



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 06:46 am GMT

PDB ID : 3S51  
Title : Structure of FANCI  
Authors : Pavletich, N.P.  
Deposited on : 2011-05-20  
Resolution : 3.30 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
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
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)	:	recalc28949

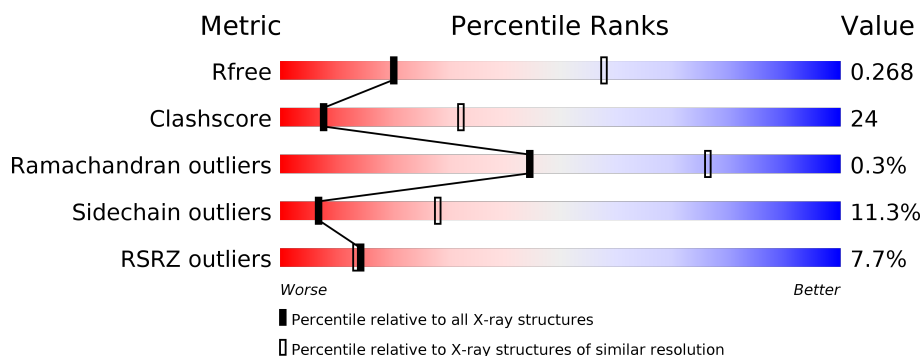
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.30 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	1034 (3.36-3.24)
Clashscore	112137	1100 (3.36-3.24)
Ramachandran outliers	110173	1081 (3.36-3.24)
Sidechain outliers	110143	1080 (3.36-3.24)
RSRZ outliers	101464	1039 (3.36-3.24)

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	1308	
1	B	1308	
1	C	1308	
1	D	1308	

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 34594 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 Fanconi anemia group I protein homolog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1134	Total	C	N	O	S	0	0	0
			8960	5762	1489	1656	53			
1	B	1071	Total	C	N	O	S	0	0	0
			8487	5468	1409	1559	51			
1	C	1134	Total	C	N	O	S	0	0	0
			8960	5762	1489	1656	53			
1	D	1034	Total	C	N	O	S	0	0	0
			8187	5277	1358	1504	48			

There are 24 discrepancies between the modelled and reference sequences:

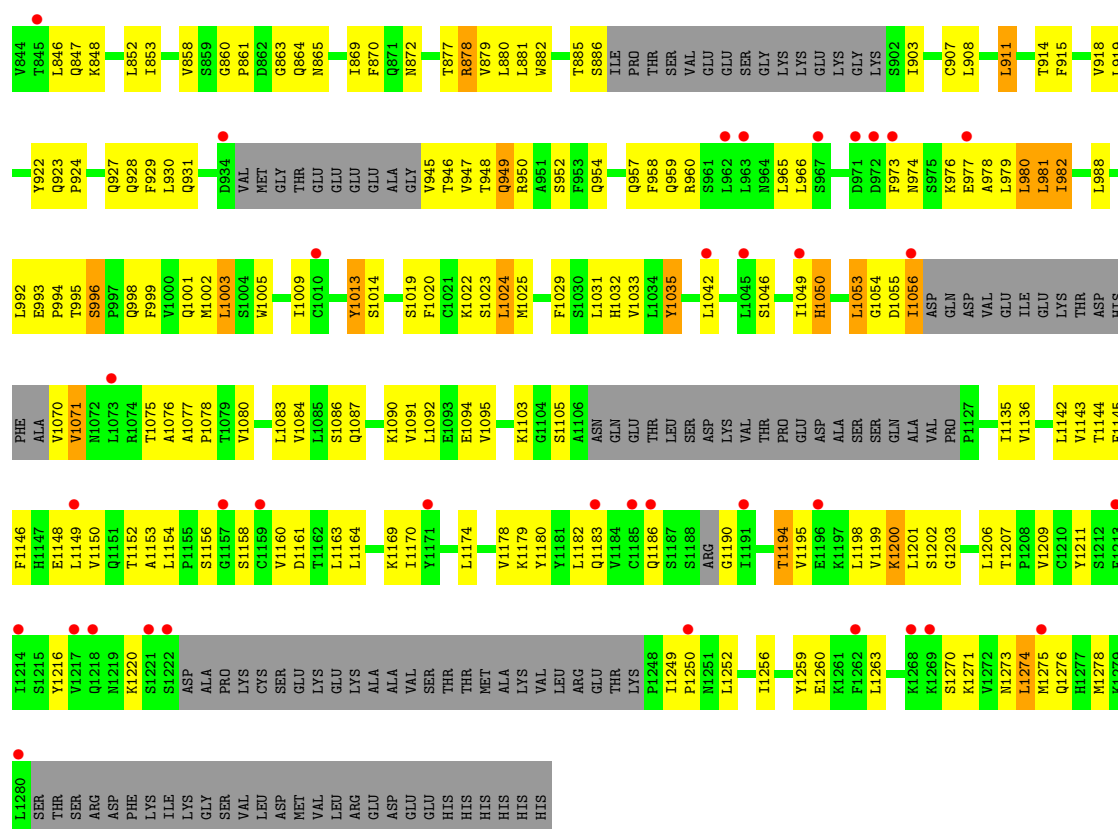
Chain	Residue	Modelled	Actual	Comment	Reference
A	1303	HIS	-	EXPRESSION TAG	UNP Q8K368
A	1304	HIS	-	EXPRESSION TAG	UNP Q8K368
A	1305	HIS	-	EXPRESSION TAG	UNP Q8K368
A	1306	HIS	-	EXPRESSION TAG	UNP Q8K368
A	1307	HIS	-	EXPRESSION TAG	UNP Q8K368
A	1308	HIS	-	EXPRESSION TAG	UNP Q8K368
B	1303	HIS	-	EXPRESSION TAG	UNP Q8K368
B	1304	HIS	-	EXPRESSION TAG	UNP Q8K368
B	1305	HIS	-	EXPRESSION TAG	UNP Q8K368
B	1306	HIS	-	EXPRESSION TAG	UNP Q8K368
B	1307	HIS	-	EXPRESSION TAG	UNP Q8K368
B	1308	HIS	-	EXPRESSION TAG	UNP Q8K368
C	1303	HIS	-	EXPRESSION TAG	UNP Q8K368
C	1304	HIS	-	EXPRESSION TAG	UNP Q8K368
C	1305	HIS	-	EXPRESSION TAG	UNP Q8K368
C	1306	HIS	-	EXPRESSION TAG	UNP Q8K368
C	1307	HIS	-	EXPRESSION TAG	UNP Q8K368
C	1308	HIS	-	EXPRESSION TAG	UNP Q8K368
D	1303	HIS	-	EXPRESSION TAG	UNP Q8K368
D	1304	HIS	-	EXPRESSION TAG	UNP Q8K368
D	1305	HIS	-	EXPRESSION TAG	UNP Q8K368

