wwPDB X-ray Structure Validation Summary Report (i)

Nov 5, 2023 – 05:31 PM EST

PDB ID	:	5W4L
Title	:	Crystal structure of the non-neutralizing and ADCC-potent C11-like antibody N12-i3 in complex with HIV-1 clade A/E gp120, the CD4 mimetic M48U1, and the antibody N5-i5.
Authors	:	Tolbert, W.D.; Gohain, N.; Pazgier, M.
Deposited on	:	2017-06-12
Resolution	:	2.92 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:

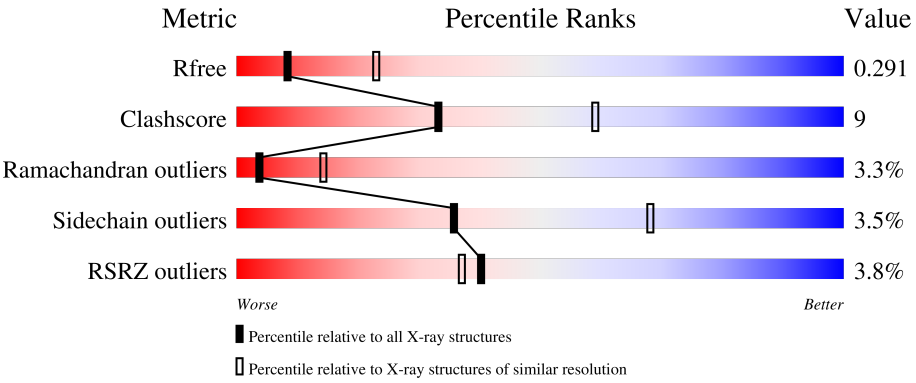
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.92 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	2307 (2.94-2.90)
Clashscore	141614	2531 (2.94-2.90)
Ramachandran outliers	138981	2462 (2.94-2.90)
Sidechain outliers	138945	2464 (2.94-2.90)
RSRZ outliers	127900	2248 (2.94-2.90)

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	A	385	% 80% 10% · 8%
1	G	385	81% 11% 7%
2	M	28	75% 18% 7%
2	N	28	86% 7% 7%
3	F	226	20% 70% 22% · 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	H	226	5% 76% 17% 6%
4	I	217	8% 77% 19%
4	L	217	5% 85% 12%
5	B	225	63% 26% 6%
5	D	225	73% 20% 6%
6	C	215	% 80% 16%
6	E	215	% 64% 29% 6%

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	NAG	A	602	-	-	-	X
7	NAG	A	610	-	-	-	X

2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 19111 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 clade A/E 93TH057 HIV-1 gp120 core.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	G	357	Total	C	N	O	S	0	0	0
			2792	1756	481	532	23			
1	A	356	Total	C	N	O	S	0	0	0
			2786	1753	480	530	23			

