



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 26, 2016 – 12:43 AM EDT

PDB ID : 5HJ3
Title : Crystal structure of host-primed Ebola virus GP, GPcl.
Authors : Bornholdt, Z.A.; Fusco, M.L.; Saphire, E.O.
Deposited on : 2016-01-12
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
<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:

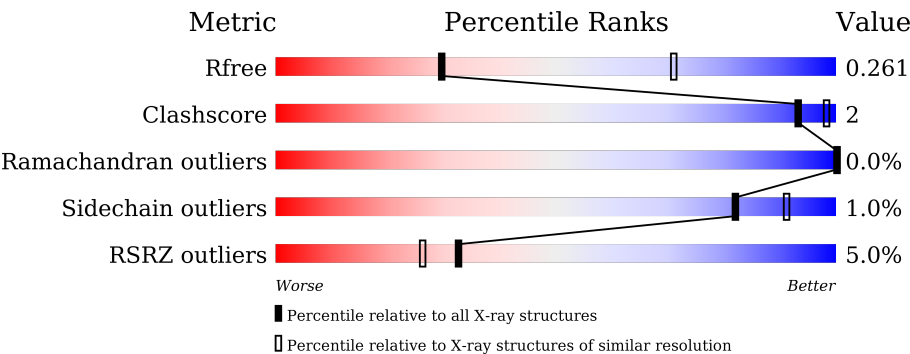
MolProbity : 4.02b-467
Mogul : 1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20027257
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20027257

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}	91344	2060 (3.40-3.20)
Clashscore	102246	1058 (3.38-3.22)
Ramachandran outliers	100387	1038 (3.38-3.22)
Sidechain outliers	100360	1037 (3.38-3.22)
RSRZ outliers	91569	2070 (3.40-3.20)

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. 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	C	179	<div><div>2%</div><div><div></div><div>82%</div><div>7%</div><div>11%</div></div></div>
1	G	179	<div><div>2%</div><div><div></div><div>84%</div><div>•</div><div>12%</div></div></div>
1	K	179	<div><div>%</div><div><div></div><div>84%</div><div>•</div><div>12%</div></div></div>
1	O	179	<div><div>%</div><div><div></div><div>81%</div><div>6%</div><div>•</div><div>12%</div></div></div>
2	D	136	<div><div>2%</div><div><div></div><div>72%</div><div>8%</div><div>•</div><div>19%</div></div></div>
2	H	136	<div><div>4%</div><div><div></div><div>76%</div><div>•</div><div>•</div><div>19%</div></div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	L	136	
2	P	136	
3	A	226	
3	E	226	
3	I	226	
3	M	226	
4	B	217	
4	F	217	
4	J	217	
4	N	217	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
7	MAN	D	706	-	-	-	X
7	MAN	P	706	-	-	-	X

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 41146 atoms, of which 20270 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 Envelope glycoprotein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	C	159	Total	C	H	N	O	S	0	0	0
			2410	770	1195	215	225	5			
1	G	158	Total	C	H	N	O	S	0	0	0
			2400	767	1190	214	224	5			
1	K	157	Total	C	H	N	O	S	0	0	0
			2376	761	1177	210	223	5			
1	O	157	Total	C	H	N	O	S	0	0	0
			2376	761	1177	210	223	5			

There are 64 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	16	TYR	-	expression tag	UNP Q05320
C	17	PRO	-	expression tag	UNP Q05320
C	18	TYR	-	expression tag	UNP Q05320
C	19	ASP	-	expression tag	UNP Q05320
C	20	VAL	-	expression tag	UNP Q05320
C	21	PRO	-	expression tag	UNP Q05320
C	22	ASP	-	expression tag	UNP Q05320
C	23	TYR	-	expression tag	UNP Q05320
C	24	ALA	-	expression tag	UNP Q05320
C	25	ILE	-	expression tag	UNP Q05320
C	26	GLU	-	expression tag	UNP Q05320
C	27	GLY	-	expression tag	UNP Q05320
C	28	ARG	-	expression tag	UNP Q05320
C	29	GLY	-	expression tag	UNP Q05320
C	30	ALA	-	expression tag	UNP Q05320
C	31	ARG	-	expression tag	UNP Q05320
G	16	TYR	-	expression tag	UNP Q05320
G	17	PRO	-	expression tag	UNP Q05320
G	18	TYR	-	expression tag	UNP Q05320
G	19	ASP	-	expression tag	UNP Q05320
G	20	VAL	-	expression tag	UNP Q05320

