



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:53 PM GMT

PDB ID : 4ATZ  
Title : Ad5 knob in complex with a designed ankyrin repeat protein  
Authors : Mittl, P.R.E.; Hess, C.; Dreier, B.  
Deposited on : 2012-05-11  
Resolution : 1.95 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

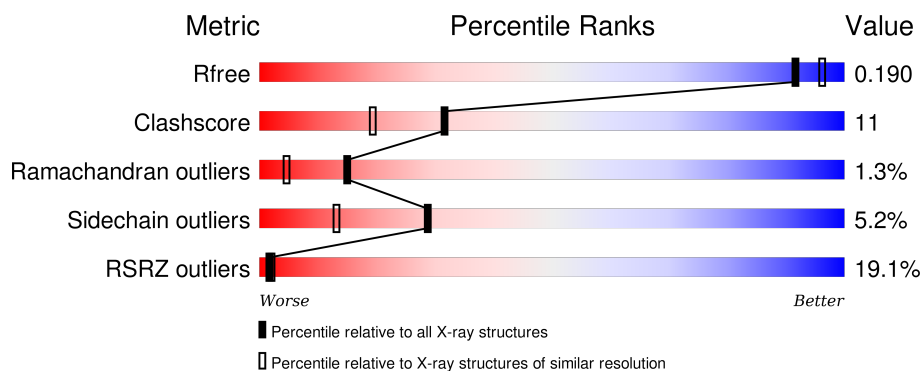
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.95 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	1833 (1.96-1.96)
Clashscore	102246	1953 (1.96-1.96)
Ramachandran outliers	100387	1936 (1.96-1.96)
Sidechain outliers	100360	1936 (1.96-1.96)
RSRZ outliers	91569	1835 (1.96-1.96)

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	201	<div> <div>11%</div> <div>69%</div> <div>17%</div> <div>9%</div> </div>
1	B	201	<div> <div>14%</div> <div>67%</div> <div>16%</div> <div>15%</div> </div>
1	C	201	<div> <div>15%</div> <div>72%</div> <div>12%</div> <div>12%</div> </div>
2	D	154	<div> <div>10%</div> <div>77%</div> <div>20%</div> <div>5%</div> </div>
2	E	154	<div> <div>16%</div> <div>83%</div> <div>15%</div> <div>5%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	154	<div><div></div><div>44%</div><div></div><div>75%</div><div></div><div>21%</div><div></div><div>• •</div></div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 8492 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 FIBER PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	182	Total	C	N	O	S	0	14	0
			1499	949	250	296	4			
1	B	171	Total	C	N	O	S	0	11	0
			1402	892	233	273	4			
1	C	176	Total	C	N	O	S	0	3	0
			1374	873	226	271	4			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	381	MET	-	EXPRESSION TAG	UNP P11818
A	382	ALA	-	EXPRESSION TAG	UNP P11818
A	383	HIS	-	EXPRESSION TAG	UNP P11818
A	384	HIS	-	EXPRESSION TAG	UNP P11818
A	385	HIS	-	EXPRESSION TAG	UNP P11818
A	386	HIS	-	EXPRESSION TAG	UNP P11818
A	387	HIS	-	EXPRESSION TAG	UNP P11818
A	388	HIS	-	EXPRESSION TAG	UNP P11818
A	389	GLY	-	EXPRESSION TAG	UNP P11818
A	390	SER	-	EXPRESSION TAG	UNP P11818
A	.	-	THR	DELETION	UNP P11818
A	.	-	ALA	DELETION	UNP P11818
A	.	-	TYR	DELETION	UNP P11818
A	.	-	THR	DELETION	UNP P11818
B	377	MET	-	EXPRESSION TAG	UNP P11818
B	378	ALA	-	EXPRESSION TAG	UNP P11818
B	379	HIS	-	EXPRESSION TAG	UNP P11818
B	380	HIS	-	EXPRESSION TAG	UNP P11818
B	381	HIS	-	EXPRESSION TAG	UNP P11818
B	382	HIS	-	EXPRESSION TAG	UNP P11818
B	383	HIS	-	EXPRESSION TAG	UNP P11818
B	384	HIS	-	EXPRESSION TAG	UNP P11818
B	385	GLY	-	EXPRESSION TAG	UNP P11818