There are 66 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	31	SER	-	see sequence details	UNP A0A0M3KKW9
G	32	ASP	-	see sequence details	UNP A0A0M3KKW9
G	33	ASN	-	see sequence details	UNP A0A0M3KKW9
G	34	LEU	-	see sequence details	UNP A0A0M3KKW9
G	35	TRP	-	see sequence details	UNP A0A0M3KKW9
G	36	VAL	-	see sequence details	UNP A0A0M3KKW9
G	37	THR	-	see sequence details	UNP A0A0M3KKW9
G	38	VAL	-	see sequence details	UNP A0A0M3KKW9
G	39	TYR	-	see sequence details	UNP A0A0M3KKW9
G	40	TYR	-	see sequence details	UNP A0A0M3KKW9
G	41	GLY	-	see sequence details	UNP A0A0M3KKW9
G	42	VAL	-	see sequence details	UNP A0A0M3KKW9
G	43	PRO	-	see sequence details	UNP A0A0M3KKW9
G	375	SER	HIS	engineered mutation	UNP A0A0M3KKW9
G	493	PRO	-	see sequence details	UNP A0A0M3KKW9
G	494	LEU	-	see sequence details	UNP A0A0M3KKW9
G	495	GLY	-	see sequence details	UNP A0A0M3KKW9
G	496	ILE	-	see sequence details	UNP A0A0M3KKW9
G	497	ALA	-	see sequence details	UNP A0A0M3KKW9
G	498	PRO	-	see sequence details	UNP A0A0M3KKW9
G	499	THR	-	see sequence details	UNP A0A0M3KKW9
G	500	LYS	-	see sequence details	UNP A0A0M3KKW9
G	501	ALA	-	see sequence details	UNP A0A0M3KKW9
G	502	LYS	-	see sequence details	UNP A0A0M3KKW9
G	503	ARG	-	see sequence details	UNP A0A0M3KKW9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	504	ARG	-	see sequence details	UNP A0A0M3KKW9
G	505	VAL	-	see sequence details	UNP A0A0M3KKW9
G	506	VAL	-	see sequence details	UNP A0A0M3KKW9
G	507	GLN	-	see sequence details	UNP A0A0M3KKW9
G	508	ARG	-	see sequence details	UNP A0A0M3KKW9
G	509	GLU	-	see sequence details	UNP A0A0M3KKW9
G	510	LYS	-	see sequence details	UNP A0A0M3KKW9
G	511	ARG	-	see sequence details	UNP A0A0M3KKW9
A	31	SER	-	see sequence details	UNP A0A0M3KKW9
A	32	ASP	-	see sequence details	UNP A0A0M3KKW9
A	33	ASN	-	see sequence details	UNP A0A0M3KKW9
A	34	LEU	-	see sequence details	UNP A0A0M3KKW9
A	35	TRP	-	see sequence details	UNP A0A0M3KKW9
A	36	VAL	-	see sequence details	UNP A0A0M3KKW9
A	37	THR	-	see sequence details	UNP A0A0M3KKW9
A	38	VAL	-	see sequence details	UNP A0A0M3KKW9
A	39	TYR	-	see sequence details	UNP A0A0M3KKW9
A	40	TYR	-	see sequence details	UNP A0A0M3KKW9
A	41	GLY	-	see sequence details	UNP A0A0M3KKW9
A	42	VAL	-	see sequence details	UNP A0A0M3KKW9
A	43	PRO	-	see sequence details	UNP A0A0M3KKW9
A	375	SER	HIS	engineered mutation	UNP A0A0M3KKW9
A	493	PRO	-	see sequence details	UNP A0A0M3KKW9
A	494	LEU	-	see sequence details	UNP A0A0M3KKW9
A	495	GLY	-	see sequence details	UNP A0A0M3KKW9
A	496	ILE	-	see sequence details	UNP A0A0M3KKW9
A	497	ALA	-	see sequence details	UNP A0A0M3KKW9
A	498	PRO	-	see sequence details	UNP A0A0M3KKW9
A	499	THR	-	see sequence details	UNP A0A0M3KKW9
A	500	LYS	-	see sequence details	UNP A0A0M3KKW9
A	501	ALA	-	see sequence details	UNP A0A0M3KKW9
A	502	LYS	-	see sequence details	UNP A0A0M3KKW9
A	503	ARG	-	see sequence details	UNP A0A0M3KKW9
A	504	ARG	-	see sequence details	UNP A0A0M3KKW9
A	505	VAL	-	see sequence details	UNP A0A0M3KKW9
A	506	VAL	-	see sequence details	UNP A0A0M3KKW9
A	507	GLN	-	see sequence details	UNP A0A0M3KKW9
A	508	ARG	-	see sequence details	UNP A0A0M3KKW9
A	509	GLU	-	see sequence details	UNP A0A0M3KKW9
A	510	LYS	-	see sequence details	UNP A0A0M3KKW9
A	511	ARG	-	see sequence details	UNP A0A0M3KKW9

- Molecule 2 is a protein called CD4 mimetic peptide M48U1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	N	28	Total	C	N	O	S	0	1	1
			217	137	40	34	6			
2	M	28	Total	C	N	O	S	0	1	1
			219	139	41	33	6			

- Molecule 3 is a protein called Antibody N5-i5 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	H	212	Total	C	N	O	S	0	0	0
			1573	988	269	310	6			
3	F	212	Total	C	N	O	S	0	0	0
			1573	988	269	310	6			

- Molecule 4 is a protein called Antibody N5-i5 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	L	212	Total	C	N	O	S	0	0	0
			1574	981	264	324	5			
4	I	212	Total	C	N	O	S	0	1	0
			1585	990	265	325	5			

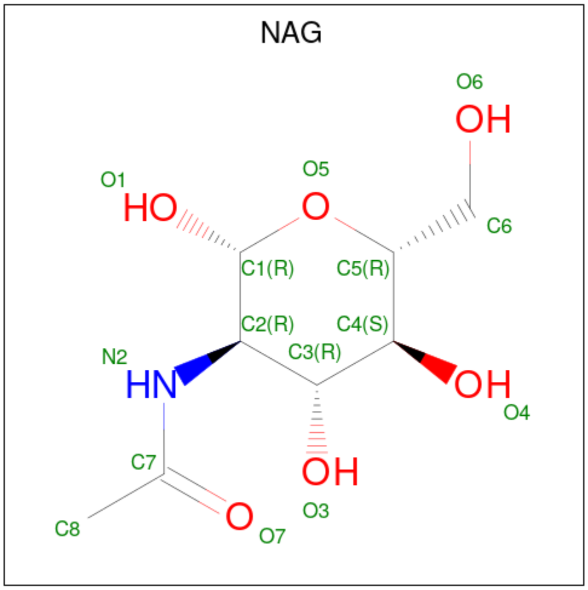
- Molecule 5 is a protein called Antibody N12-i3 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	D	211	Total	C	N	O	S	0	0	0
			1601	1011	270	314	6			
5	B	211	Total	C	N	O	S	0	1	0
			1611	1017	273	315	6			

