wwPDB X-ray Structure Validation Summary Report (i)

Jan 18, 2024 – 12:10 PM JST

PDB ID	:	8JNK
Title	:	Crystal structure of human ALKBH3 bound to ssDNA through active site crosslink
Authors	:	Zhang, L.
Deposited on	:	2023-06-06
Resolution	:	2.69 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:

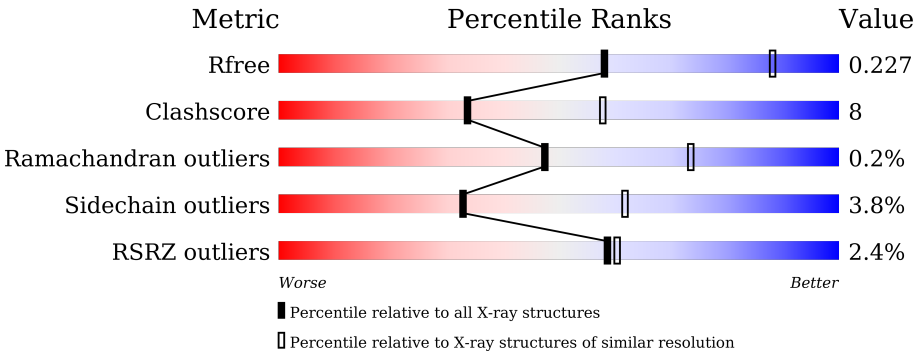
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 2.69 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.



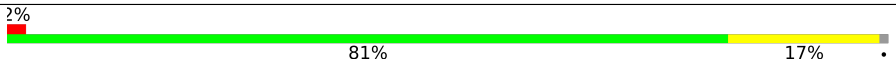
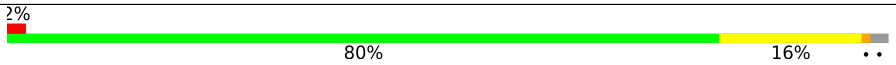
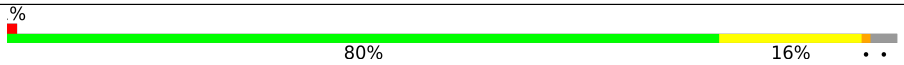
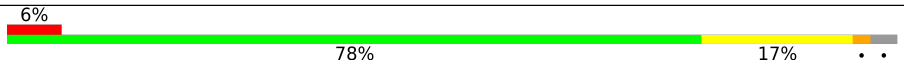
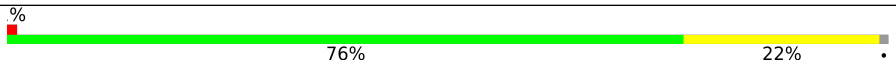
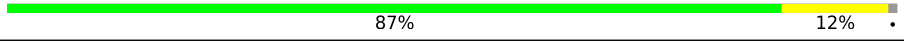
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	232	2% 65% 21% 12%
1	C	232	5% 63% 22% 12%
1	E	232	1% 66% 20% 12%
1	G	232	6% 63% 21% 13%
2	B	6	50% 50%
2	D	6	17% 33% 33% 17%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	6	
2	H	6	
3	I	216	
3	J	216	
3	K	216	
3	L	216	
4	M	217	
4	N	217	
4	O	217	
4	P	217	

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 20379 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 Alpha-ketoglutarate-dependent dioxygenase alkB homolog 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	203	Total	C	N	O	S	0	0	0
			1674	1061	299	308	6			
1	C	205	Total	C	N	O	S	1	0	0
			1694	1073	299	316	6			
1	E	203	Total	C	N	O	S	1	0	0
			1679	1065	297	311	6			
1	G	201	Total	C	N	O	S	2	0	0
			1654	1050	294	304	6			

There are 96 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	49	MET	-	initiating methionine	UNP Q96Q83
A	50	GLY	-	expression tag	UNP Q96Q83
A	51	SER	-	expression tag	UNP Q96Q83
A	52	SER	-	expression tag	UNP Q96Q83
A	53	HIS	-	expression tag	UNP Q96Q83
A	54	HIS	-	expression tag	UNP Q96Q83
A	55	HIS	-	expression tag	UNP Q96Q83
A	56	HIS	-	expression tag	UNP Q96Q83
A	57	HIS	-	expression tag	UNP Q96Q83
A	58	HIS	-	expression tag	UNP Q96Q83
A	59	SER	-	expression tag	UNP Q96Q83
A	60	SER	-	expression tag	UNP Q96Q83
A	61	GLY	-	expression tag	UNP Q96Q83
A	62	LEU	-	expression tag	UNP Q96Q83
A	63	VAL	-	expression tag	UNP Q96Q83
A	64	PRO	-	expression tag	UNP Q96Q83
A	65	ARG	-	expression tag	UNP Q96Q83
A	66	GLY	-	expression tag	UNP Q96Q83
A	67	SER	-	expression tag	UNP Q96Q83
A	68	HIS	-	expression tag	UNP Q96Q83
A	69	MET	-	expression tag	UNP Q96Q83

