WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jan 6, 2026 – 02:35 PM EST

PDB ID	:	9MNS / pdb_00009mns
Title	:	Structure of Hepatitis C virus envelope glycoprotein E2 core from genotype 6a bound to broadly neutralizing antibody RM1-36
Authors	:	Nguyen, T.K.Y.; Stanfield, R.L.; Wilson, I.A.
Deposited on	:	2024-12-23
Resolution	:	2.89 Å(reported)

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

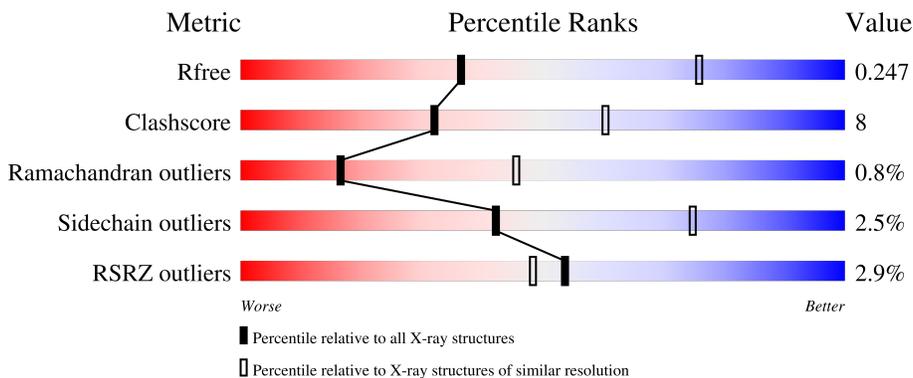|		
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.47

1 Overall quality at a glance (i)

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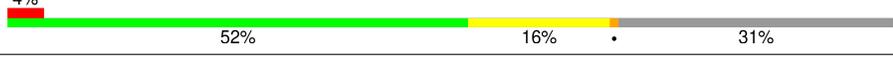*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.89 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	Percentile Ranks	Value
Rfree		0.247
Clashscore		8
Ramachandran outliers		0.8%
Sidechain outliers		2.5%
RSRZ outliers		2.9%

Worse   *Better*

■ Percentile relative to all X-ray structures

❘ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	2335 (2.90-2.90)
Clashscore	180529	2564 (2.90-2.90)
Ramachandran outliers	177936	2514 (2.90-2.90)
Sidechain outliers	177891	2516 (2.90-2.90)
RSRZ outliers	164620	2337 (2.90-2.90)

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	H	222	2% / 79% 18% •
1	h	222	2% / 82% 17% •
2	L	211	84% 16%
2	l	211	2% / 83% 16%
3	B	231	4% / 52% 16% • 31%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	E	231	

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 9140 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 RM1-36 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	H	222	Total 1672	C 1067	N 276	O 322	S 7	0	0	0
1	h	222	Total 1672	C 1067	N 276	O 322	S 7	0	0	0

- Molecule 2 is a protein called RM1-36 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	211	Total 1623	C 1014	N 268	O 336	S 5	0	0	0
2	l	211	Total 1623	C 1014	N 268	O 336	S 5	0	0	0

- Molecule 3 is a protein called Envelope glycoprotein E2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	160	Total 1238	C 787	N 211	O 227	S 13	0	0	0
3	B	159	Total 1228	C 781	N 208	O 226	S 13	0	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	448	ASP	ASN	conflict	UNP B9V0E2
E	569D	PRO	TYR	conflict	UNP B9V0E2
E	583	THR	GLN	conflict	UNP B9V0E2
E	584	ASP	ARG	conflict	UNP B9V0E2
E	591	GLY	CYS	conflict	UNP B9V0E2
B	448	ASP	ASN	conflict	UNP B9V0E2
B	582	PRO	TYR	conflict	UNP B9V0E2
B	583	THR	GLN	conflict	UNP B9V0E2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	584	ASP	ARG	conflict	UNP B9V0E2
B	591	GLY	CYS	conflict	UNP B9V0E2

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (CCD ID: NAG) (formula: $C_8H_{15}NO_6$).

