wwPDB X-ray Structure Validation Summary Report (i)

Jun 5, 2024 – 02:08 PM JST

PDB ID	:	8WSU
Title	:	Crystal structure of SFTSV Gc and antibody
Authors	:	Chang, Z.; Gao, F.; Wu, Y.
Deposited on	:	2023-10-17
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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

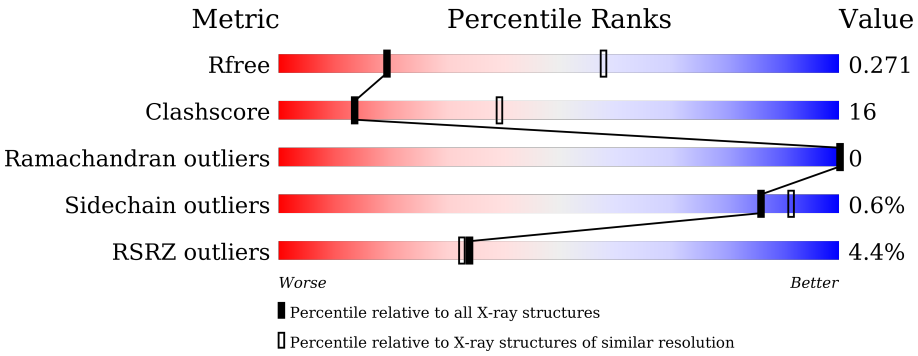
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

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}	130704	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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 of the lower bar indicate the fraction of residues that contain outliers for >=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 <=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	439	3% 30% 12% 57%
1	D	439	4% 25% 18% 57%
2	B	224	4% 64% 36%
2	E	224	4% 67% 32%
3	C	214	2% 71% 28%
3	F	214	2% 72% 27%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 9442 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 Glycoprotein C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	188	Total	C	N	O	S	0	0	0
			1409	877	248	266	18			
1	D	188	Total	C	N	O	S	0	0	0
			1409	877	248	266	18			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	587	VAL	ILE	conflict	UNP A0A0B5A886
A	726	VAL	ALA	conflict	UNP A0A0B5A886
A	960	THR	ILE	conflict	UNP A0A0B5A886
A	996	HIS	-	expression tag	UNP A0A0B5A886
A	997	HIS	-	expression tag	UNP A0A0B5A886
A	998	HIS	-	expression tag	UNP A0A0B5A886
A	999	HIS	-	expression tag	UNP A0A0B5A886
A	1000	HIS	-	expression tag	UNP A0A0B5A886
A	1001	HIS	-	expression tag	UNP A0A0B5A886
D	587	VAL	ILE	conflict	UNP A0A0B5A886
D	726	VAL	ALA	conflict	UNP A0A0B5A886
D	960	THR	ILE	conflict	UNP A0A0B5A886
D	996	HIS	-	expression tag	UNP A0A0B5A886
D	997	HIS	-	expression tag	UNP A0A0B5A886
D	998	HIS	-	expression tag	UNP A0A0B5A886
D	999	HIS	-	expression tag	UNP A0A0B5A886
D	1000	HIS	-	expression tag	UNP A0A0B5A886
D	1001	HIS	-	expression tag	UNP A0A0B5A886

- Molecule 2 is a protein called Ab-H.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	224	Total	C	N	O	S	0	0	0
			1686	1060	288	332	6			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	224	Total	C	N	O	S	0	0	0
			1686	1060	288	332	6			

- Molecule 3 is a protein called Ab-L.