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	110	SER	CYS	engineered mutation	UNP Q96Q83
A	195	CYS	GLU	conflict	UNP Q96Q83
A	201	SER	CYS	engineered mutation	UNP Q96Q83
C	49	MET	-	initiating methionine	UNP Q96Q83
C	50	GLY	-	expression tag	UNP Q96Q83
C	51	SER	-	expression tag	UNP Q96Q83
C	52	SER	-	expression tag	UNP Q96Q83
C	53	HIS	-	expression tag	UNP Q96Q83
C	54	HIS	-	expression tag	UNP Q96Q83
C	55	HIS	-	expression tag	UNP Q96Q83
C	56	HIS	-	expression tag	UNP Q96Q83
C	57	HIS	-	expression tag	UNP Q96Q83
C	58	HIS	-	expression tag	UNP Q96Q83
C	59	SER	-	expression tag	UNP Q96Q83
C	60	SER	-	expression tag	UNP Q96Q83
C	61	GLY	-	expression tag	UNP Q96Q83
C	62	LEU	-	expression tag	UNP Q96Q83
C	63	VAL	-	expression tag	UNP Q96Q83
C	64	PRO	-	expression tag	UNP Q96Q83
C	65	ARG	-	expression tag	UNP Q96Q83
C	66	GLY	-	expression tag	UNP Q96Q83
C	67	SER	-	expression tag	UNP Q96Q83
C	68	HIS	-	expression tag	UNP Q96Q83
C	69	MET	-	expression tag	UNP Q96Q83
C	110	SER	CYS	engineered mutation	UNP Q96Q83
C	195	CYS	GLU	conflict	UNP Q96Q83
C	201	SER	CYS	engineered mutation	UNP Q96Q83
E	49	MET	-	initiating methionine	UNP Q96Q83
E	50	GLY	-	expression tag	UNP Q96Q83
E	51	SER	-	expression tag	UNP Q96Q83
E	52	SER	-	expression tag	UNP Q96Q83
E	53	HIS	-	expression tag	UNP Q96Q83
E	54	HIS	-	expression tag	UNP Q96Q83
E	55	HIS	-	expression tag	UNP Q96Q83
E	56	HIS	-	expression tag	UNP Q96Q83
E	57	HIS	-	expression tag	UNP Q96Q83
E	58	HIS	-	expression tag	UNP Q96Q83
E	59	SER	-	expression tag	UNP Q96Q83
E	60	SER	-	expression tag	UNP Q96Q83
E	61	GLY	-	expression tag	UNP Q96Q83
E	62	LEU	-	expression tag	UNP Q96Q83
E	63	VAL	-	expression tag	UNP Q96Q83

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	64	PRO	-	expression tag	UNP Q96Q83
E	65	ARG	-	expression tag	UNP Q96Q83
E	66	GLY	-	expression tag	UNP Q96Q83
E	67	SER	-	expression tag	UNP Q96Q83
E	68	HIS	-	expression tag	UNP Q96Q83
E	69	MET	-	expression tag	UNP Q96Q83
E	110	SER	CYS	engineered mutation	UNP Q96Q83
E	195	CYS	GLU	conflict	UNP Q96Q83
E	201	SER	CYS	engineered mutation	UNP Q96Q83
G	49	MET	-	initiating methionine	UNP Q96Q83
G	50	GLY	-	expression tag	UNP Q96Q83
G	51	SER	-	expression tag	UNP Q96Q83
G	52	SER	-	expression tag	UNP Q96Q83
G	53	HIS	-	expression tag	UNP Q96Q83
G	54	HIS	-	expression tag	UNP Q96Q83
G	55	HIS	-	expression tag	UNP Q96Q83
G	56	HIS	-	expression tag	UNP Q96Q83
G	57	HIS	-	expression tag	UNP Q96Q83
G	58	HIS	-	expression tag	UNP Q96Q83
G	59	SER	-	expression tag	UNP Q96Q83
G	60	SER	-	expression tag	UNP Q96Q83
G	61	GLY	-	expression tag	UNP Q96Q83
G	62	LEU	-	expression tag	UNP Q96Q83
G	63	VAL	-	expression tag	UNP Q96Q83
G	64	PRO	-	expression tag	UNP Q96Q83
G	65	ARG	-	expression tag	UNP Q96Q83
G	66	GLY	-	expression tag	UNP Q96Q83
G	67	SER	-	expression tag	UNP Q96Q83
G	68	HIS	-	expression tag	UNP Q96Q83
G	69	MET	-	expression tag	UNP Q96Q83
G	110	SER	CYS	engineered mutation	UNP Q96Q83
G	195	CYS	GLU	conflict	UNP Q96Q83
G	201	SER	CYS	engineered mutation	UNP Q96Q83

