



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 26, 2020 – 11:41 pm BST

PDB ID : 4F37  
Title : Structure of the tethered N-terminus of Alzheimer's disease A peptide  
Authors : Nisbet, R.M.; Nuttall, S.D.; Caine, J.M.; Rober, R.; Hittaki, M.; Pearce, L.A.; Davydova, N.; Masters, C.L.; Varghese, J.N.; Streltsov, V.A.  
Deposited on : 2012-05-09  
Resolution : 2.57 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

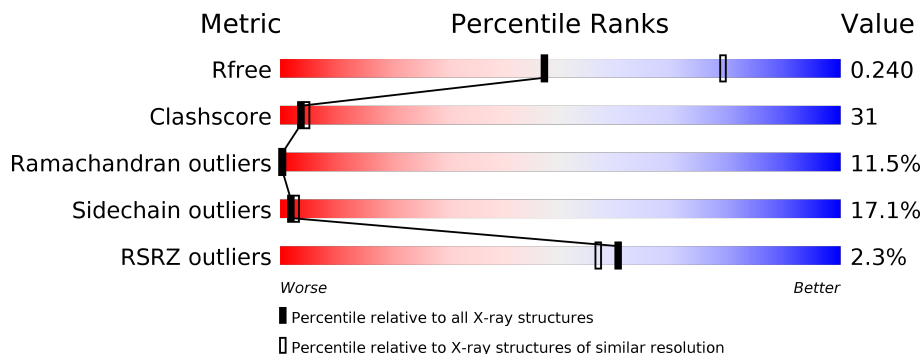
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.57 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	3676 (2.60-2.56)
Clashscore	141614	4049 (2.60-2.56)
Ramachandran outliers	138981	3979 (2.60-2.56)
Sidechain outliers	138945	3979 (2.60-2.56)
RSRZ outliers	127900	3614 (2.60-2.56)

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	124	
1	B	124	
2	F	228	
2	H	228	
3	K	219	
3	L	219	

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 8769 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 Colicin-E7 immunity protein.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
1	A	102	829	520	138	171	0	0	0
1	B	102	829	520	138	171	0	0	0

There are 76 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	ASP	-	EXPRESSION TAG	UNP Q03708
A	2	ALA	-	EXPRESSION TAG	UNP Q03708
A	3	GLU	-	EXPRESSION TAG	UNP Q03708
A	4	PHE	-	EXPRESSION TAG	UNP Q03708
A	5	ARG	-	EXPRESSION TAG	UNP Q03708
A	6	HIS	-	EXPRESSION TAG	UNP Q03708
A	7	ASP	-	EXPRESSION TAG	UNP Q03708
A	8	SER	-	EXPRESSION TAG	UNP Q03708
A	9	GLY	-	EXPRESSION TAG	UNP Q03708
A	10	TYR	-	EXPRESSION TAG	UNP Q03708
A	11	GLU	-	EXPRESSION TAG	UNP Q03708
A	12	VAL	-	EXPRESSION TAG	UNP Q03708
A	13	HIS	-	EXPRESSION TAG	UNP Q03708
A	14	HIS	-	EXPRESSION TAG	UNP Q03708
A	15	GLN	-	EXPRESSION TAG	UNP Q03708
A	16	LYS	-	EXPRESSION TAG	UNP Q03708
A	17	SER	-	EXPRESSION TAG	UNP Q03708
A	104	ALA	-	EXPRESSION TAG	UNP Q03708
A	105	ALA	-	EXPRESSION TAG	UNP Q03708
A	106	ALA	-	EXPRESSION TAG	UNP Q03708
A	107	ASP	-	EXPRESSION TAG	UNP Q03708
A	108	TYR	-	EXPRESSION TAG	UNP Q03708
A	109	LYS	-	EXPRESSION TAG	UNP Q03708
A	110	ASP	-	EXPRESSION TAG	UNP Q03708
A	111	ASP	-	EXPRESSION TAG	UNP Q03708