- Molecule 6 is a protein called Antibody N12-i3 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	E	213	Total	C	N	O	S	0	0	0
			1634	1019	281	329	5			
6	C	212	Total	C	N	O	S	0	0	0
			1630	1017	280	328	5			

- Molecule 7 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	G	1	Total C N O 14 8 1 5	0	0
7	G	1	Total C N O 14 8 1 5	0	0
7	G	1	Total C N O 14 8 1 5	0	0
7	G	1	Total C N O 14 8 1 5	0	0
7	G	1	Total C N O 14 8 1 5	0	0
7	G	1	Total C N O 14 8 1 5	0	0
7	G	1	Total C N O 14 8 1 5	0	0
7	G	1	Total C N O 14 8 1 5	0	0
7	G	1	Total C N O 14 8 1 5	0	0
7	G	1	Total C N O 14 8 1 5	0	0
7	G	1	Total C N O 14 8 1 5	0	0
7	A	1	Total C N O 14 8 1 5	0	0
7	A	1	Total C N O 14 8 1 5	0	0
7	A	1	Total C N O 14 8 1 5	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total C N O 14 8 1 5	0	0
7	A	1	Total C N O 14 8 1 5	0	0
7	A	1	Total C N O 14 8 1 5	0	0
7	A	1	Total C N O 14 8 1 5	0	0
7	A	1	Total C N O 14 8 1 5	0	0
7	A	1	Total C N O 14 8 1 5	0	0
7	A	1	Total C N O 14 8 1 5	0	0

- Molecule 8 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	N	1	Total Cl 1 1	0	0
8	A	1	Total Cl 1 1	0	0

- Molecule 9 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	G	4	Total O 4 4	0	0
9	H	1	Total O 1 1	0	0
9	L	2	Total O 2 2	0	0
9	D	3	Total O 3 3	0	0
9	E	1	Total O 1 1	0	0
9	A	6	Total O 6 6	0	0
9	I	1	Total O 1 1	0	0
9	B	1	Total O 1 1	0	0

Continued on next page...

Continued from previous page...

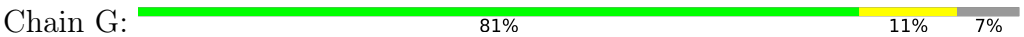
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	C	1	Total O 1 1	0	0

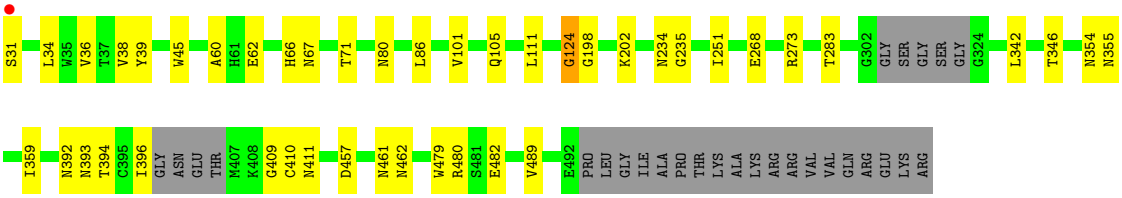
3 Residue-property plots (i)

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: clade A/E 93TH057 HIV-1 gp120 core

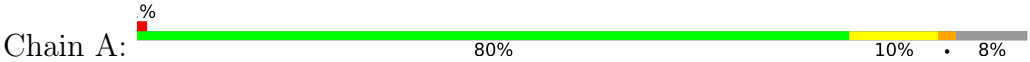
Chain G:

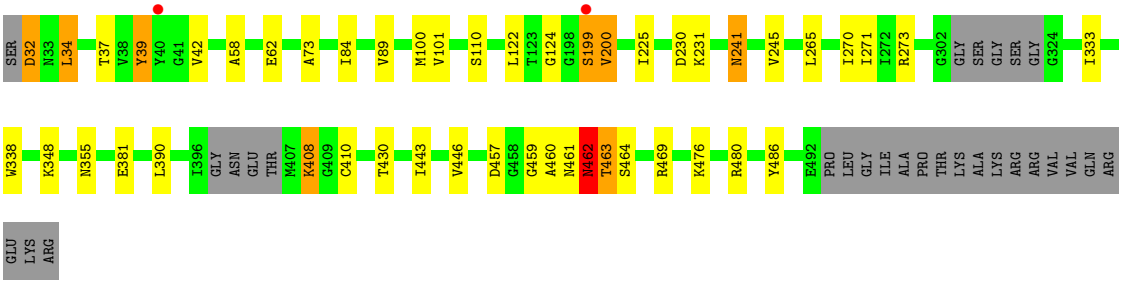




• Molecule 1: clade A/E 93TH057 HIV-1 gp120 core

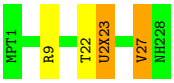
Chain A:





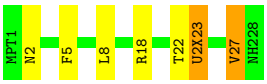
• Molecule 2: CD4 mimetic peptide M48U1

Chain N:



