W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

May 22, 2020 – 07:00 pm BST

PDB ID	:	5NT2
Title	:	Complex of influenza A NS1 with TRIM25 coiled coil domain
Authors	:	Koliopoulos, M.G.; Rittinger, K.
Deposited on	:	2017-04-27
Resolution	:	4.26 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 following versions of software and data (see references (1)) 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 (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 4.26 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.
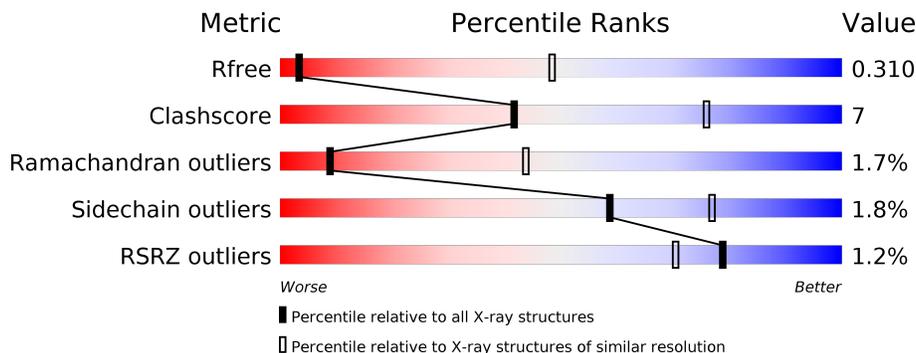
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1017 (4.72-3.78)
Clashscore	141614	1059 (4.72-3.80)
Ramachandran outliers	138981	1014 (4.72-3.80)
Sidechain outliers	138945	1018 (4.72-3.78)
RSRZ outliers	127900	1072 (4.80-3.70)

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 >=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	C	233	
1	D	233	
1	E	233	
1	F	233	
2	A	193	
2	I	193	

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	N	193	4% ▉ 75% 13% · 10%
2	V	193	2% ▉ 73% 16% · 10%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 10516 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 Non-structural protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	F	115	Total 894	C 567	N 152	O 169	S 6	0	0	0
1	C	115	Total 895	C 567	N 152	O 170	S 6	0	0	0
1	D	200	Total 1559	C 979	N 274	O 297	S 9	0	0	0
1	E	201	Total 1566	C 984	N 275	O 298	S 9	0	0	0

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-2	GLY	-	expression tag	UNP P03496
F	-1	PRO	-	expression tag	UNP P03496
F	0	GLY	-	expression tag	UNP P03496
F	38	ALA	ARG	engineered mutation	UNP P03496
F	41	ALA	LYS	engineered mutation	UNP P03496
F	101	GLU	ASP	variant	UNP P03496
F	187	ALA	TRP	engineered mutation	UNP P03496
C	-2	GLY	-	expression tag	UNP P03496
C	-1	PRO	-	expression tag	UNP P03496
C	0	GLY	-	expression tag	UNP P03496
C	38	ALA	ARG	engineered mutation	UNP P03496
C	41	ALA	LYS	engineered mutation	UNP P03496
C	101	GLU	ASP	variant	UNP P03496
C	187	ALA	TRP	engineered mutation	UNP P03496
D	-2	GLY	-	expression tag	UNP P03496
D	-1	PRO	-	expression tag	UNP P03496
D	0	GLY	-	expression tag	UNP P03496
D	38	ALA	ARG	engineered mutation	UNP P03496
D	41	ALA	LYS	engineered mutation	UNP P03496
D	101	GLU	ASP	variant	UNP P03496
D	187	ALA	TRP	engineered mutation	UNP P03496

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	-2	GLY	-	expression tag	UNP P03496
E	-1	PRO	-	expression tag	UNP P03496
E	0	GLY	-	expression tag	UNP P03496
E	38	ALA	ARG	engineered mutation	UNP P03496
E	41	ALA	LYS	engineered mutation	UNP P03496
E	101	GLU	ASP	variant	UNP P03496
E	187	ALA	TRP	engineered mutation	UNP P03496

- Molecule 2 is a protein called E3 ubiquitin/ISG15 ligase TRIM25.