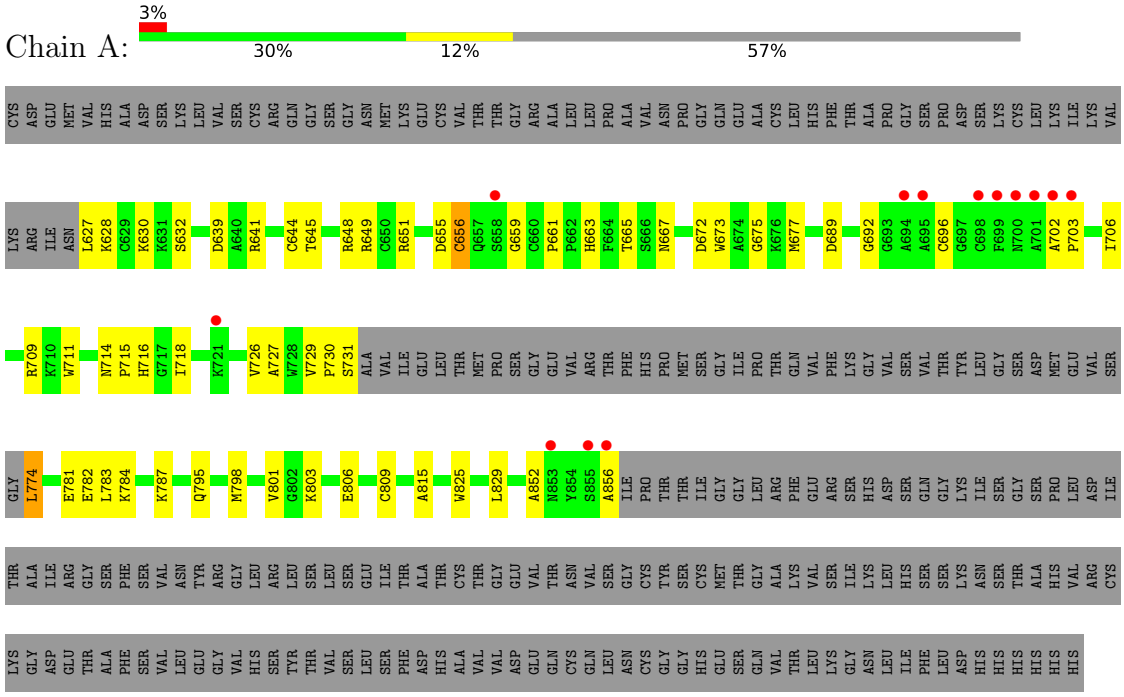
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	214	Total	C	N	O	S	0	0	0
			1626	1014	278	328	6			
3	F	214	Total	C	N	O	S	0	0	0
			1626	1014	278	328	6			

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

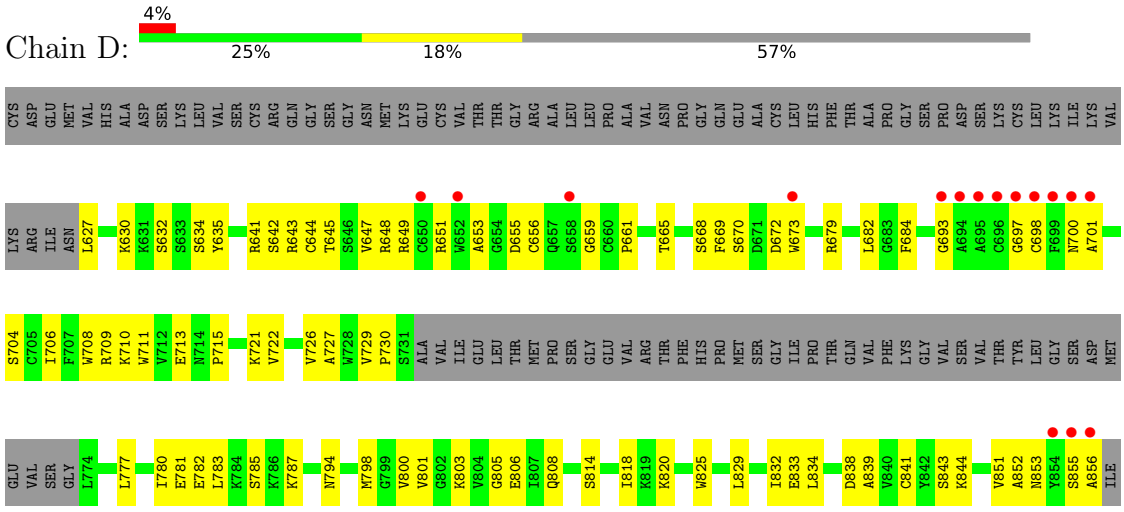
• Molecule 1: Glycoprotein C

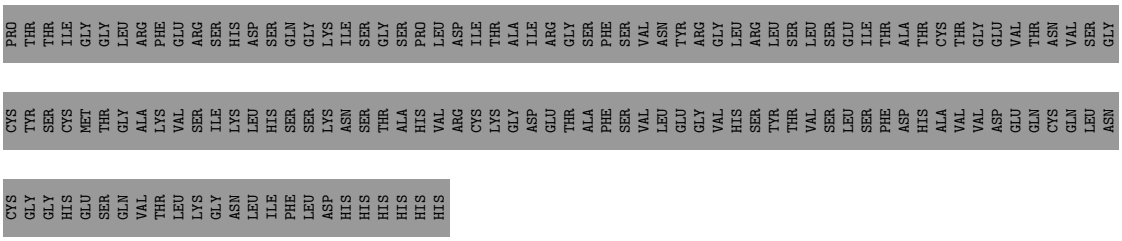
Chain A:



• Molecule 1: Glycoprotein C

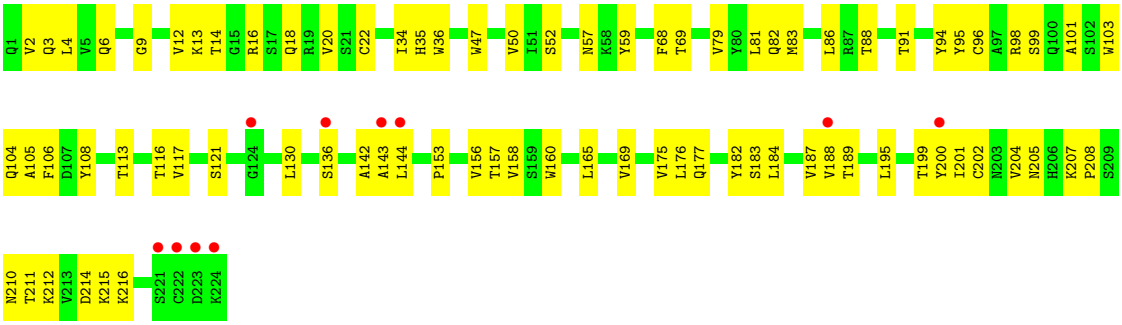
Chain D:





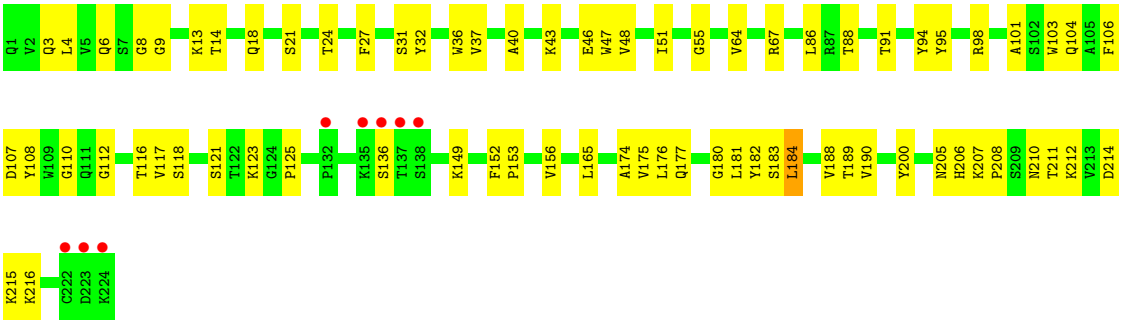
- Molecule 2: Ab-H

Chain B:



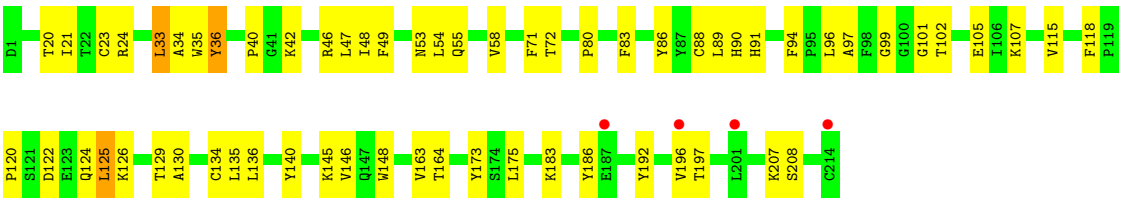
- Molecule 2: Ab-H

Chain E:



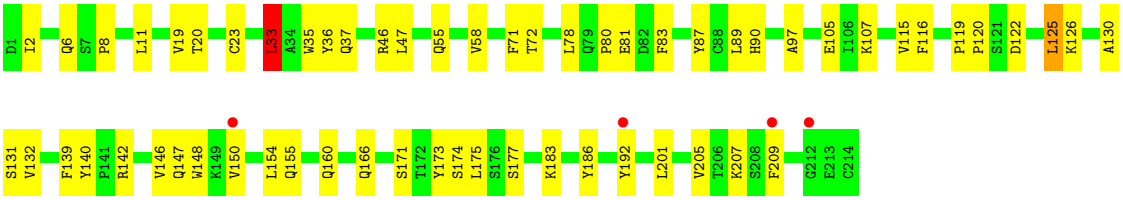
- Molecule 3: Ab-L

Chain C:



- Molecule 3: Ab-L

Chain F:


D1
I2
Q6
S7
P8
L11
V19
T20
C23
L33
A34
W35
Y36
Q37
R46
L47
Q55
V58
F71
T72
L78
Q79
P80
E81
D82
F83
Y87
C88
L89
H90
A97
E105
I106
K107
V115
F116
P119
P120
S121
D122
L125
K126
A130
S131
V132
F139
Y140
P141
R142
V146
Q147
W148
K149
V150
L154
Q155
Q160
Q166
S171
T172
Y173
S174
L175
S176
S177
K183
Y186
Y192
L201
V205
T206
K207
S208
F209
G212
E213
C214

