



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 26, 2022 – 03:46 PM EDT

PDB ID : 7S7I
Title : Crystal structure of Fab in complex with MICA alpha3 domain
Authors : Lee, P.S.; Strop, P.
Deposited on : 2021-09-16
Resolution : 2.40 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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 : 2.31.2
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) : 2.31.2

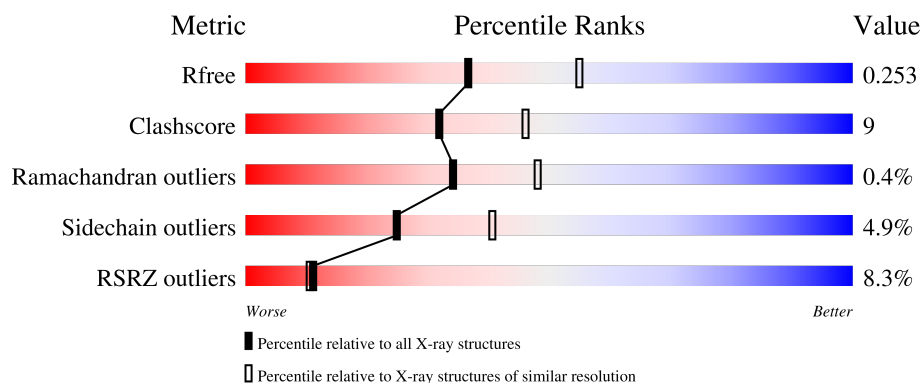
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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.






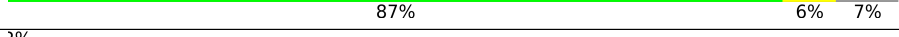
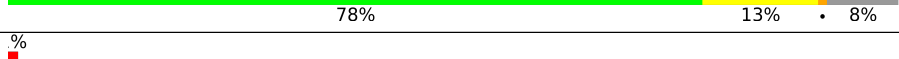
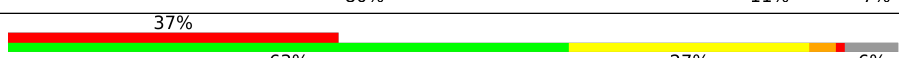


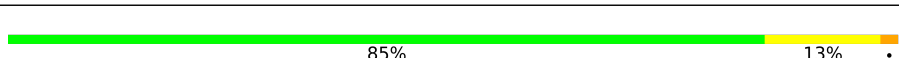
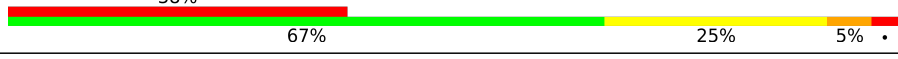


Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 $\leq 5\%$. 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	103	
1	B	103	
1	C	103	
1	D	103	
1	E	103	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	103	
2	H	230	
2	I	230	
2	J	230	
2	U	230	
2	W	230	
2	Y	230	
3	L	214	
3	M	214	
3	N	214	
3	V	214	
3	X	214	
3	Z	214	

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
5	EDO	B	402	-	-	X	-
5	EDO	E	402	-	-	X	-
5	EDO	U	301	-	-	-	X