*Continued on next page...*

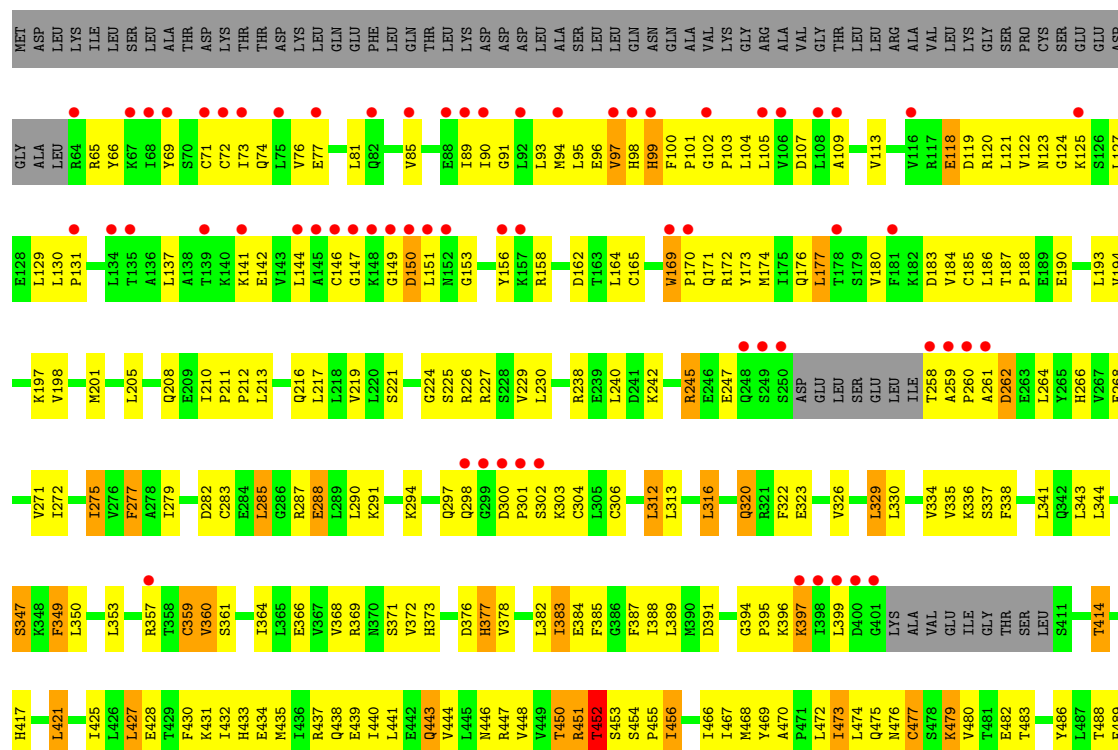
*Continued from previous page...*

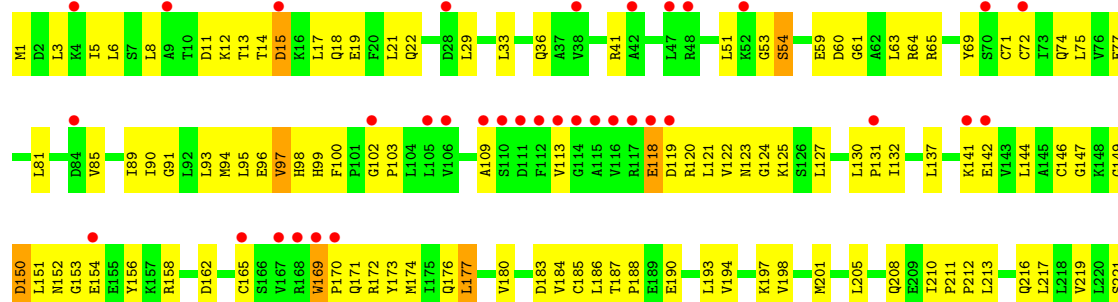
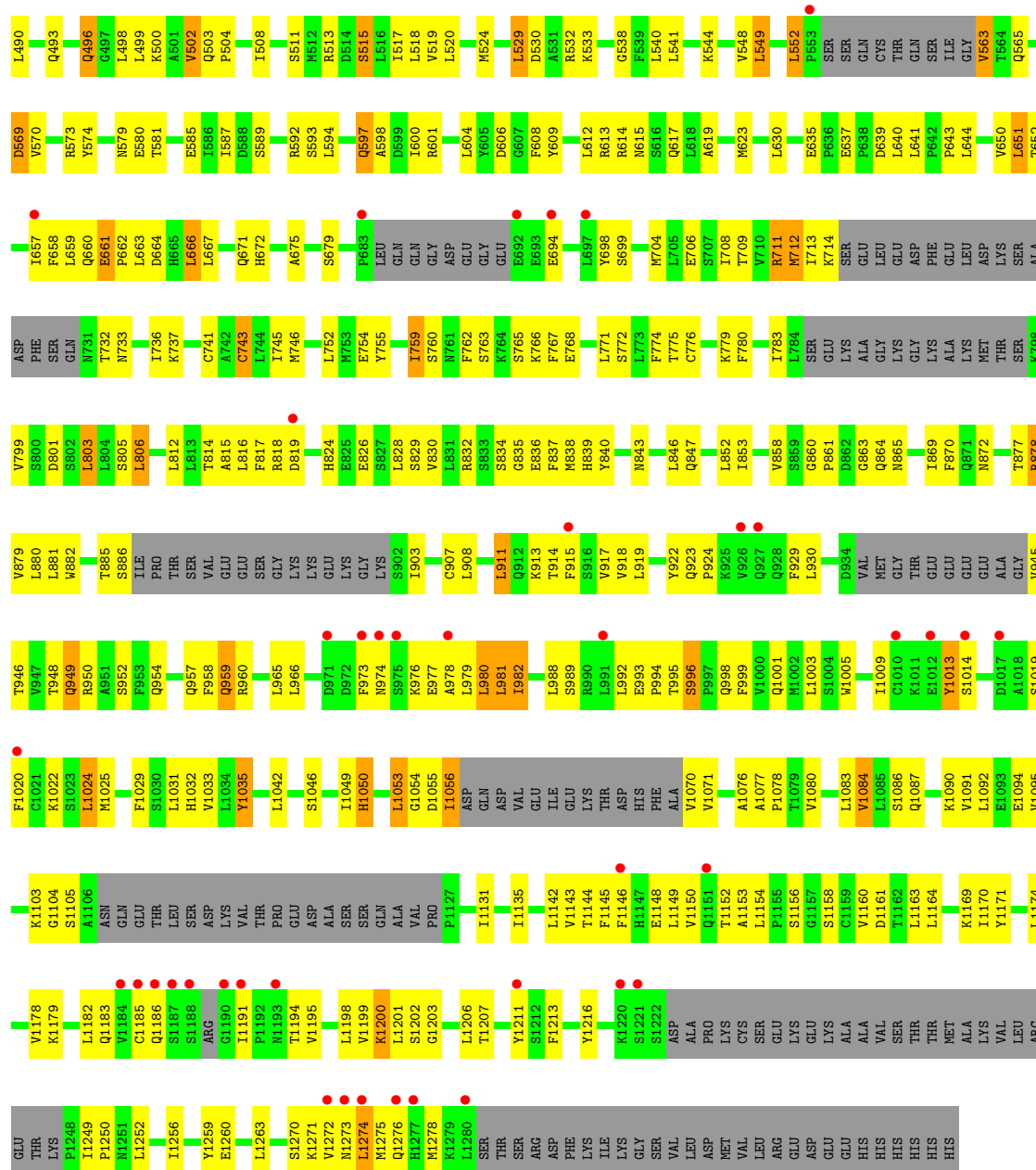
Chain	Residue	Modelled	Actual	Comment	Reference
D	1306	HIS	-	EXPRESSION TAG	UNP Q8K368
D	1307	HIS	-	EXPRESSION TAG	UNP Q8K368
D	1308	HIS	-	EXPRESSION TAG	UNP Q8K368