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

3 Residue-property plots ⓘ

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: RM1-36 Fab heavy chain
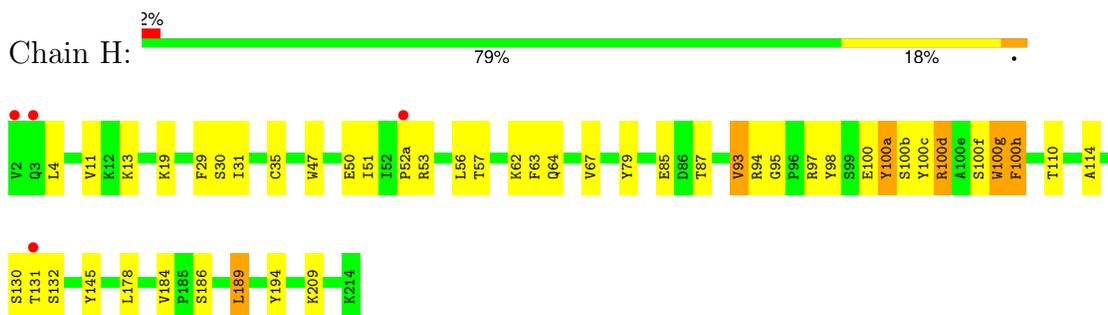
• Molecule 1: RM1-36 Fab heavy chain
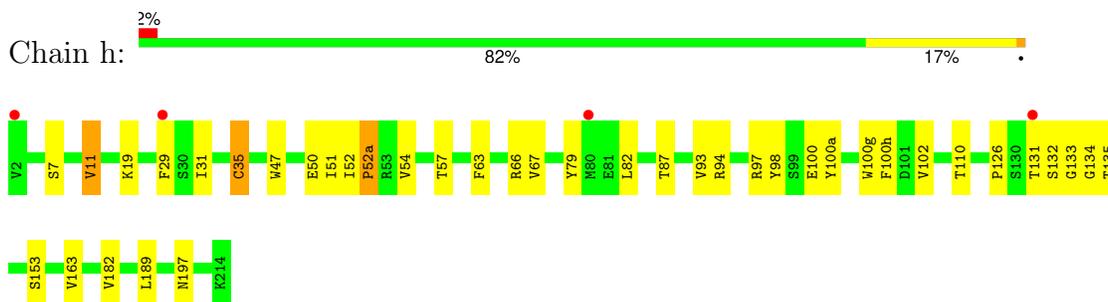
• Molecule 2: RM1-36 Fab light chain

• Molecule 2: RM1-36 Fab light chain

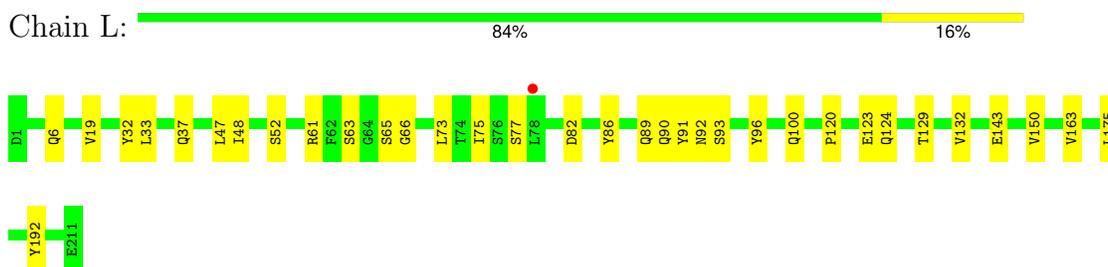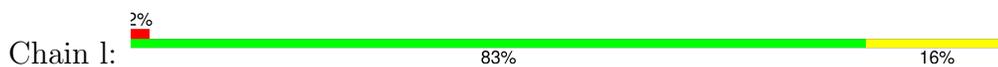D1 R18 L33 L47 I48 N53 R61 S65 G66 S76 S77 L78 Q89 Q90 Y91 N92 S93 D94 P113 S114 P120 Q124 L125 T129 V132 E143 L154 N158 V163 L175 L179 K183 A184 D185 Y186 E187 K188 H189 Y192

Q198 R209 G210 E211

• Molecule 3: Envelope glycoprotein E2

Chain E:

5%

51% 18% 31%

H421 I422 N423 R424 C429 L433 Q434 T439 N446 V447 C452 PRO GLU ARG MET ALA ALA CYS LYS PRO LEU ALA ASP PHE ARG GLN GLY TRP GLY GLN ILE THR TYR LYS VAL ASN ILE SER GLY PRO SER ASP ASP ARG PRO TYR CYS TRP HIS TYR ALA PRO R492 D495