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 24458 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 MHC class I chain-related protein A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	94	Total	C	N	O	S	0	0	0
			741	455	136	145	5			
1	B	95	Total	C	N	O	S	0	0	0
			747	458	137	147	5			
1	C	93	Total	C	N	O	S	0	0	0
			735	452	135	143	5			
1	D	93	Total	C	N	O	S	0	0	0
			735	452	135	143	5			
1	E	94	Total	C	N	O	S	0	0	0
			741	455	136	145	5			
1	F	94	Total	C	N	O	S	0	0	0
			741	455	136	145	5			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	202	GLY	-	expression tag	UNP H9CTV0
A	203	SER	-	expression tag	UNP H9CTV0
A	299	HIS	-	expression tag	UNP H9CTV0
A	300	HIS	-	expression tag	UNP H9CTV0
A	301	HIS	-	expression tag	UNP H9CTV0
A	302	HIS	-	expression tag	UNP H9CTV0
A	303	HIS	-	expression tag	UNP H9CTV0
A	304	HIS	-	expression tag	UNP H9CTV0
B	202	GLY	-	expression tag	UNP H9CTV0
B	203	SER	-	expression tag	UNP H9CTV0
B	299	HIS	-	expression tag	UNP H9CTV0
B	300	HIS	-	expression tag	UNP H9CTV0
B	301	HIS	-	expression tag	UNP H9CTV0
B	302	HIS	-	expression tag	UNP H9CTV0
B	303	HIS	-	expression tag	UNP H9CTV0
B	304	HIS	-	expression tag	UNP H9CTV0
C	202	GLY	-	expression tag	UNP H9CTV0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	203	SER	-	expression tag	UNP H9CTV0
C	299	HIS	-	expression tag	UNP H9CTV0
C	300	HIS	-	expression tag	UNP H9CTV0
C	301	HIS	-	expression tag	UNP H9CTV0
C	302	HIS	-	expression tag	UNP H9CTV0
C	303	HIS	-	expression tag	UNP H9CTV0
C	304	HIS	-	expression tag	UNP H9CTV0
D	202	GLY	-	expression tag	UNP H9CTV0
D	203	SER	-	expression tag	UNP H9CTV0
D	299	HIS	-	expression tag	UNP H9CTV0
D	300	HIS	-	expression tag	UNP H9CTV0
D	301	HIS	-	expression tag	UNP H9CTV0
D	302	HIS	-	expression tag	UNP H9CTV0
D	303	HIS	-	expression tag	UNP H9CTV0
D	304	HIS	-	expression tag	UNP H9CTV0
E	202	GLY	-	expression tag	UNP H9CTV0
E	203	SER	-	expression tag	UNP H9CTV0
E	299	HIS	-	expression tag	UNP H9CTV0
E	300	HIS	-	expression tag	UNP H9CTV0
E	301	HIS	-	expression tag	UNP H9CTV0
E	302	HIS	-	expression tag	UNP H9CTV0
E	303	HIS	-	expression tag	UNP H9CTV0
E	304	HIS	-	expression tag	UNP H9CTV0
F	202	GLY	-	expression tag	UNP H9CTV0
F	203	SER	-	expression tag	UNP H9CTV0
F	299	HIS	-	expression tag	UNP H9CTV0
F	300	HIS	-	expression tag	UNP H9CTV0
F	301	HIS	-	expression tag	UNP H9CTV0
F	302	HIS	-	expression tag	UNP H9CTV0
F	303	HIS	-	expression tag	UNP H9CTV0
F	304	HIS	-	expression tag	UNP H9CTV0

- Molecule 2 is a protein called Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	213	Total	C	N	O	S	0	0	0
			1618	1023	274	315	6			
2	I	211	Total	C	N	O	S	0	0	0
			1606	1017	272	311	6			
2	J	214	Total	C	N	O	S	0	0	0
			1622	1025	275	316	6			
2	U	212	Total	C	N	O	S	0	0	0
			1610	1019	273	312	6			

Continued on next page...

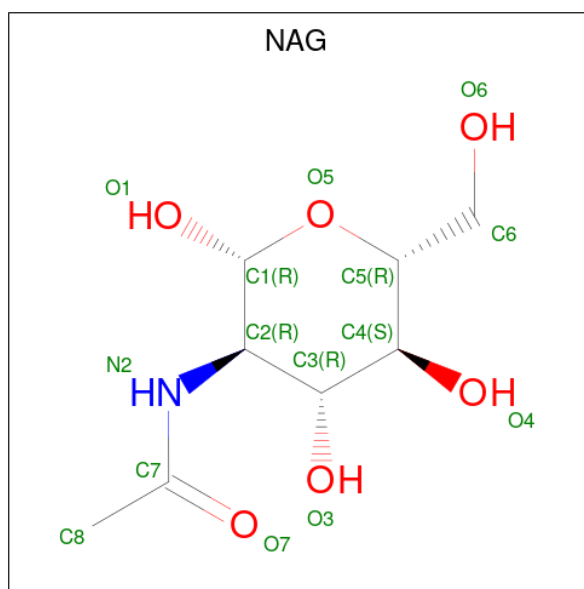
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	W	215	Total	C	N	O	S	0	0	0
			1633	1032	277	318	6			
2	Y	217	Total	C	N	O	S	0	0	0
			1641	1036	279	320	6			

- Molecule 3 is a protein called Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	213	Total	C	N	O	S	0	0	0
			1630	1020	275	331	4			
3	M	213	Total	C	N	O	S	0	0	0
			1630	1020	275	331	4			
3	N	213	Total	C	N	O	S	0	0	0
			1630	1020	275	331	4			
3	V	213	Total	C	N	O	S	0	0	0
			1630	1020	275	331	4			
3	X	213	Total	C	N	O	S	0	0	0
			1630	1020	275	331	4			
3	Z	213	Total	C	N	O	S	0	0	0
			1630	1020	275	331	4			

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



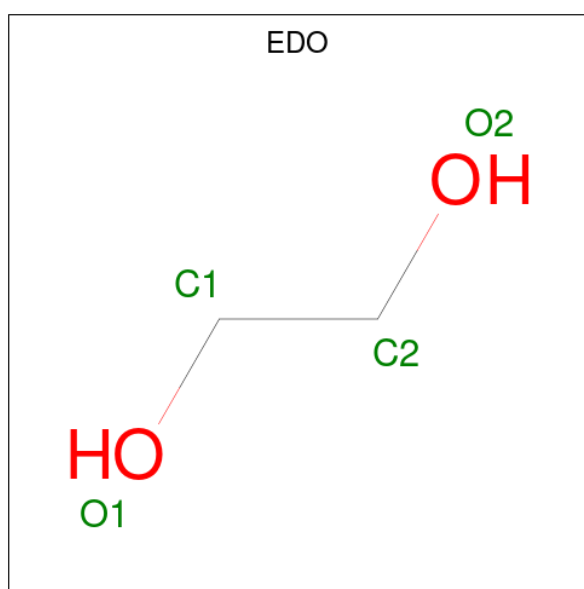
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...

