



Full wwPDB X-ray Structure Validation Report ⓘ

Sep 6, 2022 – 01:20 PM EDT

PDB ID : 7SAI
Title : Crystal Structure of Lag30 Nanobody bound to eGFP
Authors : Cong, A.T.Q.; Schellenberg, M.J.
Deposited on : 2021-09-22
Resolution : 2.23 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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	:	2.29
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)	:	2.29

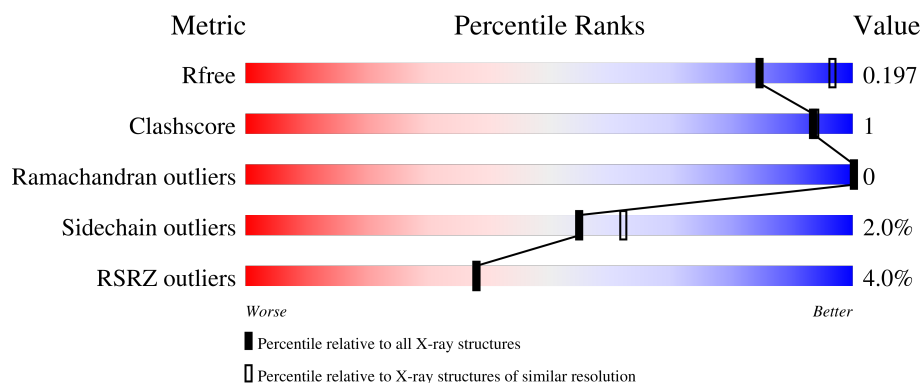
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.23 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	2391 (2.26-2.22)
Clashscore	141614	2539 (2.26-2.22)
Ramachandran outliers	138981	2489 (2.26-2.22)
Sidechain outliers	138945	2490 (2.26-2.22)
RSRZ outliers	127900	2353 (2.26-2.22)

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	237	<div> <div>5%</div> <div> <div></div> <div>90%</div> <div>6%</div> <div>.</div> </div> </div>
2	C	147	<div> <div>%</div> <div> <div></div> <div>82%</div> <div>14%</div> <div>.</div> </div> </div>

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
4	MES	A	303	-	-	-	X

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 3054 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 Green fluorescent protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	227	Total	C	N	O	S	0	2	0
			1835	1167	308	354	6			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	expression tag	UNP P42212
A	1	VAL	-	expression tag	UNP P42212
A	64	LEU	PHE	conflict	UNP P42212
A	66	CRO	SER	chromophore	UNP P42212
A	66	CRO	TYR	chromophore	UNP P42212
A	66	CRO	GLY	chromophore	UNP P42212
A	231	LEU	HIS	conflict	UNP P42212

- Molecule 2 is a protein called LAG30 Nanobody.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	126	Total	C	N	O	S	0	1	0
			972	599	176	192	5			

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



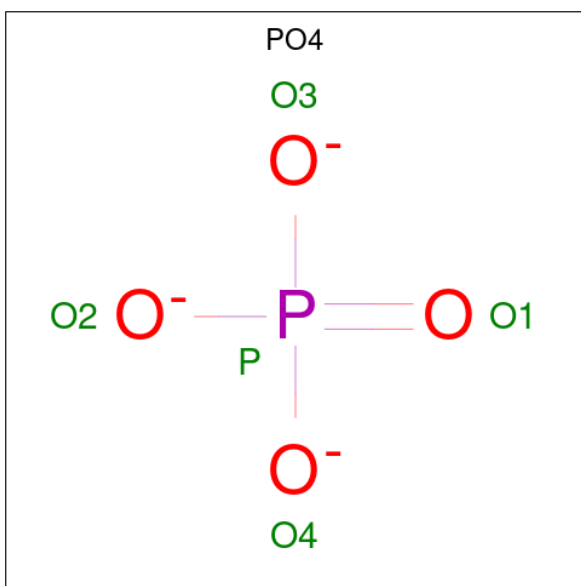
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		
3	A	1	Total	C	O	0	0
			6	3	3		
3	C	1	Total	C	O	0	0
			6	3	3		
3	C	1	Total	C	O	0	0
			6	3	3		
3	C	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: $C_6H_{13}NO_4S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

- Molecule 5 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	O	P	0	0
			5	4	1		
5	A	1	Total	O	P	0	0
			5	4	1		
5	C	1	Total	O	P	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	C	1	Total	O	P	0	0
			5	4	1		