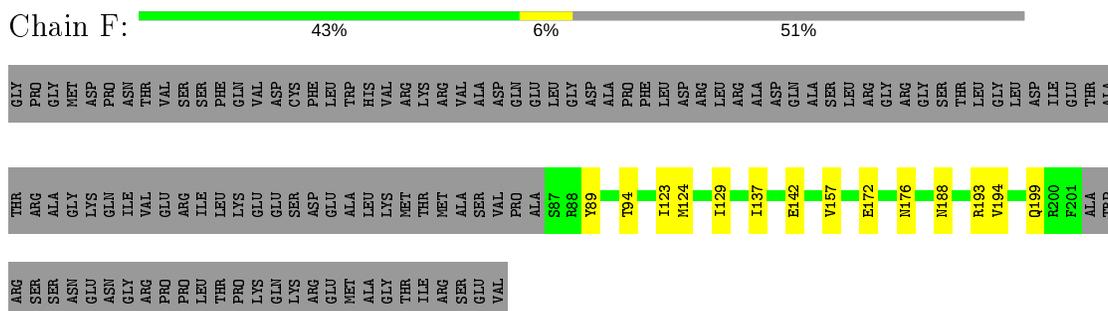| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	I	173	Total 1399	C 868	N 250	O 277	S 4	0	0	0
2	A	173	Total 1401	C 869	N 250	O 278	S 4	0	0	0
2	V	173	Total 1401	C 869	N 250	O 278	S 4	0	0	0
2	N	173	Total 1401	C 869	N 250	O 278	S 4	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

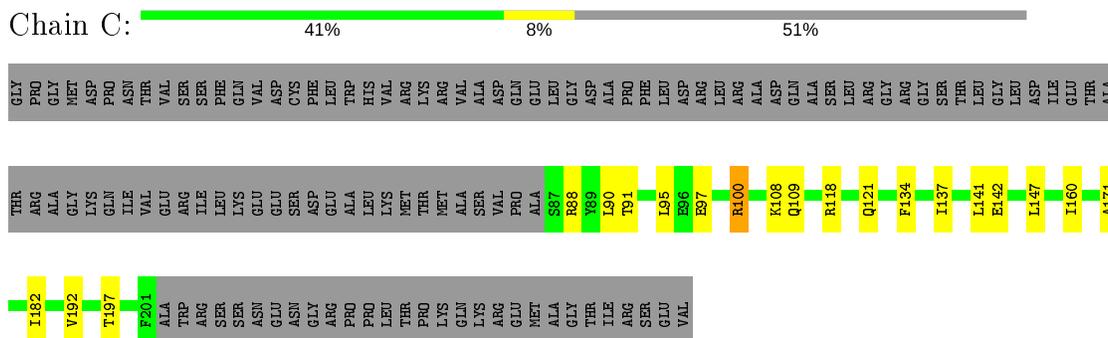| Chain	Residue	Modelled	Actual	Comment	Reference
I	187	GLY	-	expression tag	UNP Q14258
I	188	PRO	-	expression tag	UNP Q14258
I	189	GLY	-	expression tag	UNP Q14258
I	358	LEU	PRO	variant	UNP Q14258
A	187	GLY	-	expression tag	UNP Q14258
A	188	PRO	-	expression tag	UNP Q14258
A	189	GLY	-	expression tag	UNP Q14258
A	358	LEU	PRO	variant	UNP Q14258
V	187	GLY	-	expression tag	UNP Q14258
V	188	PRO	-	expression tag	UNP Q14258
V	189	GLY	-	expression tag	UNP Q14258
V	358	LEU	PRO	variant	UNP Q14258
N	187	GLY	-	expression tag	UNP Q14258
N	188	PRO	-	expression tag	UNP Q14258
N	189	GLY	-	expression tag	UNP Q14258
N	358	LEU	PRO	variant	UNP Q14258