Continued from previous page...

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

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	L	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	U	1	Total	C	O	0	0
			4	2	2		
5	E	1	Total	C	O	0	0
			4	2	2		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	28	Total	O	0	0
			28	28		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	H	58	Total 58	O 58	0	0
6	L	22	Total 22	O 22	0	0
6	B	28	Total 28	O 28	0	0
6	I	54	Total 54	O 54	0	0
6	M	13	Total 13	O 13	0	0
6	C	9	Total 9	O 9	0	0
6	J	52	Total 52	O 52	0	0
6	N	16	Total 16	O 16	0	0
6	D	6	Total 6	O 6	0	0
6	U	31	Total 31	O 31	0	0
6	V	11	Total 11	O 11	0	0
6	E	22	Total 22	O 22	0	0
6	W	30	Total 30	O 30	0	0
6	X	12	Total 12	O 12	0	0
6	F	14	Total 14	O 14	0	0
6	Y	16	Total 16	O 16	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: MHC class I chain-related protein A

Chain A: 




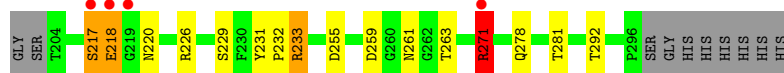
- Molecule 1: MHC class I chain-related protein A

Chain B: 



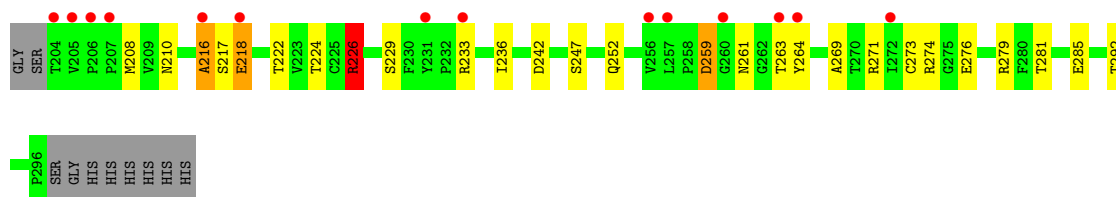
- Molecule 1: MHC class I chain-related protein A

Chain C: 



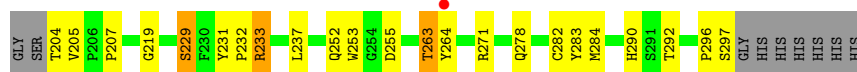
- Molecule 1: MHC class I chain-related protein A

Chain D: 

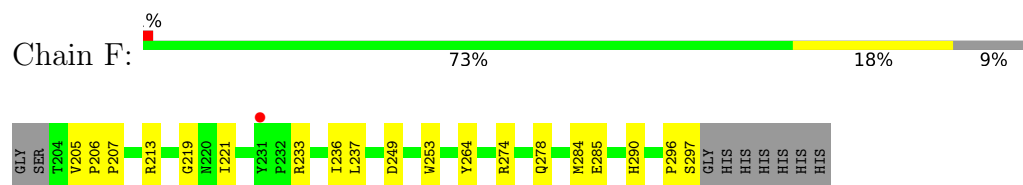


- Molecule 1: MHC class I chain-related protein A

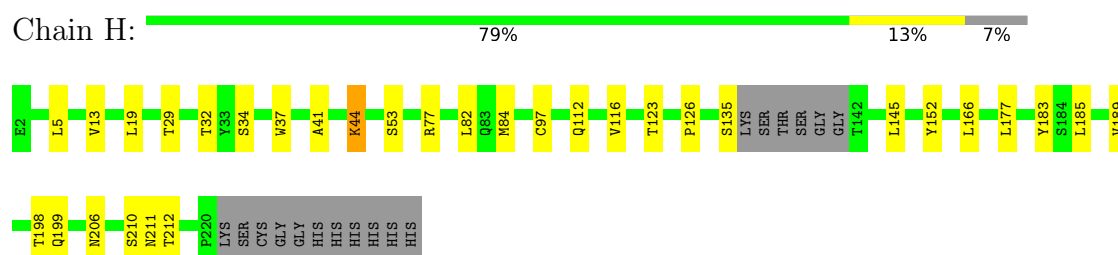
Chain E: 