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	21	PRO	-	expression tag	UNP Q05320
G	22	ASP	-	expression tag	UNP Q05320
G	23	TYR	-	expression tag	UNP Q05320
G	24	ALA	-	expression tag	UNP Q05320
G	25	ILE	-	expression tag	UNP Q05320
G	26	GLU	-	expression tag	UNP Q05320
G	27	GLY	-	expression tag	UNP Q05320
G	28	ARG	-	expression tag	UNP Q05320
G	29	GLY	-	expression tag	UNP Q05320
G	30	ALA	-	expression tag	UNP Q05320
G	31	ARG	-	expression tag	UNP Q05320
K	16	TYR	-	expression tag	UNP Q05320
K	17	PRO	-	expression tag	UNP Q05320
K	18	TYR	-	expression tag	UNP Q05320
K	19	ASP	-	expression tag	UNP Q05320
K	20	VAL	-	expression tag	UNP Q05320
K	21	PRO	-	expression tag	UNP Q05320
K	22	ASP	-	expression tag	UNP Q05320
K	23	TYR	-	expression tag	UNP Q05320
K	24	ALA	-	expression tag	UNP Q05320
K	25	ILE	-	expression tag	UNP Q05320
K	26	GLU	-	expression tag	UNP Q05320
K	27	GLY	-	expression tag	UNP Q05320
K	28	ARG	-	expression tag	UNP Q05320
K	29	GLY	-	expression tag	UNP Q05320
K	30	ALA	-	expression tag	UNP Q05320
K	31	ARG	-	expression tag	UNP Q05320
O	16	TYR	-	expression tag	UNP Q05320
O	17	PRO	-	expression tag	UNP Q05320
O	18	TYR	-	expression tag	UNP Q05320
O	19	ASP	-	expression tag	UNP Q05320
O	20	VAL	-	expression tag	UNP Q05320
O	21	PRO	-	expression tag	UNP Q05320
O	22	ASP	-	expression tag	UNP Q05320
O	23	TYR	-	expression tag	UNP Q05320
O	24	ALA	-	expression tag	UNP Q05320
O	25	ILE	-	expression tag	UNP Q05320
O	26	GLU	-	expression tag	UNP Q05320
O	27	GLY	-	expression tag	UNP Q05320
O	28	ARG	-	expression tag	UNP Q05320
O	29	GLY	-	expression tag	UNP Q05320
O	30	ALA	-	expression tag	UNP Q05320

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
O	31	ARG	-	expression tag	UNP Q05320

- Molecule 2 is a protein called Envelope glycoprotein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	D	110	Total	C	H	N	O	S	0	0	0
			1697	549	835	151	156	6			
2	H	110	Total	C	H	N	O	S	0	0	0
			1698	549	836	151	156	6			
2	L	110	Total	C	H	N	O	S	0	0	0
			1698	549	836	151	156	6			
2	P	110	Total	C	H	N	O	S	0	0	0
			1697	549	835	151	156	6			

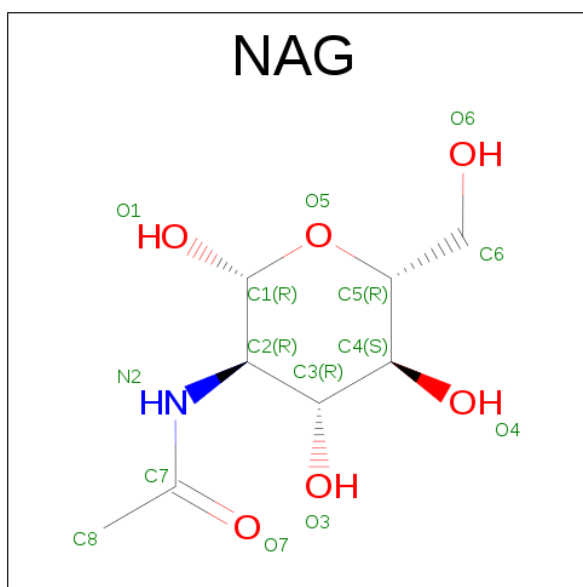
- Molecule 3 is a protein called KZ52 Antibody Fragment.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	E	204	Total	C	H	N	O	S	0	0	0
			3025	962	1495	260	300	8			
3	I	199	Total	C	H	N	O	S	0	0	0
			2973	948	1470	256	291	8			
3	M	203	Total	C	H	N	O	S	0	0	0
			3021	960	1495	260	298	8			
3	A	199	Total	C	H	N	O	S	0	0	0
			2982	950	1476	256	292	8			

- Molecule 4 is a protein called KZ52 Antibody Fragment.

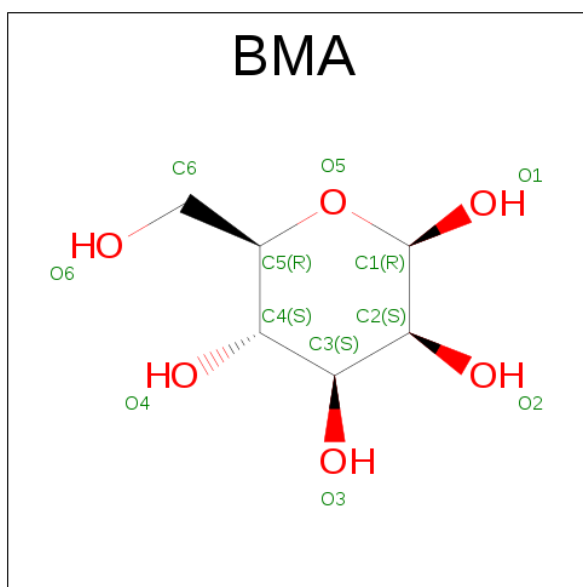
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
4	F	189	Total	C	H	N	O	S	0	0	0
			2891	927	1423	242	295	4			
4	J	209	Total	C	H	N	O	S	0	0	0
			3172	1007	1566	266	328	5			
4	N	203	Total	C	H	N	O	S	0	0	0
			3066	989	1495	258	319	5			
4	B	203	Total	C	H	N	O	S	0	0	0
			3052	975	1489	263	320	5			

- Molecule 5 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: C₈H₁₅NO₆).