3 Residue-property plots (i)

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.
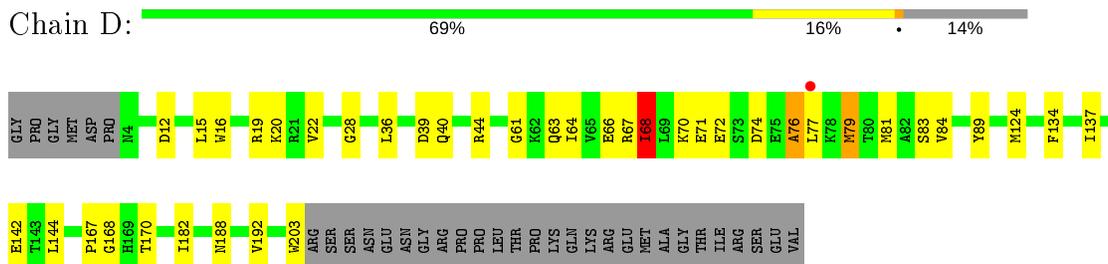
- Molecule 1: Non-structural protein 1

Chain F:
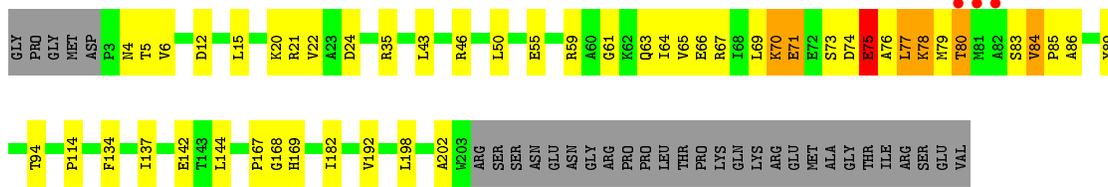
- Molecule 1: Non-structural protein 1
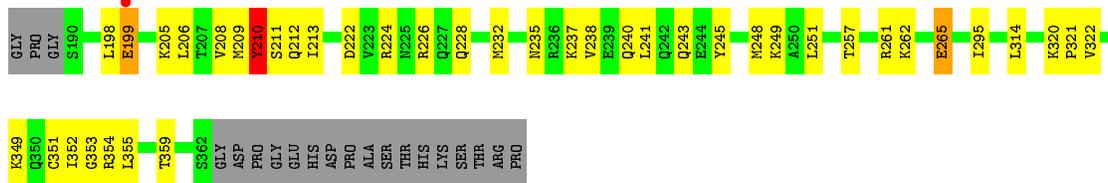
Chain C:

- Molecule 1: Non-structural protein 1
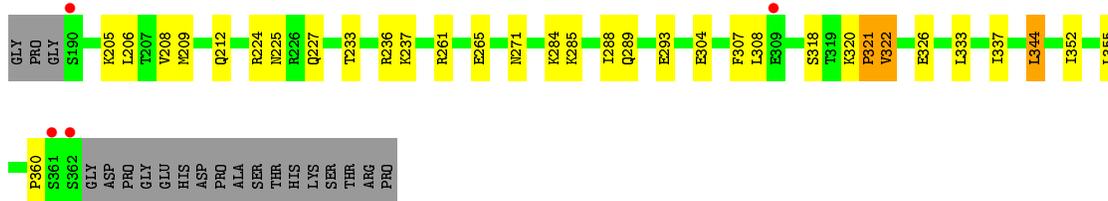
Chain D:

- Molecule 1: Non-structural protein 1

Chain E:

GLY PRO GLY MET ASP P3 N4 T5 V6 D12 L15 K20 R21 V22 A23 D24 R35 L43 R46 L50 E55 R59 A60 G61 K62 Q63 I64 V65 E66 R67 I68 L69 K70 E71 E72 S73 D74 E75 A76 L77 K78 M79 T80 M81 A82 S83 V84 P85 A86 Y89