- Molecule 2 is a DNA chain called DNA.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	6	Total	C	N	O	P	S	0	0	0
			124	60	20	37	6	1			
2	D	5	Total	C	N	O	P	S	0	0	0
			104	50	18	30	5	1			
2	F	6	Total	C	N	O	P	S	0	0	0
			124	60	20	37	6	1			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	H	5	Total	C	N	O	P	S	0	0	0
			104	50	18	30	5	1			

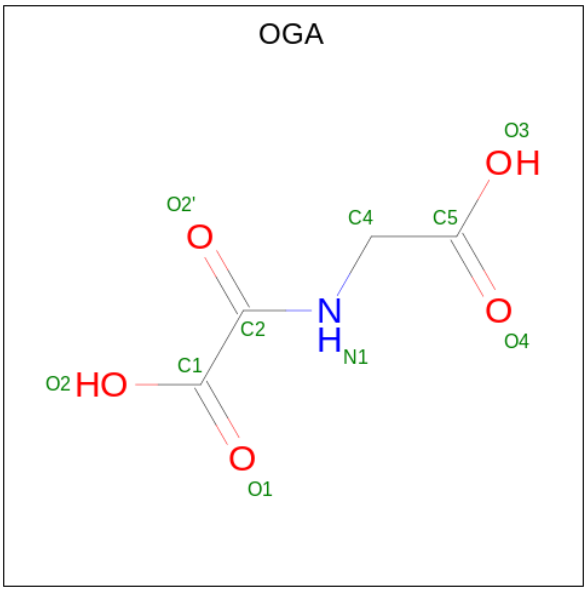
- Molecule 3 is a protein called Synthetic antibody heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	I	214	Total	C	N	O	S	0	0	0
			1601	1014	264	317	6			
3	J	212	Total	C	N	O	S	1	0	0
			1590	1008	262	314	6			
3	K	210	Total	C	N	O	S	0	0	0
			1572	997	259	310	6			
3	L	209	Total	C	N	O	S	0	0	0
			1566	994	258	308	6			

- Molecule 4 is a protein called Synthetic antibody light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	M	214	Total	C	N	O	S	1	0	0
			1643	1032	274	332	5			
4	N	214	Total	C	N	O	S	1	0	0
			1643	1032	274	332	5			
4	O	214	Total	C	N	O	S	2	0	0
			1643	1032	274	332	5			
4	P	214	Total	C	N	O	S	2	0	0
			1643	1032	274	332	5			

- Molecule 5 is N-OXALYLGLYCINE (three-letter code: OGA) (formula: $C_4H_5NO_5$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C N O 10 4 1 5	0	0
5	C	1	Total C N O 10 4 1 5	0	0
5	E	1	Total C N O 10 4 1 5	0	0
5	G	1	Total C N O 10 4 1 5	0	0

- Molecule 6 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total Mn 1 1	0	0
6	C	1	Total Mn 1 1	0	0
6	E	1	Total Mn 1 1	0	0
6	G	1	Total Mn 1 1	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	34	Total O 34 34	0	0

Continued on next page...

Continued from previous page...

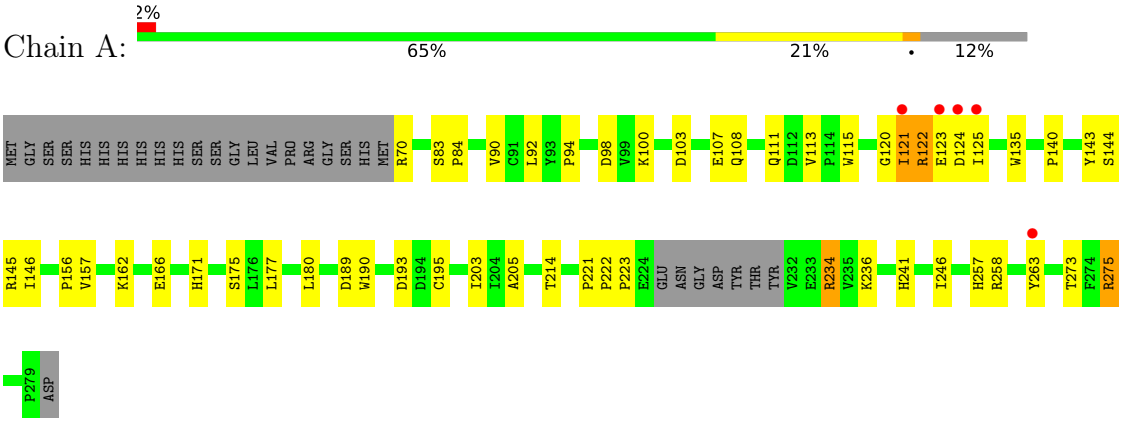
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	B	2	Total O 2 2	0	0
7	C	4	Total O 4 4	0	0
7	E	26	Total O 26 26	0	0
7	F	2	Total O 2 2	0	0
7	G	12	Total O 12 12	0	0
7	H	1	Total O 1 1	0	0
7	I	41	Total O 41 41	0	0
7	J	30	Total O 30 30	0	0
7	K	24	Total O 24 24	0	0
7	L	7	Total O 7 7	0	0
7	M	26	Total O 26 26	0	0
7	N	20	Total O 20 20	0	0
7	O	30	Total O 30 30	0	0
7	P	18	Total O 18 18	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.

