



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 24, 2024 – 05:05 PM EDT

PDB ID : 7A40  
Title : Nucleotide-free OSM-3 kinesin motor domain  
Authors : Varela, F.P.; Menetrey, J.; Gigant, B.  
Deposited on : 2020-08-19  
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](mailto: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 : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.20.1  
EDS : 2.37.1  
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) : 2.37.1

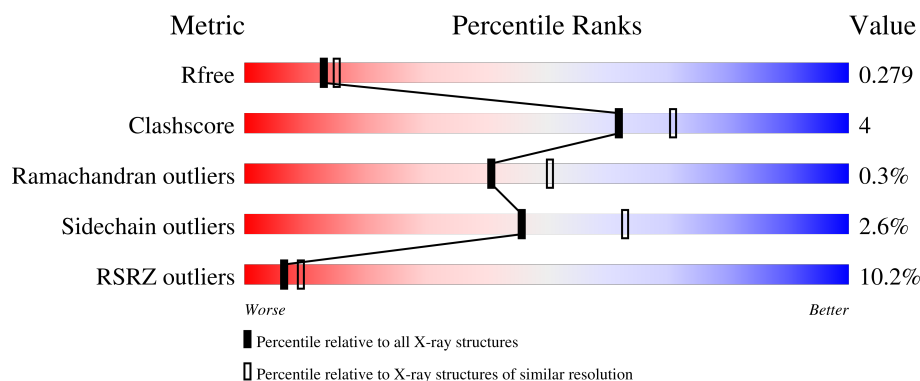
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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  $\geq 3$ , 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	346	<div> <div>10%</div> <div>79%</div> <div>11%</div> <div>10%</div> </div>
1	B	346	<div> <div>8%</div> <div>80%</div> <div>11%</div> <div>9%</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4633 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 Osmotic avoidance abnormal protein 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	313	Total	C	N	O	S	0	0	0
			2276	1413	390	461	12			
1	B	315	Total	C	N	O	S	0	0	0
			2270	1411	394	453	12			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP P46873
A	1	ALA	-	expression tag	UNP P46873
A	338	LEU	-	expression tag	UNP P46873
A	339	GLU	-	expression tag	UNP P46873
A	340	HIS	-	expression tag	UNP P46873
A	341	HIS	-	expression tag	UNP P46873
A	342	HIS	-	expression tag	UNP P46873
A	343	HIS	-	expression tag	UNP P46873
A	344	HIS	-	expression tag	UNP P46873
A	345	HIS	-	expression tag	UNP P46873
B	0	MET	-	initiating methionine	UNP P46873
B	1	ALA	-	expression tag	UNP P46873
B	338	LEU	-	expression tag	UNP P46873
B	339	GLU	-	expression tag	UNP P46873
B	340	HIS	-	expression tag	UNP P46873
B	341	HIS	-	expression tag	UNP P46873
B	342	HIS	-	expression tag	UNP P46873
B	343	HIS	-	expression tag	UNP P46873
B	344	HIS	-	expression tag	UNP P46873
B	345	HIS	-	expression tag	UNP P46873

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



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

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		

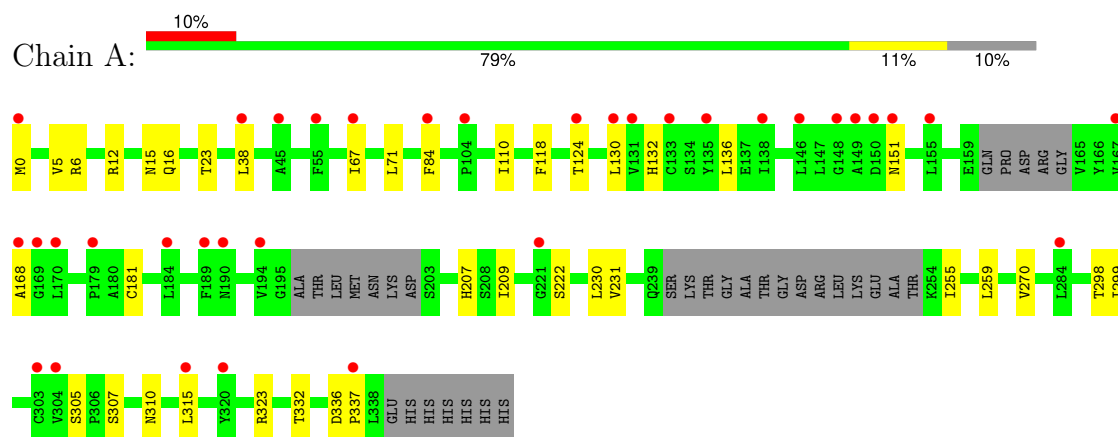
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	31	Total 31	O 31	0	0
4	B	34	Total 34	O 34	0	0

