Full wwPDB X-ray Structure Validation Report ⓘ

Aug 25, 2021 – 02:02 pm BST

PDB ID	:	7PHY
Title	:	Vaccinia virus E2
Authors	:	Gao, W.N.D.; Gao, C.; Graham, S.C.
Deposited on	:	2021-08-19
Resolution	:	2.30 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 following versions of software and data (see references ⓘ) were used in the production of this report:

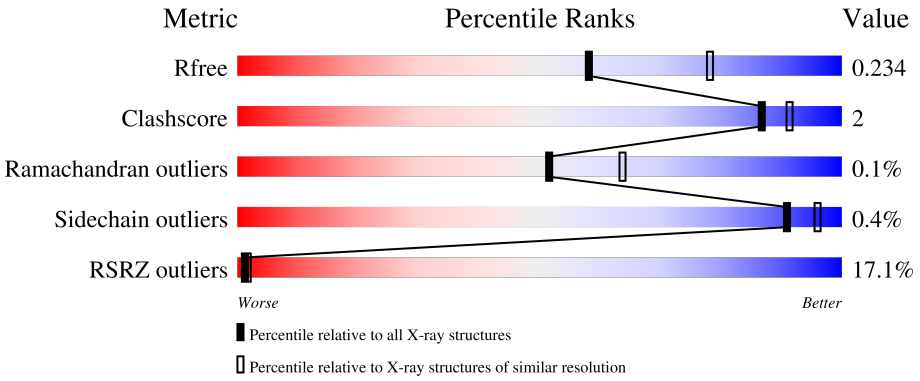
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.1

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 2.30 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	750	17% 91% 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
2	GOL	A	804	-	-	-	X

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6295 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.

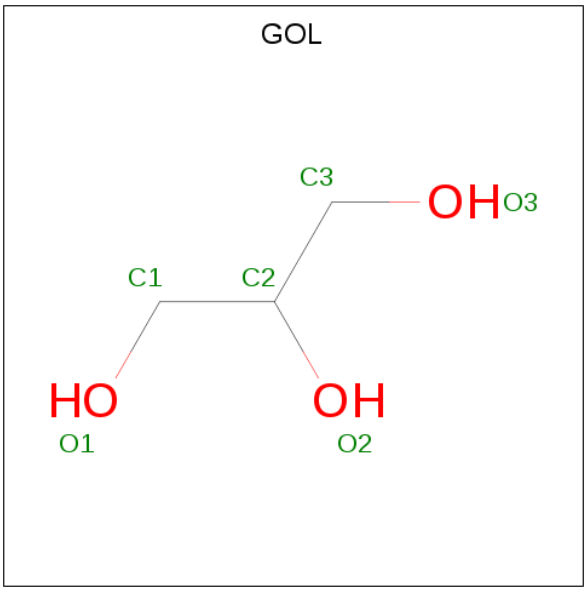
- Molecule 1 is a protein called Protein E2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	732	Total	C	N	O	S	0	3	0
			6016	3893	961	1124	38			

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	738	ALA	-	expression tag	UNP P21604
A	739	ALA	-	expression tag	UNP P21604
A	740	ALA	-	expression tag	UNP P21604
A	741	HIS	-	expression tag	UNP P21604
A	742	HIS	-	expression tag	UNP P21604
A	743	HIS	-	expression tag	UNP P21604
A	744	HIS	-	expression tag	UNP P21604
A	745	HIS	-	expression tag	UNP P21604
A	746	HIS	-	expression tag	UNP P21604
A	747	HIS	-	expression tag	UNP P21604
A	748	HIS	-	expression tag	UNP P21604
A	749	HIS	-	expression tag	UNP P21604
A	750	HIS	-	expression tag	UNP P21604

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 3 is water.