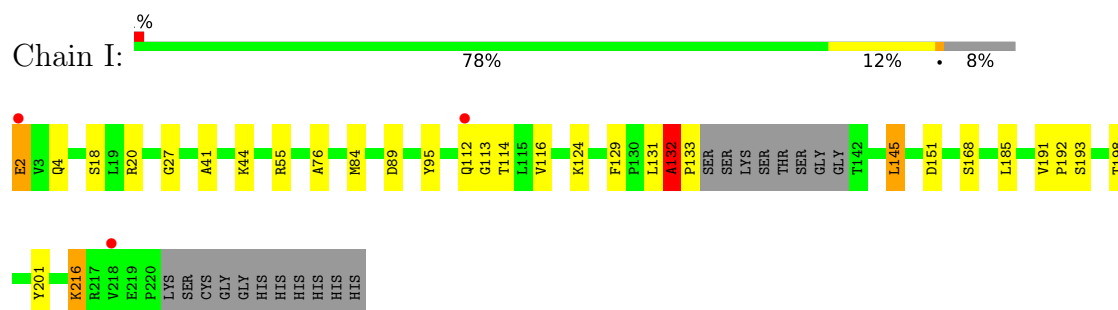
- Molecule 1: MHC class I chain-related protein A



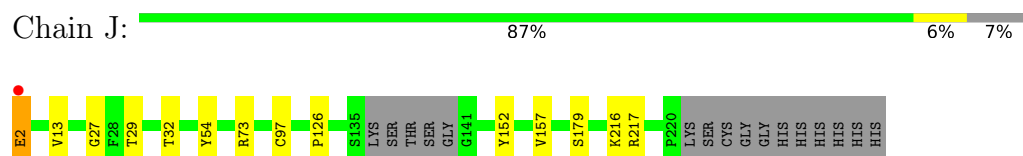
- Molecule 2: Fab heavy chain



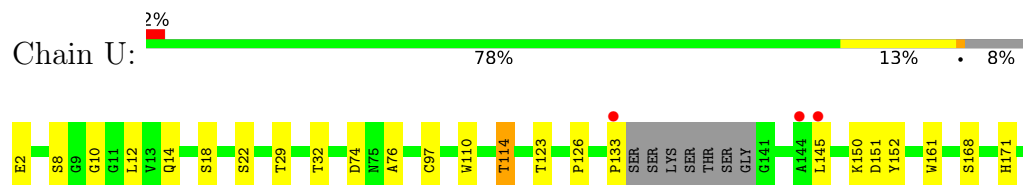
- Molecule 2: Fab heavy chain



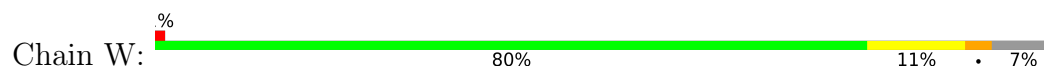
- Molecule 2: Fab heavy chain

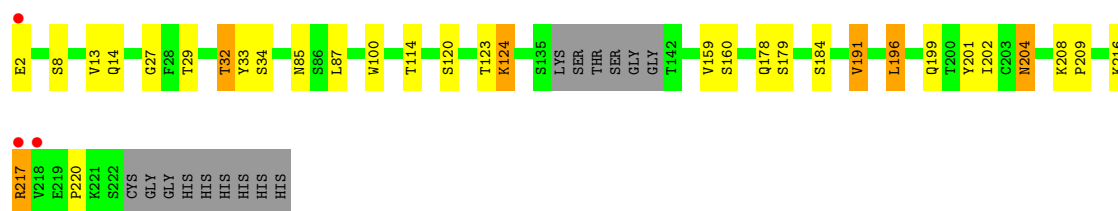


- Molecule 2: Fab heavy chain

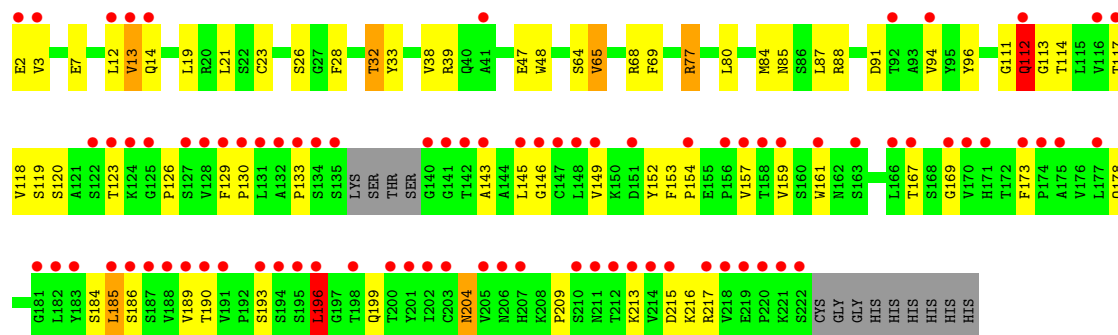


- Molecule 2: Fab heavy chain

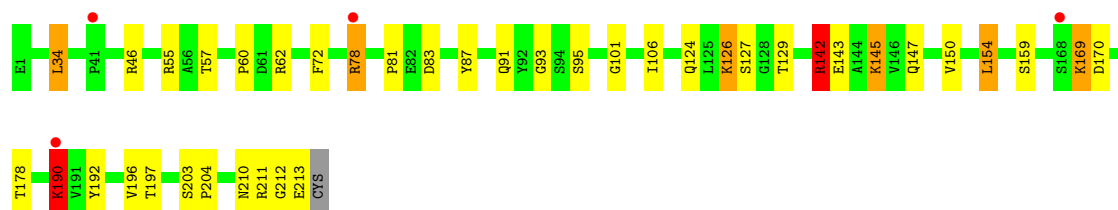