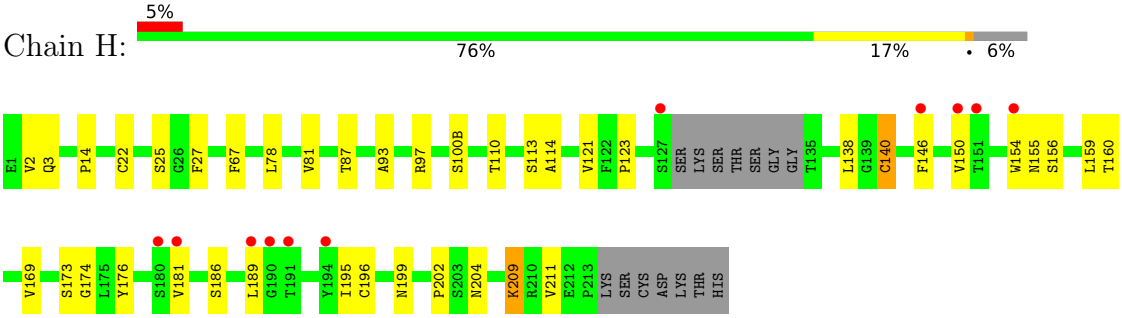
• Molecule 2: CD4 mimetic peptide M48U1

Chain M:



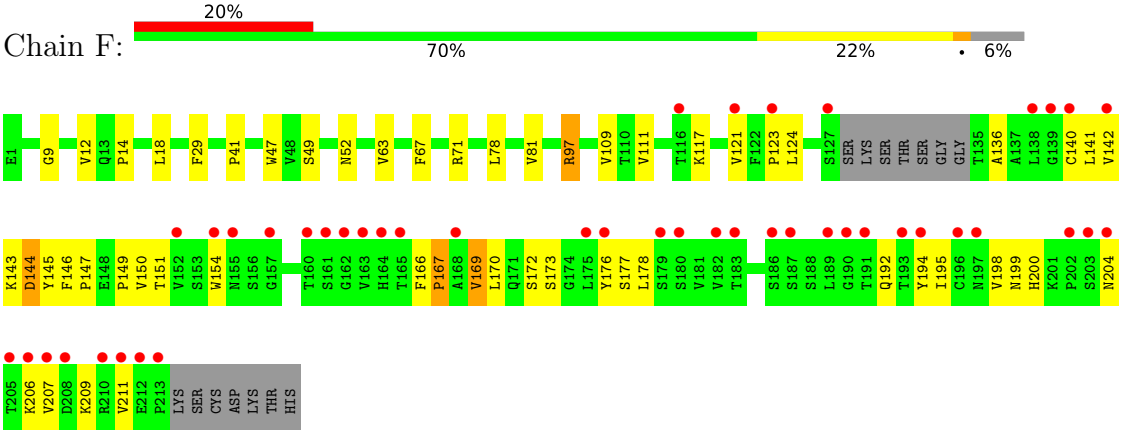
• Molecule 3: Antibody N5-i5 Fab heavy chain

Chain H:



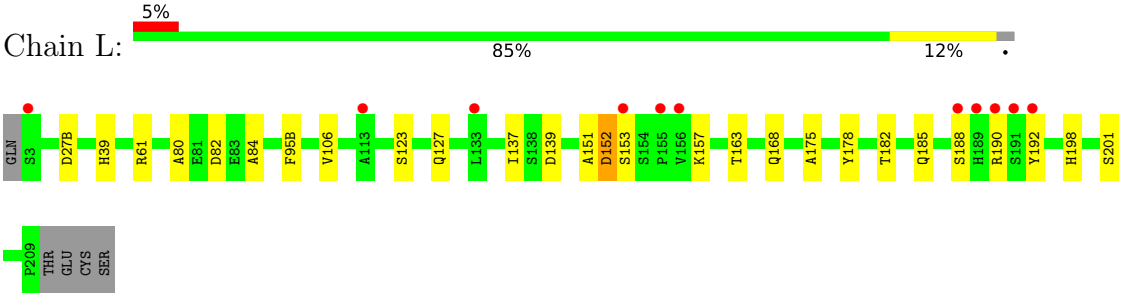
• Molecule 3: Antibody N5-i5 Fab heavy chain

Chain F:



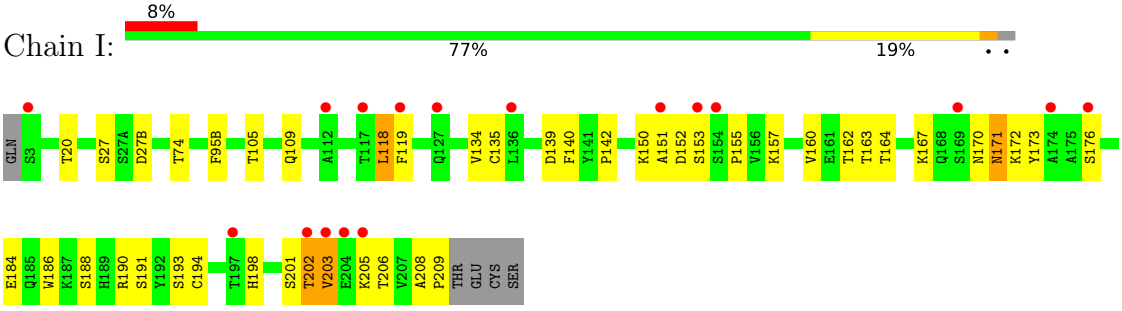
• Molecule 4: Antibody N5-i5 light chain