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	386	SER	-	EXPRESSION TAG	UNP P11818
B	.	-	THR	DELETION	UNP P11818
B	.	-	ALA	DELETION	UNP P11818
B	.	-	TYR	DELETION	UNP P11818
B	.	-	THR	DELETION	UNP P11818
C	377	MET	-	EXPRESSION TAG	UNP P11818
C	378	ALA	-	EXPRESSION TAG	UNP P11818
C	379	HIS	-	EXPRESSION TAG	UNP P11818
C	380	HIS	-	EXPRESSION TAG	UNP P11818
C	381	HIS	-	EXPRESSION TAG	UNP P11818
C	382	HIS	-	EXPRESSION TAG	UNP P11818
C	383	HIS	-	EXPRESSION TAG	UNP P11818
C	384	HIS	-	EXPRESSION TAG	UNP P11818
C	385	GLY	-	EXPRESSION TAG	UNP P11818
C	386	SER	-	EXPRESSION TAG	UNP P11818
C	.	-	THR	DELETION	UNP P11818
C	.	-	ALA	DELETION	UNP P11818
C	.	-	TYR	DELETION	UNP P11818
C	.	-	THR	DELETION	UNP P11818

- Molecule 2 is a protein called DESIGNED ANKYRIN REPEAT PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	154	Total	C	N	O	S	0	2	0
			1172	742	200	228	2			
2	E	154	Total	C	N	O	S	0	1	0
			1167	739	199	227	2			
2	F	152	Total	C	N	O	S	0	0	0
			1143	721	197	223	2			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	201	Total	O	0	0
			201	201		
3	B	178	Total	O	0	0
			178	178		
3	C	132	Total	O	0	0
			132	132		
3	D	102	Total	O	0	0
			102	102		
3	E	74	Total	O	0	0
			74	74		

*Continued on next page...*

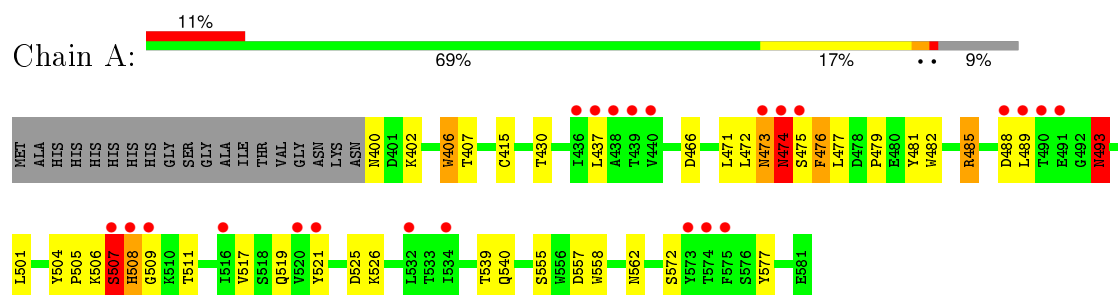
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	F	48	Total	O	0	0
			48	48		

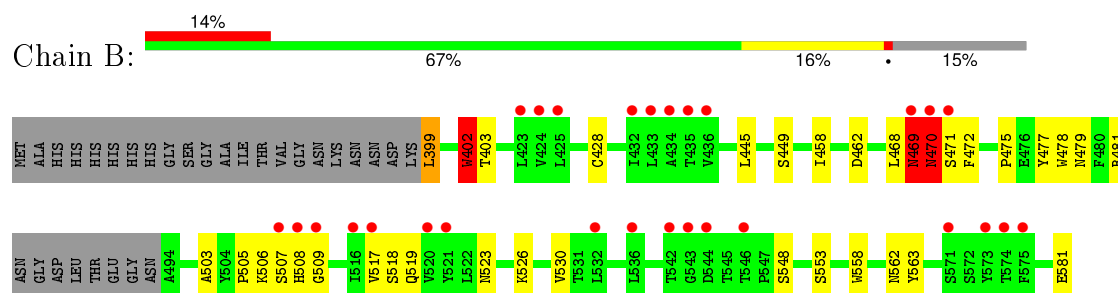
### 3 Residue-property plots

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 ( $\text{RSRZ} > 2$ ). 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.