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	112	ASP	-	EXPRESSION TAG	UNP Q03708
A	113	ASP	-	EXPRESSION TAG	UNP Q03708
A	114	LYS	-	EXPRESSION TAG	UNP Q03708
A	115	ALA	-	EXPRESSION TAG	UNP Q03708
A	116	ALA	-	EXPRESSION TAG	UNP Q03708
A	117	ASP	-	EXPRESSION TAG	UNP Q03708
A	118	TYR	-	EXPRESSION TAG	UNP Q03708
A	119	LYS	-	EXPRESSION TAG	UNP Q03708
A	120	ASP	-	EXPRESSION TAG	UNP Q03708
A	121	ASP	-	EXPRESSION TAG	UNP Q03708
A	122	ASP	-	EXPRESSION TAG	UNP Q03708
A	123	ASP	-	EXPRESSION TAG	UNP Q03708
A	124	LYS	-	EXPRESSION TAG	UNP Q03708
B	1	ASP	-	EXPRESSION TAG	UNP Q03708
B	2	ALA	-	EXPRESSION TAG	UNP Q03708
B	3	GLU	-	EXPRESSION TAG	UNP Q03708
B	4	PHE	-	EXPRESSION TAG	UNP Q03708
B	5	ARG	-	EXPRESSION TAG	UNP Q03708
B	6	HIS	-	EXPRESSION TAG	UNP Q03708
B	7	ASP	-	EXPRESSION TAG	UNP Q03708
B	8	SER	-	EXPRESSION TAG	UNP Q03708
B	9	GLY	-	EXPRESSION TAG	UNP Q03708
B	10	TYR	-	EXPRESSION TAG	UNP Q03708
B	11	GLU	-	EXPRESSION TAG	UNP Q03708
B	12	VAL	-	EXPRESSION TAG	UNP Q03708
B	13	HIS	-	EXPRESSION TAG	UNP Q03708
B	14	HIS	-	EXPRESSION TAG	UNP Q03708
B	15	GLN	-	EXPRESSION TAG	UNP Q03708
B	16	LYS	-	EXPRESSION TAG	UNP Q03708
B	17	SER	-	EXPRESSION TAG	UNP Q03708
B	104	ALA	-	EXPRESSION TAG	UNP Q03708
B	105	ALA	-	EXPRESSION TAG	UNP Q03708
B	106	ALA	-	EXPRESSION TAG	UNP Q03708
B	107	ASP	-	EXPRESSION TAG	UNP Q03708
B	108	TYR	-	EXPRESSION TAG	UNP Q03708
B	109	LYS	-	EXPRESSION TAG	UNP Q03708
B	110	ASP	-	EXPRESSION TAG	UNP Q03708
B	111	ASP	-	EXPRESSION TAG	UNP Q03708
B	112	ASP	-	EXPRESSION TAG	UNP Q03708
B	113	ASP	-	EXPRESSION TAG	UNP Q03708
B	114	LYS	-	EXPRESSION TAG	UNP Q03708
B	115	ALA	-	EXPRESSION TAG	UNP Q03708

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	116	ALA	-	EXPRESSION TAG	UNP Q03708
B	117	ASP	-	EXPRESSION TAG	UNP Q03708
B	118	TYR	-	EXPRESSION TAG	UNP Q03708
B	119	LYS	-	EXPRESSION TAG	UNP Q03708
B	120	ASP	-	EXPRESSION TAG	UNP Q03708
B	121	ASP	-	EXPRESSION TAG	UNP Q03708
B	122	ASP	-	EXPRESSION TAG	UNP Q03708
B	123	ASP	-	EXPRESSION TAG	UNP Q03708
B	124	LYS	-	EXPRESSION TAG	UNP Q03708

- Molecule 2 is a protein called Im7 immunity protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	F	225	Total	C	N	O	S	0	0	0
			1730	1094	292	339	5			
2	H	225	Total	C	N	O	S	0	0	0
			1730	1094	292	339	5			

- Molecule 3 is a protein called Fab WO2 anti-amyloid-beta antibody Fab fragment.