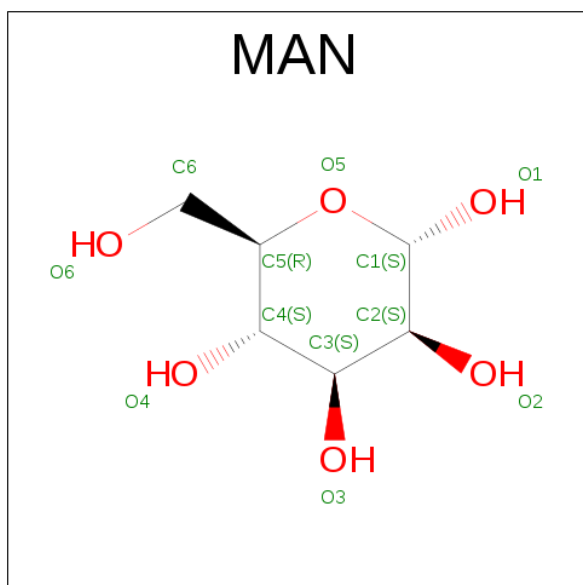
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	D	1	Total	C	H	N	O	0	0
			26	8	12	1	5		
5	D	1	Total	C	H	N	O	0	0
			26	8	12	1	5		
5	H	1	Total	C	H	N	O	0	0
			26	8	12	1	5		
5	H	1	Total	C	H	N	O	0	0
			26	8	12	1	5		
5	L	1	Total	C	H	N	O	0	0
			26	8	12	1	5		
5	L	1	Total	C	H	N	O	0	0
			26	8	12	1	5		
5	P	1	Total	C	H	N	O	0	0
			26	8	12	1	5		
5	P	1	Total	C	H	N	O	0	0
			26	8	12	1	5		

- Molecule 6 is BETA-D-MANNOSE (three-letter code: BMA) (formula: C₆H₁₂O₆).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	D	1	Total	C	H	O	0	0
			19	6	8	5		
6	H	1	Total	C	H	O	0	0
			19	6	8	5		
6	L	1	Total	C	H	O	0	0
			19	6	8	5		
6	P	1	Total	C	H	O	0	0
			19	6	8	5		

- Molecule 7 is ALPHA-D-MANNOSE (three-letter code: MAN) (formula: C₆H₁₂O₆).

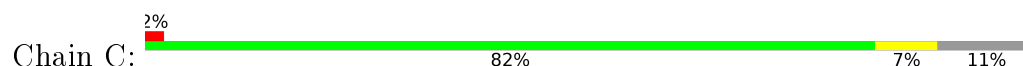


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	D	1	Total 20	C 6	H 9	O 5	0	0
7	D	1	Total 20	C 6	H 9	O 5	0	0
7	D	1	Total 21	C 6	H 10	O 5	0	0
7	D	1	Total 21	C 6	H 10	O 5	0	0
7	H	1	Total 20	C 6	H 9	O 5	0	0
7	H	1	Total 20	C 6	H 9	O 5	0	0
7	H	1	Total 21	C 6	H 10	O 5	0	0
7	H	1	Total 21	C 6	H 10	O 5	0	0
7	L	1	Total 20	C 6	H 9	O 5	0	0
7	L	1	Total 20	C 6	H 9	O 5	0	0
7	L	1	Total 21	C 6	H 10	O 5	0	0
7	L	1	Total 21	C 6	H 10	O 5	0	0
7	P	1	Total 20	C 6	H 9	O 5	0	0
7	P	1	Total 20	C 6	H 9	O 5	0	0
7	P	1	Total 21	C 6	H 10	O 5	0	0
7	P	1	Total 21	C 6	H 10	O 5	0	0

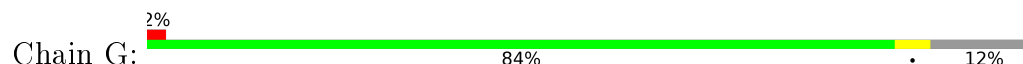
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