Chain L:



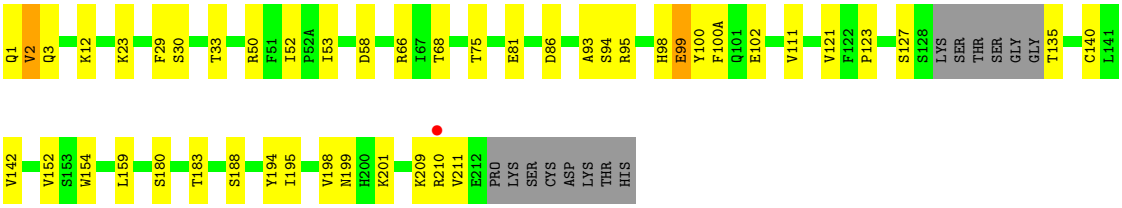
• Molecule 4: Antibody N5-i5 light chain

Chain I:



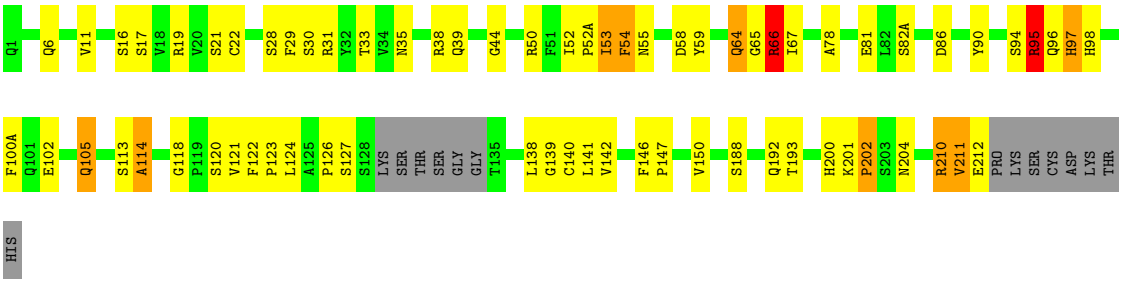
• Molecule 5: Antibody N12-i3 Fab heavy chain

Chain D:



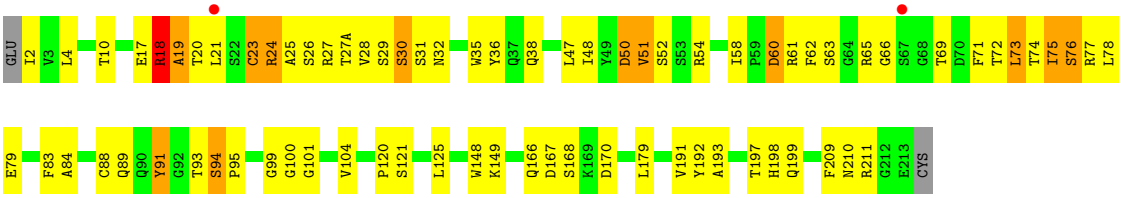
• Molecule 5: Antibody N12-i3 Fab heavy chain

Chain B:



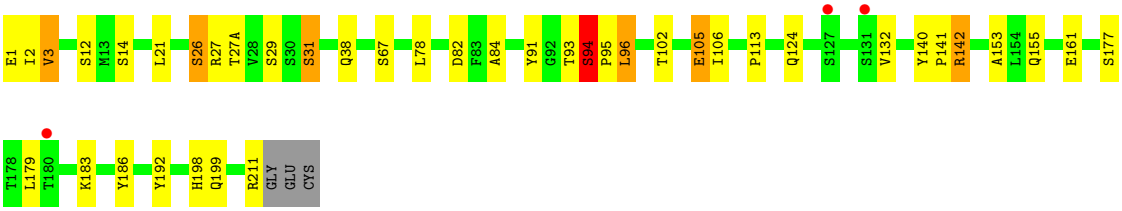
• Molecule 6: Antibody N12-i3 light chain

Chain E:



• Molecule 6: Antibody N12-i3 light chain

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	311.02Å 53.29Å 223.58Å 90.00° 128.89° 90.00°	Depositor
Resolution (Å)	50.00 – 2.92 47.18 – 2.92	Depositor EDS
% Data completeness (in resolution range)	98.5 (50.00-2.92) 98.6 (47.18-2.92)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.10 (at 2.91Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.231 , 0.292 0.232 , 0.291	Depositor DCC
R_{free} test set	3079 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	53.4	Xtriage
Anisotropy	0.326	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 36.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	19111	wwPDB-VP
Average B, all atoms (Å^2)	72.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.83% 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: DPR, MPT, NH2, CL, U2X, 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	A	0.38	0/2846	0.73	6/3866 (0.2%)
1	G	0.37	0/2852	0.67	6/3874 (0.2%)
2	M	0.41	0/187	0.76	0/246
2	N	0.41	0/184	0.71	0/242
3	F	0.41	0/1607	0.65	0/2187
3	H	0.40	0/1607	0.66	0/2187
4	I	0.41	0/1624	0.63	0/2212
4	L	0.37	0/1612	0.60	0/2196
5	B	0.46	0/1652	0.76	2/2253 (0.1%)
5	D	0.39	0/1641	0.65	0/2238
6	C	0.43	0/1664	0.73	1/2259 (0.0%)
6	E	0.48	0/1668	0.80	3/2264 (0.1%)
All	All	0.41	0/19144	0.69	18/26024 (0.1%)

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	A	0	1
2	M	0	2
2	N	0	2
5	B	0	1
6	C	0	1
All	All	0	7

There are no bond length outliers.

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	124	GLY	O-C-N	-14.42	98.69	123.20
1	A	124	GLY	CA-C-N	12.44	141.07	116.20
1	A	124	GLY	C-N-CA	11.02	145.45	122.30
1	G	124	GLY	C-N-CA	9.95	143.20	122.30
1	G	124	GLY	O-C-N	-9.87	106.42	123.20

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	408	LYS	Peptide
2	M	22	THR	Mainchain
2	M	23	U2X	Mainchain
2	N	22	THR	Mainchain
2	N	23	U2X	Mainchain

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	2786	0	2712	33	0
1	G	2792	0	2716	18	0
2	M	219	0	218	0	0
2	N	217	0	217	0	0
3	F	1573	0	1546	39	0
3	H	1573	0	1546	26	0
4	I	1585	0	1525	30	0
4	L	1574	0	1517	15	0
5	B	1611	0	1565	53	0
5	D	1601	0	1559	28	0
6	C	1630	0	1596	27	0
6	E	1634	0	1596	76	0
7	A	140	0	130	3	0
7	G	154	0	143	0	0
8	A	1	0	0	0	0
8	N	1	0	0	0	0
9	A	6	0	0	0	0
9	B	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	C	1	0	0	0	0
9	D	3	0	0	0	0
9	E	1	0	0	0	0
9	G	4	0	0	0	0
9	H	1	0	0	0	0
9	I	1	0	0	0	0
9	L	2	0	0	0	0
All	All	19111	0	18586	325	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 9.

The worst 5 of 325 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)
6:E:35:TRP:CE3	6:E:73:LEU:HD11	1.53	1.41
6:E:35:TRP:CG	6:E:73:LEU:HD21	1.82	1.15
6:E:35:TRP:CE3	6:E:73:LEU:CD1	2.30	1.14
6:E:35:TRP:CD2	6:E:73:LEU:HD21	1.88	1.08
5:B:66:ARG:NH2	5:B:86:ASP:OD1	1.94	0.99

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	350/385 (91%)	324 (93%)	20 (6%)	6 (2%)	9 29
1	G	351/385 (91%)	320 (91%)	26 (7%)	5 (1%)	11 34
2	M	25/28 (89%)	23 (92%)	0	2 (8%)	1 2
2	N	25/28 (89%)	23 (92%)	1 (4%)	1 (4%)	3 10

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
3	F	208/226 (92%)	183 (88%)	16 (8%)	9 (4%)	2 9
3	H	208/226 (92%)	181 (87%)	21 (10%)	6 (3%)	4 17
4	I	211/217 (97%)	175 (83%)	30 (14%)	6 (3%)	5 17
4	L	210/217 (97%)	190 (90%)	17 (8%)	3 (1%)	11 34
5	B	208/225 (92%)	185 (89%)	12 (6%)	11 (5%)	2 5
5	D	207/225 (92%)	183 (88%)	19 (9%)	5 (2%)	6 21
6	C	210/215 (98%)	186 (89%)	16 (8%)	8 (4%)	3 11
6	E	211/215 (98%)	175 (83%)	18 (8%)	18 (8%)	1 2
All	All	2424/2592 (94%)	2148 (89%)	196 (8%)	80 (3%)	4 14

5 of 80 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	D	2	VAL
6	E	18	ARG
6	E	19	ALA
6	E	27(A)	THR
6	E	30	SER

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	317/339 (94%)	306 (96%)	11 (4%)	36 68
1	G	318/339 (94%)	311 (98%)	7 (2%)	52 80
2	M	21/20 (105%)	17 (81%)	4 (19%)	1 4
2	N	21/20 (105%)	19 (90%)	2 (10%)	8 24
3	F	177/189 (94%)	172 (97%)	5 (3%)	43 75
3	H	177/189 (94%)	174 (98%)	3 (2%)	60 84
4	I	178/182 (98%)	173 (97%)	5 (3%)	43 75
4	L	177/182 (97%)	175 (99%)	2 (1%)	73 91