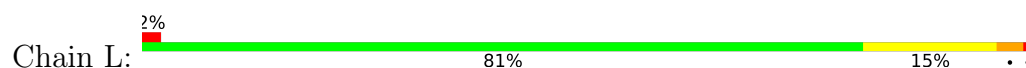




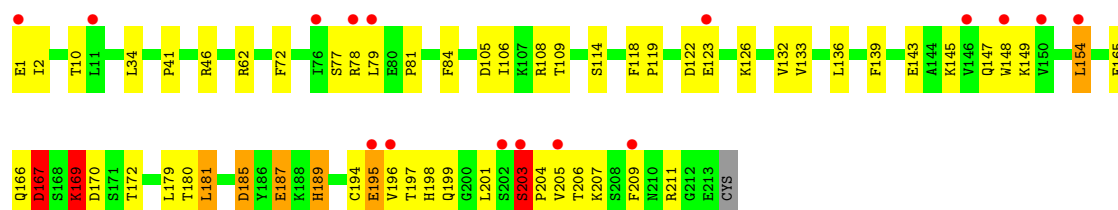
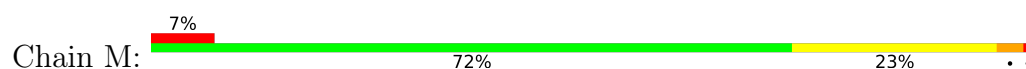
• Molecule 2: Fab heavy chain



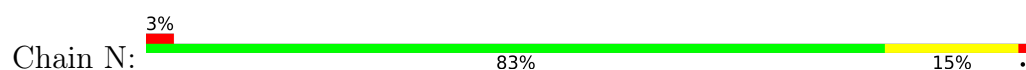
• Molecule 3: Fab light chain

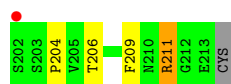


• Molecule 3: Fab light chain

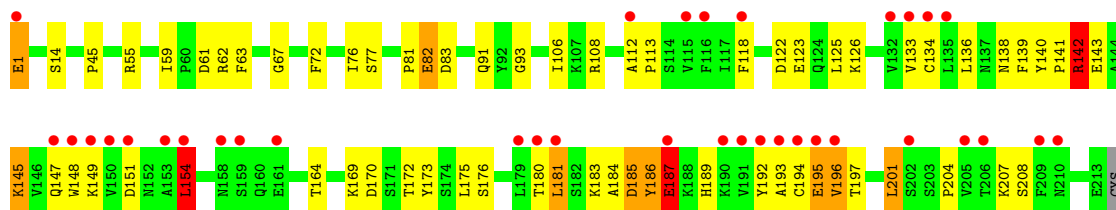


• Molecule 3: Fab light chain

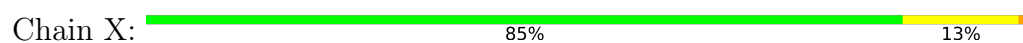




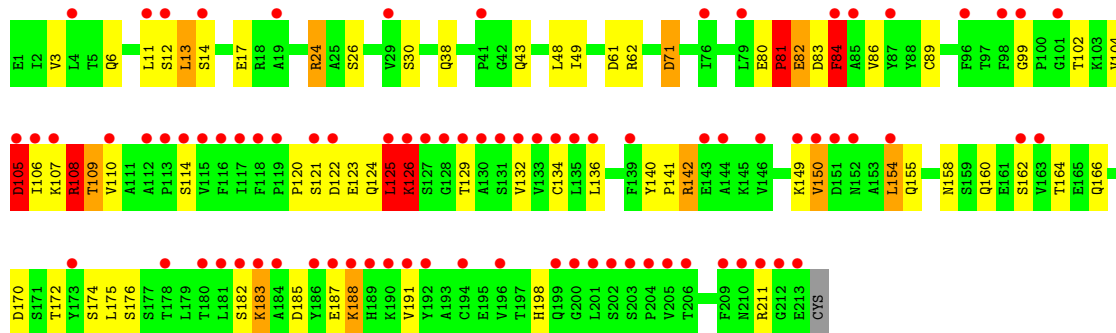
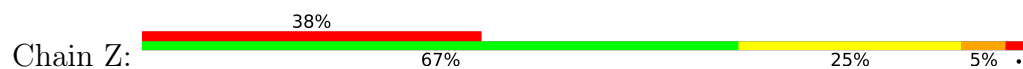
• Molecule 3: Fab light chain