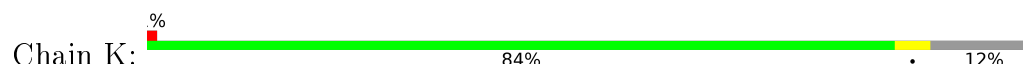
- Molecule 1: Envelope glycoprotein



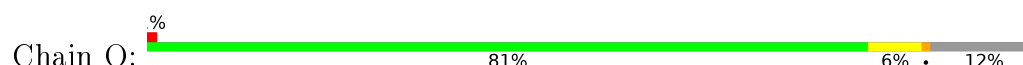
- Molecule 1: Envelope glycoprotein



- Molecule 1: Envelope glycoprotein



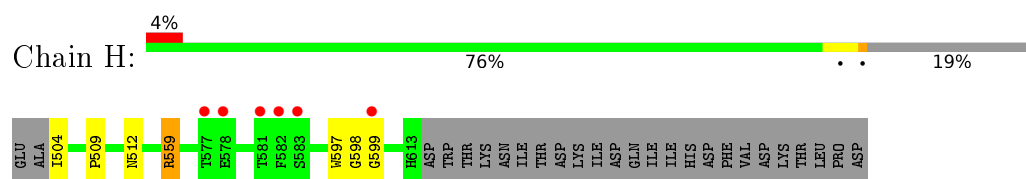
- Molecule 1: Envelope glycoprotein



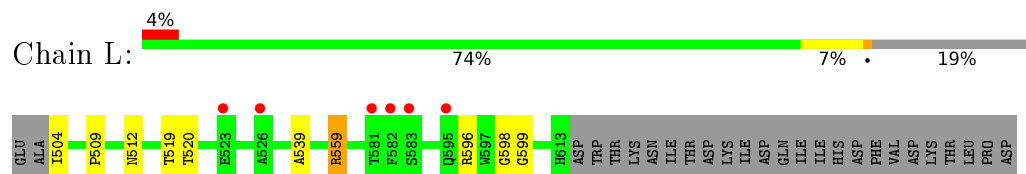
- Molecule 2: Envelope glycoprotein



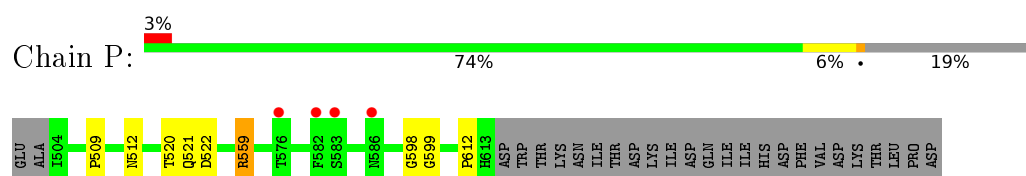
- Molecule 2: Envelope glycoprotein



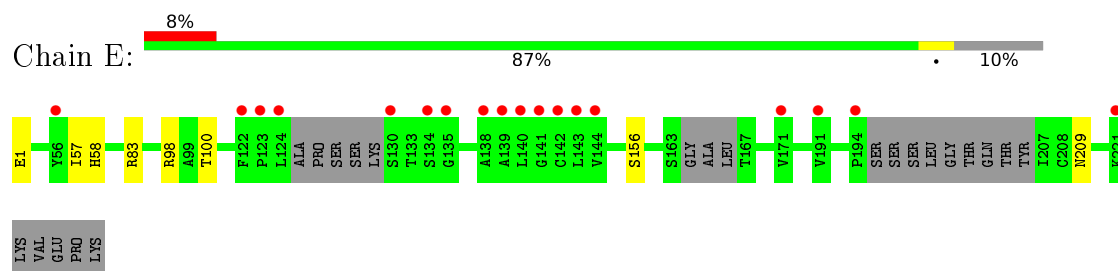
- Molecule 2: Envelope glycoprotein



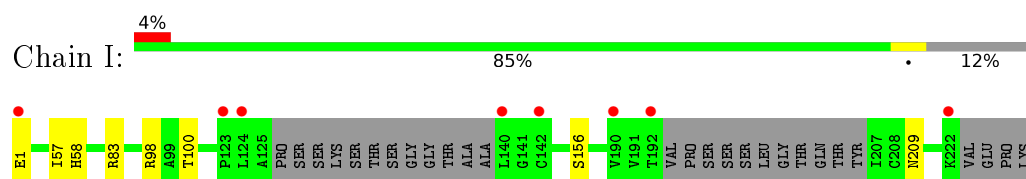
- Molecule 2: Envelope glycoprotein



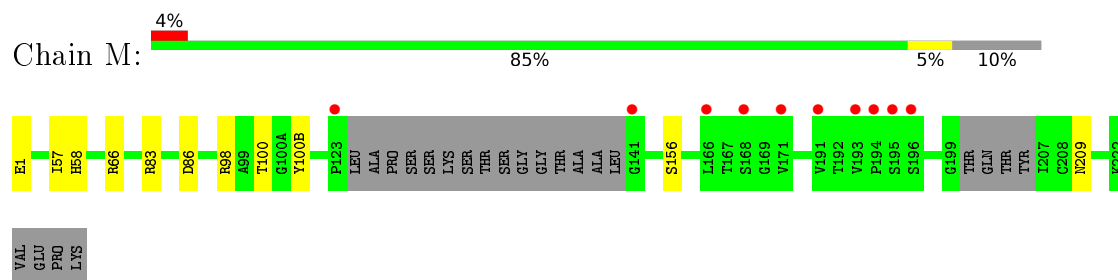
- Molecule 3: KZ52 Antibody Fragment



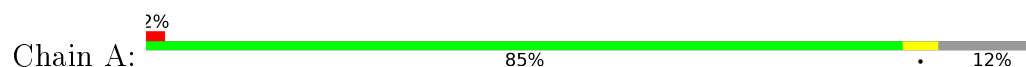
- Molecule 3: KZ52 Antibody Fragment

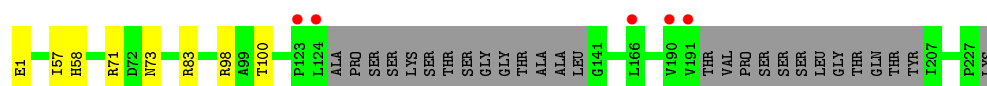


- Molecule 3: KZ52 Antibody Fragment

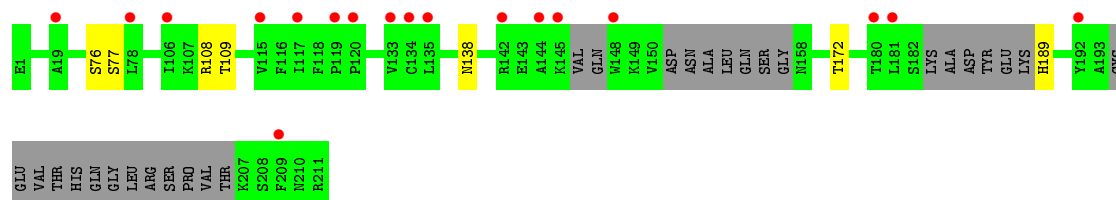
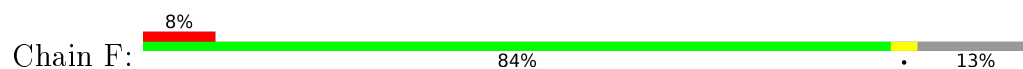


- Molecule 3: KZ52 Antibody Fragment

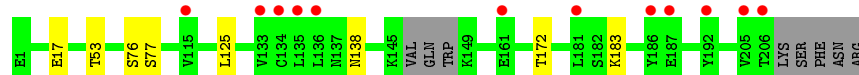
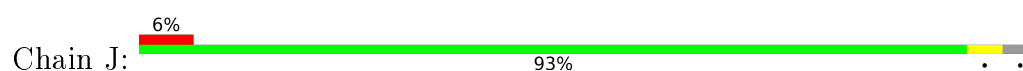




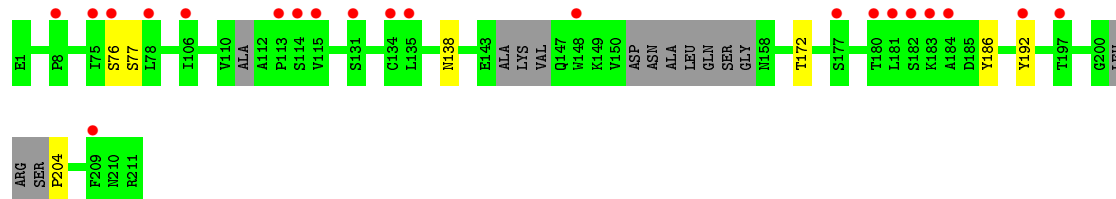
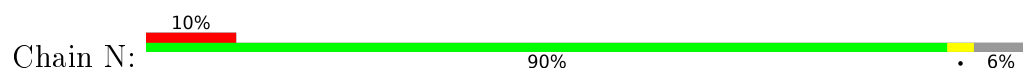
• Molecule 4: KZ52 Antibody Fragment



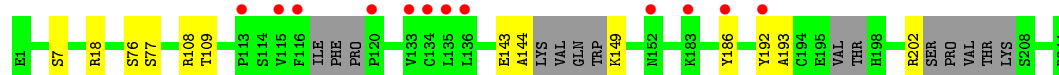
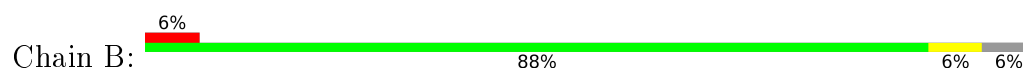
• Molecule 4: KZ52 Antibody Fragment



• Molecule 4: KZ52 Antibody Fragment



• Molecule 4: KZ52 Antibody Fragment