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
5	B	182/193 (94%)	171 (94%)	11 (6%)	19 47
5	D	181/193 (94%)	176 (97%)	5 (3%)	43 75
6	C	185/187 (99%)	177 (96%)	8 (4%)	29 61
6	E	185/187 (99%)	175 (95%)	10 (5%)	22 52
All	All	2119/2220 (96%)	2046 (97%)	73 (3%)	36 69

5 of 73 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	B	66	ARG
6	C	106	ILE
5	B	95	ARG
6	C	3	VAL
6	E	54	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

4 non-standard protein/DNA/RNA residues 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	$\#\|Z\| > 2$
2	U2X	M	23	2	19,20,21	1.25	1 (5%)	22,25,27	0.89	1 (4%)
2	U2X	N	23	2	19,20,21	1.25	1 (5%)	22,25,27	0.83	1 (4%)

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
2	U2X	M	23	2	-	2/10/19/21	0/2/2/2
2	U2X	N	23	2	-	2/10/19/21	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	M	23	U2X	OH-CZ	3.49	1.45	1.37
2	N	23	U2X	OH-CZ	3.43	1.45	1.37

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	M	23	U2X	C7-OH-CZ	-2.25	113.16	117.93
2	N	23	U2X	C7-OH-CZ	-2.09	113.50	117.93

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	N	23	U2X	CE1-CZ-OH-C7
2	N	23	U2X	CE2-CZ-OH-C7
2	M	23	U2X	CE1-CZ-OH-C7
2	M	23	U2X	CE2-CZ-OH-C7

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 23 ligands modelled in this entry, 2 are monoatomic - leaving 21 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	$\#\|Z\| > 2$
7	NAG	G	607	1	14,14,15	0.25	0	17,19,21	0.57	0
7	NAG	A	602	1	14,14,15	0.68	1 (7%)	17,19,21	1.10	1 (5%)
7	NAG	G	602	1	14,14,15	0.44	0	17,19,21	1.54	3 (17%)
7	NAG	A	601	1	14,14,15	0.26	0	17,19,21	0.65	0
7	NAG	G	609	1	14,14,15	0.23	0	17,19,21	0.90	1 (5%)
7	NAG	G	611	1	14,14,15	0.41	0	17,19,21	0.90	1 (5%)
7	NAG	A	610	1	14,14,15	0.39	0	17,19,21	0.80	1 (5%)
7	NAG	A	608	1	14,14,15	0.22	0	17,19,21	0.66	0
7	NAG	A	604	1	14,14,15	0.38	0	17,19,21	1.17	1 (5%)
7	NAG	G	610	1	14,14,15	0.35	0	17,19,21	0.57	0
7	NAG	G	601	1	14,14,15	0.45	0	17,19,21	1.20	2 (11%)
7	NAG	G	608	1	14,14,15	0.30	0	17,19,21	1.01	1 (5%)
7	NAG	G	604	1	14,14,15	0.28	0	17,19,21	0.89	1 (5%)
7	NAG	A	605	1	14,14,15	0.31	0	17,19,21	1.27	2 (11%)
7	NAG	A	607	1	14,14,15	0.30	0	17,19,21	0.69	0
7	NAG	G	603	1	14,14,15	0.57	0	17,19,21	1.13	1 (5%)
7	NAG	G	606	1	14,14,15	0.31	0	17,19,21	1.46	1 (5%)
7	NAG	A	606	1	14,14,15	0.31	0	17,19,21	0.72	0
7	NAG	A	603	1	14,14,15	0.27	0	17,19,21	0.90	1 (5%)
7	NAG	A	609	1	14,14,15	0.38	0	17,19,21	0.89	1 (5%)
7	NAG	G	605	1	14,14,15	0.32	0	17,19,21	0.87	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
7	NAG	G	607	1	-	0/6/23/26	0/1/1/1
7	NAG	A	602	1	-	2/6/23/26	0/1/1/1
7	NAG	G	602	1	-	0/6/23/26	0/1/1/1
7	NAG	A	601	1	-	2/6/23/26	0/1/1/1
7	NAG	G	609	1	-	0/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	NAG	G	611	1	-	0/6/23/26	0/1/1/1
7	NAG	A	610	1	-	0/6/23/26	0/1/1/1
7	NAG	A	608	1	-	0/6/23/26	0/1/1/1
7	NAG	A	604	1	-	2/6/23/26	0/1/1/1
7	NAG	G	610	1	-	0/6/23/26	0/1/1/1
7	NAG	G	601	1	-	0/6/23/26	0/1/1/1
7	NAG	G	608	1	-	0/6/23/26	0/1/1/1
7	NAG	G	604	1	-	2/6/23/26	0/1/1/1
7	NAG	A	605	1	-	0/6/23/26	0/1/1/1
7	NAG	A	607	1	-	0/6/23/26	0/1/1/1
7	NAG	G	603	1	-	0/6/23/26	0/1/1/1
7	NAG	G	606	1	-	2/6/23/26	0/1/1/1
7	NAG	A	606	1	-	0/6/23/26	0/1/1/1
7	NAG	A	603	1	-	0/6/23/26	0/1/1/1
7	NAG	A	609	1	-	0/6/23/26	0/1/1/1
7	NAG	G	605	1	-	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)
7	A	602	NAG	C1-C2	2.15	1.55	1.52