• Molecule 3: Fab light chain



• Molecule 3: Fab light chain



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	92.78Å 170.43Å 148.89Å 90.00° 105.81° 90.00°	Depositor
Resolution (Å)	44.63 – 2.40 143.26 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.4 (44.63-2.40) 99.4 (143.26-2.40)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.38 (at 2.40Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.220 , 0.260 0.214 , 0.253	Depositor DCC
R_{free} test set	8731 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å ²)	45.8	Xtriage
Anisotropy	0.246	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 41.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.012 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	24458	wwPDB-VP
Average B, all atoms (Å ²)	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.59% 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

5.1 Standard geometry

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

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	0.56	0/761	0.73	0/1040
1	B	0.55	0/767	0.73	0/1048
1	C	0.53	0/755	0.87	3/1032 (0.3%)
1	D	0.46	0/755	0.84	1/1032 (0.1%)
1	E	0.48	0/761	0.70	0/1040
1	F	0.47	0/761	0.66	0/1040
2	H	0.57	1/1659 (0.1%)	0.72	3/2260 (0.1%)
2	I	0.58	0/1647	0.87	5/2244 (0.2%)
2	J	0.56	1/1663 (0.1%)	0.69	2/2265 (0.1%)
2	U	0.49	1/1651 (0.1%)	0.70	0/2249
2	W	0.47	0/1674	0.78	3/2279 (0.1%)
2	Y	0.57	0/1682	1.01	10/2289 (0.4%)
3	L	0.51	0/1666	0.96	11/2262 (0.5%)
3	M	0.53	0/1666	1.05	11/2262 (0.5%)
3	N	0.51	1/1666 (0.1%)	0.88	10/2262 (0.4%)
3	V	0.61	3/1666 (0.2%)	1.04	15/2262 (0.7%)
3	X	0.52	0/1666	0.91	8/2262 (0.4%)
3	Z	0.53	2/1666 (0.1%)	1.26	19/2262 (0.8%)
All	All	0.53	9/24532 (0.0%)	0.89	101/33390 (0.3%)

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	C	0	3
1	D	0	2
2	I	0	2
2	Y	0	2
3	L	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
3	M	0	3
3	N	0	1
3	V	0	3
3	X	0	3
3	Z	0	4
All	All	0	24

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	V	154	LEU	CG-CD1	7.52	1.79	1.51
2	J	97	CYS	CB-SG	-6.28	1.71	1.82
3	V	142	ARG	CB-CG	-6.15	1.35	1.52
2	H	97	CYS	CB-SG	-5.92	1.72	1.81
3	N	199	GLN	CG-CD	5.43	1.63	1.51

The worst 5 of 101 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	Z	154	LEU	CB-CG-CD1	-28.41	62.71	111.00
3	Z	154	LEU	CB-CG-CD2	18.15	141.85	111.00
3	M	154	LEU	CB-CG-CD2	-17.89	80.59	111.00
2	Y	196	LEU	CB-CG-CD2	-17.42	81.39	111.00
3	V	154	LEU	CB-CG-CD2	15.10	136.67	111.00

There are no chirality outliers.

5 of 24 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	I	132	ALA	Peptide
2	I	2	GLU	Peptide
3	L	142	ARG	Sidechain
3	M	167	ASP	Sidechain
3	M	187	GLU	Peptide