T94 P114 F134 I137 E142 T143 L144 P167 G168 H169 I182 V192 L198 A202 W203 ARG SER ASN GLU ASN GLY ARG PRO LEU THR PRO LYS GLN LYS ARG GLU MET ALA GLY THR ILE ARG SER GLU VAL

• **Molecule 2: E3 ubiquitin/ISG15 ligase TRIM25**

Chain I: 68% 20% 10%

GLY PRO GLY S190 L198 E199 K205 L206 T207 V208 M209 Y210 S211 Q212 I213 D222 V223 R224 N225 R226 Q227 Q228 M232 N235 R236 K237 V238 E239 Q240 L241 Q242 Q243 E244 Y245 M248 K249 A250 L251 T257 R261 K262 E265 I295 L314 K320 P321 V322

K349 Q350 C351 I352 G353 R354 L355 T359 S362 GLY ASP PRO GLY GLU HIS ASP PRO ALA SER THR HIS LYS SER THR ARG PRO

• **Molecule 2: E3 ubiquitin/ISG15 ligase TRIM25**

Chain A: 74% 15% 10%

GLY PRO GLY S190 L198 K205 V208 M209 V223 R224 N225 R226 Q227 V230 R231 M232 T233 K237 Y245 M248 K249 L252 T260 T276 I280 K284 I288 E296 P321 V322 E326 L333 D343 L344 E347 T359 S362 GLY

ASP PRO GLY GLU HIS ASP PRO ALA SER THR HIS LYS SER THR ARG PRO

• **Molecule 2: E3 ubiquitin/ISG15 ligase TRIM25**

Chain V: 73% 16% 10%

GLY PRO GLY S190 K205 L206 T207 V208 M209 Q212 R224 N225 R226 Q227 T233 R236 K237 R261 E265 N271 K284 K285 I288 Q289 E293 E304 F307 L308 E309 S318 T319 K320 P321 V322 E326 L333 I337 L344 I352 L355

P360 S361 S362 GLY ASP PRO GLY GLU HIS ASP PRO ALA SER THR HIS LYS SER THR ARG PRO

• **Molecule 2: E3 ubiquitin/ISG15 ligase TRIM25**

Chain N: 75% 13% 10%

GLY PRO GLY S190 R203 H204 K205 V208 Q212 D229 R261 E265 I277 Y278 K284 L291 K292 I295 E296 Q297 S298 R302 D303 E304 F307 L308 E309 K310 A311 S312 K313 L314 R315 G316 I317 K320 P321 V322 K332 I337 L344 S361 S362

GLY ASP PRO GLY GLU HIS ASP PRO ALA SER THR HIS LYS SER THR ARG PRO

4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	73.26Å 76.30Å 92.07Å 100.04° 93.71° 111.14°	Depositor
Resolution (Å)	61.68 – 4.26 61.68 – 4.26	Depositor EDS
% Data completeness (in resolution range)	98.1 (61.68-4.26) 98.0 (61.68-4.26)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.43 (at 4.29Å)	Xtriage
Refinement program	PHENIX (1.12_2829: ???)	Depositor
R, R_{free}	0.270 , 0.309 0.270 , 0.310	Depositor DCC
R_{free} test set	635 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å^2)	122.2	Xtriage
Anisotropy	0.403	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 110.9	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	0.90	EDS
Total number of atoms	10516	wwPDB-VP
Average B, all atoms (Å^2)	161.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.23% 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 (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	C	0.25	0/907	0.51	0/1226
1	D	0.26	0/1581	0.50	0/2136
1	E	0.27	0/1589	0.52	1/2147 (0.0%)
1	F	0.25	0/906	0.49	0/1224
2	A	0.29	0/1413	0.50	0/1892
2	I	0.30	0/1411	0.55	1/1889 (0.1%)
2	N	0.29	0/1413	0.50	0/1892
2	V	0.29	0/1413	0.49	0/1892
All	All	0.28	0/10633	0.51	2/14298 (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	E	0	2

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	84	VAL	C-N-CD	-6.33	106.68	120.60
2	I	210	TYR	CA-CB-CG	6.06	124.92	113.40

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	71	GLU	Peptide
1	E	73	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	C	895	0	921	11	0
1	D	1559	0	1586	28	0
1	E	1566	0	1594	34	0
1	F	894	0	921	8	0
2	A	1401	0	1450	24	0
2	I	1399	0	1446	30	0
2	N	1401	0	1450	26	0
2	V	1401	0	1450	26	0
All	All	10516	0	10818	149	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 7.