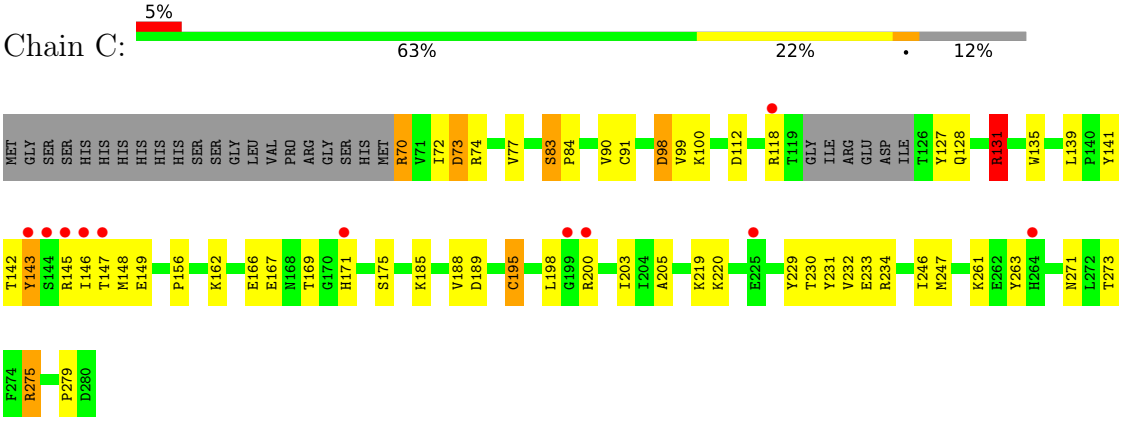
• Molecule 1: Alpha-ketoglutarate-dependent dioxygenase alkB homolog 3

Chain A:



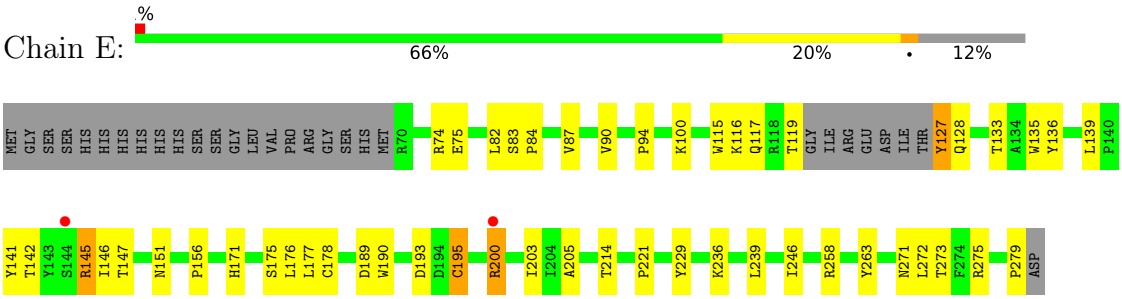
• Molecule 1: Alpha-ketoglutarate-dependent dioxygenase alkB homolog 3

Chain C:



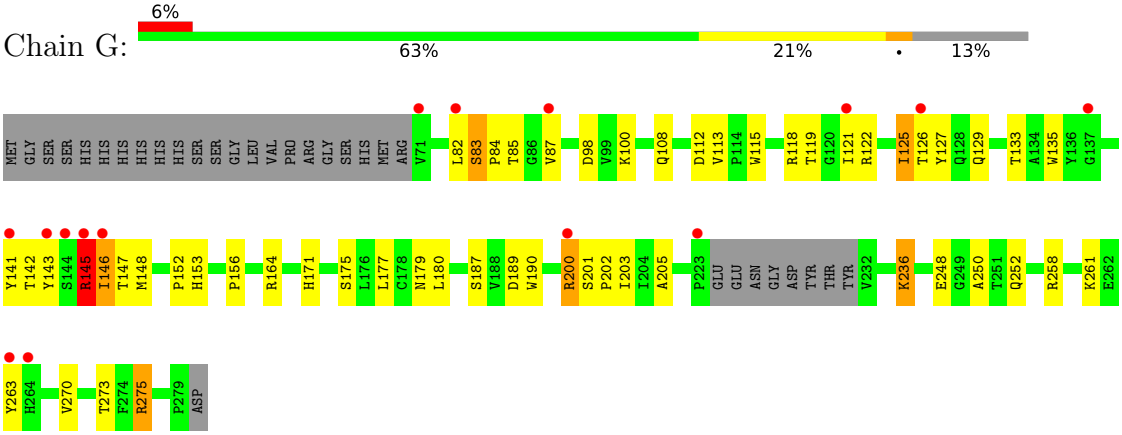
• Molecule 1: Alpha-ketoglutarate-dependent dioxygenase alkB homolog 3

