



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 7, 2020 – 09:46 PM BST

PDB ID : 3N44
Title : Crystal structure of the mature envelope glycoprotein complex (trypsin cleavage; Osmium soak) of Chikungunya virus.
Authors : Voss, J.; Vaney, M.C.; Duquerroy, S.; Rey, F.A.
Deposited on : 2010-05-21
Resolution : 2.35 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 ⓘ symbol.

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

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
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)	:	2.13.1

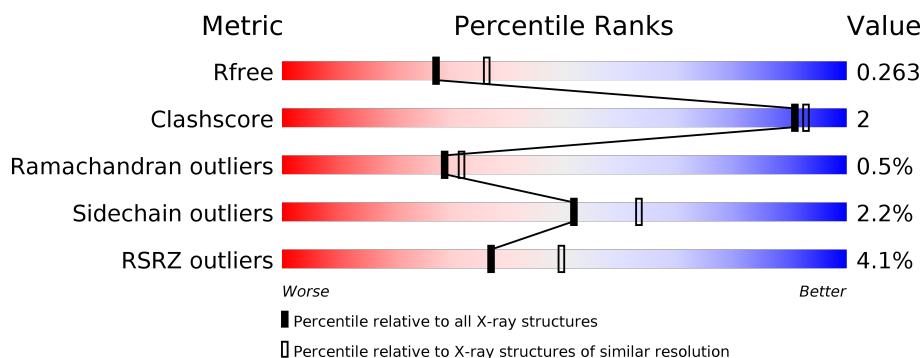
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.35 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	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)
RSRZ outliers	127900	1150 (2.36-2.36)

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 $\leq 5\%$. 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	65	<div> <div>9%</div> <div> <div></div> <div>78%</div> <div>5%</div> <div>17%</div> </div> </div>
2	B	360	<div> <div>5%</div> <div> <div></div> <div>88%</div> <div>6%</div> <div>6%</div> </div> </div>
3	F	473	<div> <div>2%</div> <div> <div></div> <div>76%</div> <div>7%</div> <div>17%</div> </div> </div>

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 6398 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 E3 envelope glycoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	54	Total	C	N	O	S	0	0	0
			420	263	66	82	9			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	expression tag	UNP Q1H8W5

- Molecule 2 is a protein called E2 envelope glycoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	338	Total	C	N	O	S	0	3	0
			2683	1673	485	505	20			

- Molecule 3 is a protein called E1 envelope glycoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	F	391	Total	C	N	O	S	0	0	0
			2981	1886	499	572	24			

There are 61 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-21	GLU	-	expression tag	UNP Q1H8W5
F	-20	LEU	-	expression tag	UNP Q1H8W5
F	-19	TYR	-	expression tag	UNP Q1H8W5
F	-18	GLY	-	expression tag	UNP Q1H8W5
F	-17	GLY	-	expression tag	UNP Q1H8W5
F	-16	GLY	-	expression tag	UNP Q1H8W5
F	-15	GLY	-	expression tag	UNP Q1H8W5
F	-14	SER	-	expression tag	UNP Q1H8W5
F	-13	GLY	-	expression tag	UNP Q1H8W5

Continued on next page...

Continued from previous page...