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	A	741	0	688	12	0
1	B	747	0	693	13	0
1	C	735	0	684	12	0
1	D	735	0	684	19	0
1	E	741	0	688	18	0
1	F	741	0	687	14	0
2	H	1618	0	1562	16	0
2	I	1606	0	1552	14	0
2	J	1622	0	1565	7	0
2	U	1610	0	1555	23	0
2	W	1633	0	1580	24	0
2	Y	1641	0	1585	56	0
3	L	1630	0	1581	30	0
3	M	1630	0	1581	35	0
3	N	1630	0	1581	24	0
3	V	1630	0	1581	49	0
3	X	1630	0	1581	16	0
3	Z	1630	0	1581	60	0
4	A	14	0	13	0	0
4	B	14	0	13	0	0
4	E	14	0	13	0	0
4	F	28	0	26	0	0
5	B	4	0	6	7	0
5	E	4	0	6	5	0
5	L	4	0	6	1	0
5	U	4	0	6	3	0
6	A	28	0	0	0	0
6	B	28	0	0	1	0
6	C	9	0	0	1	0
6	D	6	0	0	0	0
6	E	22	0	0	1	0
6	F	14	0	0	1	0
6	H	58	0	0	0	0
6	I	54	0	0	0	0
6	J	52	0	0	2	0
6	L	22	0	0	1	0
6	M	13	0	0	1	0
6	N	16	0	0	1	0
6	U	31	0	0	0	0
6	V	11	0	0	0	0
6	W	30	0	0	1	0
6	X	12	0	0	1	0
6	Y	16	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	24458	0	23098	419	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 419 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:V:154:LEU:CG	3:V:154:LEU:CD1	1.79	1.59
3:V:142:ARG:HH22	3:V:175:LEU:HD12	1.13	1.09
5:U:301:EDO:H21	2:Y:32:THR:HG21	1.37	1.04
2:J:27:GLY:H	1:E:278:GLN:HE22	1.09	0.99
3:Z:108:ARG:HH12	3:Z:172:THR:HG23	1.28	0.97

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	92/103 (89%)	90 (98%)	2 (2%)	0	100	100
1	B	93/103 (90%)	92 (99%)	1 (1%)	0	100	100
1	C	91/103 (88%)	89 (98%)	2 (2%)	0	100	100
1	D	91/103 (88%)	89 (98%)	2 (2%)	0	100	100
1	E	92/103 (89%)	90 (98%)	2 (2%)	0	100	100
1	F	92/103 (89%)	90 (98%)	2 (2%)	0	100	100
2	H	209/230 (91%)	205 (98%)	4 (2%)	0	100	100
2	I	207/230 (90%)	200 (97%)	5 (2%)	2 (1%)	15	23
2	J	210/230 (91%)	204 (97%)	6 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	U	208/230 (90%)	204 (98%)	4 (2%)	0	100	100
2	W	211/230 (92%)	205 (97%)	6 (3%)	0	100	100
2	Y	213/230 (93%)	203 (95%)	9 (4%)	1 (0%)	29	41
3	L	211/214 (99%)	204 (97%)	7 (3%)	0	100	100
3	M	211/214 (99%)	200 (95%)	9 (4%)	2 (1%)	17	25
3	N	211/214 (99%)	203 (96%)	7 (3%)	1 (0%)	29	41
3	V	211/214 (99%)	200 (95%)	10 (5%)	1 (0%)	29	41
3	X	211/214 (99%)	201 (95%)	10 (5%)	0	100	100
3	Z	211/214 (99%)	201 (95%)	6 (3%)	4 (2%)	8	10
All	All	3075/3282 (94%)	2970 (97%)	94 (3%)	11 (0%)	34	48

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	I	129	PHE
2	I	132	ALA
3	M	203	SER
3	V	187	GLU
2	Y	112	GLN

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	84/91 (92%)	81 (96%)	3 (4%)	35	54
1	B	85/91 (93%)	81 (95%)	4 (5%)	26	42
1	C	83/91 (91%)	79 (95%)	4 (5%)	25	41
1	D	83/91 (91%)	75 (90%)	8 (10%)	8	12
1	E	84/91 (92%)	79 (94%)	5 (6%)	19	31
1	F	84/91 (92%)	82 (98%)	2 (2%)	49	68
2	H	179/192 (93%)	175 (98%)	4 (2%)	52	71

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	I	177/192 (92%)	168 (95%)	9 (5%)	24	39
2	J	179/192 (93%)	174 (97%)	5 (3%)	43	63
2	U	177/192 (92%)	169 (96%)	8 (4%)	27	44
2	W	181/192 (94%)	174 (96%)	7 (4%)	32	50
2	Y	181/192 (94%)	170 (94%)	11 (6%)	18	30
3	L	184/185 (100%)	175 (95%)	9 (5%)	25	40
3	M	184/185 (100%)	170 (92%)	14 (8%)	13	20
3	N	184/185 (100%)	181 (98%)	3 (2%)	62	79
3	V	184/185 (100%)	170 (92%)	14 (8%)	13	20
3	X	184/185 (100%)	177 (96%)	7 (4%)	33	51
3	Z	184/185 (100%)	169 (92%)	15 (8%)	11	17
All	All	2681/2808 (96%)	2549 (95%)	132 (5%)	25	40

5 of 132 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	Y	204	ASN
3	Z	43	GLN
3	Z	162	SER
2	J	157	VAL
2	J	13	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
2	U	178	GLN
3	V	138	ASN
3	Z	160	GLN
3	V	137	ASN
1	E	278	GLN

5.3.3 RNA ⓘ

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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