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	229.51Å 72.05Å 112.36Å 90.00° 111.47° 90.00°	Depositor
Resolution (Å)	36.03 – 3.30 36.03 – 3.30	Depositor EDS
% Data completeness (in resolution range)	92.8 (36.03-3.30) 92.8 (36.03-3.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.74 (at 3.32Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487, PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.207 , 0.267 0.209 , 0.271	Depositor DCC
R_{free} test set	1204 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	83.2	Xtriage
Anisotropy	0.629	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 49.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	9442	wwPDB-VP
Average B, all atoms (Å^2)	110.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.72% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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	0.76	1/1440 (0.1%)	0.96	1/1946 (0.1%)
1	D	0.60	1/1440 (0.1%)	0.89	0/1946
2	B	0.61	0/1726	0.85	1/2349 (0.0%)
2	E	0.52	0/1726	0.80	1/2349 (0.0%)
3	C	0.62	1/1661 (0.1%)	0.85	3/2254 (0.1%)
3	F	0.53	0/1661	0.82	3/2254 (0.1%)
All	All	0.61	3/9654 (0.0%)	0.86	9/13098 (0.1%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	644	CYS	CB-SG	-9.50	1.66	1.82
1	D	659	GLY	C-N	-5.38	1.21	1.34
3	C	36	TYR	CD1-CE1	-5.11	1.31	1.39

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	E	184	LEU	CA-CB-CG	7.96	133.61	115.30
2	B	184	LEU	CA-CB-CG	6.65	130.60	115.30
1	A	774	LEU	CA-CB-CG	-6.61	100.09	115.30
3	F	33	LEU	CB-CG-CD1	-6.23	100.40	111.00
3	C	24	ARG	CG-CD-NE	-6.05	99.09	111.80

There are no chirality outliers.

There are no planarity outliers.

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	1409	0	1350	44	0
1	D	1409	0	1350	56	1
2	B	1686	0	1650	70	1
2	E	1686	0	1650	58	0
3	C	1626	0	1575	48	0
3	F	1626	0	1575	49	0
All	All	9442	0	9150	297	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 16.

The worst 5 of 297 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:B:121:SER:HB3	3:F:126:LYS:HD3	1.30	1.08
1:A:692:GLY:HA3	1:A:703:PRO:HA	1.37	1.07
1:D:649:ARG:NH1	1:D:655:ASP:OD2	1.97	0.98
2:B:211:THR:OG1	2:E:215:LYS:NZ	2.10	0.84
2:B:121:SER:HB3	3:F:126:LYS:CD	2.08	0.82

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)
2:B:205:ASN:OD1	1:D:700:ASN:ND2[2_555]	1.98	0.22

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	184/439 (42%)	164 (89%)	20 (11%)	0	100	100
1	D	184/439 (42%)	166 (90%)	18 (10%)	0	100	100
2	B	222/224 (99%)	204 (92%)	18 (8%)	0	100	100
2	E	222/224 (99%)	203 (91%)	19 (9%)	0	100	100
3	C	212/214 (99%)	203 (96%)	9 (4%)	0	100	100
3	F	212/214 (99%)	199 (94%)	13 (6%)	0	100	100
All	All	1236/1754 (70%)	1139 (92%)	97 (8%)	0	100	100

There are no Ramachandran outliers to report.

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	153/368 (42%)	151 (99%)	2 (1%)	69	82
1	D	153/368 (42%)	151 (99%)	2 (1%)	69	82
2	B	190/190 (100%)	190 (100%)	0	100	100
2	E	190/190 (100%)	189 (100%)	1 (0%)	88	93
3	C	184/184 (100%)	184 (100%)	0	100	100
3	F	184/184 (100%)	183 (100%)	1 (0%)	88	93
All	All	1054/1484 (71%)	1048 (99%)	6 (1%)	86	91

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	841	CYS
2	E	205	ASN
3	F	33	LEU
1	A	825	TRP
1	A	656	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	205	ASN
2	E	6	GLN

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 monosaccharides 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

6.1 Protein, DNA and RNA chains

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^{th} 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	188/439 (42%)	0.22	13 (6%)	16	16	54, 95, 196, 288	0
1	D	188/439 (42%)	0.32	16 (8%)	10	10	74, 115, 231, 313	0
2	B	224/224 (100%)	0.10	10 (4%)	33	32	53, 99, 170, 305	0
2	E	224/224 (100%)	0.10	8 (3%)	42	40	75, 107, 180, 297	0
3	C	214/214 (100%)	0.01	4 (1%)	66	65	56, 95, 159, 235	0
3	F	214/214 (100%)	0.09	4 (1%)	66	65	73, 124, 180, 255	0
All	All	1252/1754 (71%)	0.13	55 (4%)	34	33	53, 104, 191, 313	0

The worst 5 of 55 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	700	ASN	10.2
2	B	223	ASP	10.0
1	D	699	PHE	8.4
2	E	136	SER	8.0
1	A	701	ALA	7.4

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.