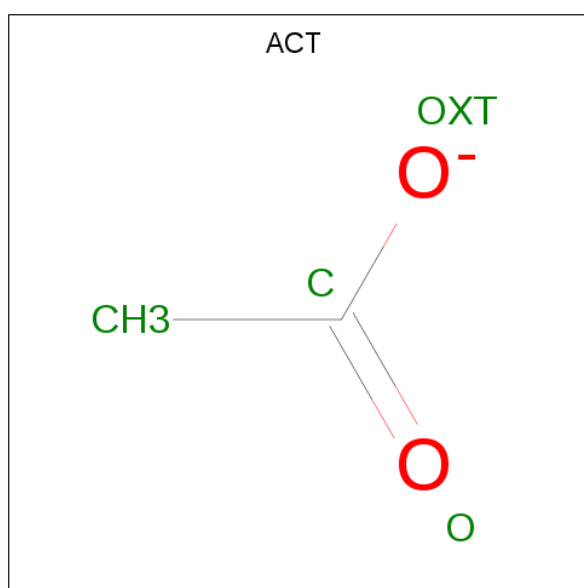
Chain	Residue	Modelled	Actual	Comment	Reference
F	-12	GLY	-	expression tag	UNP Q1H8W5
F	-11	GLY	-	expression tag	UNP Q1H8W5
F	-10	GLY	-	expression tag	UNP Q1H8W5
F	-9	SER	-	expression tag	UNP Q1H8W5
F	-8	GLY	-	expression tag	UNP Q1H8W5
F	-7	GLY	-	expression tag	UNP Q1H8W5
F	-6	GLY	-	expression tag	UNP Q1H8W5
F	-5	GLY	-	expression tag	UNP Q1H8W5
F	-4	SER	-	expression tag	UNP Q1H8W5
F	-3	GLY	-	expression tag	UNP Q1H8W5
F	-2	GLY	-	expression tag	UNP Q1H8W5
F	-1	GLY	-	expression tag	UNP Q1H8W5
F	0	GLY	-	expression tag	UNP Q1H8W5
F	413	GLY	-	expression tag	UNP Q1H8W5
F	414	PRO	-	expression tag	UNP Q1H8W5
F	415	PHE	-	expression tag	UNP Q1H8W5
F	416	GLU	-	expression tag	UNP Q1H8W5
F	417	ASP	-	expression tag	UNP Q1H8W5
F	418	ASP	-	expression tag	UNP Q1H8W5
F	419	ASP	-	expression tag	UNP Q1H8W5
F	420	ASP	-	expression tag	UNP Q1H8W5
F	421	LYS	-	expression tag	UNP Q1H8W5
F	422	ALA	-	expression tag	UNP Q1H8W5
F	423	GLY	-	expression tag	UNP Q1H8W5
F	424	TRP	-	expression tag	UNP Q1H8W5
F	425	SER	-	expression tag	UNP Q1H8W5
F	426	HIS	-	expression tag	UNP Q1H8W5
F	427	PRO	-	expression tag	UNP Q1H8W5
F	428	GLN	-	expression tag	UNP Q1H8W5
F	429	PHE	-	expression tag	UNP Q1H8W5
F	430	GLU	-	expression tag	UNP Q1H8W5
F	431	LYS	-	expression tag	UNP Q1H8W5
F	432	GLY	-	expression tag	UNP Q1H8W5
F	433	GLY	-	expression tag	UNP Q1H8W5
F	434	GLY	-	expression tag	UNP Q1H8W5
F	435	SER	-	expression tag	UNP Q1H8W5
F	436	GLY	-	expression tag	UNP Q1H8W5
F	437	GLY	-	expression tag	UNP Q1H8W5
F	438	GLY	-	expression tag	UNP Q1H8W5
F	439	SER	-	expression tag	UNP Q1H8W5
F	440	GLY	-	expression tag	UNP Q1H8W5
F	441	GLY	-	expression tag	UNP Q1H8W5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	442	GLY	-	expression tag	UNP Q1H8W5
F	443	SER	-	expression tag	UNP Q1H8W5
F	444	TRP	-	expression tag	UNP Q1H8W5
F	445	SER	-	expression tag	UNP Q1H8W5
F	446	HIS	-	expression tag	UNP Q1H8W5
F	447	PRO	-	expression tag	UNP Q1H8W5
F	448	GLN	-	expression tag	UNP Q1H8W5
F	449	PHE	-	expression tag	UNP Q1H8W5
F	450	GLU	-	expression tag	UNP Q1H8W5
F	451	LYS	-	expression tag	UNP Q1H8W5

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 4 2 2	0	0
4	F	1	Total C O 4 2 2	0	0
4	F	1	Total C O 4 2 2	0	0

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	1	Total	C	N	O	0	0
			14	8	1	5		
5	F	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 6 is OSMIUM ION (three-letter code: OS) (formula: Os).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	4	Total	Os	0	0
			4	4		
6	F	2	Total	Os	0	0
			2	2		