9 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	EDO	L	301	-	3,3,3	0.64	0	2,2,2	0.10	0
5	EDO	E	402	-	3,3,3	0.57	0	2,2,2	1.09	0
4	NAG	E	401	1	14,14,15	0.83	1 (7%)	17,19,21	0.76	0
4	NAG	B	401	1	14,14,15	0.37	0	17,19,21	0.71	0
4	NAG	F	402	1	14,14,15	0.57	0	17,19,21	0.72	0
5	EDO	B	402	-	3,3,3	0.62	0	2,2,2	0.49	0
4	NAG	F	401	1	14,14,15	0.39	0	17,19,21	0.46	0
4	NAG	A	401	1	14,14,15	0.84	1 (7%)	17,19,21	0.91	1 (5%)
5	EDO	U	301	-	3,3,3	0.75	0	2,2,2	0.40	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
5	EDO	L	301	-	-	0/1/1/1	-
5	EDO	E	402	-	-	1/1/1/1	-
4	NAG	E	401	1	-	0/6/23/26	0/1/1/1
4	NAG	B	401	1	-	0/6/23/26	0/1/1/1
4	NAG	F	402	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
5	EDO	B	402	-	-	0/1/1/1	-
4	NAG	F	401	1	-	1/6/23/26	0/1/1/1
4	NAG	A	401	1	-	0/6/23/26	0/1/1/1
5	EDO	U	301	-	-	1/1/1/1	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	E	401	NAG	O5-C1	-2.72	1.39	1.43
4	A	401	NAG	O5-C1	2.27	1.47	1.43

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	401	NAG	C1-O5-C5	2.06	114.99	112.19

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	U	301	EDO	O1-C1-C2-O2
5	E	402	EDO	O1-C1-C2-O2
4	F	401	NAG	C4-C5-C6-O6

There are no ring outliers.

4 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	L	301	EDO	1	0
5	E	402	EDO	5	0
5	B	402	EDO	7	0
5	U	301	EDO	3	0

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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	A	94/103 (91%)	0.28	0 100 100	25, 37, 70, 90	0
1	B	95/103 (92%)	0.26	0 100 100	24, 37, 68, 94	0
1	C	93/103 (90%)	0.49	4 (4%) 35 33	33, 62, 88, 109	0
1	D	93/103 (90%)	1.03	14 (15%) 2 1	37, 74, 109, 121	0
1	E	94/103 (91%)	0.30	1 (1%) 80 79	30, 49, 84, 109	0
1	F	94/103 (91%)	0.22	1 (1%) 80 79	29, 49, 92, 107	0
2	H	213/230 (92%)	0.32	0 100 100	24, 45, 90, 117	0
2	I	211/230 (91%)	0.34	3 (1%) 75 73	22, 40, 85, 96	0
2	J	214/230 (93%)	0.19	1 (0%) 91 89	26, 42, 81, 97	0
2	U	212/230 (92%)	0.35	4 (1%) 66 64	27, 49, 91, 97	0
2	W	215/230 (93%)	0.23	3 (1%) 75 73	28, 53, 97, 113	0
2	Y	217/230 (94%)	2.04	84 (38%) 0 0	32, 79, 140, 155	0
3	L	213/214 (99%)	0.23	4 (1%) 66 64	30, 61, 85, 102	0
3	M	213/214 (99%)	0.58	16 (7%) 14 13	29, 74, 108, 132	0
3	N	213/214 (99%)	0.25	6 (2%) 53 51	34, 61, 85, 101	0
3	V	213/214 (99%)	1.04	35 (16%) 1 1	38, 77, 120, 133	0
3	X	213/214 (99%)	0.37	1 (0%) 91 89	41, 59, 81, 103	0
3	Z	213/214 (99%)	1.80	82 (38%) 0 0	54, 102, 143, 162	0
All	All	3123/3282 (95%)	0.61	259 (8%) 11 10	22, 59, 117, 162	0

The worst 5 of 259 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	Y	157	VAL	11.4
2	Y	188	VAL	9.9
3	Z	209	PHE	9.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	Y	140	GLY	8.9
3	V	148	TRP	8.9

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(Å ²)	Q<0.9
5	EDO	U	301	4/4	0.72	0.47	42,42,45,53	0
5	EDO	L	301	4/4	0.82	0.20	46,47,49,51	0
5	EDO	E	402	4/4	0.84	0.34	32,33,34,34	0
4	NAG	F	401	14/15	0.86	0.22	89,96,101,101	0
4	NAG	E	401	14/15	0.92	0.21	57,60,65,65	0
4	NAG	A	401	14/15	0.93	0.16	35,38,45,48	0
4	NAG	B	401	14/15	0.93	0.14	31,39,42,47	0
4	NAG	F	402	14/15	0.93	0.18	59,63,69,73	0
5	EDO	B	402	4/4	0.95	0.34	28,29,31,31	0

6.5 Other polymers [i](#)

There are no such residues in this entry.