4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α , β , γ	193.10Å 193.10Å 350.32Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.28 – 3.30 48.28 – 3.30	Depositor EDS
% Data completeness (in resolution range)	99.9 (48.28-3.30) 99.9 (48.28-3.30)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.54 (at 3.33Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.227 , 0.258 0.232 , 0.261	Depositor DCC
R_{free} test set	3686 reflections (5.04%)	DCC
Wilson B-factor (Å ²)	124.8	Xtriage
Anisotropy	0.167	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 91.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	0.000 for h,-h-k,-l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	41146	wwPDB-VP
Average B, all atoms (Å ²)	144.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.35% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: BMA, NAG, MAN

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	0.26	0/1244	0.40	0/1689
1	G	0.24	0/1239	0.39	0/1682
1	K	0.24	0/1228	0.39	0/1668
1	O	0.26	0/1228	0.40	0/1668
2	D	0.25	0/884	0.43	0/1205
2	H	0.23	0/884	0.42	0/1205
2	L	0.24	0/884	0.40	0/1205
2	P	0.24	0/884	0.41	0/1205
3	A	0.25	0/1539	0.38	0/2089
3	E	0.23	0/1562	0.38	0/2121
3	I	0.30	0/1535	0.40	0/2083
3	M	0.23	0/1559	0.38	0/2116
4	B	0.25	0/1591	0.39	0/2150
4	F	0.23	0/1498	0.38	0/2028
4	J	0.25	0/1639	0.39	0/2226
4	N	0.23	0/1603	0.40	1/2171 (0.0%)
All	All	0.24	0/21001	0.40	1/28511 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	N	204	PRO	N-CA-CB	5.91	110.39	103.30