The worst 5 of 149 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:I:261:ARG:NH2	2:N:361:SER:OG	2.17	0.77
1:D:68:ILE:HA	1:D:70:LYS:H	1.52	0.74
2:I:224:ARG:NH2	2:N:278:TYR:OH	2.23	0.72
2:N:315:ARG:HG3	2:N:316:GLY:H	1.56	0.70
1:D:19:ARG:NH1	1:D:39:ASP:OD2	2.24	0.69

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	C	113/233 (48%)	112 (99%)	1 (1%)	0	100	100
1	D	198/233 (85%)	174 (88%)	19 (10%)	5 (2%)	5	35
1	E	199/233 (85%)	174 (87%)	18 (9%)	7 (4%)	3	29
1	F	113/233 (48%)	112 (99%)	1 (1%)	0	100	100
2	A	171/193 (89%)	166 (97%)	3 (2%)	2 (1%)	13	50
2	I	171/193 (89%)	165 (96%)	4 (2%)	2 (1%)	13	50
2	N	171/193 (89%)	164 (96%)	3 (2%)	4 (2%)	6	37
2	V	171/193 (89%)	166 (97%)	3 (2%)	2 (1%)	13	50
All	All	1307/1704 (77%)	1233 (94%)	52 (4%)	22 (2%)	9	44

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	78	LYS
1	E	85	PRO
2	I	321	PRO
2	A	321	PRO
2	V	321	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	C	100/198 (50%)	97 (97%)	3 (3%)	41	63
1	D	170/198 (86%)	167 (98%)	3 (2%)	59	77
1	E	171/198 (86%)	165 (96%)	6 (4%)	36	60
1	F	100/198 (50%)	100 (100%)	0	100	100
2	A	160/175 (91%)	160 (100%)	0	100	100
2	I	159/175 (91%)	155 (98%)	4 (2%)	47	68
2	N	160/175 (91%)	157 (98%)	3 (2%)	57	75
2	V	160/175 (91%)	158 (99%)	2 (1%)	69	82

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1180/1492 (79%)	1159 (98%)	21 (2%)	59 77

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	75	GLU
1	E	114	PRO
2	V	344	LEU
1	E	74	ASP
2	N	261	ARG

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

Mol	Chain	Res	Type
2	A	271	ASN
2	V	227	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 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, 95th 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	C	115/233 (49%)	-0.12	0	100	100	112, 139, 177, 203	0
1	D	200/233 (85%)	-0.15	1 (0%)	91	86	115, 169, 241, 260	0
1	E	201/233 (86%)	-0.12	3 (1%)	73	64	113, 154, 208, 240	0
1	F	115/233 (49%)	-0.19	0	100	100	112, 140, 178, 196	0
2	A	173/193 (89%)	-0.08	0	100	100	112, 141, 241, 259	0
2	I	173/193 (89%)	-0.08	1 (0%)	89	84	118, 147, 252, 284	0
2	N	173/193 (89%)	-0.01	7 (4%)	38	30	109, 147, 327, 404	0
2	V	173/193 (89%)	-0.05	4 (2%)	60	51	111, 148, 260, 285	0
All	All	1323/1704 (77%)	-0.10	16 (1%)	79	70	109, 149, 241, 404	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	V	362	SER	5.9
2	N	314	LEU	3.7
2	V	190	SER	3.6
2	N	311	ALA	3.1
1	E	81	MET	2.8

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.