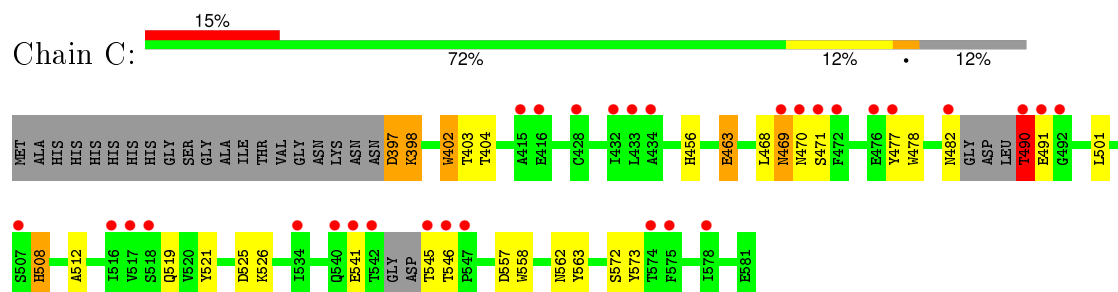
#### • Molecule 1: FIBER PROTEIN



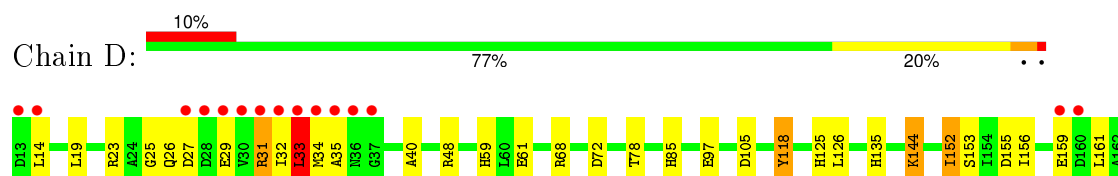
#### • Molecule 1: FIBER PROTEIN



#### • Molecule 1: FIBER PROTEIN

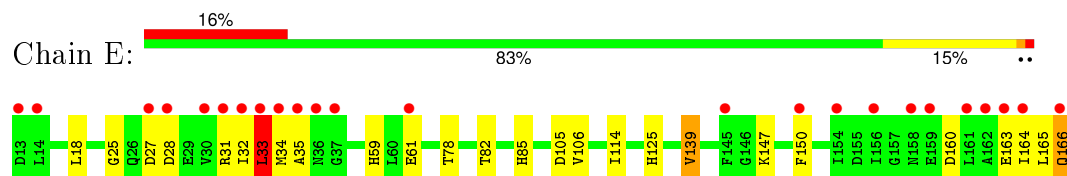


#### • Molecule 2: DESIGNED ANKYRIN REPEAT PROTEIN

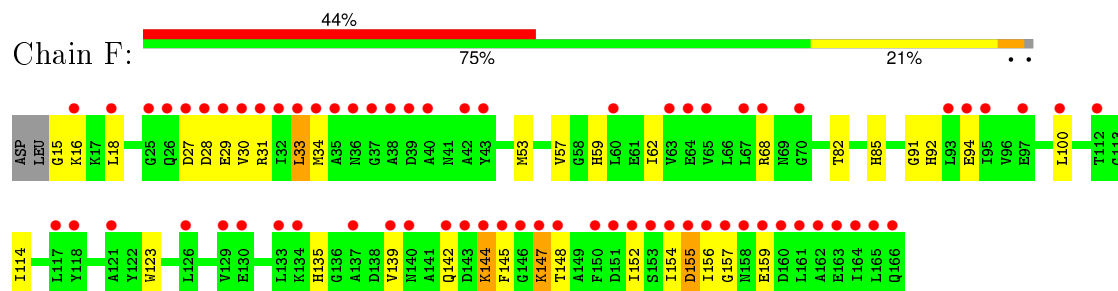




• Molecule 2: DESIGNED ANKYRIN REPEAT PROTEIN



• Molecule 2: DESIGNED ANKYRIN REPEAT PROTEIN





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	109.60Å 112.10Å 129.67Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.57 – 1.95 46.57 – 1.95	Depositor EDS
% Data completeness (in resolution range)	100.0 (46.57-1.95) 100.0 (46.57-1.95)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.75 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.161 , 0.193 0.157 , 0.190	Depositor DCC
$R_{free}$ test set	5802 reflections (5.23%)	DCC
Wilson B-factor (Å <sup>2</sup> )	35.3	Xtriage
Anisotropy	0.015	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 62.9	EDS
Estimated twinning fraction	0.019 for k,h,l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 116677 reflections	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	8492	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	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

### 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	1.51	10/1540 (0.6%)	1.31	6/2096 (0.3%)
1	B	1.48	7/1439 (0.5%)	1.24	6/1959 (0.3%)
1	C	1.37	8/1411 (0.6%)	1.19	3/1922 (0.2%)
2	D	1.24	2/1193 (0.2%)	1.09	8/1623 (0.5%)
2	E	1.10	0/1189	1.04	4/1617 (0.2%)
2	F	0.94	1/1161 (0.1%)	0.86	0/1579
All	All	1.31	28/7933 (0.4%)	1.15	27/10796 (0.3%)

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	3
1	B	0	1
1	C	0	2
All	All	0	6