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	C	1215	1195	1195	6	0
1	G	1210	1190	1190	4	0
1	K	1199	1177	1177	4	0
1	O	1199	1177	1177	7	0
2	D	862	835	835	8	0
2	H	862	836	835	5	0
2	L	862	836	835	8	0
2	P	862	835	835	9	0
3	A	1506	1476	1476	4	0
3	E	1530	1495	1495	4	0
3	I	1503	1470	1477	4	0
3	M	1526	1495	1495	6	0
4	B	1563	1489	1509	8	1
4	F	1468	1423	1430	5	0
4	J	1606	1566	1566	4	1
4	N	1571	1495	1503	3	0
5	D	28	24	24	0	0
5	H	28	24	24	0	0
5	L	28	24	24	0	0
5	P	28	24	24	0	0
6	D	11	8	8	0	0
6	H	11	8	8	0	0
6	L	11	8	8	0	0
6	P	11	8	8	0	0
7	D	44	38	38	2	0
7	H	44	38	38	0	0
7	L	44	38	38	0	0
7	P	44	38	38	1	0
All	All	20876	20270	20310	78	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 2.

The worst 5 of 78 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:D:512:ASN:O	2:D:559:ARG:NH1	2.06	0.89
2:P:512:ASN:O	2:P:559:ARG:NH1	2.10	0.85
2:H:512:ASN:O	2:H:559:ARG:NH1	2.14	0.81
2:L:512:ASN:O	2:L:559:ARG:NH1	2.15	0.80
2:D:598:GLY:N	2:D:599:GLY:HA3	2.11	0.66

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)
4:J:17:GLU:OE1	4:B:7:SER:OG[9_664]	2.08	0.12

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	157/179 (88%)	150 (96%)	7 (4%)	0	100	100
1	G	156/179 (87%)	149 (96%)	7 (4%)	0	100	100
1	K	155/179 (87%)	146 (94%)	9 (6%)	0	100	100
1	O	155/179 (87%)	149 (96%)	6 (4%)	0	100	100
2	D	108/136 (79%)	102 (94%)	6 (6%)	0	100	100
2	H	108/136 (79%)	102 (94%)	6 (6%)	0	100	100
2	L	108/136 (79%)	101 (94%)	7 (6%)	0	100	100
2	P	108/136 (79%)	97 (90%)	10 (9%)	1 (1%)	21	60
3	A	193/226 (85%)	185 (96%)	8 (4%)	0	100	100
3	E	196/226 (87%)	191 (97%)	5 (3%)	0	100	100
3	I	193/226 (85%)	186 (96%)	7 (4%)	0	100	100
3	M	197/226 (87%)	190 (96%)	7 (4%)	0	100	100
4	B	193/217 (89%)	185 (96%)	8 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	F	179/217 (82%)	169 (94%)	10 (6%)	0	100	100
4	J	205/217 (94%)	197 (96%)	8 (4%)	0	100	100
4	N	193/217 (89%)	185 (96%)	8 (4%)	0	100	100
All	All	2604/3032 (86%)	2484 (95%)	119 (5%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	P	612	PRO

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	C	130/146 (89%)	126 (97%)	4 (3%)	47	79
1	G	130/146 (89%)	129 (99%)	1 (1%)	86	93
1	K	129/146 (88%)	128 (99%)	1 (1%)	86	93
1	O	129/146 (88%)	126 (98%)	3 (2%)	58	83
2	D	90/115 (78%)	88 (98%)	2 (2%)	60	84
2	H	90/115 (78%)	88 (98%)	2 (2%)	60	84
2	L	90/115 (78%)	88 (98%)	2 (2%)	60	84
2	P	90/115 (78%)	89 (99%)	1 (1%)	80	90
3	A	169/190 (89%)	167 (99%)	2 (1%)	78	90
3	E	172/190 (90%)	170 (99%)	2 (1%)	78	90
3	I	168/190 (88%)	166 (99%)	2 (1%)	78	90
3	M	172/190 (90%)	171 (99%)	1 (1%)	90	95
4	B	176/191 (92%)	176 (100%)	0	100	100
4	F	167/191 (87%)	167 (100%)	0	100	100
4	J	183/191 (96%)	183 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	N	177/191 (93%)	177 (100%)	0	100	100
All	All	2262/2568 (88%)	2239 (99%)	23 (1%)	82	91

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	559	ARG
3	I	100	THR
3	A	83	ARG
3	I	83	ARG
1	K	64	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
4	F	189	HIS
4	J	28	ASN
4	N	28	ASN
4	N	89	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](#)