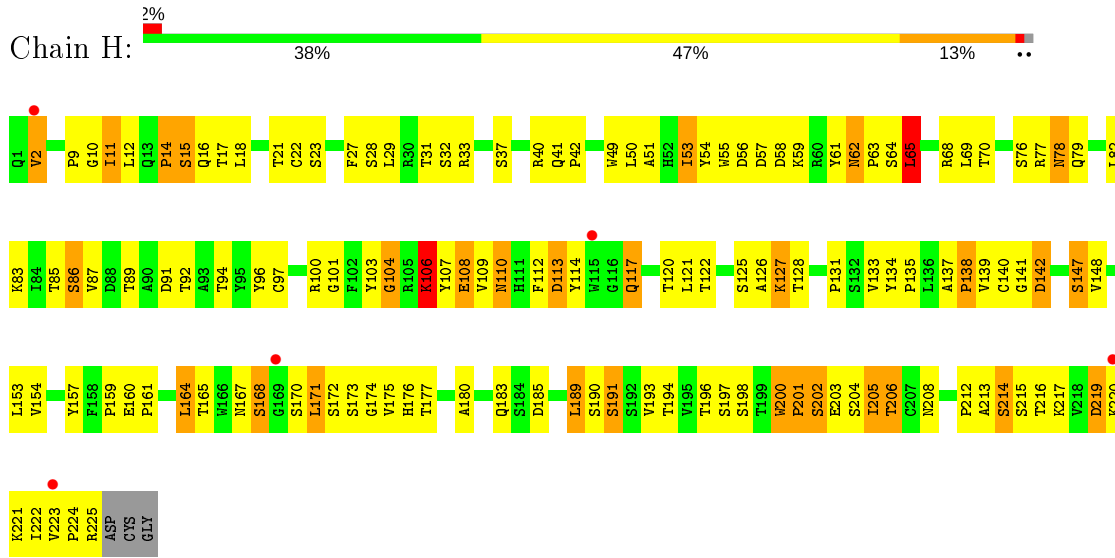
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	K	218	Total	C	N	O	S	0	0	0
			1693	1059	287	341	6			
3	L	218	Total	C	N	O	S	0	0	0
			1693	1059	287	341	6			

- Molecule 4 is water.

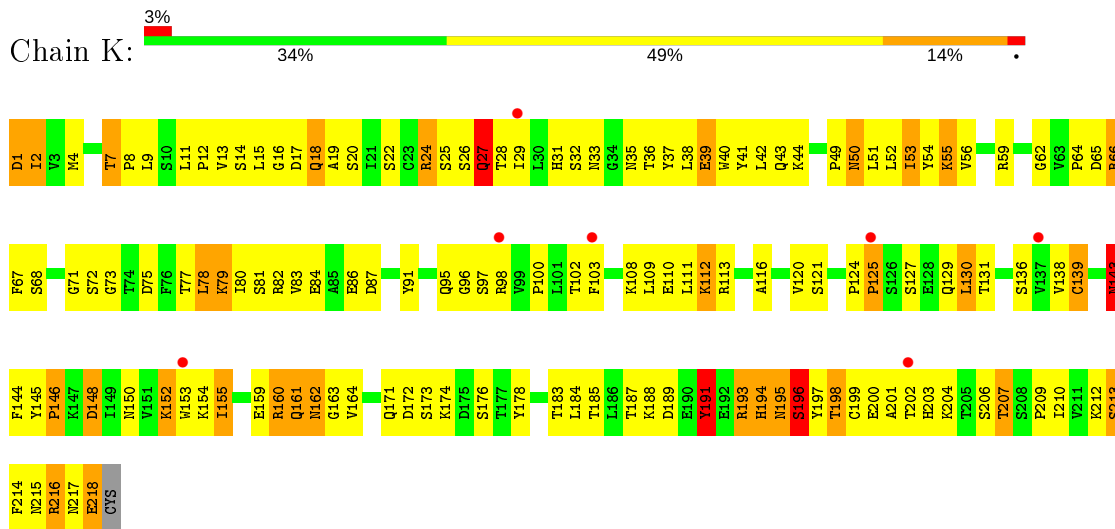
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	32	Total	O	0	0
			32	32		
4	B	34	Total	O	0	0
			34	34		
4	F	53	Total	O	0	0
			53	53		
4	H	50	Total	O	0	0
			50	50		
4	K	42	Total	O	0	0
			42	42		
4	L	54	Total	O	0	0
			54	54		