V496 V497 S498 A499 R500 T501 V502 C503 G504 P505 F509 T519 T526 W529 G530 E531 N532 E533 T534 D535 V536 F537 M538 L539 E540 R543 P544 T546 G547 G548 S557 P568 CYS GLN ILE VAL PRO GLY ASP TYR ASN SER SER ALA ASN GLU LEU LEU CYS PRO THR ASP

CYS PHE ARG LYS HIS PRO GLU ALA THR PRO T583 G600 P601 L608 Y613 R614 P619 C620 T621 V622 V629 R630 G634 G635 I636 R639 A643 C644 N645

• Molecule 3: Envelope glycoprotein E2

Chain B:

4%

52% 16% • 31%

HIS I422 N423 R424 C429 N446 V447 C452 PRO GLU ARG MET ALA ALA CYS LYS PRO LEU ALA ASP PHE ARG GLN GLY TRP GLY GLN ILE THR TYR LYS VAL ASN ILE SER GLY PRO SER ASP ASP ARG PRO TYR CYS TRP HIS TYR ALA PRO R492 D495 S498 A499 R500 T501

V502 C503 F509 T519 P525 T526 E533 T534 D535 L539 E540 S541 L542 R543 P544 P545 T546 S557 T561 K562 P567 PRO CYS GLN ILE VAL PRO GLY ASP TYR ASN SER SER ALA ASN GLU LEU LEU CYS PRO THR ASP CYS PHE ARG LYS HIS PRO GLU ALA THR P582

S599 G600 P601 P605 R606 C607 L608 V609 R614 L615 P619 V622 V629 R630 M631 F632 V633 G634 E637 H638 R639 F640 A643 C644 N645

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	59.75Å 223.54Å 73.30Å 90.00° 93.58° 90.00°	Depositor
Resolution (Å)	44.87 – 2.89 44.87 – 2.89	Depositor EDS
% Data completeness (in resolution range)	98.1 (44.87-2.89) 99.0 (44.87-2.89)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.17 (at 2.91Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158, PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.232 , 0.250 0.228 , 0.247	Depositor DCC
R_{free} test set	2065 reflections (4.83%)	wwPDB-VP
Wilson B-factor (Å^2)	86.0	Xtriage
Anisotropy	0.095	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 48.9	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	9140	wwPDB-VP
Average B, all atoms (Å^2)	92.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.97% 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 ⓘ

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

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	H	0.30	0/1714	0.70	0/2339
1	h	0.32	0/1714	0.70	0/2339
2	L	0.25	0/1658	0.58	0/2251
2	l	0.24	0/1658	0.58	0/2251
3	B	0.27	0/1266	0.69	0/1729
3	E	0.27	0/1277	0.71	0/1745
All	All	0.28	0/9287	0.66	0/12654

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 ⓘ

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	H	1672	0	1658	29	0
1	h	1672	0	1658	26	0
2	L	1623	0	1563	21	0
2	l	1623	0	1563	22	0
3	B	1228	0	1153	25	0
3	E	1238	0	1159	28	0
4	B	42	0	39	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	E	42	0	39	2	0
All	All	9140	0	8832	142	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 8.

The worst 5 of 142 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:H:64:GLN:HE21	2:l:154:LEU:HD23	1.27	0.97
1:H:31:ILE:HD11	1:H:98:TYR:HB2	1.71	0.73
3:B:546:THR:HG21	3:B:634:GLY:HA3	1.70	0.73
1:H:209:LYS:NZ	2:L:123:GLU:OE2	2.21	0.72
3:B:557:SER:HB2	4:B:701:NAG:H82	1.72	0.72

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	H	218/222 (98%)	204 (94%)	13 (6%)	1 (0%)	25	56
1	h	218/222 (98%)	202 (93%)	14 (6%)	2 (1%)	14	43
2	L	209/211 (99%)	201 (96%)	7 (3%)	1 (0%)	25	56
2	l	209/211 (99%)	200 (96%)	9 (4%)	0	100	100
3	B	153/231 (66%)	138 (90%)	12 (8%)	3 (2%)	6	23
3	E	154/231 (67%)	141 (92%)	11 (7%)	2 (1%)	10	33
All	All	1161/1328 (87%)	1086 (94%)	66 (6%)	9 (1%)	16	45

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	131	THR
1	h	134	GLY
3	E	600	GLY
3	B	600	GLY
3	E	540	GLU