28 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
5	NAG	D	701	2,5	14,14,15	0.41	0	15,19,21	0.31	0
5	NAG	D	702	5,6	14,14,15	0.23	0	15,19,21	0.49	0
6	BMA	D	703	5,7	11,11,12	0.60	0	15,15,17	0.79	0
7	MAN	D	704	7,6	11,11,12	0.83	1 (9%)	15,15,17	1.19	2 (13%)
7	MAN	D	705	7,6	11,11,12	0.82	0	15,15,17	1.08	2 (13%)
7	MAN	D	706	7	11,11,12	0.68	0	15,15,17	1.03	2 (13%)
7	MAN	D	707	7	11,11,12	0.75	0	15,15,17	0.85	1 (6%)
5	NAG	H	701	2,5	14,14,15	0.36	0	15,19,21	0.34	0
5	NAG	H	702	5,6	14,14,15	0.15	0	15,19,21	0.48	0
6	BMA	H	703	5,7	11,11,12	0.65	0	15,15,17	0.76	0
7	MAN	H	704	7,6	11,11,12	0.82	1 (9%)	15,15,17	1.18	2 (13%)
7	MAN	H	705	7,6	11,11,12	0.86	0	15,15,17	1.07	2 (13%)
7	MAN	H	706	7	11,11,12	0.68	0	15,15,17	1.04	2 (13%)
7	MAN	H	707	7	11,11,12	0.80	0	15,15,17	0.84	1 (6%)
5	NAG	L	701	2,5	14,14,15	0.31	0	15,19,21	0.30	0
5	NAG	L	702	5,6	14,14,15	0.23	0	15,19,21	0.43	0
6	BMA	L	703	5,7	11,11,12	0.67	0	15,15,17	0.84	0
7	MAN	L	704	7,6	11,11,12	0.87	1 (9%)	15,15,17	1.16	1 (6%)
7	MAN	L	705	7,6	11,11,12	0.88	1 (9%)	15,15,17	1.05	2 (13%)
7	MAN	L	706	7	11,11,12	0.69	0	15,15,17	1.06	2 (13%)
7	MAN	L	707	7	11,11,12	0.78	0	15,15,17	0.84	1 (6%)
5	NAG	P	701	2,5	14,14,15	0.29	0	15,19,21	0.33	0
5	NAG	P	702	5,6	14,14,15	0.17	0	15,19,21	0.49	0
6	BMA	P	703	5,7	11,11,12	0.65	0	15,15,17	0.75	0
7	MAN	P	704	7,6	11,11,12	0.80	1 (9%)	15,15,17	1.18	2 (13%)
7	MAN	P	705	7,6	11,11,12	0.83	0	15,15,17	1.06	2 (13%)
7	MAN	P	706	7	11,11,12	0.70	0	15,15,17	1.06	2 (13%)
7	MAN	P	707	7	11,11,12	0.78	0	15,15,17	0.84	1 (6%)

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	D	701	2,5	-	0/6/23/26	0/1/1/1
5	NAG	D	702	5,6	-	0/6/23/26	0/1/1/1
6	BMA	D	703	5,7	-	0/2/19/22	0/1/1/1
7	MAN	D	704	7,6	-	0/2/19/22	0/1/1/1
7	MAN	D	705	7,6	-	0/2/19/22	0/1/1/1
7	MAN	D	706	7	-	0/2/19/22	1/1/1/1
7	MAN	D	707	7	-	0/2/19/22	0/1/1/1
5	NAG	H	701	2,5	-	0/6/23/26	0/1/1/1
5	NAG	H	702	5,6	-	0/6/23/26	0/1/1/1
6	BMA	H	703	5,7	-	0/2/19/22	0/1/1/1
7	MAN	H	704	7,6	-	0/2/19/22	0/1/1/1
7	MAN	H	705	7,6	-	0/2/19/22	0/1/1/1
7	MAN	H	706	7	-	0/2/19/22	1/1/1/1
7	MAN	H	707	7	-	0/2/19/22	0/1/1/1
5	NAG	L	701	2,5	-	0/6/23/26	0/1/1/1
5	NAG	L	702	5,6	-	0/6/23/26	0/1/1/1
6	BMA	L	703	5,7	-	0/2/19/22	0/1/1/1
7	MAN	L	704	7,6	-	0/2/19/22	0/1/1/1
7	MAN	L	705	7,6	-	0/2/19/22	0/1/1/1
7	MAN	L	706	7	-	0/2/19/22	1/1/1/1
7	MAN	L	707	7	-	0/2/19/22	0/1/1/1
5	NAG	P	701	2,5	-	0/6/23/26	0/1/1/1
5	NAG	P	702	5,6	-	0/6/23/26	0/1/1/1
6	BMA	P	703	5,7	-	0/2/19/22	0/1/1/1
7	MAN	P	704	7,6	-	0/2/19/22	0/1/1/1
7	MAN	P	705	7,6	-	0/2/19/22	0/1/1/1
7	MAN	P	706	7	-	0/2/19/22	1/1/1/1
7	MAN	P	707	7	-	0/2/19/22	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	L	704	MAN	O5-C1	-2.52	1.39	1.43
7	D	704	MAN	O5-C1	-2.42	1.39	1.43
7	H	704	MAN	O5-C1	-2.39	1.39	1.43
7	P	704	MAN	O5-C1	-2.35	1.39	1.43
7	L	705	MAN	O5-C1	-2.06	1.40	1.43

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	D	704	MAN	O2-C2-C3	-3.17	103.80	110.19
7	P	704	MAN	O2-C2-C3	-3.10	103.94	110.19

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	L	704	MAN	O2-C2-C3	-3.08	103.99	110.19
7	H	704	MAN	O2-C2-C3	-3.04	104.06	110.19
7	L	707	MAN	O2-C2-C3	-2.38	105.39	110.19

There are no chirality outliers.