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	G	606	NAG	C1-O5-C5	5.67	119.88	112.19
7	A	604	NAG	C1-O5-C5	3.96	117.56	112.19
7	G	602	NAG	C1-C2-N2	3.50	116.47	110.49
7	A	605	NAG	C1-O5-C5	3.44	116.85	112.19
7	G	601	NAG	C3-C4-C5	3.25	116.04	110.24

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	A	602	NAG	C4-C5-C6-O6
7	A	601	NAG	C4-C5-C6-O6
7	A	601	NAG	O5-C5-C6-O6
7	A	602	NAG	O5-C5-C6-O6
7	A	604	NAG	O5-C5-C6-O6

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	602	NAG	1	0
7	A	610	NAG	1	0
7	A	606	NAG	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, 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	356/385 (92%)	-0.28	2 (0%) 89 89	38, 61, 98, 122	0
1	G	357/385 (92%)	-0.27	1 (0%) 94 94	32, 55, 92, 113	0
2	M	24/28 (85%)	-0.06	0 100 100	46, 54, 60, 66	0
2	N	24/28 (85%)	0.02	0 100 100	46, 61, 65, 77	0
3	F	212/226 (93%)	0.79	45 (21%) 0 0	36, 84, 171, 176	0
3	H	212/226 (93%)	0.07	11 (5%) 27 24	27, 58, 133, 144	0
4	I	212/217 (97%)	0.21	17 (8%) 12 10	45, 86, 128, 138	0
4	L	212/217 (97%)	0.01	11 (5%) 27 24	33, 56, 120, 137	0
5	B	211/225 (93%)	-0.05	0 100 100	52, 77, 99, 108	0
5	D	211/225 (93%)	-0.25	1 (0%) 91 91	41, 59, 94, 117	0
6	C	212/215 (98%)	-0.07	3 (1%) 75 76	49, 68, 108, 125	0
6	E	213/215 (99%)	-0.17	2 (0%) 84 84	48, 76, 120, 134	0
All	All	2456/2592 (94%)	-0.03	93 (3%) 40 37	27, 67, 130, 176	0

The worst 5 of 93 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	F	211	VAL	8.6
4	L	190	ARG	6.9
3	F	162	GLY	6.8
3	F	212	GLU	6.3
3	F	127	SER	5.9

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

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, 95^{th} 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
2	DPR	M	21	7/8	0.96	0.17	44,44,45,46	0
2	U2X	N	23	19/20	0.96	0.15	49,51,53,55	0
2	U2X	M	23	19/20	0.97	0.17	48,49,51,53	0
2	DPR	N	21	7/8	0.98	0.19	44,44,44,45	0

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, 95^{th} 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
7	NAG	A	602	14/15	0.65	0.49	89,97,103,105	0
7	NAG	A	610	14/15	0.71	0.43	78,87,90,92	0
7	NAG	G	601	14/15	0.74	0.29	77,85,87,87	0
7	NAG	G	611	14/15	0.82	0.30	81,90,93,96	0
7	NAG	G	603	14/15	0.85	0.26	71,77,85,88	0
7	NAG	A	601	14/15	0.86	0.21	65,71,75,78	0
7	NAG	A	609	14/15	0.88	0.32	75,80,85,85	0
7	NAG	A	604	14/15	0.88	0.22	64,70,72,73	0
8	CL	N	101	1/1	0.88	0.10	54,54,54,54	0
7	NAG	A	605	14/15	0.89	0.24	60,63,65,68	0
7	NAG	A	607	14/15	0.90	0.24	69,74,79,81	0
7	NAG	A	606	14/15	0.90	0.21	57,62,63,63	0
7	NAG	G	610	14/15	0.91	0.20	74,80,83,84	0
8	CL	A	611	1/1	0.91	0.09	59,59,59,59	0
7	NAG	G	606	14/15	0.92	0.15	46,50,53,54	0
7	NAG	G	602	14/15	0.92	0.18	57,62,64,66	0
7	NAG	G	605	14/15	0.92	0.17	56,61,65,66	0
7	NAG	A	608	14/15	0.92	0.18	48,51,52,54	0
7	NAG	G	609	14/15	0.94	0.25	43,45,46,46	0
7	NAG	G	608	14/15	0.95	0.28	53,57,58,58	0
7	NAG	G	604	14/15	0.95	0.14	43,45,46,48	0
7	NAG	A	603	14/15	0.95	0.14	48,51,52,54	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
7	NAG	G	607	14/15	0.96	0.09	54,57,60,60	0

6.5 Other polymers (i)

There are no such residues in this entry.