The worst 5 of 28 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	402	TRP	CD2-CE2	7.91	1.50	1.41
1	B	518	SER	CB-OG	7.88	1.52	1.42
1	A	504	TYR	CD1-CE1	6.92	1.49	1.39
1	A	485	ARG	CZ-NH1	6.87	1.42	1.33
2	F	123	TRP	CD2-CE2	6.63	1.49	1.41

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	485	ARG	NE-CZ-NH2	-11.47	114.57	120.30
1	A	466	ASP	CB-CG-OD1	7.85	125.36	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	399	LEU	CB-CG-CD2	-6.84	99.37	111.00
2	D	105	ASP	CB-CG-OD1	6.56	124.21	118.30
1	B	399	LEU	CB-CG-CD1	6.54	122.11	111.00

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	472[B]	LEU	Peptide
1	A	474[B]	ASN	Peptide
1	A	507[B]	SER	Peptide
1	B	469[B]	ASN	Peptide
1	C	397	ASP	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	1499	0	1467	35	0
1	B	1402	0	1373	29	0
1	C	1374	0	1348	19	0
2	D	1172	0	1166	40	0
2	E	1167	0	1154	17	0
2	F	1143	0	1130	30	0
3	A	201	0	0	10	1
3	B	178	0	0	13	1
3	C	132	0	0	6	0
3	D	102	0	0	8	1
3	E	74	0	0	1	1
3	F	48	0	0	7	0
All	All	8492	0	7638	165	2

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 165 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)
2:D:166:GLN:HG3	3:D:2091:HOH:O	1.40	1.19
1:B:481:ARG:C	3:B:2079:HOH:O	1.84	1.13
2:D:126[B]:LEU:HD13	2:D:164:ILE:CD1	1.80	1.12
1:A:473[B]:ASN:HB3	3:A:2082:HOH:O	1.49	1.09
1:A:474[B]:ASN:OD1	1:A:475[B]:SER:O	1.71	1.08

All (2) 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)
3:B:2170:HOH:O	3:E:2038:HOH:O[3_555]	2.14	0.06
3:A:2183:HOH:O	3:D:2078:HOH:O[4_555]	2.17	0.03

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	194/201 (96%)	175 (90%)	9 (5%)	10 (5%)	2	0
1	B	178/201 (89%)	164 (92%)	8 (4%)	6 (3%)	5	0
1	C	173/201 (86%)	162 (94%)	11 (6%)	0	100	100
2	D	154/154 (100%)	151 (98%)	1 (1%)	2 (1%)	15	4
2	E	153/154 (99%)	152 (99%)	1 (1%)	0	100	100
2	F	150/154 (97%)	138 (92%)	9 (6%)	3 (2%)	9	2
All	All	1002/1065 (94%)	942 (94%)	39 (4%)	21 (2%)	15	2

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	474[A]	ASN
1	A	474[B]	ASN
1	A	476[A]	PHE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	476[B]	PHE
1	A	507[A]	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	169/171 (99%)	162 (96%)	7 (4%)	37	22
1	B	158/171 (92%)	151 (96%)	7 (4%)	35	19
1	C	156/171 (91%)	147 (94%)	9 (6%)	25	10
2	D	119/117 (102%)	115 (97%)	4 (3%)	44	30
2	E	118/117 (101%)	109 (92%)	9 (8%)	16	5
2	F	115/117 (98%)	107 (93%)	8 (7%)	19	6
All	All	835/864 (97%)	791 (95%)	44 (5%)	29	13

5 of 44 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	491	GLU
2	D	31	ARG
2	F	147	LYS
1	C	508	HIS
1	C	562	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	562	ASN
2	D	85	HIS
2	F	85	HIS
2	D	26	GLN
2	D	59	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	182/201 (90%)	0.74	23 (12%) 5 8	21, 28, 57, 94	0
1	B	171/201 (85%)	0.92	28 (16%) 2 3	21, 30, 60, 75	0
1	C	176/201 (87%)	0.90	30 (17%) 2 3	22, 35, 89, 116	0
2	D	154/154 (100%)	0.48	16 (10%) 8 13	26, 39, 76, 99	0
2	E	154/154 (100%)	0.65	24 (15%) 3 4	28, 46, 82, 104	0
2	F	152/154 (98%)	2.36	68 (44%) 0 0	37, 59, 110, 122	0
All	All	989/1065 (92%)	0.99	189 (19%) 2 2	21, 37, 89, 122	0

The worst 5 of 189 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	161	LEU	10.6
2	F	145	PHE	10.5
2	F	165	LEU	9.9
2	F	164	ILE	9.0
2	F	154	ILE	8.4

### 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.