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	H	188/188 (100%)	176 (94%)	12 (6%)	14	41
1	h	188/188 (100%)	184 (98%)	4 (2%)	48	78
2	L	185/185 (100%)	184 (100%)	1 (0%)	86	96
2	l	185/185 (100%)	182 (98%)	3 (2%)	58	84
3	B	137/199 (69%)	133 (97%)	4 (3%)	37	72
3	E	138/199 (69%)	136 (99%)	2 (1%)	62	86
All	All	1021/1144 (89%)	995 (98%)	26 (2%)	42	75

5 of 26 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	l	93	SER
1	h	35	CYS
3	B	533	GLU
1	h	11	VAL
1	h	52	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	h	164	HIS
3	E	446	ASN
3	B	446	ASN
2	L	79	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	H	64	GLN

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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

6 ligands are modelled in this entry.

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	$\#|Z| > 2$
4	NAG	E	703	3	14,14,15	0.67	0	17,19,21	0.82	0
4	NAG	B	701	3	14,14,15	0.69	0	17,19,21	1.22	2 (11%)
4	NAG	E	701	3	14,14,15	0.76	1 (7%)	17,19,21	2.21	4 (23%)
4	NAG	E	702	3	14,14,15	0.69	0	17,19,21	0.77	0
4	NAG	B	702	3	14,14,15	0.62	0	17,19,21	1.53	1 (5%)
4	NAG	B	703	3	14,14,15	0.76	0	17,19,21	0.95	0

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
4	NAG	E	703	3	-	2/6/23/26	0/1/1/1
4	NAG	B	701	3	-	1/6/23/26	0/1/1/1
4	NAG	E	701	3	-	0/6/23/26	0/1/1/1
4	NAG	E	702	3	-	1/6/23/26	0/1/1/1
4	NAG	B	702	3	-	3/6/23/26	0/1/1/1
4	NAG	B	703	3	-	2/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	E	701	NAG	C1-C2	2.19	1.55	1.52

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	E	701	NAG	C1-C2-N2	5.14	118.54	110.43
4	B	702	NAG	C2-N2-C7	4.53	128.97	122.90
4	E	701	NAG	C1-O5-C5	4.31	117.96	112.19
4	E	701	NAG	C2-N2-C7	3.39	127.45	122.90
4	E	701	NAG	O5-C1-C2	-3.38	106.07	111.29

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	702	NAG	C1-C2-N2-C7
4	E	703	NAG	O5-C5-C6-O6
4	B	703	NAG	O5-C5-C6-O6
4	E	703	NAG	C4-C5-C6-O6
4	B	703	NAG	C4-C5-C6-O6

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	701	NAG	1	0
4	E	701	NAG	2	0

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	h	1
1	H	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	h	126:PRO	C	130:SER	N	6.95
1	H	126:PRO	C	130:SER	N	6.26

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, 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(Å²)	Q<0.9
1	H	222/222 (100%)	0.14	4 (1%)	67	61	46, 76, 119, 140	0
1	h	222/222 (100%)	0.06	4 (1%)	67	61	48, 77, 109, 133	0
2	L	211/211 (100%)	0.11	1 (0%)	87	84	53, 87, 116, 132	0
2	l	211/211 (100%)	0.05	4 (1%)	66	60	49, 83, 109, 116	0
3	B	159/231 (68%)	0.79	9 (5%)	30	26	79, 125, 151, 167	0
3	E	160/231 (69%)	0.76	12 (7%)	22	18	73, 118, 151, 164	0
All	All	1185/1328 (89%)	0.27	34 (2%)	54	48	46, 89, 138, 167	0

The worst 5 of 34 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	2	VAL	4.3
3	E	583	THR	3.4
1	h	131	THR	3.3
3	E	434	GLN	3.1
3	B	422	ILE	3.0

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 oligosaccharides in this entry.

6.4 Ligands ⓘ

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(Å^2)	Q<0.9
4	NAG	B	702	14/15	0.55	0.14	131,151,156,158	0
4	NAG	B	703	14/15	0.55	0.15	121,131,144,145	0
4	NAG	E	703	14/15	0.57	0.16	146,165,170,170	0
4	NAG	E	702	14/15	0.67	0.13	131,140,146,151	0
4	NAG	B	701	14/15	0.70	0.10	110,119,127,130	0
4	NAG	E	701	14/15	0.73	0.14	109,127,137,138	0

6.5 Other polymers ⓘ

There are no such residues in this entry.