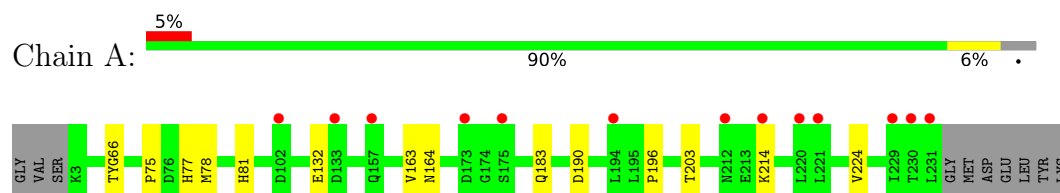
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	97	Total	O	0	0
			97	97		
6	C	88	Total	O	0	0
			88	88		

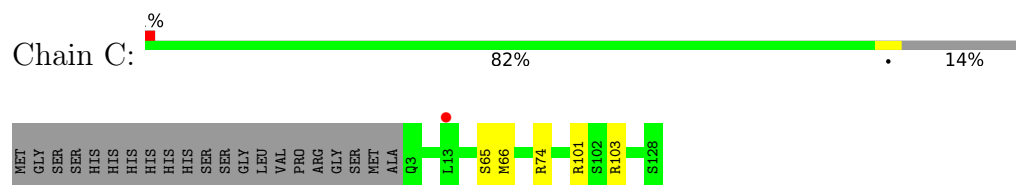
3 Residue-property plots [i](#)

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: Green fluorescent protein



- Molecule 2: LAG30 Nanobody



4 Data and refinement statistics

Property	Value	Source
Space group	I 4 2 2	Depositor
Cell constants a, b, c, α , β , γ	173.17Å 173.17Å 92.29Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.94 – 2.23 48.94 – 2.23	Depositor EDS
% Data completeness (in resolution range)	99.2 (48.94-2.23) 90.7 (48.94-2.23)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.22 (at 2.22Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.171 , 0.197 0.171 , 0.197	Depositor DCC
R_{free} test set	1714 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	45.1	Xtriage
Anisotropy	0.237	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 62.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	3054	wwPDB-VP
Average B, all atoms (Å ²)	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.64% 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: CRO, GOL, PO4, MES

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	0.45	0/1853	0.56	0/2504
2	C	0.51	0/987	0.60	0/1329
All	All	0.47	0/2840	0.57	0/3833

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1835	0	1778	6	0
2	C	972	0	947	2	0
3	A	12	0	16	1	0
3	C	18	0	24	1	0
4	A	12	0	13	2	0
5	A	10	0	0	0	0
5	C	10	0	0	0	0
6	A	97	0	0	0	0
6	C	88	0	0	0	0
All	All	3054	0	2778	8	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 1.

All (8) 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)
4:A:303:MES:H61	2:C:103:ARG:HD3	1.82	0.59
1:A:203:THR:HG23	1:A:224:VAL:HG22	1.88	0.55
1:A:190:ASP:CG	4:A:303:MES:H51	2.28	0.54
1:A:75:PRO:HB3	3:A:301:GOL:H32	1.93	0.50
1:A:163:VAL:HB	1:A:183:GLN:HB3	1.96	0.47
1:A:77:HIS:CD2	1:A:78:MET:HG3	2.54	0.43
1:A:81:HIS:O	1:A:196:PRO:HB3	2.19	0.42
2:C:101:ARG:HB2	3:C:201:GOL:H12	2.00	0.42

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	224/237 (94%)	220 (98%)	4 (2%)	0	100	100
2	C	125/147 (85%)	123 (98%)	2 (2%)	0	100	100
All	All	349/384 (91%)	343 (98%)	6 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	200/206 (97%)	197 (98%)	3 (2%)	65	72
2	C	104/120 (87%)	101 (97%)	3 (3%)	42	48
All	All	304/326 (93%)	298 (98%)	6 (2%)	55	62

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	132	GLU
1	A	164	ASN
1	A	214	LYS
2	C	65	SER
2	C	66	MET
2	C	74	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	C	5	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	$\# Z > 2$
1	CRO	A	66	1	23,23,24	2.77	6 (26%)	30,32,34	2.12	8 (26%)

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
1	CRO	A	66	1	-	2/12/31/32	0/2/2/2

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	66	CRO	C1-N2	7.47	1.43	1.32
1	A	66	CRO	CA2-C2	6.32	1.54	1.48
1	A	66	CRO	C1-N3	5.53	1.46	1.37
1	A	66	CRO	C2-N3	3.64	1.48	1.39
1	A	66	CRO	CG2-CB2	3.28	1.53	1.46
1	A	66	CRO	CA2-N2	2.87	1.44	1.38

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	66	CRO	O2-C2-CA2	-6.41	127.36	130.96
1	A	66	CRO	CA2-C2-N3	5.85	106.14	103.37
1	A	66	CRO	CA2-N2-C1	3.54	108.38	105.77
1	A	66	CRO	O3-C3-CA3	-2.81	117.89	126.39
1	A	66	CRO	C2-N3-C1	-2.72	106.59	107.97
1	A	66	CRO	CA1-C1-N3	-2.44	121.82	124.75
1	A	66	CRO	C2-CA2-N2	-2.18	107.41	108.93
1	A	66	CRO	CA1-C1-N2	2.01	126.70	123.89