Chain E:



• Molecule 1: Alpha-ketoglutarate-dependent dioxygenase alkB homolog 3

Chain G:



• Molecule 2: DNA

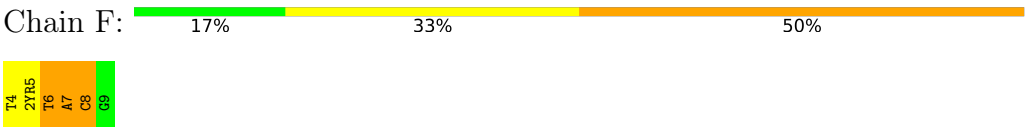
Chain B:

• Molecule 2: DNA

Chain D:

• Molecule 2: DNA

Chain F:

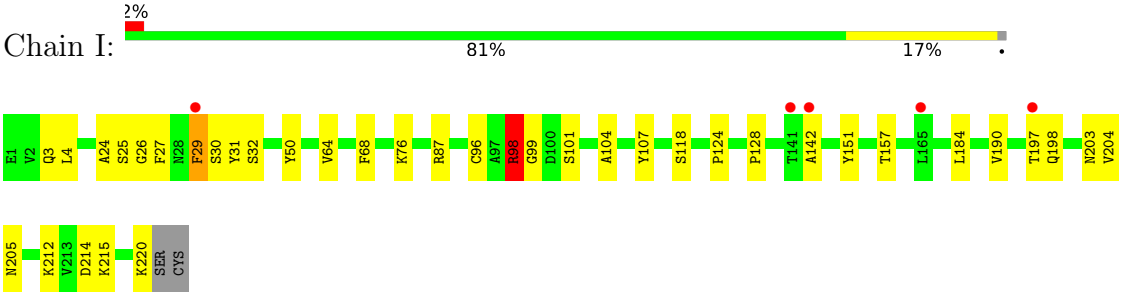


• Molecule 2: DNA

Chain H:

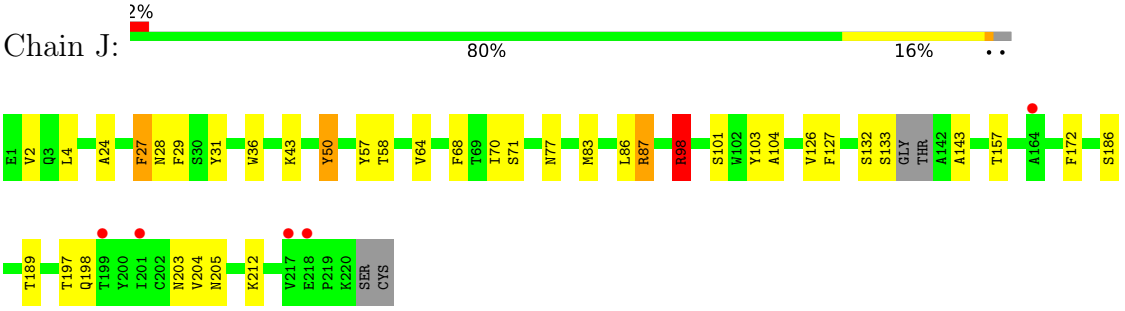
• Molecule 3: Synthetic antibody heavy chain

Chain I:



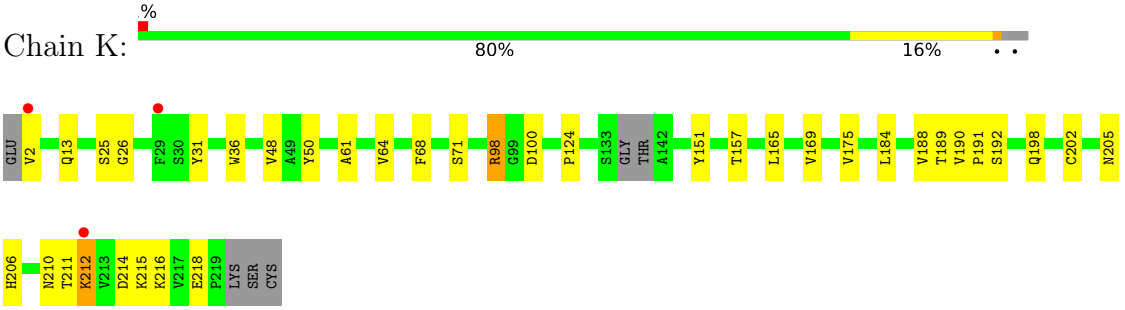
• Molecule 3: Synthetic antibody heavy chain

Chain J:



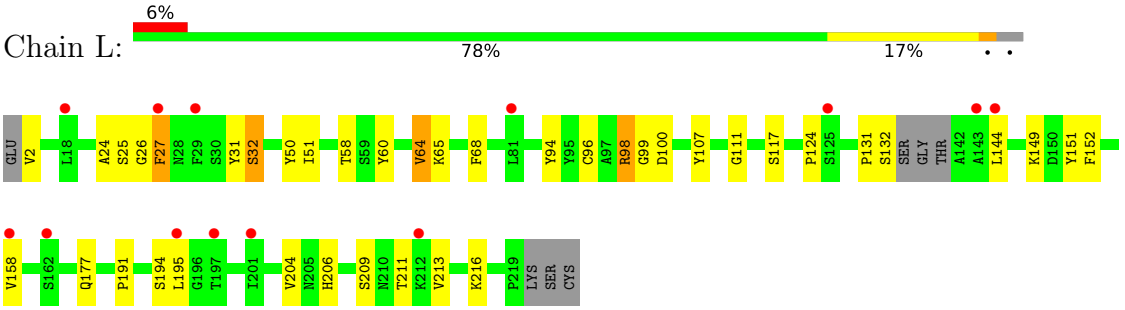
• Molecule 3: Synthetic antibody heavy chain

Chain K:



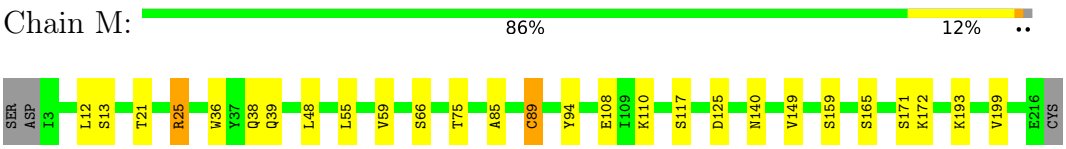
• Molecule 3: Synthetic antibody heavy chain

Chain L:



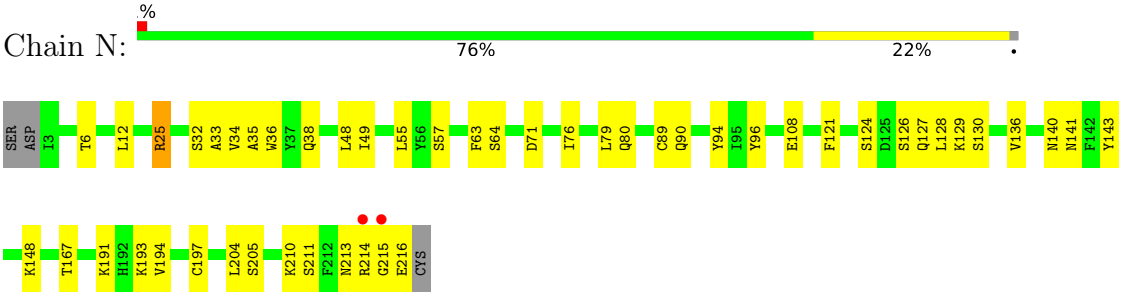
• Molecule 4: Synthetic antibody light chain

Chain M:



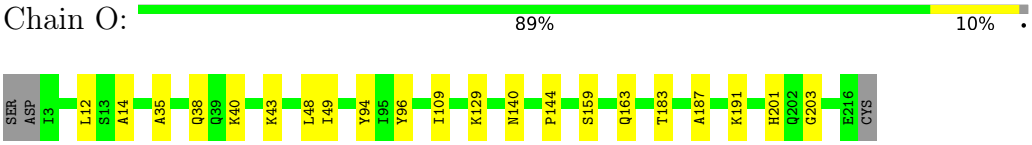
• Molecule 4: Synthetic antibody light chain

Chain N:



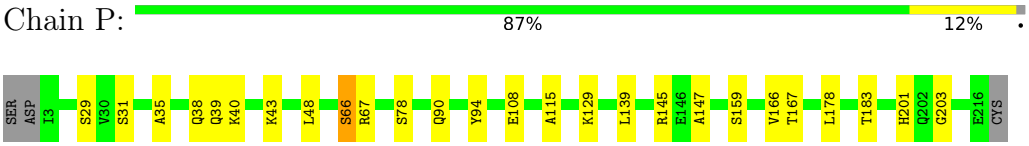
• Molecule 4: Synthetic antibody light chain

Chain O:



• Molecule 4: Synthetic antibody light chain

Chain P:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	129.46Å 136.15Å 203.39Å 90.00° 107.11° 90.00°	Depositor
Resolution (Å)	46.90 – 2.69 46.90 – 2.69	Depositor EDS
% Data completeness (in resolution range)	99.1 (46.90-2.69) 99.2 (46.90-2.69)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.32 (at 2.69Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, R_{free}	0.218 , 0.229 0.216 , 0.227	Depositor DCC
R_{free} test set	4687 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	52.2	Xtriage
Anisotropy	0.373	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 40.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.008 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	20379	wwPDB-VP
Average B, all atoms (Å^2)	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.33% 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: 2YR, MN, OGA