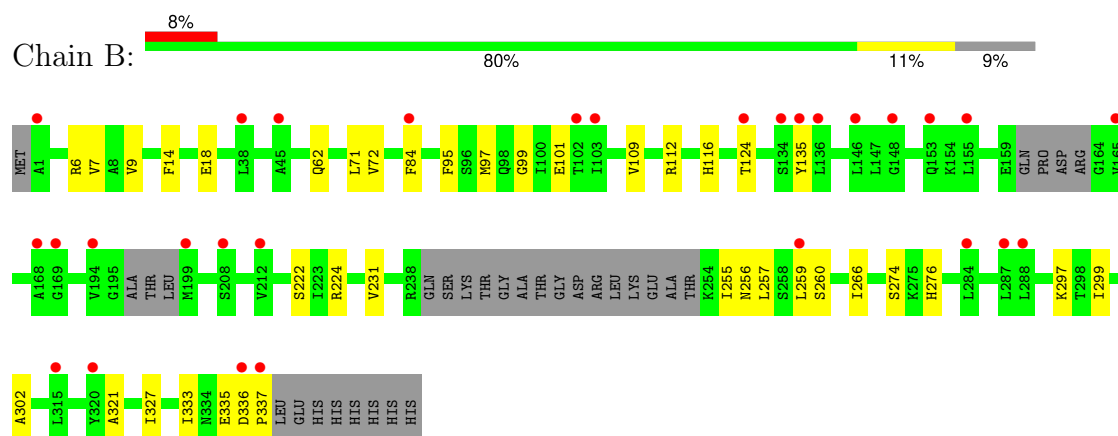
### 3 Residue-property plots [i](#)

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: Osmotic avoidance abnormal protein 3



- Molecule 1: Osmotic avoidance abnormal protein 3



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	42.77Å 155.18Å 63.72Å 90.00° 108.97° 90.00°	Depositor
Resolution (Å)	47.59 – 2.30 47.59 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.6 (47.59-2.30) 99.6 (47.59-2.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.24 (at 2.29Å)	Xtriage
Refinement program	BUSTER 2.10.3 (6-FEB-2020)	Depositor
R, $R_{free}$	0.225 , 0.254 0.244 , 0.279	Depositor DCC
$R_{free}$ test set	1740 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	57.2	Xtriage
Anisotropy	0.389	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 66.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.023 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	4633	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	86.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.76% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: SO4, 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	0/2310	0.59	0/3145
1	B	0.41	0/2304	0.60	0/3137
All	All	0.41	0/4614	0.60	0/6282

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	2276	0	2100	23	0
1	B	2270	0	2092	16	0
2	A	6	0	8	0	0
2	B	6	0	8	2	0
3	A	5	0	0	0	0
3	B	5	0	0	0	0
4	A	31	0	0	0	0
4	B	34	0	0	0	0
All	All	4633	0	4208	39	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 4.



All (39) 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:23:THR:HG21	1:A:307:SER:HB2	1.30	1.10
1:A:23:THR:CG2	1:A:307:SER:HB2	2.04	0.86
1:A:84:PHE:HB3	1:A:231:VAL:HB	1.65	0.78
1:A:0:MET:HA	1:A:332:THR:HG22	1.73	0.69
1:A:110:ILE:HG12	1:A:230:LEU:HD13	1.75	0.69
1:A:84:PHE:CB	1:A:231:VAL:HB	2.25	0.65
1:A:130:LEU:HD21	1:A:132:HIS:NE2	2.12	0.65
1:A:71:LEU:HD22	1:A:299:ILE:HD12	1.81	0.62
1:A:270:VAL:HG21	1:A:323:ARG:HG2	1.81	0.61
1:B:97:MET:HA	1:B:109:VAL:HG22	1.85	0.59
1:B:255:ILE:HG13	1:B:256:ASN:H	1.69	0.58
1:A:6:ARG:HD3	1:A:67:ILE:HD12	1.86	0.56
1:B:72:VAL:HB	1:B:116:HIS:CE1	2.40	0.56
1:B:71:LEU:HD22	1:B:299:ILE:HD12	1.89	0.54
1:A:38:LEU:HD23	1:A:315:LEU:HD13	1.92	0.51
1:B:6:ARG:HE	2:B:401:GOL:H31	1.76	0.51
1:B:257:LEU:HA	1:B:260:SER:HB2	1.94	0.50
1:A:207:HIS:HE1	1:A:259:LEU:HG	1.78	0.49
1:A:305:SER:OG	1:A:310:ASN:ND2	2.43	0.48
1:A:222:SER:HB3	1:A:337:PRO:HA	1.94	0.48
1:A:118:PHE:HE2	1:A:181:CYS:SG	2.38	0.47
1:B:84:PHE:CB	1:B:231:VAL:HB	2.45	0.47
1:B:95:PHE:O	1:B:99:GLY:HA2	2.15	0.47
1:B:6:ARG:HE	2:B:401:GOL:C3	2.28	0.46
1:A:12:ARG:HH12	1:A:15:ASN:ND2	2.14	0.45
1:A:136:LEU:HD11	1:A:209:ILE:HD12	1.98	0.45
1:A:130:LEU:CD2	1:A:132:HIS:NE2	2.79	0.44
1:A:6:ARG:HD2	1:A:67:ILE:HG23	2.01	0.43
1:A:12:ARG:NH1	1:A:15:ASN:HD22	2.16	0.42
1:A:15:ASN:O	1:A:16:GLN:C	2.56	0.42
1:A:336:ASP:HA	1:A:337:PRO:HD3	1.92	0.42
1:B:336:ASP:HA	1:B:337:PRO:HD3	1.86	0.42
1:B:14:PHE:HD1	1:B:18:GLU:HB3	1.84	0.42
1:B:333:ILE:HG22	1:B:335:GLU:HB2	2.02	0.42
1:B:266:ILE:HG23	1:B:327:ILE:HD11	2.01	0.42
1:B:7:VAL:HG11	1:B:321:ALA:HB1	2.02	0.42
1:B:71:LEU:HD11	1:B:297:LYS:HB2	2.01	0.41
1:B:9:VAL:HG22	1:B:302:ALA:HB3	2.03	0.41
1:A:5:VAL:HG22	1:A:298:THR:HB	2.04	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	305/346 (88%)	295 (97%)	9 (3%)	1 (0%)	41	50
1	B	307/346 (89%)	294 (96%)	12 (4%)	1 (0%)	41	50
All	All	612/692 (88%)	589 (96%)	21 (3%)	2 (0%)	41	50