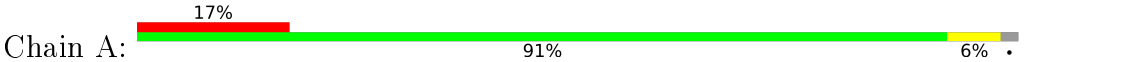
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	249	Total O 249 249	0	0

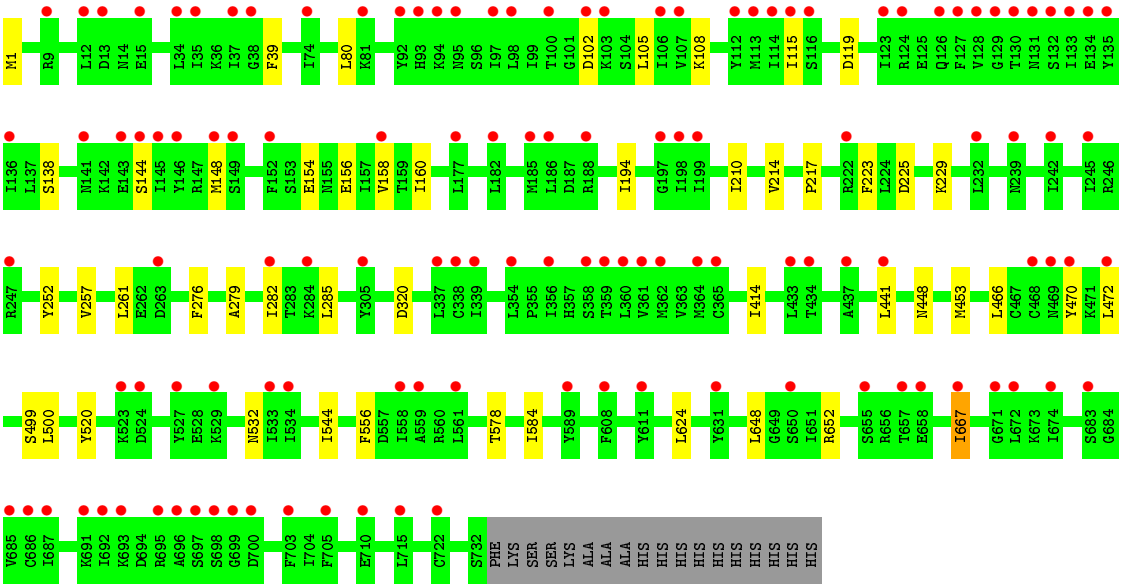
3 Residue-property plots (i)

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: Protein E2

Chain A:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	77.17Å 90.92Å 147.21Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.52 – 2.30 39.17 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.6 (35.52-2.30) 99.9 (39.17-2.30)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.03 (at 2.29Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.192 , 0.237 0.190 , 0.234	Depositor DCC
R_{free} test set	2402 reflections (5.14%)	wwPDB-VP
Wilson B-factor (Å^2)	66.7	Xtriage
Anisotropy	0.129	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 57.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	6295	wwPDB-VP
Average B, all atoms (Å^2)	75.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.47% 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: AME, GOL

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	0.41	0/6135	0.57	0/8313

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	6016	0	6104	26	0
2	A	30	0	40	1	0
3	A	249	0	0	1	0
All	All	6295	0	6144	26	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 2.

All (26) 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:A:194:ILE:HD12	1:A:223:PHE:HB2	1.72	0.70