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	0.39	0/1722	0.74	0/2345
1	C	0.41	0/1744	0.82	4/2377 (0.2%)
1	E	0.35	0/1729	0.75	3/2356 (0.1%)
1	G	0.41	0/1702	0.77	0/2319
2	B	0.57	0/112	1.67	4/168 (2.4%)
2	D	0.54	0/91	1.54	3/138 (2.2%)
2	F	0.60	0/112	1.54	3/168 (1.8%)
2	H	0.61	0/91	1.56	3/138 (2.2%)
3	I	0.37	0/1643	0.62	2/2242 (0.1%)
3	J	0.57	0/1631	0.77	0/2224
3	K	0.47	0/1613	0.63	2/2201 (0.1%)
3	L	0.35	0/1607	0.57	0/2193
4	M	0.29	0/1680	0.52	0/2283
4	N	0.27	0/1680	0.50	0/2283
4	O	0.35	0/1680	0.52	0/2283
4	P	0.27	0/1680	0.54	0/2283
All	All	0.39	0/20517	0.69	24/28001 (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
1	C	0	2
1	G	0	2
3	I	0	1
3	J	0	2
All	All	0	8

There are no bond length outliers.

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	6	DT	P-O3'-C3'	-10.69	106.88	119.70
2	F	6	DT	P-O3'-C3'	-9.20	108.66	119.70
2	D	7	DA	P-O3'-C3'	-8.92	109.00	119.70
2	F	7	DA	P-O3'-C3'	-8.57	109.42	119.70
2	H	7	DA	P-O3'-C3'	-8.14	109.93	119.70

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	234	ARG	Sidechain
1	C	131	ARG	Sidechain
1	C	275	ARG	Sidechain
1	G	145	ARG	Sidechain
1	G	200	ARG	Sidechain

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	1674	0	1634	42	0
1	C	1694	0	1633	49	0
1	E	1679	0	1622	36	0
1	G	1654	0	1614	48	0
2	B	124	0	74	12	0
2	D	104	0	61	10	0
2	F	124	0	72	7	0
2	H	104	0	61	6	0
3	I	1601	0	1548	19	0
3	J	1590	0	1537	28	0
3	K	1572	0	1515	17	0
3	L	1566	0	1510	36	0
4	M	1643	0	1603	15	0
4	N	1643	0	1603	30	0
4	O	1643	0	1603	15	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	P	1643	0	1603	17	0
5	A	10	0	3	0	0
5	C	10	0	3	0	0
5	E	10	0	3	0	0
5	G	10	0	3	1	0
6	A	1	0	0	0	0
6	C	1	0	0	0	0
6	E	1	0	0	0	0
6	G	1	0	0	0	0
7	A	34	0	0	0	0
7	B	2	0	0	0	0
7	C	4	0	0	0	0
7	E	26	0	0	2	0
7	F	2	0	0	0	0
7	G	12	0	0	0	0
7	H	1	0	0	0	0
7	I	41	0	0	1	0
7	J	30	0	0	0	0
7	K	24	0	0	0	0
7	L	7	0	0	0	0
7	M	26	0	0	0	0
7	N	20	0	0	0	0
7	O	30	0	0	0	0
7	P	18	0	0	0	0
All	All	20379	0	19305	334	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 8.

The worst 5 of 334 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:141:TYR:CE1	2:D:5:2YR:H16	1.65	1.31
1:A:195:CYS:SG	2:B:5:2YR:S	2.35	1.25
1:C:141:TYR:CZ	2:D:5:2YR:H16	1.80	1.17
3:L:2:VAL:HG22	3:L:26:GLY:HA3	1.46	0.98
1:E:195:CYS:N	2:F:5:2YR:S	2.33	0.95

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	199/232 (86%)	196 (98%)	2 (1%)	1 (0%)	29 54
1	C	201/232 (87%)	194 (96%)	7 (4%)	0	100 100
1	E	199/232 (86%)	194 (98%)	5 (2%)	0	100 100
1	G	197/232 (85%)	195 (99%)	2 (1%)	0	100 100
3	I	212/216 (98%)	196 (92%)	15 (7%)	1 (0%)	29 54
3	J	208/216 (96%)	199 (96%)	9 (4%)	0	100 100
3	K	206/216 (95%)	193 (94%)	13 (6%)	0	100 100
3	L	205/216 (95%)	190 (93%)	14 (7%)	1 (0%)	29 54
4	M	212/217 (98%)	200 (94%)	12 (6%)	0	100 100
4	N	212/217 (98%)	200 (94%)	11 (5%)	1 (0%)	29 54
4	O	212/217 (98%)	206 (97%)	6 (3%)	0	100 100
4	P	212/217 (98%)	197 (93%)	15 (7%)	0	100 100
All	All	2475/2660 (93%)	2360 (95%)	111 (4%)	4 (0%)	47 73

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	I	26	GLY
4	N	33	ALA
1	A	140	PRO
3	L	64	VAL

