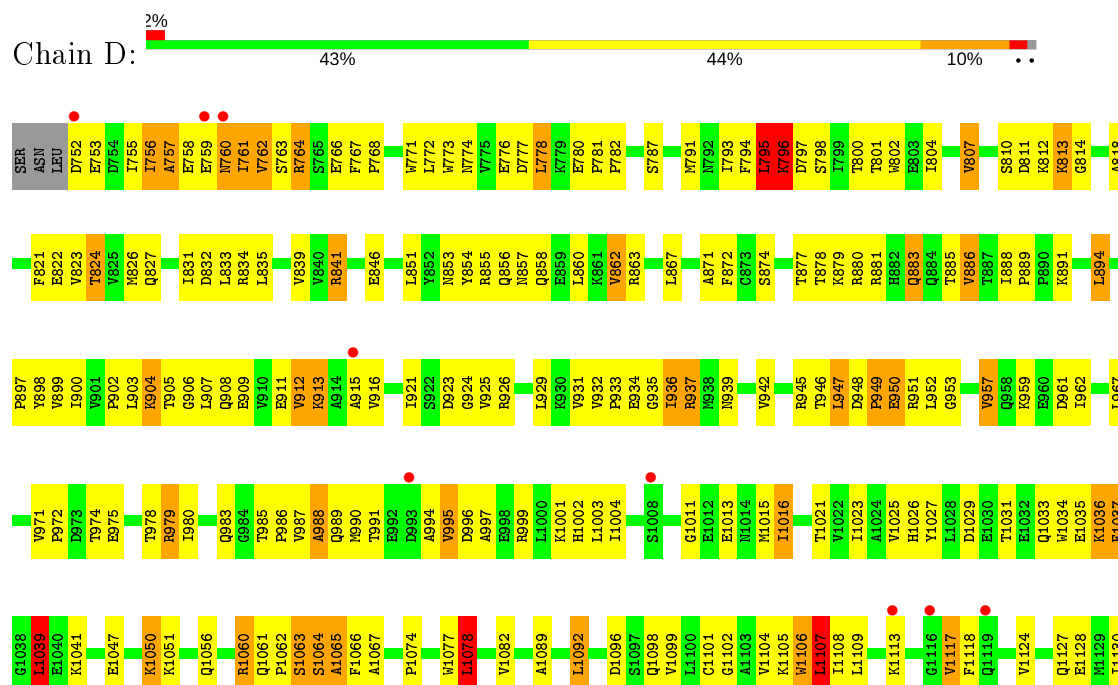
#### • Molecule 1: COMPLEMENT C3 BETA CHAIN

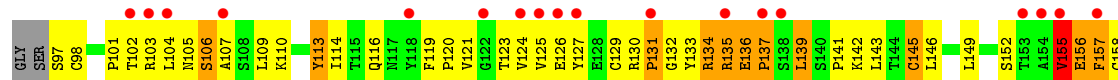


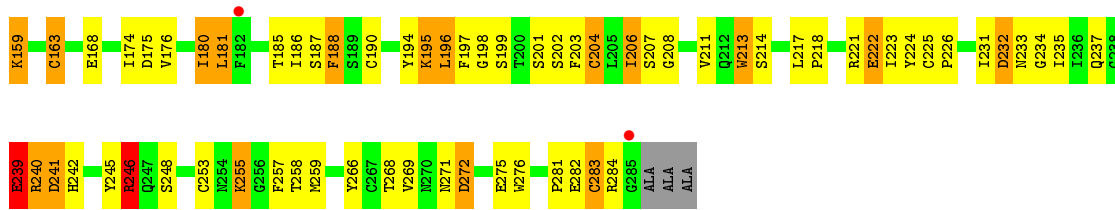
#### • Molecule 1: COMPLEMENT C3 BETA CHAIN











## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	117.43 Å   142.38 Å   323.70 Å 90.00°   90.00°   90.00°	Depositor
Resolution (Å)	110.39 – 4.19 110.39 – 4.19	Depositor EDS
% Data completeness (in resolution range)	95.5 (110.39-4.19) 89.6 (110.39-4.19)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.79 (at 4.15 Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.270   ,   0.307 0.272   ,   0.308	Depositor DCC
$R_{free}$ test set	1955 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	124.1	Xtriage
Anisotropy	0.389	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 115.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.82	EDS
Total number of atoms	27275	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	124.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.49% 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, 0.333 respectively for untwinned datasets, and 0.375, 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	0.58	2/5103 (0.0%)	0.90	16/6934 (0.2%)
1	C	0.63	5/5103 (0.1%)	0.95	21/6934 (0.3%)
2	B	0.71	16/7319 (0.2%)	1.07	55/9912 (0.6%)
2	D	0.71	10/7306 (0.1%)	1.09	49/9896 (0.5%)
3	E	0.82	5/1506 (0.3%)	1.07	11/2048 (0.5%)
3	F	0.68	1/1506 (0.1%)	1.00	11/2048 (0.5%)
All	All	0.68	39/27843 (0.1%)	1.02	163/37772 (0.4%)