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:138:SER:HA	1:A:160:ILE:HG12	1.75	0.66
1:A:225:ASP:OD1	1:A:229:LYS:NZ	2.38	0.57
1:A:556:PHE:HB2	1:A:667:ILE:HG23	1.86	0.57
1:A:282:ILE:HA	1:A:285:LEU:HD12	1.87	0.56
1:A:500:LEU:H	1:A:500:LEU:HD23	1.73	0.54
1:A:80:LEU:O	1:A:108:LYS:NZ	2.36	0.52
1:A:466:LEU:HB3	1:A:584:ILE:HG22	1.90	0.52
1:A:414:ILE:HD11	1:A:441:LEU:HD12	1.92	0.51
1:A:39:PHE:CE1	2:A:803:GOL:H12	2.46	0.50
1:A:144:SER:O	1:A:148:MET:HB2	2.12	0.49
1:A:257:VAL:HG22	1:A:261:LEU:HD12	1.95	0.49
1:A:448:ASN:HB3	1:A:453:MET:HG3	1.95	0.48
1:A:276:PHE:HB2	1:A:279:ALA:HB2	1.95	0.48
1:A:210:ILE:O	1:A:214:VAL:HG13	2.14	0.47
1:A:472:LEU:HD23	1:A:472:LEU:HA	1.83	0.44
1:A:115:ILE:HG23	1:A:119:ASP:HB2	2.00	0.44
1:A:102:ASP:HB3	1:A:105:LEU:HB3	1.99	0.43
1:A:156:GLU:O	1:A:160:ILE:HG13	2.18	0.43
1:A:154:GLU:O	1:A:158:VAL:HG23	2.18	0.43
1:A:544:ILE:HG23	3:A:1098:HOH:O	2.17	0.43
1:A:652:ARG:HD3	1:A:652:ARG:HA	1.72	0.43
1:A:499:SER:HA	1:A:532:ASN:O	2.21	0.41
1:A:648:LEU:HD23	1:A:648:LEU:HA	1.82	0.41
1:A:217:PRO:HA	1:A:252:TYR:CD2	2.56	0.41
1:A:578:THR:HG21	1:A:624:LEU:HD21	2.02	0.40

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	733/750 (98%)	718 (98%)	14 (2%)	1 (0%)	51 64

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	470	TYR

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	693/705 (98%)	690 (100%)	3 (0%)	91 96

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	320	ASP
1	A	520	TYR
1	A	667	ILE

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)

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	AME	A	1	1	9,10,11	1.50	1 (11%)	9,11,13	1.57	3 (33%)

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	AME	A	1	1	-	3/9/10/12	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1	AME	CA-N	-3.68	1.41	1.46

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1	AME	CA-N-CT1	-2.49	118.56	123.15
1	A	1	AME	CE-SD-CG	2.39	108.61	100.40
1	A	1	AME	CT2-CT1-N	2.11	119.67	116.10

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	1	AME	CA-CB-CG-SD
1	A	1	AME	C-CA-N-CT1
1	A	1	AME	CB-CG-SD-CE

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)

5 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	$\#\|Z\| > 2$
2	GOL	A	803	-	5,5,5	1.49	0	5,5,5	0.76	0
2	GOL	A	802	-	5,5,5	0.92	0	5,5,5	1.13	0
2	GOL	A	801	-	5,5,5	1.05	0	5,5,5	1.00	0
2	GOL	A	804	-	5,5,5	0.91	0	5,5,5	0.88	0
2	GOL	A	805	-	5,5,5	0.99	0	5,5,5	1.20	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
2	GOL	A	803	-	-	0/4/4/4	-
2	GOL	A	802	-	-	0/4/4/4	-
2	GOL	A	801	-	-	0/4/4/4	-
2	GOL	A	804	-	-	2/4/4/4	-
2	GOL	A	805	-	-	0/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	804	GOL	O1-C1-C2-C3
2	A	804	GOL	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	803	GOL	1	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, 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	731/750 (97%)	1.00	125 (17%) 1 2	49, 73, 108, 175	0