There are no torsion outliers.

All (4) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	D	706	MAN	C1-C2-C3-C4-C5-O5
7	H	706	MAN	C1-C2-C3-C4-C5-O5
7	P	706	MAN	C1-C2-C3-C4-C5-O5
7	L	706	MAN	C1-C2-C3-C4-C5-O5

4 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	D	705	MAN	1	0
7	D	706	MAN	2	0
7	P	705	MAN	1	0
7	P	706	MAN	1	0

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, 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	159/179 (88%)	0.15	3 (1%) 70 63	81, 102, 147, 173	0
1	G	158/179 (88%)	0.30	4 (2%) 61 54	96, 117, 151, 178	0
1	K	157/179 (87%)	0.24	1 (0%) 90 88	93, 117, 158, 174	0
1	O	157/179 (87%)	0.18	2 (1%) 79 74	84, 106, 154, 163	0
2	D	110/136 (80%)	0.48	3 (2%) 58 51	83, 99, 158, 255	0
2	H	110/136 (80%)	0.51	6 (5%) 29 23	88, 113, 180, 231	0
2	L	110/136 (80%)	0.59	6 (5%) 29 23	88, 109, 189, 280	0
2	P	110/136 (80%)	0.43	4 (3%) 46 39	90, 106, 170, 273	0
3	A	199/226 (88%)	0.14	5 (2%) 61 54	81, 102, 184, 227	0
3	E	204/226 (90%)	0.47	18 (8%) 12 10	101, 118, 218, 310	0
3	I	199/226 (88%)	0.34	8 (4%) 42 34	75, 110, 193, 227	0
3	M	203/226 (89%)	0.26	10 (4%) 33 27	93, 123, 225, 285	0
4	B	203/217 (93%)	0.41	12 (5%) 26 20	84, 118, 226, 250	0
4	F	189/217 (87%)	0.56	18 (9%) 10 9	24, 156, 230, 276	0
4	J	209/217 (96%)	0.41	12 (5%) 27 22	85, 120, 231, 258	0
4	N	203/217 (93%)	0.58	21 (10%) 9 7	109, 159, 247, 276	0
All	All	2680/3032 (88%)	0.37	133 (4%) 32 26	24, 117, 216, 310	0

The worst 5 of 133 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	E	140	LEU	8.0
4	N	209	PHE	6.2
3	E	124	LEU	5.9
3	E	143	LEU	5.6
4	J	135	LEU	5.3

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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å ²)	Q<0.9
7	MAN	D	706	11/12	0.73	0.46	5.34	243,287,344,361	0
7	MAN	P	706	11/12	0.67	0.45	5.27	243,287,344,361	0
5	NAG	L	701	14/15	0.89	0.32	1.16	87,117,156,156	0
7	MAN	H	706	11/12	0.84	0.27	0.70	243,287,345,361	0
5	NAG	D	701	14/15	0.93	0.28	0.62	86,115,155,155	0
5	NAG	P	701	14/15	0.95	0.28	0.58	85,117,155,155	0
5	NAG	H	701	14/15	0.96	0.26	-0.39	86,118,157,157	0
5	NAG	H	702	14/15	0.93	0.18	-1.36	120,144,181,188	0
5	NAG	L	702	14/15	0.90	0.16	-1.44	117,146,181,187	0
5	NAG	D	702	14/15	0.94	0.19	-1.55	117,143,180,186	0
5	NAG	P	702	14/15	0.89	0.22	-1.65	119,144,182,189	0
6	BMA	P	703	11/12	0.93	0.10	-	169,204,288,288	0
6	BMA	L	703	11/12	0.85	0.13	-	169,204,288,288	0
7	MAN	P	705	11/12	0.70	0.38	-	227,276,334,345	0
7	MAN	H	705	11/12	0.73	0.27	-	226,276,334,345	0
7	MAN	L	706	11/12	0.82	0.27	-	243,287,344,361	0
6	BMA	D	703	11/12	0.92	0.13	-	168,204,288,288	0
7	MAN	P	707	11/12	0.71	0.36	-	247,276,331,340	0
7	MAN	P	704	11/12	0.89	0.17	-	193,231,266,277	0
7	MAN	H	707	11/12	0.63	0.44	-	247,276,331,341	0
7	MAN	H	704	11/12	0.81	0.17	-	193,231,266,278	0
7	MAN	D	707	11/12	0.77	0.28	-	247,276,331,340	0
7	MAN	D	704	11/12	0.86	0.17	-	193,231,267,278	0
7	MAN	L	704	11/12	0.77	0.27	-	194,231,267,278	0
7	MAN	L	705	11/12	0.66	0.30	-	227,276,334,345	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
7	MAN	D	705	11/12	0.75	0.33	-	227,276,334,345	0
7	MAN	L	707	11/12	0.24	0.68	-	245,276,331,340	0
6	BMA	H	703	11/12	0.79	0.11	-	169,204,288,288	0

6.5 Other polymers [i](#)

There are no such residues in this entry.