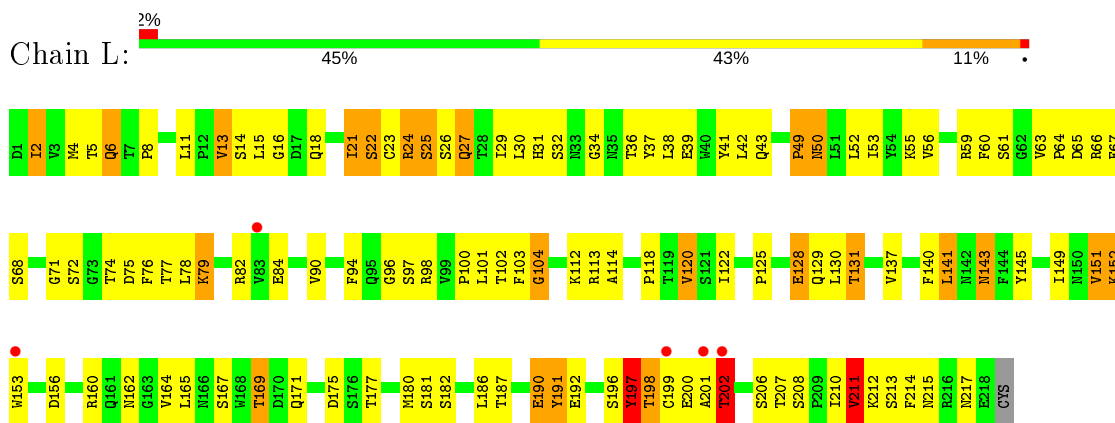
• Molecule 2: Im7 immunity protein



• Molecule 3: Fab WO2 anti-amyloid-beta antibody Fab fragment



• Molecule 3: Fab WO2 anti-amyloid-beta antibody Fab fragment



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	36.64Å 82.84Å 89.20Å 90.05° 92.51° 90.00°	Depositor
Resolution (Å)	44.46 – 2.57 44.56 – 2.50	Depositor EDS
% Data completeness (in resolution range)	77.9 (44.46-2.57) 85.8 (44.56-2.50)	Depositor EDS
$R_{merge}$	0.17	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.11 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.5.0110	Depositor
R, $R_{free}$	0.223 , 0.269 0.245 , 0.240	Depositor DCC
$R_{free}$ test set	1548 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	54.0	Xtriage
Anisotropy	0.154	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 24.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.52$ , $\langle L^2 \rangle = 0.36$	Xtriage
Estimated twinning fraction	0.000 for h,-k,-l 0.168 for -h,k,-l 0.000 for -h,-k,l	Xtriage
Reported twinning fraction	0.489 for H, K, L 0.391 for -H, K, -L 0.120 for -h,-k,l	Depositor
Outliers	0 of 31156 reflections	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	8769	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.36% 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 [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	0.47	0/847	0.70	0/1145
1	B	0.48	0/847	0.80	2/1145 (0.2%)
2	F	0.52	0/1775	0.72	0/2425
2	H	0.52	0/1775	0.71	0/2425
3	K	0.53	0/1730	0.76	0/2347
3	L	0.55	0/1730	0.75	0/2347
All	All	0.52	0/8704	0.74	2/11834 (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	B	0	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	B	80	SER	C-N-CD	-9.61	99.45	120.60
1	B	80	SER	C-N-CA	5.46	144.93	122.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	80	SER	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	829	0	782	50	0
1	B	829	0	782	39	1
2	F	1730	0	1699	132	1
2	H	1730	0	1699	114	0
3	K	1693	0	1645	114	0
3	L	1693	0	1645	111	0
4	A	32	0	0	2	0
4	B	34	0	0	1	0
4	F	53	0	0	2	0
4	H	50	0	0	4	0
4	K	42	0	0	1	0
4	L	54	0	0	1	1
All	All	8769	0	8252	525	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 31.