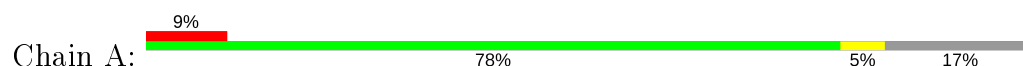
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	4	Total	O	0	0
			4	4		
7	B	89	Total	O	0	0
			89	89		
7	F	175	Total	O	0	0
			175	175		

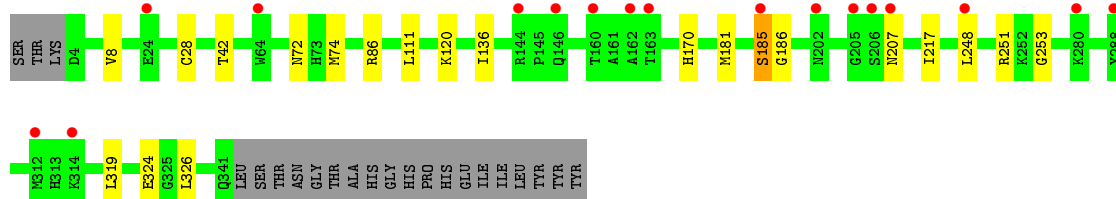
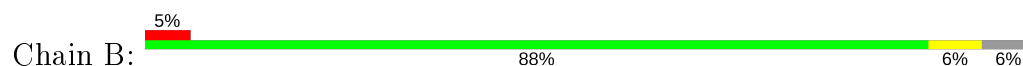
3 Residue-property plots [i](#)

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.

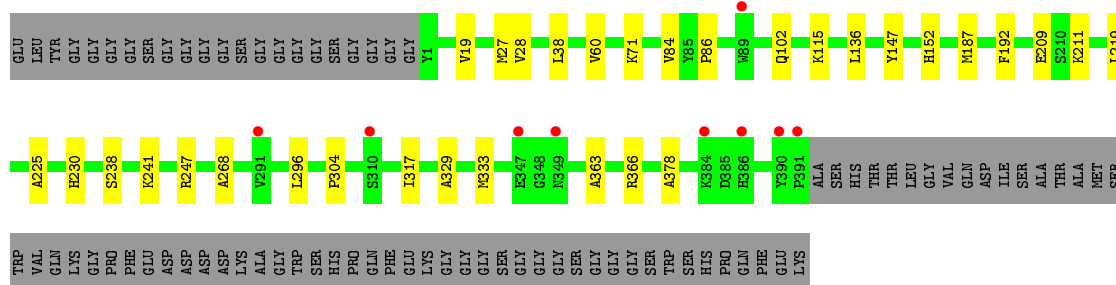
- Molecule 1: E3 envelope glycoprotein



- Molecule 2: E2 envelope glycoprotein



- Molecule 3: E1 envelope glycoprotein



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	58.09Å 89.95Å 174.02Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.30 – 2.35 79.91 – 2.06	Depositor EDS
% Data completeness (in resolution range)	91.2 (48.30-2.35) 69.9 (79.91-2.06)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.00 (at 2.07Å)	Xtriage
Refinement program	BUSTER 2.9.3	Depositor
R, R_{free}	0.229 , 0.250 0.238 , 0.263	Depositor DCC
R_{free} test set	2020 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	36.8	Xtriage
Anisotropy	0.773	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 42.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	6398	wwPDB-VP
Average B, all atoms (Å ²)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.43% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: OS, NAG, ACT

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	0.38	0/430	0.58	0/586
2	B	0.35	0/2755	0.57	0/3751
3	F	0.37	0/3058	0.57	0/4172
All	All	0.36	0/6243	0.57	0/8509

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	420	0	403	2	0
2	B	2683	0	2594	7	0
3	F	2981	0	2887	14	0
4	A	4	0	3	0	0
4	F	8	0	6	0	0
5	B	14	0	13	0	0
5	F	14	0	13	0	0
6	B	4	0	0	0	0
6	F	2	0	0	0	0
7	A	4	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	B	89	0	0	0	0
7	F	175	0	0	0	0
All	All	6398	0	5919	21	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 2.