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	66	CRO	N2-CA2-CB2-CG2
1	A	66	CRO	C2-CA2-CB2-CG2

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry

10 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
3	GOL	C	201	-	5,5,5	1.09	0	5,5,5	0.76	0
3	GOL	C	202	-	5,5,5	0.97	0	5,5,5	1.08	0
3	GOL	C	203	-	5,5,5	0.96	0	5,5,5	0.99	0
5	PO4	A	305	-	4,4,4	1.76	1 (25%)	6,6,6	0.81	0
5	PO4	C	204	-	4,4,4	0.94	0	6,6,6	0.43	0
3	GOL	A	301	-	5,5,5	1.07	0	5,5,5	1.06	0
4	MES	A	303	-	12,12,12	2.16	1 (8%)	14,16,16	1.71	4 (28%)
3	GOL	A	302	-	5,5,5	0.89	0	5,5,5	1.05	0
5	PO4	C	205	-	4,4,4	1.24	0	6,6,6	0.72	0
5	PO4	A	304	-	4,4,4	1.50	0	6,6,6	0.45	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
3	GOL	C	201	-	-	2/4/4/4	-
3	GOL	C	202	-	-	0/4/4/4	-
3	GOL	C	203	-	-	2/4/4/4	-
3	GOL	A	301	-	-	2/4/4/4	-
4	MES	A	303	-	-	6/6/14/14	0/1/1/1
3	GOL	A	302	-	-	2/4/4/4	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	303	MES	C8-S	-7.42	1.67	1.77
5	A	305	PO4	P-O4	-2.47	1.47	1.54

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	303	MES	O1S-S-C8	3.64	111.30	106.92
4	A	303	MES	O2S-S-C8	3.06	110.59	106.92
4	A	303	MES	O3S-S-C8	2.23	109.38	105.77
4	A	303	MES	O2S-S-O1S	-2.04	106.89	113.95

There are no chirality outliers.

All (14) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	302	GOL	C1-C2-C3-O3
3	A	302	GOL	O2-C2-C3-O3
4	A	303	MES	C7-C8-S-O1S
4	A	303	MES	C7-C8-S-O2S
4	A	303	MES	C7-C8-S-O3S
3	A	301	GOL	O1-C1-C2-C3
4	A	303	MES	N4-C7-C8-S
3	A	301	GOL	O1-C1-C2-O2
3	C	203	GOL	O2-C2-C3-O3
3	C	201	GOL	O2-C2-C3-O3
4	A	303	MES	C8-C7-N4-C3
4	A	303	MES	C8-C7-N4-C5
3	C	201	GOL	C1-C2-C3-O3
3	C	203	GOL	C1-C2-C3-O3

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	201	GOL	1	0
3	A	301	GOL	1	0
4	A	303	MES	2	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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	226/237 (95%)	0.21	13 (5%) 23 22	41, 62, 111, 205	0
2	C	126/147 (85%)	-0.28	1 (0%) 86 86	39, 49, 84, 121	0
All	All	352/384 (91%)	0.03	14 (3%) 38 38	39, 58, 105, 205	0

All (14) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	231	LEU	8.5
1	A	230	THR	8.2
1	A	229	ILE	4.5
1	A	157	GLN	3.9
1	A	133	ASP	3.2
1	A	214	LYS	2.9
1	A	220	LEU	2.9
2	C	13	LEU	2.6
1	A	175	SER	2.5
1	A	221	LEU	2.3
1	A	173	ASP	2.3
1	A	212	ASN	2.3
1	A	102	ASP	2.2
1	A	194	LEU	2.1

6.2 Non-standard residues in protein, DNA, RNA chains [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(\AA^2)	Q<0.9
1	CRO	A	66	22/23	0.97	0.18	38,50,55,57	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, 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(\AA^2)	Q<0.9
3	GOL	A	301	6/6	0.58	0.22	96,99,100,101	0
4	MES	A	303	12/12	0.75	0.44	153,156,161,163	0
5	PO4	C	205	5/5	0.76	0.31	175,176,176,178	0
5	PO4	A	305	5/5	0.82	0.21	119,122,123,127	0
3	GOL	C	203	6/6	0.82	0.44	97,101,103,103	0
5	PO4	C	204	5/5	0.86	0.41	172,173,174,174	0
3	GOL	C	202	6/6	0.86	0.20	69,85,93,101	0
3	GOL	A	302	6/6	0.92	0.15	94,99,101,104	0
5	PO4	A	304	5/5	0.95	0.23	91,91,91,93	5
3	GOL	C	201	6/6	0.96	0.12	59,62,63,67	0

6.5 Other polymers [i](#)

There are no such residues in this entry.