The worst 5 of 525 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)
3:L:201:ALA:HB2	3:L:210:ILE:HB	1.28	1.14
2:H:205:ILE:HG23	2:H:221:LYS:HB2	1.16	1.11
2:H:109:VAL:HG23	3:L:55:LYS:HE3	1.24	1.10
2:F:205:ILE:HG21	2:F:222:ILE:HG12	1.32	1.07
2:H:103:TYR:HB2	2:H:109:VAL:HG11	1.36	1.07

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)
4:L:321:HOH:O	4:L:354:HOH:O[1_655]	1.90	0.30
1:B:45:ALA:O	2:F:114:TYR:OH[1_655]	2.08	0.12

## 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	100/124 (81%)	53 (53%)	26 (26%)	21 (21%)	0	0
1	B	100/124 (81%)	62 (62%)	20 (20%)	18 (18%)	0	0
2	F	223/228 (98%)	145 (65%)	55 (25%)	23 (10%)	0	0
2	H	223/228 (98%)	155 (70%)	43 (19%)	25 (11%)	0	0
3	K	216/219 (99%)	153 (71%)	42 (19%)	21 (10%)	0	0
3	L	216/219 (99%)	163 (76%)	37 (17%)	16 (7%)	1	1
All	All	1078/1142 (94%)	731 (68%)	223 (21%)	124 (12%)	0	0

5 of 124 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	11	GLU
1	A	14	HIS
1	A	40	LYS
1	A	42	ASN
1	A	73	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	91/107 (85%)	82 (90%)	9 (10%)	8	14
1	B	91/107 (85%)	82 (90%)	9 (10%)	8	14
2	F	199/201 (99%)	161 (81%)	38 (19%)	1	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	H	199/201 (99%)	167 (84%)	32 (16%)	2	3
3	K	195/196 (100%)	156 (80%)	39 (20%)	1	2
3	L	195/196 (100%)	156 (80%)	39 (20%)	1	2
All	All	970/1008 (96%)	804 (83%)	166 (17%)	2	3

5 of 166 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	168	SER
3	K	24	ARG
3	L	165	LEU
2	H	185	ASP
2	H	205	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
2	H	62	ASN
3	K	27	GLN
3	L	161	GLN
2	H	176	HIS
2	H	208	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, 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	102/124 (82%)	0.01	2 (1%) 65 62	22, 28, 31, 32	0
1	B	102/124 (82%)	0.15	3 (2%) 51 47	23, 29, 30, 31	0
2	F	225/228 (98%)	0.04	3 (1%) 77 74	18, 22, 29, 34	0
2	H	225/228 (98%)	0.08	5 (2%) 62 58	17, 23, 29, 31	0
3	K	218/219 (99%)	0.05	7 (3%) 47 43	16, 20, 30, 32	0
3	L	218/219 (99%)	0.04	5 (2%) 60 57	15, 18, 30, 33	0
All	All	1090/1142 (95%)	0.06	25 (2%) 60 57	15, 24, 30, 34	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	L	153	TRP	4.8
3	L	199	CYS	4.6
2	H	223	VAL	4.1
2	H	2	VAL	4.0
3	L	202	THR	3.9

### 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

There are no such residues in this entry.