The worst 5 of 21 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:F:225:ALA:H	3:F:230:HIS:HE1	1.41	0.67
3:F:333:MET:HB2	3:F:366:ARG:HB2	1.85	0.58
3:F:28:VAL:HG23	3:F:329:ALA:HB1	1.89	0.54
2:B:42:THR:HB	2:B:136:ILE:HD11	1.91	0.53
3:F:60:VAL:HG22	3:F:102:GLN:HG3	1.92	0.52

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	52/65 (80%)	49 (94%)	2 (4%)	1 (2%)	8	6
2	B	338/360 (94%)	318 (94%)	17 (5%)	3 (1%)	17	17
3	F	389/473 (82%)	375 (96%)	14 (4%)	0	100	100
All	All	779/898 (87%)	742 (95%)	33 (4%)	4 (0%)	29	32

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	11	ALA

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	185	SER
2	B	207	ASN
2	B	72	ASN

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	51/61 (84%)	51 (100%)	0	100	100
2	B	302/319 (95%)	294 (97%)	8 (3%)	46	56
3	F	328/378 (87%)	321 (98%)	7 (2%)	53	65
All	All	681/758 (90%)	666 (98%)	15 (2%)	52	63

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	217	ILE
2	B	324	GLU
3	F	219	LEU
2	B	181	MET
3	F	211	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
3	F	102	GLN
3	F	373	GLN
3	F	152	HIS
2	B	236	GLN
3	F	230	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 11 ligands modelled in this entry, 6 are monoatomic - leaving 5 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
5	NAG	B	2001	2	14,14,15	1.25	1 (7%)	17,19,21	1.54	5 (29%)
4	ACT	F	452	-	1,3,3	1.84	0	0,3,3	0.00	-
4	ACT	F	453	-	1,3,3	1.90	0	0,3,3	0.00	-
5	NAG	F	1001	3	14,14,15	1.19	1 (7%)	17,19,21	1.09	1 (5%)
4	ACT	A	65	-	1,3,3	1.83	0	0,3,3	0.00	-

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	B	2001	2	-	1/6/23/26	0/1/1/1
5	NAG	F	1001	3	-	0/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	F	1001	NAG	C1-C2	2.93	1.56	1.52
5	B	2001	NAG	C1-C2	2.75	1.56	1.52

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	2001	NAG	C4-C3-C2	3.54	116.21	111.02
5	F	1001	NAG	C1-O5-C5	3.52	116.97	112.19
5	B	2001	NAG	C1-O5-C5	2.58	115.69	112.19
5	B	2001	NAG	C1-C2-N2	2.40	114.59	110.49
5	B	2001	NAG	O5-C5-C4	2.20	116.19	110.83

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	2001	NAG	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

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	A	54/65 (83%)	0.56	6 (11%) 5 8	48, 62, 90, 98	0
2	B	338/360 (93%)	0.19	17 (5%) 28 41	31, 53, 76, 103	0
3	F	391/473 (82%)	-0.06	9 (2%) 60 70	25, 40, 74, 110	0
All	All	783/898 (87%)	0.09	32 (4%) 37 49	25, 49, 79, 110	0

The worst 5 of 32 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	F	390	TYR	4.8
2	B	162	ALA	4.6
2	B	24	GLU	4.0
1	A	15	PHE	4.0
2	B	314	LYS	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 monosaccharides in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
4	ACT	F	452	4/4	0.48	0.33	88,88,88,88	0
5	NAG	B	2001	14/15	0.68	0.29	70,74,76,76	0
4	ACT	A	65	4/4	0.75	0.21	87,88,88,88	0
6	OS	B	361	1/1	0.81	0.09	81,81,81,81	1
6	OS	B	362	1/1	0.81	0.11	126,126,126,126	1
6	OS	B	364	1/1	0.86	0.06	89,89,89,89	1
6	OS	B	363	1/1	0.87	0.13	113,113,113,113	1
6	OS	F	454	1/1	0.90	0.08	131,131,131,131	1
5	NAG	F	1001	14/15	0.91	0.11	52,55,58,58	0
6	OS	F	455	1/1	0.94	0.05	70,70,70,70	1
4	ACT	F	453	4/4	0.96	0.14	43,43,44,44	0

6.5 Other polymers [i](#)

There are no such residues in this entry.