All (125) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	127	PHE	13.2
1	A	145	ILE	8.9
1	A	687	ILE	8.2
1	A	128	VAL	8.0
1	A	133	ILE	7.9
1	A	123	ILE	6.8
1	A	135	TYR	6.7
1	A	129	GLY	6.4
1	A	132	SER	6.1
1	A	693	LYS	5.5
1	A	696	ALA	5.1
1	A	107	VAL	4.9
1	A	144	SER	4.8
1	A	691	LYS	4.8
1	A	126	GLN	4.8
1	A	534	ILE	4.7
1	A	148	MET	4.6
1	A	685	VAL	4.6
1	A	146	TYR	4.5
1	A	106	ILE	4.3
1	A	95	ASN	4.2
1	A	437	ALA	4.1
1	A	589	TYR	4.1
1	A	199	ILE	4.0
1	A	699	GLY	4.0
1	A	130	THR	4.0
1	A	81	LYS	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	698	SER	3.9
1	A	158	VAL	3.8
1	A	686	CYS	3.8
1	A	527	TYR	3.7
1	A	222	ARG	3.6
1	A	115	ILE	3.6
1	A	141	ASN	3.6
1	A	186	LEU	3.5
1	A	103	LYS	3.4
1	A	114	ILE	3.4
1	A	672	LEU	3.4
1	A	338	CYS	3.4
1	A	284	LYS	3.3
1	A	657	THR	3.3
1	A	131	ASN	3.2
1	A	305	TYR	3.2
1	A	263	ASP	3.2
1	A	356	ILE	3.2
1	A	124	ARG	3.2
1	A	361	VAL	3.2
1	A	93	HIS	3.2
1	A	94	LYS	3.2
1	A	116	SER	3.1
1	A	434	THR	3.1
1	A	136	ILE	3.1
1	A	282	ILE	3.1
1	A	134	GLU	3.1
1	A	470	TYR	3.1
1	A	143	GLU	3.1
1	A	113	MET	3.0
1	A	472	LEU	3.0
1	A	703	PHE	3.0
1	A	695	ARG	3.0
1	A	12	LEU	2.9
1	A	34	LEU	2.9
1	A	177	LEU	2.9
1	A	364	MET	2.9
1	A	469	ASN	2.8
1	A	611	TYR	2.8
1	A	37	ILE	2.8
1	A	710	GLU	2.8
1	A	242	ILE	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	655	SER	2.8
1	A	658	GLU	2.8
1	A	15	GLU	2.7
1	A	149	SER	2.7
1	A	631	TYR	2.7
1	A	360	LEU	2.7
1	A	9	ARG	2.6
1	A	102	ASP	2.6
1	A	523	LYS	2.6
1	A	97	ILE	2.6
1	A	692	ILE	2.6
1	A	433	LEU	2.5
1	A	608	PHE	2.5
1	A	358	SER	2.5
1	A	232	LEU	2.5
1	A	667	ILE	2.5
1	A	700	ASP	2.5
1	A	98	LEU	2.4
1	A	182	LEU	2.4
1	A	671	GLY	2.4
1	A	705	PHE	2.4
1	A	112	TYR	2.4
1	A	354	LEU	2.4
1	A	650	SER	2.4
1	A	524	ASP	2.3
1	A	197	GLY	2.3
1	A	558	ILE	2.3
1	A	92	TYR	2.3
1	A	13	ASP	2.3
1	A	100	THR	2.3
1	A	38	GLY	2.3
1	A	245	ILE	2.3
1	A	529	LYS	2.3
1	A	74	ILE	2.3
1	A	339	ILE	2.3
1	A	359	THR	2.2
1	A	337	LEU	2.2
1	A	188	ARG	2.2
1	A	365	CYS	2.2
1	A	239	ASN	2.2
1	A	715	LEU	2.2
1	A	722	CYS	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	468	CYS	2.1
1	A	697	SER	2.1
1	A	561	LEU	2.1
1	A	198	ILE	2.1
1	A	35	ILE	2.1
1	A	559	ALA	2.1
1	A	683	SER	2.1
1	A	247	ARG	2.0
1	A	185	MET	2.0
1	A	152	PHE	2.0
1	A	674	ILE	2.0
1	A	441	LEU	2.0
1	A	533	ILE	2.0
1	A	362	MET	2.0

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, 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
1	AME	A	1	11/12	0.95	0.19	53,59,64,64	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
2	GOL	A	804	6/6	0.79	1.20	87,93,108,108	0
2	GOL	A	801	6/6	0.85	0.20	116,116,117,117	0
2	GOL	A	805	6/6	0.86	0.32	91,94,96,105	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	GOL	A	803	6/6	0.87	0.23	75,76,80,83	0
2	GOL	A	802	6/6	0.93	0.46	89,91,93,93	0

6.5 Other polymers (i)

There are no such residues in this entry.