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	2
2	B	0	4
2	D	0	4
3	F	0	2
All	All	0	12

The worst 5 of 39 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1313	TRP	CE3-CZ3	-15.41	1.12	1.38
3	E	276	TRP	CB-CG	-9.69	1.32	1.50
2	D	813	LYS	CD-CE	-8.91	1.28	1.51
1	A	37	GLU	CG-CD	-8.84	1.38	1.51
3	E	158	CYS	CB-SG	-8.41	1.68	1.82

The worst 5 of 163 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	1554	LEU	CB-CG-CD1	-12.82	89.20	111.00

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	83	MET	CG-SD-CE	-12.34	80.46	100.20
2	D	1649	LEU	CB-CG-CD2	-12.30	90.09	111.00
2	B	1292	LEU	CA-CB-CG	12.19	143.33	115.30
2	B	1300	LEU	CB-CG-CD1	-11.78	90.98	111.00

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	183	GLN	Mainchain
1	A	622	LYS	Mainchain
2	B	1292	LEU	Mainchain
2	B	1311	ILE	Mainchain
2	B	841	ARG	Mainchain

## 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	5002	0	5058	352	2
1	C	5002	0	5058	343	1
2	B	7177	0	7086	567	2
2	D	7164	0	7066	667	1
3	E	1465	0	1389	153	1
3	F	1465	0	1389	142	0
All	All	27275	0	27046	2109	4

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 39.

The worst 5 of 2109 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:67:LEU:CG	1:C:67:LEU:CD2	1.78	1.61
1:C:32:ASN:HD21	1:C:657:ARG:HD3	1.11	1.16

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:466:ARG:HG3	1:A:556:LYS:HD3	1.28	1.15
2:D:978:THR:HG22	2:D:1346:THR:HG22	1.31	1.10
3:F:157:PHE:O	3:F:159:LYS:NZ	1.84	1.09

All (4) 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:A:77:THR:OG1	2:B:954:ARG:NH2[3_455]	2.06	0.14
1:C:413:THR:O	1:C:512:GLN:NE2[2_755]	2.15	0.05
2:B:1577:GLN:NE2	3:E:133:TYR:OH[3_565]	2.15	0.05
1:A:435:GLU:OE2	2:D:1431:LYS:NZ[2_755]	2.19	0.01

## 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	640/645 (99%)	598 (93%)	37 (6%)	5 (1%)	19	59
1	C	640/645 (99%)	599 (94%)	34 (5%)	7 (1%)	14	51
2	B	898/915 (98%)	822 (92%)	58 (6%)	18 (2%)	7	40
2	D	898/915 (98%)	818 (91%)	57 (6%)	23 (3%)	5	35
3	E	187/194 (96%)	165 (88%)	17 (9%)	5 (3%)	5	34
3	F	187/194 (96%)	165 (88%)	17 (9%)	5 (3%)	5	34
All	All	3450/3508 (98%)	3167 (92%)	220 (6%)	63 (2%)	8	42

5 of 63 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	950	GLU
2	B	1137	ASN

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
2	B	1286	ALA
2	B	1287	PRO
2	B	1573	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	564/567 (100%)	522 (93%)	42 (7%)	13	40
1	C	564/567 (100%)	527 (93%)	37 (7%)	16	43
2	B	791/810 (98%)	719 (91%)	72 (9%)	9	32
2	D	787/810 (97%)	715 (91%)	72 (9%)	9	32
3	E	166/167 (99%)	138 (83%)	28 (17%)	2	13
3	F	166/167 (99%)	138 (83%)	28 (17%)	2	13
All	All	3038/3088 (98%)	2759 (91%)	279 (9%)	9	31

5 of 279 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	106	VAL
2	D	824	THR
3	F	155	VAL
1	C	191	LEU
1	C	413	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 40 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	50	GLN
1	C	215	GLN
3	E	237	GLN
1	C	126	GLN
1	C	589	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, 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	642/645 (99%)	0.15	21 (3%)	46	37	19, 107, 194, 287	0
1	C	642/645 (99%)	0.14	10 (1%)	72	62	16, 91, 178, 269	0
2	B	902/915 (98%)	0.22	23 (2%)	57	47	21, 143, 235, 308	0
2	D	902/915 (98%)	0.20	17 (1%)	66	58	16, 124, 224, 358	0
3	E	189/194 (97%)	0.39	12 (6%)	20	16	45, 129, 218, 254	0
3	F	189/194 (97%)	0.45	20 (10%)	6	6	38, 139, 231, 279	0
All	All	3466/3508 (98%)	0.21	103 (2%)	50	39	16, 118, 221, 358	0

The worst 5 of 103 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	985	THR	6.6
3	F	154	ALA	6.3
1	A	99	GLU	5.7
1	A	396	THR	4.8
2	B	1438	PHE	4.7

### 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](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.