• Molecule 1: Fanconi anemia group I protein homolog

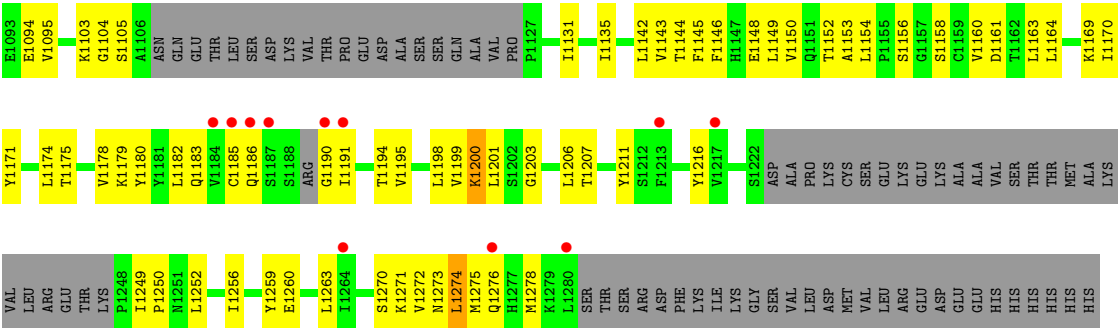












## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	114.70Å 136.50Å 149.70Å 115.90° 106.00° 95.00°	Depositor
Resolution (Å)	39.82 – 3.30 39.82 – 3.28	Depositor EDS
% Data completeness (in resolution range)	82.8 (39.82-3.30) 73.3 (39.82-3.28)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.52 (at 3.25Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.1_743)	Depositor
R, $R_{free}$	0.259 , 0.278 0.244 , 0.268	Depositor DCC
$R_{free}$ test set	1931 reflections (2.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	93.7	Xtriage
Anisotropy	0.486	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 119.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	34594	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	192.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.11% 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	0.26	2/9099 (0.0%)	0.46	0/12286
1	B	0.29	4/8624 (0.0%)	0.46	0/11646
1	C	0.26	2/9099 (0.0%)	0.46	0/12286
1	D	0.26	2/8319 (0.0%)	0.46	0/11234
All	All	0.27	10/35141 (0.0%)	0.46	0/47452

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	B	0	1
1	C	0	1
1	D	0	1
All	All	0	4

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	959	GLN	CD-NE2	-9.98	1.07	1.32
1	B	959	GLN	CD-OE1	-7.80	1.06	1.24
1	C	320	GLN	CD-NE2	-6.76	1.16	1.32
1	D	320	GLN	CD-NE2	-6.59	1.16	1.32
1	A	320	GLN	CD-NE2	-6.15	1.17	1.32

There are no bond angle outliers.

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1053	LEU	Peptide
1	B	1053	LEU	Peptide
1	C	1053	LEU	Peptide
1	D	1053	LEU	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	8960	0	9275	454	0
1	B	8487	0	8776	449	0
1	C	8960	0	9275	435	0
1	D	8187	0	8476	411	0
All	All	34594	0	35802	1705	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 24.

The worst 5 of 1705 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:799:VAL:O	1:C:847:GLN:NE2	1.87	1.07
1:A:799:VAL:O	1:A:847:GLN:NE2	1.88	1.07
1:B:799:VAL:O	1:B:847:GLN:NE2	1.88	1.06
1:D:799:VAL:O	1:D:847:GLN:NE2	1.88	1.05
1:A:489:PHE:HB3	1:C:450:THR:HG21	1.35	1.04

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	1108/1308 (85%)	1023 (92%)	82 (7%)	3 (0%)	44	76
1	B	1045/1308 (80%)	969 (93%)	72 (7%)	4 (0%)	38	71
1	C	1108/1308 (85%)	1022 (92%)	83 (8%)	3 (0%)	44	76
1	D	1008/1308 (77%)	933 (93%)	71 (7%)	4 (0%)	38	71
All	All	4269/5232 (82%)	3947 (92%)	308 (7%)	14 (0%)	44	76

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	122	VAL
1	A	150	ASP
1	B	122	VAL
1	B	150	ASP
1	C	122	VAL

### 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	1032/1188 (87%)	917 (89%)	115 (11%)	7	29
1	B	979/1188 (82%)	867 (89%)	112 (11%)	7	28
1	C	1032/1188 (87%)	917 (89%)	115 (11%)	7	29
1	D	945/1188 (80%)	838 (89%)	107 (11%)	7	28
All	All	3988/4752 (84%)	3539 (89%)	449 (11%)	7	28

5 of 449 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	878	ARG
1	C	288	GLU
1	D	765	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	981	LEU
1	C	19	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	433	HIS
1	B	672	HIS
1	D	433	HIS
1	B	377	HIS
1	D	377	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	1134/1308 (86%)	0.23	85 (7%) 15 14	85, 192, 306, 410	0
1	B	1071/1308 (81%)	0.25	104 (9%) 8 8	87, 188, 306, 415	0
1	C	1134/1308 (86%)	0.23	65 (5%) 24 23	81, 191, 306, 410	0
1	D	1034/1308 (79%)	0.26	84 (8%) 13 12	84, 185, 302, 416	0
All	All	4373/5232 (83%)	0.24	338 (7%) 14 13	81, 189, 304, 416	0

The worst 5 of 338 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	109	ALA	12.1
1	D	135	THR	11.7
1	D	250	SER	9.3
1	D	299	GLY	8.2
1	D	101	PRO	7.7

### 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.