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	187/212 (88%)	179 (96%)	8 (4%)	29	57
1	C	189/212 (89%)	181 (96%)	8 (4%)	30	58
1	E	187/212 (88%)	184 (98%)	3 (2%)	62	85
1	G	185/212 (87%)	177 (96%)	8 (4%)	29	57
3	I	177/179 (99%)	170 (96%)	7 (4%)	31	60
3	J	176/179 (98%)	169 (96%)	7 (4%)	31	60
3	K	174/179 (97%)	164 (94%)	10 (6%)	20	44
3	L	173/179 (97%)	165 (95%)	8 (5%)	27	54
4	M	189/192 (98%)	182 (96%)	7 (4%)	34	63
4	N	189/192 (98%)	179 (95%)	10 (5%)	22	48
4	O	189/192 (98%)	187 (99%)	2 (1%)	73	90
4	P	189/192 (98%)	184 (97%)	5 (3%)	46	75
All	All	2204/2332 (94%)	2121 (96%)	83 (4%)	33	62

5 of 83 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	L	144	LEU
4	N	148	LYS
4	M	25	ARG
4	M	193	LYS
4	N	214	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
3	L	177	GLN
4	O	201	HIS
1	C	226	ASN
1	E	129	GLN
3	I	3	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
5	OGA	A	1001	6	9,9,9	1.12	0	10,11,11	1.34	1 (10%)
5	OGA	E	1001	6	9,9,9	1.07	0	10,11,11	1.20	1 (10%)
5	OGA	G	1001	6	9,9,9	1.09	0	10,11,11	1.27	1 (10%)
5	OGA	C	1001	6	9,9,9	1.08	0	10,11,11	1.23	1 (10%)

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	OGA	A	1001	6	-	0/8/9/9	-
5	OGA	E	1001	6	-	2/8/9/9	-
5	OGA	G	1001	6	-	0/8/9/9	-
5	OGA	C	1001	6	-	0/8/9/9	-

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	1001	OGA	O2-C1-C2	2.59	120.78	113.15
5	C	1001	OGA	O2-C1-C2	2.51	120.55	113.15
5	E	1001	OGA	O2-C1-C2	2.21	119.66	113.15
5	G	1001	OGA	O2-C1-C2	2.17	119.55	113.15

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	E	1001	OGA	N1-C4-C5-O4
5	E	1001	OGA	N1-C4-C5-O3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	G	1001	OGA	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	203/232 (87%)	0.04	5 (2%) 57 59	33, 45, 73, 92	1 (0%)
1	C	205/232 (88%)	0.30	11 (5%) 25 24	45, 64, 88, 108	0
1	E	203/232 (87%)	-0.01	2 (0%) 82 83	35, 48, 71, 86	1 (0%)
1	G	201/232 (86%)	0.47	15 (7%) 14 12	41, 66, 94, 111	1 (0%)
2	B	5/6 (83%)	-0.00	0 100 100	66, 66, 73, 93	0
2	D	4/6 (66%)	0.25	0 100 100	76, 77, 78, 82	0
2	F	5/6 (83%)	0.08	0 100 100	70, 71, 74, 93	0
2	H	4/6 (66%)	0.23	0 100 100	70, 75, 80, 80	0
3	I	214/216 (99%)	0.18	5 (2%) 60 62	30, 47, 84, 101	0
3	J	212/216 (98%)	0.15	5 (2%) 59 60	27, 47, 80, 93	1 (0%)
3	K	210/216 (97%)	0.07	3 (1%) 75 77	36, 52, 76, 86	0
3	L	209/216 (96%)	0.47	13 (6%) 20 19	38, 68, 91, 104	0
4	M	214/217 (98%)	0.04	0 100 100	28, 51, 66, 75	1 (0%)
4	N	214/217 (98%)	0.07	2 (0%) 84 85	31, 52, 70, 94	1 (0%)
4	O	214/217 (98%)	-0.03	0 100 100	35, 51, 64, 81	2 (0%)
4	P	214/217 (98%)	0.03	0 100 100	42, 60, 76, 86	2 (0%)
All	All	2531/2684 (94%)	0.15	61 (2%) 59 60	27, 54, 84, 111	10 (0%)

The worst 5 of 61 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	L	29	PHE	4.8
1	C	146	ILE	4.6
1	G	143	TYR	4.5
1	G	146	ILE	4.2
3	I	29	PHE	4.2

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, 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
5	OGA	G	1001	10/10	0.83	0.50	52,68,79,86	0
5	OGA	E	1001	10/10	0.88	0.27	44,55,58,69	0
5	OGA	A	1001	10/10	0.88	0.34	44,55,77,86	0
5	OGA	C	1001	10/10	0.91	0.35	54,66,79,83	0
6	MN	A	1002	1/1	0.98	0.14	46,46,46,46	0
6	MN	C	1002	1/1	0.98	0.11	61,61,61,61	0
6	MN	E	1002	1/1	0.99	0.13	44,44,44,44	0
6	MN	G	1002	1/1	0.99	0.10	68,68,68,68	0

6.5 Other polymers (i)

There are no such residues in this entry.