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	222	SER
1	A	168	ALA

### 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	233/295 (79%)	230 (99%)	3 (1%)	69	82
1	B	228/295 (77%)	219 (96%)	9 (4%)	32	46
All	All	461/590 (78%)	449 (97%)	12 (3%)	46	63

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	124	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	151	ASN
1	A	255	ILE
1	B	62	GLN
1	B	101	GLU
1	B	112	ARG
1	B	124	THR
1	B	135	TYR
1	B	224	ARG
1	B	259	LEU
1	B	274	SER
1	B	276	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	15	ASN
1	A	310	ASN
1	B	310	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

4 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	401	-	5,5,5	0.10	0	5,5,5	0.21	0
3	SO4	B	402	-	4,4,4	0.40	0	6,6,6	0.61	0
2	GOL	B	401	-	5,5,5	0.17	0	5,5,5	0.36	0
3	SO4	A	402	-	4,4,4	0.56	0	6,6,6	0.59	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	401	-	-	0/4/4/4	-
2	GOL	B	401	-	-	0/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	401	GOL	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

### 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	313/346 (90%)	0.86	35 (11%) 5 7	44, 80, 140, 169	0
1	B	315/346 (91%)	0.73	29 (9%) 9 12	52, 83, 132, 152	0
All	All	628/692 (90%)	0.80	64 (10%) 6 9	44, 82, 138, 169	0

All (64) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	104	PRO	7.5
1	B	259	LEU	5.9
1	A	168	ALA	5.2
1	A	169	GLY	4.3
1	A	0	MET	4.0
1	A	155	LEU	4.0
1	A	130	LEU	3.8
1	B	45	ALA	3.7
1	B	337	PRO	3.6
1	A	315	LEU	3.6
1	B	146	LEU	3.6
1	B	38	LEU	3.5
1	A	189	PHE	3.4
1	B	135	TYR	3.4
1	A	67	ILE	3.4
1	B	287	LEU	3.3
1	A	138	ILE	3.3
1	A	320	TYR	3.2
1	A	148	GLY	3.2
1	B	169	GLY	3.2
1	B	199	MET	3.2
1	A	284	LEU	3.2
1	B	148	GLY	3.2
1	A	167	VAL	3.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	131	VAL	3.0
1	A	184	LEU	2.9
1	A	45	ALA	2.9
1	A	38	LEU	2.9
1	B	168	ALA	2.8
1	B	320	TYR	2.7
1	B	1	ALA	2.7
1	B	194	VAL	2.7
1	A	124	THR	2.7
1	A	337	PRO	2.6
1	B	155	LEU	2.6
1	B	315	LEU	2.5
1	B	288	LEU	2.5
1	A	221	GLY	2.5
1	A	133	CYS	2.5
1	A	151	ASN	2.4
1	B	124	THR	2.4
1	B	103	ILE	2.4
1	B	208	SER	2.4
1	B	134	SER	2.3
1	B	136	LEU	2.3
1	B	212	VAL	2.3
1	A	170	LEU	2.2
1	A	150	ASP	2.2
1	A	84	PHE	2.2
1	B	165	VAL	2.1
1	A	303	CYS	2.1
1	B	336	ASP	2.1
1	B	84	PHE	2.1
1	B	284	LEU	2.1
1	A	135	TYR	2.1
1	B	153	GLN	2.1
1	A	149	ALA	2.1
1	A	194	VAL	2.1
1	A	55	PHE	2.1
1	A	179	PRO	2.1
1	B	102	THR	2.0
1	A	304	VAL	2.0
1	A	190	ASN	2.0
1	A	146	LEU	2.0

## 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<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
2	GOL	A	401	6/6	0.74	0.29	89,90,90,90	0
2	GOL	B	401	6/6	0.80	0.14	73,73,74,74	0
3	SO4	B	402	5/5	0.94	0.20	66,66,66,67	0
3	SO4	A	402	5/5	0.97	0.13	59,59,60,60	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.